Linkage Analysis of the Genetic Loