DISSECTING THE GENETICS OF HUMAN HIGH MYOPIA: A MOLECULAR BIOLOGIC APPROACH

BY *Terri L. Young, MD*

ABSTRACT

Purpose: Despite the plethora of experimental myopia animal studies that demonstrate biochemical factor changes in various eye tissues, and limited human studies utilizing pharmacologic agents to thwart axial elongation, we have little knowledge of the basic physiology that drives myopic development. Identifying the implicated genes for myopia susceptibility will provide a fundamental molecular understanding of how myopia occurs and may lead to directed physiologic (ie, pharmacologic, gene therapy) interventions. The purpose of this proposal is to describe the results of positional candidate gene screening of selected genes within the autosomal dominant high-grade myopia-2 locus (MYP2) on chromosome 18p11.31.

Methods: A physical map of a contracted MYP2 interval was compiled, and gene expression studies in ocular tissues using complementary DNA library screens, microarray matches, and reverse-transcription techniques aided in prioritizing gene selection for screening. The TGIF, EMLIN-2, MLCB, and CLUL1 genes were screened in DNA samples from unrelated controls and in high-myopia affected and unaffected family members from the original seven MYP2 pedigrees. All candidate genes were screened by direct base pair sequence analysis.

Results: Consistent segregation of a gene sequence alteration (polymorphism) with myopia was not demonstrated in any of the seven families. Novel single nucleotide polymorphisms were found.

Conclusion: The positional candidate genes TGIF, EMLIN-2, MLCB, and CLUL1 are not associated with MYP2-linked high-grade myopia. Base change polymorphisms discovered with base sequence screening of these genes were submitted to an Internet database. Other genes that also map within the interval are currently undergoing mutation screening.

Trans Am Ophthalmol Soc 2004;102:423-446

INTRODUCTION

The long-term objective of this research project is to uncover the molecular genetic basis of myopia. Myopia occurs when the focused image falls anterior to the retinal photoreceptor layer of the eye. Myopia is *the* most common human eye disease, and severe cases (high myopia greater than 5 diopters) may lead to blinding disorders such as premature cataracts, glaucoma, retinal detachment, and macular degeneration. Myopia can occur as an isolated finding or as a part of specific genetic syndromes. There is substantive evidence that genetic factors play a significant role in the development of nonsyndromic high myopia. We have identified multiple families with nonsyndromic high myopia and have mapped three autosomal dominant loci by linkage analysis. Myopia-2 locus (MYP2) is localized to chromosome 18p11.31, myopia-3 locus (MYP3) is localized to chromosome 12q23.1-q24, and we recently mapped another locus to chromosome 17q21-q22. *The overall goal is to positionally clone the genes responsible for high myopia.* Initial studies reviewed in this thesis have been directed at the identification of the MYP2 gene, as we have narrowed the recombinant interval within 18p11.31 to a 2.2 centimorgan (cM) region in which this gene is located. This report discusses initial findings of positional candidate gene base pair screenings for the MYP2 locus.

It is hypothesized that the identification of myopia disease genes such as the MYP2 gene will not only provide insight into the molecular basis of this significant eye disease, but will also identify pathways that are involved in eye growth and development. In addition, this information may implicate other genes as possible myopia disease gene candidates. This effort may lead to effective therapies for the severe forms of this potentially blinding eye disease.

Background and Significance

Public Health Significance

Myopia affects approximately 25% of the population of the United States[1-5] and is a significant public health prob-

From the Departments of Ophthalmology and Pediatrics, University of Pennsylvania, and the Children's Hospital of Philadelphia, Philadelphia, Pennsylvania. This work was supported by NEI grants EY014685 and EY00376 from the National Institutes of Health; Research to Prevent Blindness, Inc; the Macula Vision Research Foundation; the Mabel E. Leslie Research Endowment Funds; the Lions Eye Bank of Minnesota; and the Lions Eye Bank of Delaware Valley.

lem because it is associated with increased risk for visual loss.[1,6-10] Myopic chorioretinal degeneration is the fourth most frequent cause of blindness leading to registration for visual services and disability, and it accounted for 8.8% of all causes.[11] It has been estimated that 5.6% of blindness among school children in the United States is attributable to myopia.[11] Substantial resources are required for optical correction of myopia with spectacles, contact lenses, and, more recently, surgical procedures such as photorefractive keratectomy. The market for optical aids in the United States was estimated to exceed $8 billion in annual sales in 1990; most dollars were spent for the correction of myopia.[11,12] The development of methods for preventing the onset, or limiting the progression, of myopia would be of considerable importance.

Epidemiology and Clinical Characteristics of High Myopia

Prevalence Rates. High myopia (refractive spherical dioptric power of –5.00 or higher) is a major cause of legal blindness in many developed countries.[6,7,9,13-15] It affects 27% to 33% of all myopic eyes, corresponding to a prevalence of 1.7% to 2% in the general population of the United States.[1,5] High myopia is especially common in Asia.[13,14,16] In Japan, pathologic or high myopia reportedly affects 6% to 18% of the myopic population and 1% to 2% of the general population.[13] Comparative prevalence rates from different countries show considerable variability but confirm that myopia affects a significant proportion of the population in many countries.[2,9,13-16]

Progression of Myopia and Ocular Refractive Parameters. Juvenile-onset myopia most often develops and progresses between the ages of 10 and 16 years, whereas pathologic myopia usually begins to develop in the perinatal period and is associated with rapid refractive error myopic shifts before 10 to 12 years of age.[1,9,17,18] The key ocular parameters that determine refractive error are the refractive dioptric power of the cornea and lens, depth of the anterior chamber, and axial eye length (AEL). Several studies[1,19-22] have shown that the refractive status of an eye is determined primarily by AEL. The average refractive error at birth is approximately 1 to 2 diopters (D) of hyperopia, and the AEL measures approximately 17 mm. By adulthood, the AEL grows to about 24 mm. This results in little change in refractive error, because the radius of curvature of the cornea increases and the refractive power of the lens decreases.[1,21] Axial eye lengths of a myopic adult population may show a bimodal distribution with a second peak of increased AEL relating to high myopia (less than –6 D at 24 mm, greater than –6 D at 30 mm) when plotted as a distribution curve.[23] This suggests that myopia of –6 D or greater represents a deviation from the normal distribution of AEL and is not physiologic.

Ocular Morbidity. Many investigators have reported on the association of high myopia with cataract, glaucoma, retinal detachment, and posterior staphyloma with retinal degenerative changes.[1,8,24-39] High myopia is associated with progressive and excessive elongation of the globe, which may be accompanied by degenerative changes in the sclera, choroid, Bruch's membrane, retinal pigment epithelium, and neural retina. Various funduscopic changes within the posterior staphyloma develop in highly myopic eyes. These changes include geographic areas of atrophy of the retinal pigment epithelium and choroid, lacquer cracks in Bruch's membrane, subretinal hemorrhage, and choroidal neovascularization. Among these various fundus lesions, macular choroidal neovascularization is the most common vision-threatening complication of high myopia.[27,28,30-32,34] Clinical and histopathologic studies have documented choroidal neovascularization in 4% to 11% of highly myopic eyes. Relative to emmetropic eyes, an approximate twofold increased risk of choroidal neovascularization was estimated for eyes with 1 to 2 D of myopia, a fourfold increase with 3 to 4 D, and a ninefold increase with 5 to 6 D.[29,34,40] Poor visual outcome following choroidal neovascularization in myopic eyes is not uncommon and often affects relatively young patients.

The risk of retinal detachment is estimated to be three to seven times greater for persons with myopia greater than 5.0 D than for those with myopia of less than 5.0 D.[8,22,37] Myopia between 5.0 and 10.0 D was associated with a 15- to 35-fold greater risk of retinal detachment relative to that associated with low levels of hyperopia.[8, 22,37] The lifetime risk for retinal detachment was estimated to be 1.6% for patients with less than 3 D of myopia and 9.3% for those with more than 5 D.[7-38] A subgroup with lattice degeneration greater than 5 D of myopia had an estimated lifetime risk of 35.9%.[38] The prevalence of lattice degeneration increases with increasing levels of myopia as measured by AEL.[8,22,41,42]

Glaucoma was observed in 3% of patients with myopia who had AELs of less than 26.5 mm, in 11% with AELs between 26.5 and 33.5 mm, and in 28% of those with longer lengths.[39]

Role of Environment and Genetics in Myopic Development

Many studies report a positive correlation between parental myopia and myopia in their children, indicating a hereditary factor in myopia susceptibility.[43-46] Children with a family history of myopia had on average less hyperopia, deeper anterior chambers, and longer vitreous chambers even before becoming myopic. This implies a strong role for genetics in the initial shape and subsequent growth of the eye in myopia. Assessing the impact of genetic inheritance on myopic development may be

confounded by children adopting their parents' behavioral traits, such as higher-than-average near-work activities (eg, reading).[47]

In addition to genetics, moderate myopic development can be influenced by environmental factors. This is exemplified by experimental modulation of refractive error in the developing eyes of several animal models (mammalian and avian)[48-50] and the development of myopia in young children with media irregularities that prevent a focused retinal image.[51-53] Moreover, the prevalence of myopia in some populations appears to have increased dramatically from one generation to the next in increasingly industrialized settings, or with increased level of educational achievement.[54-58] Of course, this is an environmental effect. The identification of myopia genes may therefore provide insight into genetic-environmental interactions.

Consensus opinion regarding common, juvenile-onset myopia of moderate amounts is that its etiology is influenced by both genetic and environmental factors.[59] As a multifactorial, common, complex trait, genes or gene loci for this type of myopia have yet to be identified. Susceptibility loci contributing to common, juvenile-onset myopia may be difficult to map by classic linkage analysis because of the limited power to detect genes of intermediate or small effect using independent pedigrees.

There are multiple genetic syndromes with systemic findings that have myopia as a consistent clinical feature. For example, Stickler syndrome is an autosomal dominant connective tissue disorder characterized by ocular, orofacial, and skeletal abnormalities. Associated ocular manifestations include high myopia, glaucoma, cataracts, vitreoretinal degeneration, and retinal detachment.[60,61] Marfan syndrome is an autosomal dominant disorder with clinical features of myopia, lens dislocation, tall body habitus, and increased aortic wall distensibility.[62,63] Knobloch syndrome has an autosomal recessive high myopia presentation with vitreous degeneration and encephalocele.[64] Unlike these syndromes, however, the vast majority of individuals with myopia—moderate or severe—have no associated defects.

Determining the role of genetic factors in the development of nonsyndromic myopia has been hampered by the high prevalence, genetic heterogeneity, and clinical spectrum of this condition. The existence of a genetic contribution to any disease is based on evidence of familial aggregation and twin studies.[65] In the past, several modes of inheritance for myopia have been proposed, with no clear agreement among studies of pedigrees.[65-67] Goss and colleagues[68] reviewed a number of studies, some of which proposed an autosomal dominant mode of inheritance, others autosomal recessive, and still others, an X-linked pattern of inheritance for myopia. More recently, Naiglin and colleagues[69] performed segregation analysis on 32 French families with high myopia, and determined an autosomal dominant mode of inheritance. The λ_s for myopia (the increase in risk to siblings of a person with a disease compared to the population prevalence) has been estimated to be approximately 20 for siblings for high myopia, compared to approximately 1.5 for low myopia, suggesting a strong genetic basis for high myopia.[70]

Studies of twins provide the most compelling evidence that myopia is inherited. Multiple studies note an increased concordance of refractive error and refractive components (AEL, corneal curvature, lens power, anterior chamber depth) in monozygotic twins compared with dizygotic twins.[66,71-73] Twin studies estimate a notable heritability value (the proportion of the total phenotypic variance that is attributed to the genome) of between 0.5 and 0.96.[71,72]

Recent mapping studies provide the greatest promise for identifying implicated genes for high myopia. An X-linked recessive form of myopia was designated the first high-grade myopia-1 locus (MYP1) on chromosome Xq28.[74] The MYP1 locus eye disorder (Online Mendelian Inheritance in Man [OMIM] database [http://www.ncbi.hlm.nih.gov/omim] accession No. 310460) was named the Bornholm eye disease because the family studied was from Bornholm, Denmark. A large Minnesota family of Danish descent with X-linked-recessive myopia was genotyped. This family showed significant linkage to chromosome Xq27.3-q28. In collaboration with the researchers of the Bornholm eye disease, stored DNA from subjects of the original Bornholm eye disease study was obtained and comparative molecular genetic screenings were performed. It has been concluded that the myopia is associated with a novel cone dystrophy in both families.[75] Several medium to large multigenerational families with autosomal dominant high myopia of –5 D or greater have been genotyped, and significant linkage was found at chromosome 18p11.31 (MYP2 locus; OMIM No. 160700), chromosome 12q23.1-24 (MYP3 locus; OMIM No. 603221),[76,77] and most recently chromosome 17q21-22.[78] Additionally, the MYP2 locus has been confirmed by two outside laboratories.[79,80] A suggestive fourth locus for autosomal dominant high myopia was reported on chromosome 7q36.[81]

Candidate Gene Hypothesis

The sclera, the tough outer coat of the eye, is a typical connective tissue that provides the structural framework for the eye. As the sclera defines the shape of the eye, it is also likely to determine the AEL. The extracellular matrix of the sclera has been shown to contain collagen fibrils in close association with proteoglycans and glycoproteins.[82,83] Alterations in any of these extracellular matrix compo-

nents are likely to lead to changes in eye shape. Recent studies have shown that the scleral extracellular matrix undergoes significant changes during growth and aging[84] and is dramatically altered during the development of myopia.[84-86] The sclera of highly myopic human eyes differs considerably from normal sclera in both its physical dimensions and its biomechanical properties.[87] Many of the pathological changes seen in highly myopic human eyes are a consequence of gross scleral thinning, particularly at the posterior pole of the eye.[87,88]

Genes responsible for several syndromic genetic disorders with myopia as a consistent clinical finding have been identified: collagen 2A1 and 11A1 for Stickler syndromes type 1 and 2, respectively,[60,61] lysyl-protocollagen hydroxylase for type 4 Ehler-Danlos syndrome,[89] collagen 18A1 for Knobloch syndrome,[64] and fibrillin for Marfan syndrome.[62,63] Each of these genes is expressed in the sclera, demonstrating how knowledge of gene expression in the scleral wall is critical to our understanding of eye expansion and myopia.

It is the hypothesis of this thesis that the nonsyndromic high myopias result from distinct, but analogous, developmental defects of scleral wall growth control, and that their causative genes may be functionally or structurally related to one another and have parallel functions in the development of the visual axis. To this end, genes expressed by human sclera using both complementary DNA (cDNA) library and microarray techniques have been identified to aid in the selection of candidate genes for high myopia. In addition to the knowledge gained through direct studies of the MYP2 disease gene and its protein product, it is predicted that identification of the MYP2 disease gene will benefit genetic studies of all high-myopia genes.

Preliminary Studies

Several families with a heritable form of high myopia have been identified and ascertained. As described below, linkage analyses have led to the mapping of four high-myopia genetic loci, including MYP1, MYP2, and MYP3, and a chromosome 17q21-23 locus. A physical map was constructed with a bacterial artificial chromosome (BAC) contig (overlapping BACs that carry small DNA sequence segments for that region) spanning the contracted recombinant interval for the MYP2 locus. Multiple candidate genes have been identified within the MYP2 critical region and within the other mapped loci. Some genes have been provisionally excluded based on screening results. This study protocol was approved by the Children's Hospital of Philadelphia Institutional Review Board on Human Subjects research, and adhered to the tenets of the Declaration of Helsinki.

Clinical and Genetic Classification of Myopia Pedigrees

Myopia Study Population. A large myopia study population provides the foundation for the proposed research. The study population, summarized in Table 1, includes 622 individuals from 95 families with nonsyndromic high myopia and has been ascertained over the past 5 years through the identification of families from clinical practice, and in collaboration with other clinicians. Large families with (a) affected study participants in three or more generations or (b) affected study participants in two generations were initially sought with at least two participating offspring of affected individuals. Affection status was initially defined as a spherical refractive error of 6 D of myopia or more. This was solely based on historical consensus that the "pathologic" myopia criterion boundary began at that spherical dioptric power. Studies of myopia ocular morbidity, such as retinal detachment and choroidal neovascularization, suggest that 5 D of myopia or greater is an acceptable cutoff. Using this criterion led to the recent mapping of a novel locus on chromosome 17q21-22 of a large English/Canadian kindred.[78] There was justification in using –5.00 D as the cutoff because of the extreme level of myopic severity in most affected members (average myopia of –13.925 D, with a range from –5.50 to –50 D), with haplotype analysis confirmation. Five diopters of myopia is now used as the cutoff for affection status.

Molecular Genetic Studies of MYP2

Identification of the MYP2 Locus. Linkage analysis using seven of the families described above in the database demonstrated significant linkage to chromosome 18p11.31 with a maximum cumulative likelihood of the odds (LOD) score of 9.59 (an LOD score of 3 or greater is considered significant for linkage).[76] The 7.6 cM recombinant interval was defined distally by marker D18S59 and proximally by marker D18S1138, with recombinants in pedigrees 1, 4, and 5 (Table 2). The genetic boundaries of the MYP2 region are currently defined by linkage analysis of these seven existing MYP2 pedigrees. These seven pedigrees, in addition to any new pedigrees that may be identified, represent the group of MYP2 families to be screened for mutations at the MYP2 locus

In an effort to contract the MYP2 interval, transmission disequilibrium test (TDT) statistics[90] were obtained with the Statistical Analysis for Genetic Epidemiology Transmission Disequilibrium Test (SAGE-TDTEX)[91] and the GENEHUNTER2-Transmission Disequilibrium Test (GH2-TDT)[92] programs. TDT analysis was focused on 11 chromosome 18p polymorphic microsatellite DNA markers used for fine-mapping in the original study.[93] Both programs examine each allele separately to look for increased frequency of disequilibrium or nonrecombina-

TABLE 1. RESEARCH LABORATORY MYOPIA STUDY POPULATION DATABASE

PHENOTYPE (CHROMOSOMAL LOCATION)	NO. OF PEDIGREES	NO. OF PARTICIPATING INDIVIDUALS	NO. AFFECTED
MYP1 (Xq27-28)	2	28	9 (7 carriers)
MYP2 (18p11.31) (AD)	7	71	37
MYP3 (12q23-24)(AD)	1	22	14
MYP4 (17q21-22)(AD)	1	22	12
Autosomal dominant (other)	82	419	13
Autosomal recessive	2	60	22
Total	95	622	227

TABLE 2. MYP2 LOCUS MARKER RECOMBINANTS AND TRANSMISSION DISEQUILIBRIUM TEST (TDT) ALLELIC ASSOCIATION ANALYSIS*

MARKER DISTANCE (cM)	PEDIGREE MARKER	1	4	5	SAGE-TDTEX *P* VALUE	GH2-TDT *P* VALUE
	Telomere					
	D18S1140	+	–	–	0.036	0.083
0.1 <						
	D18S59	+	–	–	0.013	0.317
1.5 <						
	D18S476	–	–	–	0.007	0.045
0.1 <						
	D18S1146	–	–	–	0.227	0.083
4.5 <						
	D18S481	–	–	–	0.001	0.108
1.4 <						
	D18S63	–	–	–	0.062	0.034
0.1 <						
	D18S1138	–	–	+	$3.9 \times 10\text{-}4$	0.011
0.7 <						
	D18S52	–	–	+	$1.79 \times 10\text{-}6$	0.007
9.4 <						
	D18S62	–	+	+	0.141	0.479
18.6 <						
	D18S1150	+	+	+	0.018	0.096
4.1 <						
	D18S1116	+	+	+	0.214	0.683
	Centromere					

*TDT analysis was performed using two different statistical programs, Statistical Analysis for Genetic Epidemiology-TDT (SAGE-TDTEX) and GeneHunter 2 (GH2-TDT). The + sign denotes a recombinant marker; the - sign denotes a nonrecombinant marker. The dark shaded area highlights the region excluded by recombinants. The boxed lighter shaded area highlights the most significant marker associations with the myopia phenotype by TDT analysis.

tion events on disease-bearing chromosomes over normal chromosomes using a standard one-sided test (Fisher's exact test). The SAGE program also calculates a summary χ^2 for each marker, as it examines the degree of linkage disequilibrium at the marker. The significance values determined by both programs are listed in Table 2 for each marker locus in marker order for the chromosome 18p11.31 region. Markers D18S52 and D18S1138 show

the strongest statistical association with the disease phenotype. It is noted that one pedigree (No. 5) appears to exclude marker D18S52 from the initial haplotyped region. Possible explanations for this include marker order imprecision (despite cross-referencing of several databases as well as the use of CEPH [a commercial sample of standard DNA] family DNA for marker analysis); ethnic variation between families; possible phenocopy issue; possible second locus cosegregation at 18p11; and/or allele assignment error.

Critically important are the recent independent confirmations of the MYP2 locus with an Italian patient population with autosomal dominant high myopia by Heath and colleagues[79] and six families of Hong Kong Chinese descent by Lam and colleagues.[80] Their findings support directing further gene identification efforts to the centromeric region of the initial 7.6-cM recombinant interval. This combined data suggests that the MYP2 gene is likely within a 2.2-cM interval between D18S52 and D18S481. These results provide a basis for focused positional candidate gene analysis at the MYP2 locus, as the interval of interest has likely contracted significantly from the initial 7.6 cM.

Construction of a Physical Map Spanning the MYP2 Critical Region. By taking advantage of the multiple databases available in conjunction with the Human Genome Project, a physical BAC contig (overlapping BAC sequences which cover the interval) map across the MYP2 critical region was constructed, shown in Figure 1. Integration was obtained by mapping markers of different types (monomorphic, polymorphic, genes, and expressed sequence tags [ESTs]) from different public and private database sources (eg, National Center for Biotechnical Information [NCBI], Genethon, Whitehead Institute, University of California Santa Cruz [UCSC] "Golden Path," Celera).[94] The core region extends from marker D18S481 to D18S52. It ranges in depth from one to nine BACs, with an average depth of 4 BACs, and requires 19 overlapping BACs, averaging 150 to 200 kilobases (kb), to span the MYP2 region. The MYP2 critical region on the p arm is now almost a single contig; it contains one gap just centromeric to D18S481, which according to the Celera database is spanned by the myomesin (MYOM1) gene. Because of both the overlap and the gaps, it is difficult to estimate the true physical size of the region, but 1.2 megabases (Mb) would be the upper size limit. By adding up all sequence fragment lengths and estimating gap size, using the UCSC Web site, a 960-kb interval size is calculated. At this juncture, there are six known and 20 hypothetical genes that map within the interval. Other institutions, such as the genome centers at the Riken Institute in Yokahama City, Japan, the Whitehead Institute for Biomedical Research/MIT, Cambridge, Massachusetts, and the Genome Sequencing Center, Washington University in St Louis, are sequencing chromosome 18 BACs (http://www.ncbi.nlm.nih.gov/genome/seq/chr.cgi?CHR=18&SRT=ppos&MIN=0& ORG=Hs for integrated Web site information). BACs are placed on the map as they become available. Over half of the sequences, which are available from NCBI (http://www.ncbi.nlm.nih.gov/), are still phase 1 or "working draft" sequence.

Current Genes Within the MYP2 Critical Region. The direct analysis of sequence within a critical region can be the most accurate, precise, and efficient approach to disease gene identification. This is particularly true for instances where the "perfect" candidate gene (based on function or expression) does not exist within a defined critical region. It is also true for a disorder such as myopia, in which the temporal and spatial expression of the disease gene is not known and could be restricted to early development and to any eye component. All genes that map within the MYP2 critical region are candidate disease genes based on position. However, a gene triage bias was developed to prioritize study of extracellular matrix genes, because the high myopia phenotype uniformly involves scleral wall rearrangement with increased AEL. Moreover, a subset of affected patient participants (roughly 25%) have nonspecific connective tissue disorder–type findings, such as fallen arches, localized joint hypermobility, pectus excavatum, and nonsignificant heart murmur.

Initially, a possible role for the alpha subunit of laminin (LAMA) as a MYP2 candidate disease gene was discussed.[76] LAMA is a component of an extracellular matrix protein that binds microfibrils to collagen fibrils. LAMA was excluded, because it mapped outside of the initial 7.6 cM critical region by radiation hybrid analysis. Figure 2 shows reverse transcription–polymerase chain reaction (RT-PCR) transcripts for MYP2 candidate genes in eye and nonocular tissues. As described below in the "Methods" section, bidirectional sequencing of four additional genes was performed, transcription genes clusterin-like 1 (CLUL1),[95,96] elastin-microfibril located interface-protein (EMLIN-2),[97] and 5′-TG-3′ interacting factor homeobox protein (TGIF).[98-100] CLUL1 mapped within the larger 7.6 cM MYP2 region. It was not tested by RT-PCR, because it is known to be predominantly expressed in the retina. All other genes, except MLCB, show eye tissue expression by RT-PCR analysis using extracted RNA from human donor eyes in the laboratory (Table 3). TGIF is transcribed in two variant spliced isoforms, revealing an alternative transcript in retina, optic nerve, and brain by RT-PCR analysis. Interestingly, and providing greater urgency to screen this gene, TGIF has recently been implicated as the MYP2-causative gene by single nucleotide polymorphism (SNP) association studies, but

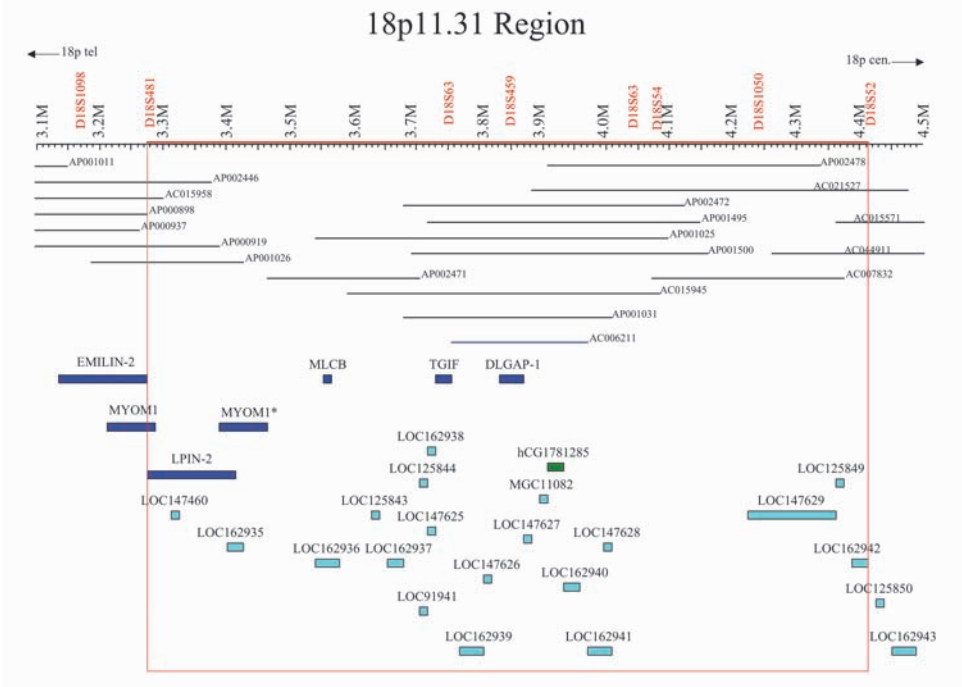


FIGURE 1

Physical map of the chromosome 18p11.31 critical region. The horizontal scale in megabases (M) is at the top. Polymorphic microsatellite DNA markers are labeled above the scale in red. Below the scale, finished (phase 3) bacterial artificial clones (BAC) clones are labeled in blue, working draft (phase 2 or 1) BAC clones are in black, known genes are in dark blue, *insilico* predicted genes by the public databases GENSCAN (http://genes.mit.edu/GENSCAN.htm) and OTTO* (http://cds.celera.com/biolib/info) are in light blue and green, respectively. The gene myomesin 1 (MYOM1) has been mapped to two different positions. In the NCBI database, MYOM1 maps distally with overlap just outside of the critical region. The Celera assembly (*) shows that MYOM1 spans the gap between BAC clones AP001024 and AP002471 in a more centromeric position.[94]

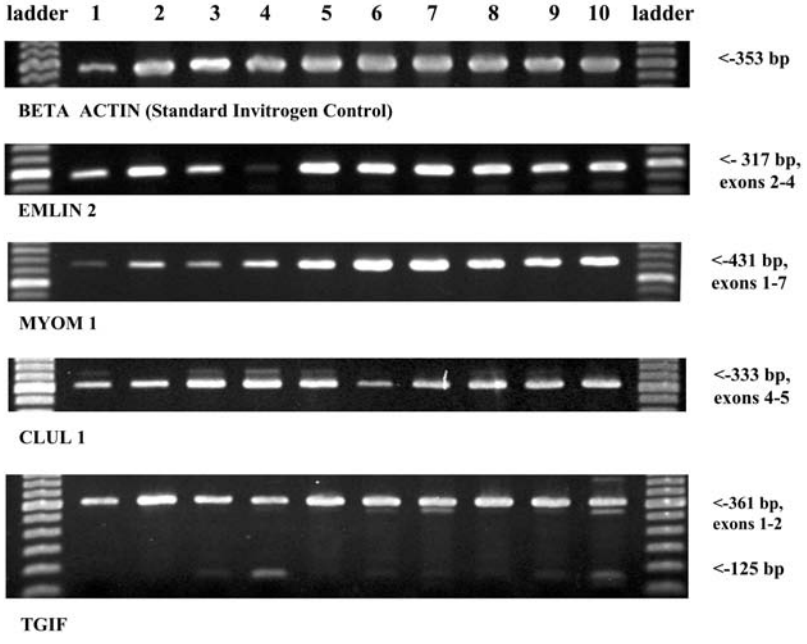


FIGURE 2

Polymerase chain reaction amplicons of MYP2 candidate gene complementary DNA (cDNA) from reverse-transcribed RNA from human ocular tissues and commercially available RNA from various human tissue types (Ambion). 1-sclera, 2-cornea, 3-optic nerve, 4-retina, 5-lung, 6-skeletal muscle, 7-heart, 8-trachea, 9-kidney, 10-brain. Expected amplicon size based on primer selection encompassing exonic sequence is shown. Note that TGIF transcripts show two variant spliced isoforms. The commercial DNA ladder has standard-sized DNA molecular weights spaced at 100 base pairs apart and is used to determine the approximate molecular weights of the test amplicons.

not by DNA sequencing.[101] TGIF mutations are associated with holoprosencephaly.[98-100]

Molecular Genetic Studies of MYP3

Significant linkage of autosomal dominant high myopia to a second locus on chromosome 12q21-q23 in a large German/Italian family was determined recently.[77] The average age at diagnosis of myopia was 5.9 years (range, 4 to 8). The average spherical component refractive error for the affected individuals was –9.47 D (range, –6.25 to –15.00 D). A representative AEL of 30.06 mm was noted in individual 5. Corneal curvature was normal. The maximum LOD scores with pairwise linkage analysis were 3.85 for markers D12S1706 and D12S327. Recombination events identified markers D12S1684 and D12S1605 as flanking markers, defining a 30.1-cM interval on chromosome 12q21-23. The coding sequences of the chromosome 12q21-23 recombinant interval candidate proteoglycans lumican, decorin, and dermatan sulfate proteoglycan-type 3 were screened for mutations first by heteroduplex analysis, and then by bidirectional sequencing. No mutations were found.[102]

Molecular Genetic Studies of a Newly Identified Locus on Chromosome 17q21-q22

One large multigenerational English/Canadian family with autosomal dominant high myopia (family MYO-68) was ascertained and genotyped (Figure 3).[78] The average age at diagnosis of myopia for affected individuals was 8.9 years (range, 2 to 11), and the average spherical component refractive error for the affected individuals was –13.925 D (range, –5.50 to –50.00 D). The representative average AEL of 35.28 mm, measured only for affected individuals 9 and 10, was significantly longer compared to adult normal values. All candidate gene loci were excluded. After a genome screen and fine-point mapping, a maximum pairwise LOD score of 3.17 with marker D17S1604 was obtained. Haplotype analysis revealed recombinant events that narrowed the critical region containing the gene to 7.71 cM, between markers D17S787 and D17S1811.

The proband (individual 9) has the highest documented level of myopic refractive error in our clinical experience, varying between –50 D and –60 D. Despite the severe myopia most affected members exhibited, there were two carriers of the putative disease haplotype, with high but less severe myopia (–4.50 to –5.50 D)—individuals 12 and 20—reflecting variability in the phenotype and possible modifying factors (individual 12 was classified as unaffected). The phenotypic variability and somewhat arbitrary assignment of affection status underscore the difficulty in mapping analyses when applying a dichotomous phenotype model to a quantitative trait. The extracellular matrix proteins collagen 1A1 (COL1A1) and proteoglycan chondroadherin (CHAD) were screened as the most promising candidate genes.[103,104] Both CHAD and COL1A1 are expressed in human sclera by RT-PCR. Sequencing of the coding regions of both genes revealed no disease-causing mutations.

TABLE 3. LIST OF KNOWN GENES THAT MAP WITHIN THE MYP2 CANDIDATE INTERVAL AS DEPICTED IN THE PHYSICAL MAP*

GENE	SYMBOL	OMIM	ACCESSION NO.	TRANSCRIPT SIZE (NO. OF EXONS)	EYE EXPRESSION
Elastin microfibril interphase located protein	EMLIN-2		NM_032048	4009 bp (8)	Yes (S,C,ON,R)
Myomesin 1	MYOM1	603508	NM_003803	4949 bp (20)	Yes (S,C,ON,R)
Lipodystrophy nuclear protein	LPIN-2	605519	NM_014646	6221 bp (20)	Yes (S,C,ON,R)
Myosin, light polypeptide, regulatory	MLCB		NM_006471	944 bp (4)	No
TG interacting factor (TALE family homeobox)	TGIF	602630	NM_003244	2214 bp (2)	Yes: 125bp-(R, ON); 361bp- (S,C,ON, R)
Disks, large (*Drosophila*) homolog- associated	DLGAP1	605445	NM_004746	1591 bp (4)	Not studied

S, sclera; C, cornea; ON, optic nerve; R, retina.

*These genes will be screened first for sequence mutations. The gene name and symbol, the reference number from the Online Mendelian Inheritance of Man (OMIM) database, the standard database gene accession number, the gene size in base pairs (bp) and number of exons (coding sequence), and the expression in eye tissue type by reverse transcription-polymerase chain reaction of human donor eye tissue in our laboratory is presented. The Ensembl database (http://www.ensembl.org/Homo_sapiens/geneview) was accessed to obtain gene transcript information and accession numbers. Ocular tissues tested for gene expression include human sclera, cornea, optic nerve, and retina.

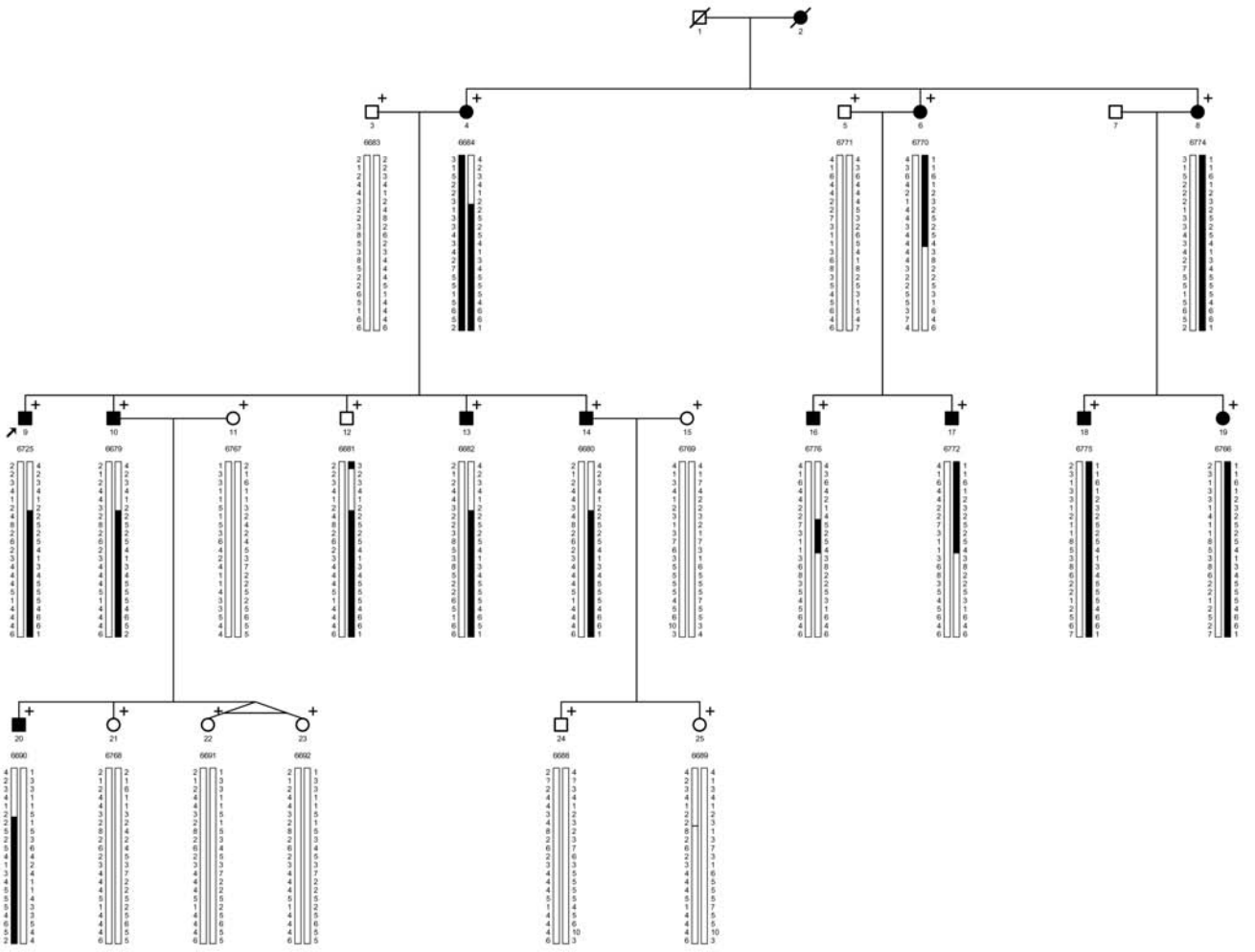


FIGURE 3
Pedigree and haplotype of family with significant linkage to a 7.71-cM interval on chromosome 17q21-q22 after a genome screen.[78]

Molecular Genetic Studies of MYP1

The Bornholm eye disease consists of X-linked high myopia, high cylinder, optic nerve hypoplasia, reduced electroretinographic flicker with abnormal photopic responses, and deuteranopia. The disease was the first designated high-grade myopia locus (MYP1). The Bornholm eye disease family phenotype had a provisional assignment to the distal part of the X chromosome at Xq28 using a limited number of markers.[74] We studied a family from Minnesota with a similar X-linked phenotype, also of Danish descent.[75] All affected males had protanopia instead of deuteranopia, high myopia, abnormal photopic electroretinographic responses, and peripapillary temporal conus. The families originated from neighboring Denmark villages; thus we explored the degree of identity in genotype and haplotype. DNA from six members of the original Bornholm eye disease family was obtained to compare haplotypes and to determine if mutations in the red or green cone pigment genes could be responsible for color vision defects and cone dysfunction. Of the 22 individuals of the Minnesota family, eight affected males and five carrier females were studied. Significant maximum LOD scores of 3.38 and 3.11 were obtained with microsatellite markers DXS8106 and DXYS154, respectively. Haplotype analysis defined a 37.8-cM interval at chromosome Xq27.3-Xq28 in the Minnesota family. Analysis of the Bornholm eye disease DNA reduced the interval to 15.3 cM. Affected Minnesota males had a red-green hybrid cone pigment gene in the first position of the cone pigment gene array consistent with a protan defect by directing sequencing and single-strand conformation polymorphism analysis. The Bornholm eye disease subjects had a first-position green-red hybrid cone pigment gene array sequence, consistent with a deutan defect. Cytogenetic and Southern blotting analysis showed no deletion abnormalities. Both families appear to have a novel form of cone dysfunction with associated myopia, and not simple high myopia. The hybrid gene mutations cosegregate with the cone dystrophy, but are common and not associated with high myopia.

Human Scleral Gene Expression Experiments

Several gene products show varied levels of expression in the sclera during the development of experimentally induced myopia in animal models.[105-108] The synthesis and accumulation of the scleral proteoglycans decorin, biglycan, and aggrecan, the matrix metalloproteinase, gelatinase A, inhibitors of metalloproteinases TIMP-1 and TIMP-2, and collagen have all shown alterations in sclera during form deprivation myopia development in chicks, tree shrews, and primates. Moreover, the expression levels of various proteoglycans, gelatinase A, and TIMP-2 normalize following restoration of normal (unrestricted) vision, indicating a direct relationship between the expression of these extracellular matrix constituents and the rate of ocular elongation. These are considered candidate genes for myopia. It is unclear, however, how experimental myopia relates to human physiological myopia and how human differs from animal sclera.

Human Scleral Complementary DNA Library. To identify positional candidate genes for human scleral disorders that may include myopia, we constructed a human scleral cDNA library from which ESTs (smaller segments of DNA known to be expressed in cellular tissues because the DNA sequence was derived from messenger RNA sequence) were generated after single-pass sequencing and characterizing randomly isolated clones.[109] The cDNA library was constructed from RNA isolated from sclera of human donor eyes, with known nonmyopic refractive history. Human donor eyes were obtained from the Lions Eye Bank of Minnesota and treated by submersion in RNA*later* Solution (Ambion, Austin, Texas) within 12 hours postmortem. Extracted RNA was submitted to Stratagene for commercial preparation of a pCMV-PCR vector cDNA library. We examined the insert sequences for similarities to genes and ESTs using the GenBank database of expressed genes, http://www.ncbi.nlm.gov/GenBank/. This was accomplished using the "basic local alignment search tool" program (BLASTN), available through the National Center for Biotechnology Information, Bethesda, Maryland, http://www.ncbi.nlm.gov/BLAST/, /GenBank/, and /dbEST/.

DNA sequences were obtained from 609 clones. With noted redundancy, 337 scleral EST sequences matched 228 known human genes. Four scleral ESTs showed sequence homology to four nonhuman genes. Of the remaining 268 scleral ESTs, 252 showed significant homology to ESTs from other cDNA libraries in GenBank. Sixteen transcripts did not match any sequences in GenBank (nonredundant database, human and mouse EST databases, mitochondrial database) and are possibly novel genes. The EST sequences were submitted to GenBank. The most abundant connective tissue–related genes were α-A crystalline (CRYAA), Xα-1 collagen, and β-5 integrin. Other extracellular matrix gene matches were biglycan, syndecan, decorin, fibromodulin, proline arginine-rich and leucine-rich repeat protein, transgelin, TIMP-1, and fibulin 1. Human scleral expression of all but decorin and biglycan has not previously been reported.

Five genes mapped to high myopia loci. The genes identified were biglycan (BGN), which maps to the MYP1 locus at chromosome Xq28; myosin regulatory light chain 2 (MYL2) and decorin, which map to the MYP3 locus; and keratin 13 (KRT 13) and transducer 1 of avian erythroblastic leukemia viral oncogene homolog 2 (ERBB2) (TOB 1), which map to the chromosome 17q21-22 locus. No gene matches were found for the MYP2 locus at chro-

TABLE 4. AFFYMETRIX OLIGONUCLEOTIDE HYBRIDIZATION RESULTS OF SIX SAMPLES OF HUMAN SCLERAL RNA WITH INDEPENDENT "PRESENT" CALLS FOR EIGHT GENES THAT MAP TO CHROMOSOME 18p11.31*

AFFYMETRIX PROBE NO.	CYTOGENETIC LOCUS	UNIGENE NO.	GENE NAME	GENE SYMBOL
36628_at	18p11.3	Hs.75447	ralA binding protein 1	RALBP1
34893_at	18p11.31-p11.2	Hs.51299	NADH dehydrogenase (ubiquinone) flavoprotein 2	NDUFV2
35137_at	18p11.31-p11.32	Hs.2504	myomesin 1 (skelemin)	MYOM1
1674_at	18p11.31-p11.21	Hs.194148	v-yes-1 Yamaguchi sarcoma viral oncogene homolog 1	YES1
33371_s_at	18p11.3	Hs.223025	member RAS oncogene family	RAB31
38805_at	18p11.3	Hs.90077	TGFB-induced factor (TALE family homeobox)	TGIF
41187_at	18p11.31	Hs.180224	myosin regulatory light chain	MLCB
37893_at	18p11.3-p11.2	Hs.82829	protein tyrosine phosphatase, nonreceptor type 2	PTPN2

*Detection *P* values for all data points were <.005. The Affymetrix probe No. is the position on the Affymetix chip that the gene occupies, cytogenetic locus is the band location on chromosome 18p11, and the Unigene No. is the standard identification number of the gene in the Unigene database (www.ncbi.nlm.nih.gov/UniGene/Hs.Home.html).

mosome 18p11.31 or the chromosome 7q36 locus. Eleven EST-BAC matches mapped specifically to the MYP3 locus at chromosome 12q23-24.

Human Scleral Microarray Analysis. We performed preliminary scleral RNA microarray absolute expression studies as a second strategy to identify myopia candidate genes.[110] Purified poly (A) mRNA was isolated from six donor sources of human sclerae, and reverse transcribed into cDNA. An in vitro transcription reaction was performed to produce biotin-labeled cRNA from the cDNA. Each cRNA was fragmented and mixed in a cocktail with probe array controls before a 16-hour incubation/hybridization to oligonucleotide probes (representing 12,625 human genes) on six Affymetrix U95A chips. The chips were scanned using GeneChip software. Array analyses were carried out with Microarray Suite, version 5.0 (Affymetrix), using the expression analysis algorithm to run an absolute analysis after cell intensities were computed. All arrays were normalized to the same target intensity using all probe sets. There were 3,789 genes with "present" calls assigned independently to all six human scleral samples. Eight genes mapped to chromosome 18p11.31 (Table 4), three of which are placed on our physical map—myomesin 1 (MYOM1), TGFB-induced factor (TALE family homeobox) (TGIF), and myosin regulatory light chain (MLCB).

These studies have identified expressed human scleral proteins that may be important in the maintenance of biochemical and biomechanical properties of the sclera. The collection of genes expressed in these studies is not indicative of all genes expressed in the sclera. Indeed, expression of genes such as collagen type I and elastin—known constituents of the sclera—was not identified in our library screening procedure. This most likely reflects the incomplete screening for all expressed genes in a cDNA library, reverse transcription bias, differences in the developmental expression of various transcripts, and the limitations of a predesigned chip with a finite number of genes (eg, candidate genes EMLIN 2, LIPIN 2, and DLGAP are all on version 2 of the U95A chip). ("The absence of proof is not the proof of absence.") The expression studies by these two methods are, to our knowledge, the first effort to establish a comprehensive list of genes involved in human scleral composition and physiology and are potentially useful for directed myopia candidate gene screening.

RESEARCH DESIGN AND METHODS

Hypothesis and Rationale

The hypothesis is that the MYP2 gene is most likely a gene involved in scleral formation or regulation. We have provisionally narrowed the candidate interval on chromosome 18p11.31 to 2.2 cM (1 to 1.2 Mb). We continually review the genes known to be located in the region and have selected the most biologically relevant gene(s) for further analyses. We have screened for gene mutations in the MYP2 families, looking for mutations that segregate with affected subjects only. This positional candidate approach has proven feasible due to the Human Genome Project. It relies on the identification of a critical region for a disease (by the presence of a translocation, deletions, or by linkage analysis) and a search of genes known to map within or near this region. It has simplified the search efforts for the disease gene in a number of disorders such as X-linked familial exudative vitreoretinopathy,[111] spastic paraplegia,[112] and congenital fibrosis of the extraocular muscle type 2.[113] To our knowledge, there are no reports of cytogenetic breakpoints due to deletions or translocations associated with high myopia as the sole phenotype. Therefore, mapping studies establishing linkage are the only avenue to initiate positional candidate screening. The flow chart (Figure 4) summarizes the steps necessary to identify the bona fide MYP2 transcript within a recombinant interval.

Methods to Identify Transcripts Within the 2.2-cM Recombinant Interval Using Available Databases

The genomic and EST/cDNA annotations at the MYP2 locus were analyzed using various databases in combination such as UniGene (http://www.ncbi.nlm.nih.gov/UniGene/), GeneMap '99 (http://www.ncbi.nlm.nih.gov/genemap99/), UCSC Human Genome Project Working Draft (University of California at Santa Cruz; http://genome.ucsc.edu), Ensembl (http://www.ensembl.org), Celera, and eGenome. This keeps us abreast of new transcripts as they are mapped to the MYP2 region, new sequence that provides the 3′ or 5′ extension of known transcripts, any new BACs that map to the region, and the continual refinement of the BAC sequences and fragment order. This effort has been significantly enhanced by the construction of a map of the whole human genome to enable the selection of clones for sequencing and for the accurate assembly of the genome sequence by the International Human Genome Mapping Consortium.[114] It is important to take advantage of the continuously and rapidly updated genomic and cDNA sequence as part of the Human Genome Project, and of the power of annotation programs to identify the transcripts within the MYP2 region.

The order of transcript analysis was based on a balance between the availability of a transcript's cDNA and genomic sequence and the strength of its functional and structural characteristics as a candidate for MYP2. Clearly, the more that is known or can be deduced about a transcript's corresponding cDNA and genomic structure, the greater the saving of time and resources. Therefore, based on information from publications and database analysis using OMIM, eGenome, and NCBI,

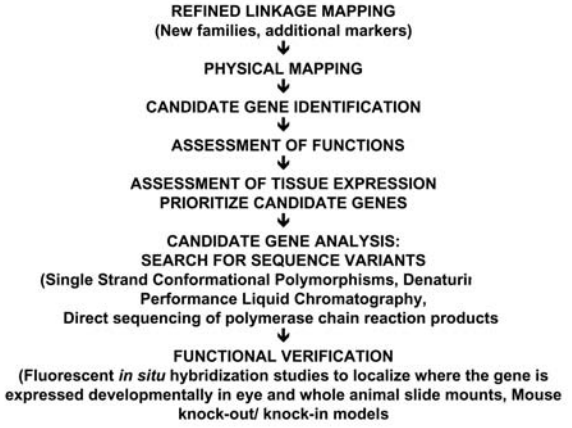


FIGURE 4
Flow chart of standard methods for candidate gene mutation identification and functional verification.

each MYP2 transcript is ranked with regard to available sequence as follows: (1) known cDNA and defined or inferred genomic structure; (2) known cDNA but inadequate genomic sequence to infer genomic structure; (3) partial cDNA or EST cluster with more than one genomic exon in the region; (4) EST cluster that maps to a single genomic location; (5) a predicted gene based on homology to known proteins; or (6) a predicted gene based on computational programs such as GENSCAN and GENIE. In addition, refinement of the MYP2 region genomic sequence will result in the continuous upgrading of a transcript's category.

Additional candidates have been added as they are mapped to the interval or identified from genomic sequence. Potential open reading frames (ORFs) will be identified using exon prediction programs (OTTO, GENSCAN, NCBI's ORF FINDER and BLAST, Metagene [http://rqd.mcw.edu/METAGENE/]). The ORF sequences will be compared to the EST, UniGene, and other databases using BLAST or FASTA, to identify known transcripts and assist in recognition of additional exons of the ORF. Previously unmapped transcripts will be added to the list of candidates and analyzed for parology to other human genes or coding sequences as well as orthology to coding sequences from other organisms, especially those with complete or emerging sequences (eg, mouse, rat, *Drosophila*, zebra fish).

There are 20 predicted genes within the interval that are partial cDNA/EST clusters represented by two or more hits within the same region of genomic sequence. Such a pattern of hits suggests the presence of multiple exons with intervening intronic sequence and classifies these transcripts as likely positional candidates. Most have start codons, and from one to 14 exons. There are also nine hypothetical genes with protein homologies to human or other species, indicating possible duplications to known genes in the interval, or alternative splice variants. There are 11 genes for which there are no corresponding cDNA sequences present in the databases. The strategy is to simultaneously use four different approaches to validate a gene: ab initio (looking at the sequence itself using algorithms such as GenScan), homology/similarity (such as BLAST), evidence of expression (such as alignment with ESTs), and comparative genomics (comparing homologies on the genomic level with genome sequence of mouse, rat, zebra fish, etc). Only genes verified by the majority of gene prediction programs, by the identification of promoters, CpG islands, poly A signals, by correlation with ESTS, and by orthology comparisons will be screened, first to confirm expression by RT-PCR in the eye and other tissues, and then by sequencing.

Alternative splicing posed another challenge for mutation screening,[115] as was shown with the candidate gene TGIF in Figure 3. There are two forms of syndromic high myopia, the Wagner and Knobloch syndromes, with implicated isoform mutations of collagens 2A1 and 18A1, respectively.[116,64] The strategy is to identify ESTs that come from the same gene and look for differences between them that are consistent with alternative splicing, such as a large insertion or deletion in one EST. Each splice form was further assessed by aligning the ESTs exactly to their gene sequence in the draft genome. This reveals candidate exons separated by candidate splices. As intronic sequences at splice junctions are highly conserved (99.24% of introns have a GT-AG at their 5′ and 3′ ends, respectively), they can be used to verify candidate splices. One way to determine if a particular splice form is worth further investigation is to note it in multiple ESTs from different libraries, which suggests that it is unlikely to be a low-frequency error product. Using large-scale EST analysis, it has been determined that the amount of alternative splicing is comparable between humans and other animals.[117] Cross-comparisons of splice variants in humans with other species also enhances validity, and we will make such comparisons with nonhuman genomes. There are a variety of alternative splicing databases for use (ASDB http://cbcg.nersc.gov/asdb, AsMAMDB http://166.111.30.65/ASMAMDB.html, ISIS http://isis.bit.uq.edu.au/). Alternative transcript expression in human eye tissues by RT-PCR relative to other organ tissue types can be confirmed.

Positional Candidate Genes Were Triaged Based on Existing Functional and Tissue Expression Data

To glean a given transcript's functional characteristics, a database and literature review of its sequence was performed to learn about DNA identity and similarity, protein sequence identity and similarity, protein structural identity and similarity, protein interactions, and protein domain identification. Transcripts were also ranked from

highest to lowest priority based on their pattern of expression. Transcripts whose protein products play a functional role in eye development and structure will be analyzed with higher priority. However, because it is surprising how many diseases appear to be caused by defects in genes that are ubiquitously expressed, prioritizing candidate genes will be done cautiously. The MYP2 gene is ultimately not required to have expression restricted to affected tissues. Expression data can be collected from sources such as the NEI Bank and UniGene Web sites (www.ncbi.nlm.nih.gov/UniGene/Hs.Home.html), as well as from the human eye tissue RT-PCR, our human scleral and other eye tissue cDNA libraries, and microarray analyses. Additionally, the MYP2 disease gene may be developmentally regulated or expressed only upon induction in specific tissues, and therefore the candidate gene may not have been shown to be expressed in adult or immature tissues, or might be missed in our adult human eye tissue RT-PCR experiments.

As discussed in the preliminary data section, there is a bias toward analyzing genes with an extracellular matrix–associated function. Obvious structural genes that map to the candidate region will be screened, such as collagens or proteoglycans. Genes found to be important for constituent organization and maintenance of connective tissue function will also be given priority. For example, a strong candidate, EMLIN-2 (elastin microfibril located interface protein), is an elastic fiber–associated glycoprotein found at the interface between amorphous elastin and microfibrils and regulates elastic fiber formation.[97] It is expressed in eye tissues based on the RT-PCR studies (see Figure 2). The exonic structure and designed PCR primers that overlap intron-exon sequence for EMLIN-2 screening are provided in Figure 5. Transforming growth β-inducing factor (TGIF) is also a strong candidate, because it showed ocular expression in both the microarray analysis of human sclera and by RT-PCR studies of eye tissue. These genes were the first to be screened.

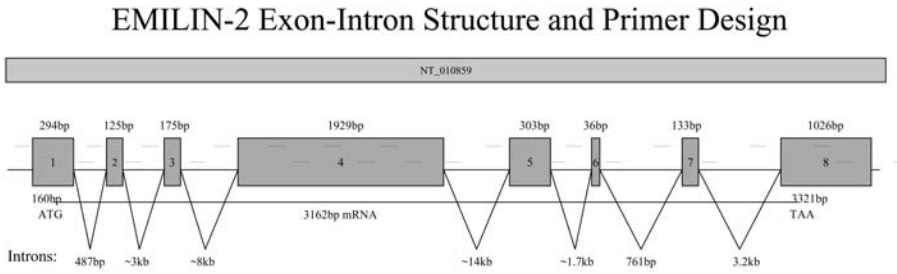
The candidate gene may also be expressed in the retina and influence scleral growth.[118] This retinal hypothesis emanates mainly from animal studies of experimental myopia. The induction of myopia in juvenile animals by deprivation of form vision demonstrates a visual feedback mechanism in eye growth control. Experimental work indicates that this neural control mechanism is at least partly localized to the retina itself, but how retinal signals directly control the growth of the outer coats of the eye is presently unknown. Genes which map to the interval will be ranked using information from retinal gene expression databases such as Ret Net (http://www.sph.uth.tmc.edu/RetNet/) and publications of retinal gene expression (cDNA library and microarray analyses).[119-122]

Mutational Sequence Analysis of Viable Candidate Genes

The MYP2 families come from a number of diverse backgrounds; all were recruited in the United States. One pedigree is of Ashkenazi Jewish descent, five are from mixed Western European ancestry, and one is from Chinese-Hawaiian descent. There is no evidence for a founder mutation, and all of these families appear unrelated. In one of the pedigrees, the oldest generation had no history of myopia, suggesting a likely de novo mutation. The lack of a founder effect should serve as a tremendous advantage in the search for, and analysis of, the MYP2 disease gene, since it increases the likelihood of finding multiple different mutations in our MYP2 families. This fact both improves the chances of finding the initial mutation by screening mechanisms and increases the ease in which changes within the MYP2 gene can be established as mutations and not benign polymorphisms. In addition, a diversity of mutations should greatly enhance the opportunity to make meaningful genotype-phenotype correlations based on severity and location of mutations within the MYP2 gene.

The primary method to elucidate the genomic organization was by computational sequence analysis.[123] The cDNA was aligned against genomic sequence, and flanking splice site motifs were identified. This alignment was performed as a NCBI BLAST function. Primers were designed to amplify the exons (coding sequence) and their adjacent splice sites using Oligo 6.6, a software program that searches for and selects oligonucleotides from a sequence file for PCR, sequencing, and other applications. The size of the primers varies from 18mer to 24mer primer pairs to span intron (noncoding sequence)-exon boundaries, and to have the most specificity and the least predicted likelihood of primer-dimer or hairpin formation.

Because this is a dominant disorder, the first approach was to search for mutations that modify or disrupt the coding regions of transcripts by direct sequencing of genomic DNA derived from blood leukocytes. Candidate genes thus far have been primarily screened for sequence alterations, such as missense/nonsense/insertion mutations and small deletions by analysis of exonic sequence amplified from affected individuals' genomic DNA using bidirectional sequencing. For direct sequencing, genomic DNA was used as a template for PCR amplification of each exon (or part of an exon). PCR amplicons were separated by electrophoresis through 1.5% agarose to verify the presence of a single band, and purified using the QIAquick PCR purification kit (Qiagen). The Thermo Sequenase Dye Terminator Cycle Sequencing Pre-Mix kit was used (Amersham Pharmacia Biotech, Piscataway, New Jersey), and sequence elements were analyzed on an automated sequencer (Applied Biosystems,



Primer	Position on NT_001036	Primer size	Amplicon Length	Primer	Position on NT_001036	Primer size	Amplicon Length
EM1cf	1390394	19		EM4df	1264376	19	
EM1cr	1390842	21	448	EM4dr	1264963	21	587
EM2f	1389721	23		EM5af	1248261	19	
EM2r	1390146	23	425	EM5ar	1248724	21	463
EM3f	1352537	23		EM6f	1246326	21	
EM3r	1352871	23	334	EM6r	1246668	19	342
EM4af	1262934	21		EM7f	1245632	23	
EM4ar	1263645	18	711	EM7r	1245945	20	313
EM4bf	1263611	19		EM8af	1241355	22	
EM4br	1264257	23	646	EM8ar	1241805	23	450
EM4cf	1264089	22		EM8bf	1241747	19	
EM4cr	1264563	23	474	EM8br	1242489	23	742

FIGURE 5

Genomic structure of the EMLIN-2 gene. The exons are in boxes, intronic sequences are shown as horizontal connecting lines. Base pair size is provided as adjoining numbers. The arrows indicate primer locations.

Model 3700, Foster City, California).

To establish that a candidate gene is the MYP2 disease gene, pathogenic mutations must be consistently identified in affected individuals. Disease-causing mutations can then be distinguished from benign polymorphisms by screening (1) dbSNP database (http://www.ncbi.nlm.nih.gov/SNP/) for known polymorphisms, (2) a panel of 100 normal unaffected and unrelated individuals of diverse geographic and ethnic background, and (3) when possible, a panel of 50 normal unaffected individuals of similar ethnic background to the affected individual. Once the MYP2 gene has been identified, we will search for mutations in all families linked to the MYP2 locus. We will also screen sporadic MYP cases and small, unlinked MYP families.

Because of the number of MYP2 families, their ethnic diversity, and the lack of evidence of a founder mutation, we believe that our collection of MYP2 pedigrees reflects mutation events that have arisen independently and thus may represent multiple different pathogenic events. Structure-function analysis of these mutations should highlight important regions of the gene and its protein product and may provide insight into phenotype-genotype correlations.

All MYP2 mutations were mapped to the cDNA of the MYP2 gene and to its predicted protein structure. Changes also found in normal control individuals were considered polymorphisms, reflecting nonpathogenic base pair changes. Changes found in affected MYP2 individuals may include deletions, nonsense mutations, and missense mutations. Clustered missense mutations may implicate a particular functional domain that is essential to the protein product's direct interaction with itself or with another protein. Alternatively, missense mutations may cause abnormal folding of the protein and not reflect a primary disturbance of a direct interaction.

Patients Studied. Probands and affected subject representatives of the 7 MYP2 families with an autosomal dominant form of high myopia were studied (Table 5). Each of the affected individuals had high myopia of –6.00 diopters sphere or more with elongated axial lengths. Clinical details regarding the complete pedigrees were published previously.[76] Controls were obtained from family marry-ins, nonmyopic family members, and unrelated subjects. Table 5 displays the family and member number of each individual, as well as controls with refractive error. A total of 20 patient samples were studied, of which 10 had high myopia and 10 were nonmyopic.

Total genomic DNA was extracted from 10 to 15 mL of venous blood from all participants after informed consent was obtained. DNA was purified from lymphocyte pellets according to standard procedures using the PUREGENE kit (Gentra Systems, Minneapolis, Minnesota) or phenol-chloroform extraction method.

Polymerase chain reactions were performed on 150-ng genomic DNA using standard methods. Amplified products were separated by agarose gel electrophoresis and visualized by staining with ethidium bromide. Amplicons were purified using QIAquick purification columns (Qiagen, Valencia, California), and sequenced using BigDye Terminator v3.1 on an ABI 3700 Genetic Analyzer (Applied Biosystems, Foster City, California).

The genes EMLIN-2, TGIF, CLUL1, and MLCB were screened first because of known expression in ocular tissues. The recent report by Lam and associates[101]

TABLE 5. LIST OF ALL SUBJECT DNA SAMPLES USED FOR THIS STUDY*

ORIGINAL MYP2 PEDIGREE NO.-INDIVIDUAL NO.	REFRACTIVE ERROR	
	OD	OS
1-15	−10.50+2.50×82	−11.50+2.75×97
1-16	−11.00+0.75×60	−12.25+1.00×70
1-19	−6.00+1.50×50	−6.50+0.75×80
1-21	Plano	Plano
1-23	Plano	Plano
2-6	−16.00 sphere	−16.00 sphere
2-10	Plano	Plano
3-6	−19.25+2.25×135	−21.00+3.00×60
4-7	−6.25 sphere	−6.00 sphere
4-8	Plano	Plano
5-6	−10.00+1.50×46	−8.25+1.75×105
5-9	+0.25+0.50×170	+0.25+0.25×170
6-4	−7.25+1.25×180	−7.25+1.25×20
6-6	Plano	Plano
7-5	−10.75+2.00×10	−10.50+2.50×70
External control-1 (C1)	Plano	Plano
External control-2 (C2)	Plano	Plano
External control-3 (C3)	Plano	Plano
External control-4 (C4)	Plano	Plano
External myopic control-5 (C5)	−9.00 sphere	−9.00 sphere

C, control; OD, right eye; OS, left eye.

*With reference to the individual's family identification relating back to the initial study describing the MYP2 myopia locus[76] and refractive errors of all control participants.

described a TGIF sequence variation study of the 3-exon transcript variant 4 using conformation specific gel-electrophoresis. One consideration is that the TGIF genetic structure studied by Lam and colleagues had 3 exons; the current sequence build is a 10-exon gene structure. Exons 1, 2, and 3 are now exons 5, 9, and 10, respectively, according to the reference sequence build 33 (http://www.ncbi.nlm.nih.gov/genome/guide/human/HsStats.html) of TGIF, which corresponds to transcript variant 4. They found 25 SNPs on exon 3 (exon 10 in our study). Six SNPs showed significant high myopia association with univariate analysis, and one showed significance with multivariate analysis. They did not sequence the full TGIF gene.

RESULTS

EMLIN-2

Using NCBI BLAST, the EMLIN 2 cDNA was aligned against the BACs that contained sequence similarity, and its genomic structure was electronically determined. The ~40 kilobase human EMLIN-2 gene is encoded by eight exons (Figure 5). Primers were designed to amplify the exons and their adjacent splice sites using Oligo 6.6, a software program that searches for and selects oligonucleotides from a sequence file for PCR, sequencing, and other applications. We designed 12 primer pairs to span intron-exon boundaries, and to overlap each other in exonic sequence. The gene was analyzed by direct sequencing.

Eleven single nucleotide polymorphisms were found after screening; all were novel and in intronic sequence or 3′ or 5′ untranslated regions (Table 6). The novel SNPs were submitted to the dbSNP database (http://www.ncbi.nlm.nih.gov/SNP/).

CLUL1

The ~53 kb human CLUL1 gene is encoded by 11 exons (Figure 6). We designed 13 primer pairs to span intron-exon boundaries and to overlap each other in exonic sequence (Table 7). The gene was analyzed by direct sequencing. No polymorphisms were identified.

MLCB

The ~8.7 kb human MLCB gene is encoded by four exons (Figure 7). We designed four primer pairs to span intron-

TABLE 6. POLYMORPHISMS DETECTED WITH DIRECT DNA SEQUENCING OF GENOMIC DNA USING EMLIN-2 PRIMERS

mRNA POSITION	OBSERVED BASE PAIR CHANGE	WILD-TYPE NUCLEOTIDE
573	T	C
1200	A	G
1512	T	C
1881	Heterozygous T/C for all affected, and one control	C
2130	A	G
2508	A	G
2866	T	C
3350	C	T
3498	A	G
3815	T	C
3837	A	G

exon boundaries and to overlap each other in exonic sequence (Table 8). The gene was analyzed by direct sequencing. No polymorphisms were identified.

TGIF

The genomic structure of TGIF, as reported in MapViewer (build 33) of the reference human genome sequence, is outlined below. The genomic structure for TGIF contains 10 exons spanning ~47.6 kb and has eight transcript variants encoding four proteins of 402 residues (variant 1), 287 residues (variant 2), 273 residues (variants 3 and 4), and 253 residues (variants 5 through 8) (Figure 8). All participant DNA samples were screened for sequence variants on exon 7, although it has no continuous open reading frame with the conserved region of exons 9 and 10, and only one known corresponding EST. Fourteen oligonucleotide primer pairs were designed to amplify the exonic sequences with 50 to 200 base pairs extensions beyond the intron-exon boundary (Table 9).

A total of 21 polymorphisms were found in the 10 exons screened for TGIF (Table 10). Of these, three were missense variances, two were silent, 10 were not translated, four were intronic, and two were homozygous deletions. The three missense allelic variants were observed at exon 10 at positions 236C→T (Pro→Leu), 244C→T(Pro→Ser), and 245C→T(Pro→Leu). Silent mutations were observed on exon 10 at positions 177A→G and 333C→T. The two deletions causing frameshift mutations were observed in exon 6 at positions 3442216 and 3442223 on NT_010859.13. Both deletions are predicted to cause early termination, yielding proteins of 132 and 141 residues, respectively. Ten polymorphisms were novel and have been submitted to the dbSNP database. Eleven polymorphisms corresponded with previously reported SNPs in public databases. None of the sequence variants cosegregated with the affected myopia phenotype. Specifically, there were no heterozygous or homozygous polymorphisms observed only in affected individuals in any MYP2 pedigree.

DISCUSSION

We sequenced the full coding regions of EMLIN2, CLUL1, MLCB, and TGIF positional candidate genes in our patient samples of individuals from pedigrees with MYP2- associated high myopia. No DNA sequence variants were noted that implicated any as the causative gene. We were especially interested in the TGIF candidate gene because of its published association with MYP2 by SNP association studies. TGIF exon 10 (exon 3 in the initial build of this gene) did not show the same level of polymorphic variants in our cohort, because we observed eight variants rather than the 25 reported by Lam and colleagues.[101] This may be due to the ethnic differences in

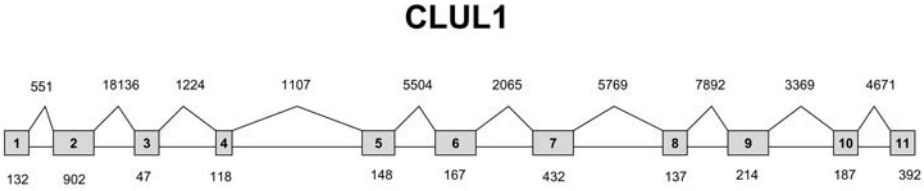


FIGURE 6

Genomic structure of the 11-exon CLUL1 gene.

TABLE 7. PRIMERS DESIGNED FOR SEQUENCE ANALYSIS OF THE CLUL1 GENE

PRIMER NAME	mRNA POSITION	PRIMER SEQUENCE 5′-3′	PRIMER SIZE (bp)	AMPLICON SIZE (bp)
CLUL E1F	AP001178:24702	CGGGGTTGGTTTCCACC	17	133
CLUL E1R	AP001178:24835	GCGAGGAGAGAAATCTGGG	19	
CLUL E2F	AP001178:34573	TGCTCACTACTTTGCAGTGTTC	22	289
CLUL E2R	AP001178:34862	TGAGATCGTGTCACTGCATTCT	22	
CLULE3F		GTAAATCTCAAAATGCGGGTTAATAG	26	277
CLULE3R	AP001178:44489	CTAACTCTTCTTCTATCATTACTC	24	
CLULE4-1F	AP001178:45596	CCCAGGTGTTTTCAATTGATGC	22	156
CLULE4-1R	AP001178:45752	AGCAGTTTTGTCCTTCCAAGTG	22	
CLULE4-2F		TGTTTATTGTGTCTGCTGTG	20	376
CLULE4-2R	AP001178:45823	GGACAACCAACATGCAAACAG	21	
CLULE5F	AP001178:46846	GTGTTTTGTAATCTGATCAGATCTC	25	237
CLULE5R	AP001178:47083	GCAGTATTTCTGGTCCAGATC	21	
CLULE6F		GGTGCACATAGATCATGAAATGG	23	317
CLULE6R	AP001178:52757	TAAGCTGAAATAGGTGCCTTAAG	23	
CLULE7-1F	AP001178:54748	TTTATTCCATTTCTGTCCCTAC	22	199
CLULE7-1R	AP001178:55038	AAGGCTCAGTTAGGTCTGTATC	22	
CLULE7-2F	AP001178:54947	CAGGAGTTTTAACGTCTTCAGAC	23	333
CLULE7-2R	AP001178:55280	GACTCAGAAATGTCTACCATTTC	23	
CLULE8F	AP001178:60921	TCTCCACTTCTTCAAAGTGC	20	232
CLULE8R	AP001178:61153	CAAAATGTACCTGAGAACTTAAAG	24	
CLULE9F	AP001178:68958	CACCTCCAAGTTTCATGGAC	20	298
CLULE9R	AP001178:69256	CAAGGTATGCACGTGTCATTTC	22	
CLULE10F	AP001178:72544	GAATGTGTATTGGGATTTAGTAAAC	25	270
CLULE10R	AP001178:72814	TTGAGAATTAACTATTCCTGTCAAC	25	
CLULLE11F	AP001178:77490	CCATCCTGGACTTTTACTCC	20	182
CLULE11R	AP001178:77672	CTTTCCTGCAACTGTGTTTATTG	23	

bp, base pair.

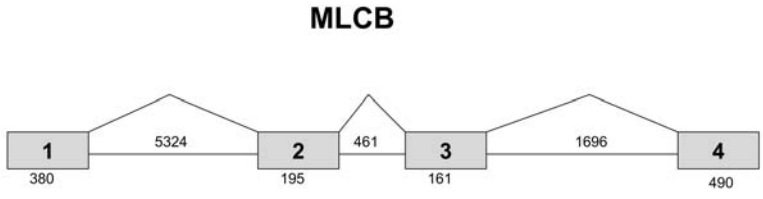


FIGURE 7

Genomic structure of the 4-exon MLCB gene.

our two sample sets, although family 1 of the MYP2 pedigrees studied was of Chinese descent. All other families were of Northern European descent. The TGIF gene was not fully screened based on the methods described in their publication, and therefore the SNP association is most likely due to a nearby gene or regulatory element. The fact that TGIF mutations cause holoprosencephaly also reduces the likelihood that it is directly associated with simplex myopia.

Information derived from this effort will be useful for submissions to the ever-growing SNP database and to other researchers also exploring candidate genes in this region, whether it is for a myopia-related project or others. Other researchers screening for myopia candidate genes in this interval may wish to avoid repeated screening of those genes that have been excluded. The molecular study of any of these genes requires PCR primers that have been optimized for the gene exonic and intronic area of interest.

Nonsyndromic high myopia is a common, complex disorder that is likely to result from alterations of multiple

TABLE 8. PRIMERS DESIGNED FOR SEQUENCE ANALYSIS OF THE MLCB GENE

PRIMER NAME	mRNA POSITION (NT_010859)	PRIMER SEQUENCE 5′-3′	PRIMER SIZE (BASE PAIR)	AMPLICON SIZE (BASE PAIR)
MLB-1F	1074792	ATCCTAAGACAGGGTCACGAA	21	660
MLB-1R	1075452	CCCTCAGCACCCCTATACT	19	
MLB-2AF	1074253	CCTGAGAGCGTTATCA	16	543
MLB-2AR	1074796	TTCGTGACCCTGTCTTA	17	
MLB-2BF	1074252	ACCTGAGAGCGTTATCA	17	556
MLB-2BR	1074808	TCACCAGCAGATTCGT	16	
MLB-3F	1072092	CTTTCTTTGGGAGATACGACT	21	831
MLB-3R	1072923	ATTACTTTGTTTAGGCATAGG	21	

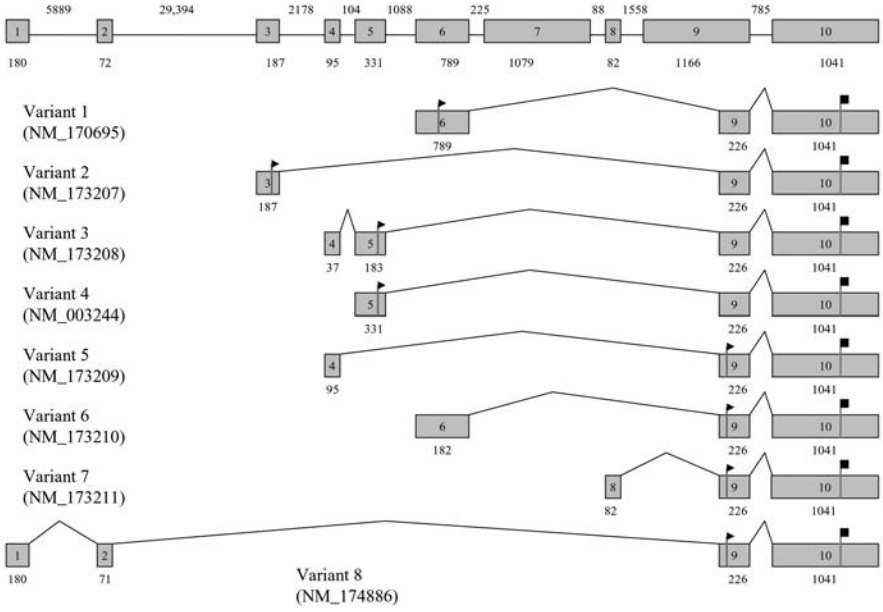


FIGURE 8

Genomic structure of the 10-exon TGIF gene. The associated ~47.6kb region of NT_010859 on chromosome 18p11.31 of TGIF, showing 10 exons with alternative start sites and splicing that generates eight transcript variants. The exons are represented as boxes, initiation codons are represented by a vertical line with arrow, and stop codons are represented by a vertical line with a black square.

genetic factors. Indeed, several loci have been mapped for nonsyndromic high myopia. We will continue our efforts to determine the gene alterations involved for the MYP2 locus and for the other known high-grade simplex myopia loci.

CONCLUSION

Mutation analysis of four encoded positional candidate genes shown to be expressed in ocular tissues for MYP2 autosomal dominant high myopia did not identify sequence alterations associated with the disease phenotype. Further studies of MYP2 candidate genes are needed to determine the gene causative for this potentially blinding disorder. Mutation screening of other genes that also map to this interval is in progress.

Considerations for further study and future directions point to two other features with disease gene identification:

1. Are we ascertaining the right families? The genes that contribute to complex or multifactorial disease—those such as diabetes, asthma, cancer, heart disease, and psychiatric illness—are notoriously difficult to identify as they typically exert small effects on disease risks.[124] The magnitude of their effects is likely to be modified by other unrelated genes as well as environmental factors. We recognize the problem of genetic heterogeneity, phenocopy, and shared environmental factors. Multiplex families with unilinear transmission of the affected phenotype provide the most unambiguous information for detecting linkage, especially for complex traits.[125] Selectively ascertaining such pedigrees with high disease severity and early age of onset biases toward a strong genetic etiology, minimizing environmental influences to the disease trait. This has

TABLE 9. TGIF GENE PRIMERS DESIGNED FOR MUTATION SCREENING*

PRIMER	EXON	PRIMER SEQUENCE 5'-3'	AMPLICON SIZE (bp)
TGIF1F	1	CAGGAAACACAGACCAGCTTATCTT	754
TGIF1R		AATGCAGTCATGCCACTCGATATA	
TGIF2F	2	CCCAAATTGTCTATCGGTGA	249
TGIF2R		ATGACTAGGTTCAAGCCAATG	
TGIF3F	3	TTTTGAGGCTGTTCCCTTTGTTACG	449
TGIF3R		TGGACTTGGCTAATGACTGCGATT	
TGIF4F	4	GAGGCGGCTGTCTCGTGCGGCTAGA	632
TGIF4R		GATCCCAGGCGCCCGCTCCTT	
TGIF5F	5	CCGCCCCCGAGGGACGAGT	662
TGIF5R		AGCCGCTGCCGTTTCAGACGC	
TGIF6F	6	GTCCCCCGAGTGTTCCGCTGT	1061
TGIF6R		CCCACCCCAGACGACTCGC	
TGIF7F	7	CCTCTGAGAAGGTGGGATTCACG	711
TGIF7R		CCTTTGAGCTGGGAGCAATGTCTCT	
TGIF7AF		GAATTCCTGCTTTGGGATTGC	796
TGIF7AR		CACAAGCTTCTTTCACGCTATCCAC	
TGIF8F	8	AGAACGTGCAGGAACAAGGTCG	329
TGIF8R		CACTGCAGTAAAACCACGTTTGCTG	
TGIF9F	9	CTCGCTCTCAGTTGTTGG	361
TGIF9R		TCACGCTCTCTTTCTTTA	
TGIF9AF		TCTTCGGGATTGGCTGTATGA	562
TGIF9AR		TCCCCATTTTAACCTATCACCTACT	
TGIF10F	10	TCAATTAGGTACCCCATAGAACAT	427
TGIF10R		TCCCACACCGACCGACTG	
TGIF10AF		TCTGCCATACCACTGTGACTG	385
TGIF10AR		AATATATAAAATAATGCAATTCATCTCTTG	
TGIF10BF		GGGAATGAGAAGATGCTAT	284
TGIF10BR		GGTGAAGGCAAGAGATGAAT	

bp, base pair.

*Exons 7, 9, and 10 primers are divided into 2, 2, and 3 segments, respectively. The primers for TGIF9F/R were initially published by Gripp and associates,[98] and primers for TGIF9AR, TGIF10F/R, and TGIF10AF/R were originally published by Chen and associates.[99]

been our ascertainment strategy for our mapping studies.

2. Are there other genotyping techniques that may be utilized? While our mapping studies are based on polymorphic microsatellite marker allele variations, we recognize that the use of SNPs for haplotype-based association studies may offer advantages over the use of conventional markers.[126] Genomic regions can be tested for association without requiring the discovery of the functional variants. SNPs are more densely distributed and abundant, occurring roughly every 1,000 bp along the human genome. SNPs are binary, and thus well suited to automated, high-throughput genotyping. Finally, in contrast to more mutable markers, such as microsatellites, SNPs have a low rate of recurrent mutation, making them stable indicators of human history. We will consider and test emerging SNP-based technologies as they become available in our institution, and if they practically improve our genotyping efforts in terms of both cost and efficiency. Irrespective of this, we can use selected SNPs as supplemental polymorphic markers for fine-point mapping after a genome-wide scan implicates a linked region. High-resolution, fully integrated maps of SNPs are publicly available at the SNP database (dbSNP, http://www.ncbi.nlm.nih.gov/SNP/).

REFERENCES

1. Curtin BJ. *The Myopias: Basic Science and Clinical Management.* New York: Harper & Rowe; 1985:237-245.

TABLE 10. LIST OF OBSERVED SEQUENCE POLYMORPHISMS FOUND IN THE TGIF GENE LABELED IN BASE PAIRS*

NT_010859.13 POSITION	WILD-TYPE NUCLEOTIDE	BASE PAIR CHANGE OBSERVED	SNP RS NO.†	ORIGINAL MYP2 FAMILY NO.-INDIVIDUAL NO.	EXON POSITION	AMINO ACID CHANGE
3402167	C	C/T, T/T	Novel	1-23,**-15**,**-16** and **-19**, 2-10, **5-6**, C-2 and 3	bp 96 of exon 1	5′UTR
3402386	C	C/T	rs8092903	C2	intron 1	
3402523	C	C/T, T/T	Novel	**5-6**, **7-5**, C2	intron 1	
3408219	T	C/T	Novel	**7-5**	intron 2	
3437725	C	C/T	238137	**1-16**, 1-21, C2	bp118 of exon 3	5′UTR
3437871	T	C/T	rs151472	**1-15**, **1-19**, **4-7**, **5-6**, **7-5**, 1-21, 1-23, 2-10, C2	intron 3	
3440455	C	C/A	rs238132	4-8, **6-4**, C-2, 3 and 4	bp 284 of exon 5	5′UTR
3441762	C	C/T, T/T	rs238533	**4-7**, **6-4**, 4-8, C4	bp172 of exon 6	5′UTR
3441896	G	G/A, A/A	Novel	**4-7**, **6-4**, 4-8	bp 306 of exon 6	5′UTR
3442978	C	C/T, T/T	2238536	**4-7**, **6-4**, 4-8	bp 374 of exon 7	UTR
3442216	C	C/C Deletion	Novel	C1	bp 626 of exon 6	frameshift
3442223	T	T/T Deletion	Novel	**3-6**, C1	bp 633 of exon 6	frameshift
3443789	T	G/T	Novel	**1-16** and 21, **2-6**	bp18 of exon 8	5′ UTR
3447539	A	G/A, G/G	rs2229337	**4-7**, **6-4**, C-2, 4 and **C5**	bp177 of exon 10	NONE
3447598	C	C/T	Novel	6-6, C1	bp236 of exon 10	PRO→LEU
3447606	C	C/T	rs4468717	3-6	bp244 of exon 10	PRO→SER
3447607	C	C/T, T/T	rs2229333	**4-7**, **6-4**, 4-8	bp245 of exon 10	PRO→LEU
3447695	C	C/T	rs2229335	C-1 and 4	bp 333 of exon 10	NONE
3447776	T	T/G	rs2229336	C-1 and 4	bp414 of exon 10	3′UTR
3448161	G	G/A	Novel	5-9	bp799 of exon 10	3′UTR
3448260	G	G/A	Novel	**4-7**, **5-6** and -9, 6-4, **7-5**, C2	bp898 of exon 10	3′UTR

bp, base pair; C, control; UTR, untranslated region.

*The wild-type sequence is derived from the scaffold sequence NT_010859 of chromosome 18p. Amino acid changes are for relevant splice variants. Affected individual sample numbers are in bold type.

†rs No. is the public reference SNP number from the dbSNP database (http://www.ncbi.nlm.nih.gov/SNP/).

2. Sperduto RD, Siegel D, Roberts J, et al. Prevalence of myopia in the United States. *Arch Ophthalmol* 1983;101:405-407.
3. Wang Q, Klein, BEK, Klein R, et al. Refractive status in the Beaver Dam Eye Study. *Invest Ophthalmol Vis Sci* 1994;35:4344-4347.
4. Sperduto RD, Siegel D, Roberts J, et al. Prevalence of myopia in the United States. *Arch Ophthalmol* 1983;101:405-407.
5. Angle J, Wissmann DA. The epidemiology of myopia. *Am J Epidemiol* 1980;111:220-228.
6. Leibowitz HM, Krueger DE, Maunder LR. The Framingham eye study monograph. *Surv Ophthalmol* 1980;24(suppl):472-479.
7. Katz J, Tielsch JM, Sommer A. Prevalence and risk factors for refractive errors in an adult inner city population. *Invest Ophthalmol Vis Sci* 1997;334-340.
8. Burton TC. The influence of refractive error and lattice degeneration on the incidence of RD. *Trans Am Ophthalmol Soc* 1990;87:143-155.
9. Curtin BJ. Myopia: a review of its etiology, pathogenesis, and treatment. *Surv Ophthalmol* 1970;15:1-17.
10. Ghafour IM, Allan D, Foulds WS. Common causes of blindness and visual handicap in the west of Scotland. *Br J Ophthalmol* 1983;67:209-213.
11. National Advisory Council, Strabismus, Amblyopia and Visual Processing Panel. *Vision Research—A National Plan: 1999-2003.* Washington, DC: National Institutes of Health; 1999.
12. Javitt JC, Chiang Y-P. The socioeconomic aspects of laser refractive surgery. *Arch Ophthalmol* 1994;112:1526-1530.
13. Tokoro T, Sato A, eds. *Results of Investigation of Pathologic Myopia in Japan. Report of Myopic Chorioretinal Atrophy*. Tokyo: Ministry of Health and Welfare; 1982:32-35.
14. Lin LL, Chen CJ, Hung PT, et al. Nation-wide survey of myopia among schoolchildren in Taiwan. *Acta Ophthalmol* 1988;66:29-33.
15. Fledelius HC. Myopia prevalence in Scandinavia: a survey, with emphasis on factors of relevance for epidemiological refraction studies in general. *Acta Ophthalmol* 1988;185:44-50.

16. Wilson A, Woo G. A review of the prevalence and causes of myopia. *Singapore Med J* 1989;30:479-484.
17. Grosvenor T. A review and a suggested classification system for myopia on the basis of age-related prevalence and age of onset. *Am J Opt Physiol Optics* 1987;64:545-554.
18. Mantyjarvi MI. Changes in refraction in schoolchildren. *Arch Ophthalmol* 1985;103:790-792.
19. Van Alphen GW. On emmetropia and ametropia. *Ophthalmologica* 1961;142(suppl):1-92.
20. Jansson F. Measurement of intraocular distances by ultrasound and comparison between optical and ultrasonic determinations of the depth of the anterior chamber. *Acta Ophthalmol* 1963;41:25-61.
21. Sorsby A, Leary GA, Richards MJ. Correlation ametropia and component ametropia. *Vision Res* 1962;2:309-313.
22. Curtin BJ, Karlin DB. Axial length measurements and fundus changes of the myopic eye. *Am J Ophthalmol* 1971;71:42-50.
23. Tron EJ. The optical elements of the refractive power of the eye. In: Ridley F, Sorsby A, eds. *Modern Trends in Ophthalmology*. 1940:245. Reprinted in Curtin.[1]
24. Grossniklaus HE, Green WR. Pathologic findings in pathologic myopia. *Retina* 1992;12:127-133.
25. Gass JDM. *Stereoscopic Atlas of Macular Diseases, 4th ed. Diagnosis and Treatment.* New York: Mosby; 1997:126-128.
26. Curtin BJ. Posterior staphyloma development in pathologic myopia. *Ann Ophthalmol* 1982;14:655-658.
27. Steidl SM, Pruett RC. Macular complications associated with posterior staphyloma. *Am J Ophthalmol* 1997;123:181-187.
28. Rabb Mf, Garoon I, LaFranco FP. Myopic macular degeneration. *Int Ophthalmol Clin* 1981;21:51-69.
29. Noble KG, Carr RE. Pathologic myopia. *Ophthalmology* 1982;89:1099-1100.
30. Jalkh AE, Weiter JJ, Trempe CL, et al. Choroidal neovascularization in degenerative myopia: role of laser photocoagulation. *Ophthalmic Surg* 1987;18:721-725.
31. Hayasaka S, Uchida M, Setogawa T. Subretinal hemorrhages with or without choroidal neovascularization in the maculae of patients with pathologic myopia. *Graefes Arch Clin Exp Ophthalmol* 1990;228:277-280.
32. Hotchkiss ML, Fine SL. Pathologic myopia and choroidal neovascularization. *Am J Ophthalmol* 1981;91:177-183.
33. Levy JH, Pollock HM, Curtin BJ. The Fuchs' spot: an ophthalmoscopic and fluorescein angiographic study. *Ann Ophthalmol* 1977;9:1433-1443.
34. Avila MP, Weiter JJ, Jalkh AE, et al. Natural history of choroidal neovascularization in degenerative myopia. *Ophthalmology* 1984;91:1573-1581.
35. Fried M, Siebert A, Meyer-Schwickerath G. A natural history of Fuchs' spot: a long-term follow-up study. *Doc Ophthalmol* 1981;28:215-221.
36. Hampton GR, Kohen D, Bird AC. Visual prognosis of disciform degeneration in myopia. *Ophthalmology* 1983;90:923-926.
37. The Eye Disease Case-Control Study Group. Risk factors for idiopathic rhegmatogenous RD. *Am J Epidemiol* 1993;137:749-757.
38. Perkins ES. Morbidity from myopia. *Sightsav Rev* 1979;49:11-19.
39. Perkins ES. Glaucoma in the younger age groups. *Arch Ophthalmol* 1960;64:882-891.
40 Vongphanit J, Mitchell P, Wang J. Prevalence and progression of myopic retinopathy in an older population. *Ophthalmology* 2002;109:704-711.
41. Pierro L, Camesasca FI, Mischi M, et al. Peripheral retinal changes and axial myopia. *Retina* 1992;12:12-17.
42. Spitznas M, Boker T. Idiopathic posterior subretinal neovascularization (IPSN) is related to myopia. *Graefes Arch Clin Exp Ophthalmol* 1991;229:536-538.
43. Goss DA, Jackson TW. Clinical findings before the onset of myopia in youth: parental history of myopia. *Optom Vis Sci* 1996;73:279-282.
44. Gwiazda J, Thorn F, Bauer J, et al. Emmetropization and the progression of manifest refraction in children followed from infancy to puberty. *Clin Vis Sci* 1993;8:337-344.
45. Zadnik K. The Glenn A. Fry Award Lecture. Myopia development in childhood. *Optom Vis Sci* 1997;74:603-608.
46. Zadnik K, Satariano WA, Mutti DO, et al. The effect of parental history of myopia on children's eye size. *JAMA* 1994;271:1323-1327.
47. Wallman J. Parental history and myopia: taking the long view [letter]. *JAMA* 1994;272:1255-1256.
48. Wallman J, Turkel JI, Trachtman J. Extreme myopia produced by modest change in visual experience. *Science* 1978;201:1249-1251.
49. Weisel TN, Raviola E. Myopia and eye enlargement after neonatal lid fusion in monkeys. *Nature* 1977;266:66-68.
50. Sherman SM, Norton TT, Casagrande VA. Myopia in the lid-sutured tree shrew (*Tupaia glis*). *Brain Res* 1997;124:154-157
51. Hoyt CS, Stone RD, Fromer C, et al. Monocular axial myopia associated with neonatal eyelid closure in human infants. *Am J Ophthalmol* 1982;91:197-200.
52. Von Norden GK, Lewis RA. Ocular axial length in unilateral congenital cataracts and blepharoptosis. *Invest Ophthalmol Vis Sci* 1987;28:750-752.
53. Twomey JM, Gilvarry A, Restori M, et al. Ocular enlargement following infantile corneal opacification. *Eye* 1990;4:497-503.
54. Lin LLK, Hung PT, Ko LS, et al. Study of myopia among aboriginal school children in Taiwan. *Acta Ophthalmol Suppl* 1988;185:34-36.
55. Zylberman R, Landau D, Berson D. The influence of study habits on myopia in Jewish teenagers. *J Pediatr Ophthalmol Strabismus* 1993;30:319-322.
56. Young FA. Myopia and personality. *Am J Optom Physiol Opt* 1987;64:136-143.
57. Young FA. Reading, measures of intelligence, and refractive errors. *Am J Optom Physiol Opt* 1963;40:257-264.

58. Rosner M, Belkin M. Intelligence, education and myopia in males. *Arch Ophthalmol* 1987;105:1508-1511.
59. Pacella R, McLellan J, Grice K, et al. Role of genetic factors in the etiology of juvenile-onset myopia based on a longitudinal study of refractive error. *Optom Vis Sci* 1999;76:381-386.
60. Knowlton RG, Weaver EJ, Struyk AF, et al. Genetic linkage analysis of hereditary arthro-ophthalmopathy (Stickler syndrome) and the type II procollagen gene. *Am J Hum Genet* 1989;45:681-688.
61. Richards AJ, Yates JRW, Williams R, et al. A family with Stickler syndrome type 2 has a mutation in the COL11A1 gene resulting in the substitution of glycine 97 by valine in alpha-1(XI) collagen. *Hum Mol Genet* 1996;5:1339-1343.
62. Dietz HC, Cutting GR, Pyeritz RE, et al. Marfan syndrome caused by a recurrent de novo missense mutation in the fibrillin gene [see comments]. *Nature* 1991;352(6333): 337-339.
63. Nijbroek G, Sood S, McIntosh I, et al. Fifteen novel FBN1 mutations causing Marfan syndrome detected by heteroduplex analysis of genomic amplicons. *Am J Hum Genet* 1995;57:8-21.
64. Sertie AL, Sossi V, Camargo AA, et al. Collagen XVIII, containing an endogenous inhibitor of angiogenesis and tumor growth, plays a critical role in the maintenance of retinal structure and in neural tube closure (Knobloch syndrome). *Hum Molec Genet* 2000;9:2051-2058.
65. Goldschmidt E. On the etiology of myopia: an epidemiological study. *Acta Ophthalmol Suppl* 1968;98:1-172.
66. Teikara JM, O'Donnell J, Kaprio J, et al. Impact of heredity in myopia. *Hum Hered* 1991;41:151-156.
67. Ashton GC. Segregation analysis of ocular refraction and myopia. *Hum Hered* 1985;35:232-239.
68. Goss DA, Hampton MJ, Wickham MG. Selected review on genetic factors in myopia. *J Am Optom Assoc* 1988;59:875-884.
69. Naiglin L, Clayton J, Gazagne Ch, et al. Familial high myopia: evidence of an autosomal dominant mode of inheritance and genetic heterogeneity. *Ann Genet* 1999;42:140-146.
70. Guggenheim JA, Kirov G, Hodson SA. The heritability of high myopia: a re-analysis of Goldschmidt's data. *J Med Genet* 2000;27:227-231.
71. Teikari JM, Kaprio J, Koskenvuo M, et al. Heritability of defects of far vision in young adults—a twin study. *Scand J Soc Med* 1992;20:73-78.
72. Sorsby A, Sheriden M, Leary GA. *Refraction and Its Components in Twins.* Medical Research Council Special Report, Series 303. London: Medical Research Council; 1962.
73. Lyhne N, Sjolie AK, Kyvik KO, et al. The importance of genes and environment for ocular refraction and its determiners: a population based study among 20-45 year old twins. *Br J Ophthalmol* 2001;85:1470-1476.
74. Schwartz M, Haim M, Skarsholm D. X-Linked myopia. Bornholm eye disease. *Clin Genet* 1990;38:281-286.
75. Young TL, Deeb SS, Ronan SM, et al. X-linked high myopia associated with cone dysfunction. *Arch Ophthalmol* 2004:122:897-908.
76. 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.
77. Young TL, Ronan SM, Alvear A, et al. A second locus for familial high myopia maps to chromosome 12q. *Am J Hum Genet* 1998;63:1419-1424.
78. Paluru P, Heon E, Devoto M, et al. A new locus for autosomal dominant high myopia maps to the long arm of chromosome 17. *Invest Ophthalmol Vis Sci* 2003;44:1830-1836.
79. Heath SC, Robeldo R, Beggs W, et al. A novel approach to search for identity by descent in small samples of patients and controls from the same Mendelian breeding unit: a pilot study on myopia. *Hum Hered* 2001;52:183-190.
80. Lam DSC, Tam POS, Fan DSP, et al. Familial high myopia linkage to chromosome 18p. *Ophthalmologica* 2003;217:115-118.
81. Naiglin L, Gazagne Ch, 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.
82. Muir H. Proteoglycans as organizers of the intercellular matrix. Seventeenth CIBA Medical Lecture. *Biochem Soc Trans* 1982;11:613-622.
83. Hassell JR, Blochberger TC, Rada JA, et al. Proteoglycan gene families, In: Bittar EE, Kleinman HK, eds. *Advances in Molecular and Cell Biology—Extracellular Matrix*. Vol 6. Greenwich, Conn: JAI Press Inc; 1993:69-113.
84. Rada JA, Achen VR, Penugonda S, et al. Proteoglycan composition in the human sclera during growth and aging. *Invest Ophthalmol Vis Sci* 2000;41:1639-1648.
85. Rada JA, McFarland AL, Cornuet PK, et al. Proteoglycan synthesis by scleral chondrocytes is modulated by a vision dependent mechanism. *Curr Eye Res* 1992;11:767-782.
86. Norton TT, Rada JA. Reduced extracellular matrix in mammalian sclera with induced myopia. *Vision Res* 1995;35:1271-1281.
87. Avetisov ES, Savitskaya NF, Vinetskaya MI, et al. A study of biochemical and biomechanical qualities of normal and myopic eye sclera in humans of different age groups. *Metab Pediatr Syst Ophthalmol* 1984;7:183-188.
88. Curtin BJ, Teng CC. Scleral changes in pathological myopia. *Trans Am Acad Ophthalmol* 1958;62:777-790.
89. Pousi B, Hautala T, Heikkinen J, et al. Alu-alu recombination results in a duplication of seven exons in the lysyl hydroxylase gene in a patient with the type VI variant of Ehlers-Danlos syndrome. *Am J Hum Genet* 1994;55:899-906.
90. Spielman RS, McGinis RE, Ewens WJ. Transmission test for linkage disequililbrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM). *Am J Hum Genet* 1993;52:506-516.
91. SAGE (computer program). Statistical Analysis for Genetic Epidemiology, Release 3.1. Cleveland, Ohio: Department of Epidemiology and Biostatistics, Rammelkamp Center for Education and Research, MetroHealth Campus, Case Western Reserve University; 1997.

92. Pratt SC, Daly MJ, Kruglyak L. Exact multi-point quantitative-trait linkage analysis in pedigrees by variance components. *Am J Hum Genet* 2000;66:1153-1157.
93. Young TL, Ronan SM, Atwood LD, et al. Further refinement of the MYP2 locus for autosomal dominant high myopia by linkage disequilibrium analysis. *Ophthalmic Genet* 2001;22:69-75.
94. Burge C, Karlin S. Prediction of complete gene structures in human genomic DNA. *J Mol Biol* 1997;268:78-94.
95. Wong P, Pfeffer BA, Bernstein SL, et al. Clusterin protein diversity in the primate eye. *Mol Vis* 2000;6:184-191.
96. Zhang Q, Ray K, Acland GM, et al. Molecular cloning, characterization and expression of a novel retinal clusterin-like protein cDNA. *Gene* 2000;243:151-160.
97. Doliana R, Bot S, Mungiguerra G, et al. Isolation and characterization of EMILIN-2, a new component of the growing EMILINs family and a member of the EMI domain-containing superfamily. *J Biol Chem* 2001;276:12003-12011.
98. Gripp KW, Wotton D, Edwards MC, et al. Mutations in TGIF cause holoprosencephaly and link NODAL signaling to human neural axis determination. *Nature Genet* 2000;25:205-208.
99. Chen C-P, Chern S-R, Du S-H, et al. Molecular diagnosis of a novel heterozygous 268CÆ T (R90C) mutation in TGIF gene in a fetus with holoprosencephaly and premaxillary agenesis. *Prenatal Diagnosis* 2002;22:5-7.
100. Overhauser J, Mitchell HF, Zackai EH, et al. Physical mapping of the holoprosencephaly critical region in 18p11.3. *Am J Hum Genet* 1995;5:1080-1085.
101. Lam DSC, Lee WS, Leung YF, et al. TGFB-induced factor: a candidate gene for high myopia. *Invest Ophthalmol Vis Sci* 2003;44:1012-1015.
102. Young TL, Roughley PJ, Ronan SM, et al. Lumican candidate gene analysis in chromosome 12q linked high myopia. *Invest Ophthalmol Vis Sci* 1999;40(Suppl):4.
103. Dalgleish R. The human type 1 collagen mutation database. *Nucleic Acids Res* 1997;25:181-187.
104. Mansson B, Wenglen C, Morgelin M, et al. Association of chondroadherin with collagen type II. *J Biol Chem* 2001;276:32883-32888.
105. Rada JA, Matthews AL, Brenza H. Regional proteoglycan synthesis in the sclera of experimentally myopic chicks. *Exp Eye Res* 1994;59:747-760.
106. Norton TT, Rada JA. Reduced extracellular matrix in mammalian sclera with induced myopia. *Vision Res* 1995;35:1271-1281.
107. Rada JA, Achen VR, Rada KG. Proteoglycan turnover in the sclera of normal and experimentally myopic chick eyes. *Invest Ophthalmol Vis Sci* 1998;39:1990-2002.
108. Rada JA, Nickla DL, Troilo D. Decreased proteoglycan synthesis associated with form deprivation myopia in mature primate eyes. *Invest Ophthalmol Vis Sci* 2000:41;2050-2058.
109. Young TL, Guo XD, King RA, et al. Identification of genes expressed in a human scleral cDNA library. *Mol Vis* 2003:9;508-514.
110. Young TL, Scavello G, Palura P, et al. Microarray analysis of gene expression in human donor sclera. *Mol Vis* 2004:10;163-176.
111. Shasty BS, Hejtmancik JF, Plager DA, et al. Linkage and candidate gene analysis of X-linked familial exudative vitreoretinopathy. *Genomics* 1995;27:341-344.
112. Zhao X, Alvarado D, Rainier S, et al. Mutations in a newly identified GTPase gene cause autosomal dominant hereditary spastic paraplegia. *Nat Genet* 2001;29:326-331.
113. Nakano M, Yamada K, Fain J, et al. Homozygous mutations in ARIX (PHOX2A) result in congenital fibrosis of the extraocular muscles type 2. *Nat Genet* 2001;29:315-320.
114. McPherson JD, Marra M, Hillier L, et al. A physical map of the human genome. *Nature* 2001;409:934-941.
115. Barmak M, Lee C. A genomic view of alternative splicing. *Nat Genet* 2002;30:13-19.
116. Brett D, Pospisil H, Valcarcel J, et al. Alternative splicing and genome complexity. *Nat Genet* 2002;30:29-30.
117. Gupta SK, Leonard BC, Damji KF, et al. A frame shift mutation in a tissue-specific alternatively spliced exon of collagen 2A1 in Wagner's vitreoretinal degeneration. *Am J Ophthalmol* 3:203-210.
118. Wallman J. Retinal control of eye growth and refraction. *Progress in Retinal Research*. Oxford, England: Pergamon Press; 1993:133-153.
119. Sohocki M, Malone K, Sullivan L, et al. Localization of retina/pineal-expressed sequences: Identification of novel candidate genes for inherited retinal disorders. *Genomics* 1999;58:29-33.
120. den Hollander A, van Driel M, de Kok Y, et al. Isolation and mapping of novel candidate genes for retinal disorders using suppression subtractive hybridization. *Genomics* 1999;58:240-249.
121. Malone K, Sohocki M, Sullivan L, et al. Identifying and mapping novel retinal-expressed ESTs from humans. *Mol Vis* 1999;5:5-8.
122. Bortoluzzi S, d'Alessi F, Danieli G. A novel resource for the study of genes expressed in the adult human retina. *Invest Ophthalmol Vis Sci* 2000;41:3305-3308.
123. Yi CH, Russ A, Brook JD. Virtual cloning and physical mapping of a human T-box gene, TBX4. *Genomics* 2000;67: 92-95.
124. Lander ES, Schork NJ. Genetic dissection of complex traits. *Science* 1994;265:2037-2048.
125. Wright AF, Carothers AD, Pirastu M. Population choice in mapping genes for complex diseases. *Nature Genet* 1999;23:397-404.
126. The International SNP Map Working Group. A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. *Nature* 2001;409:928-933.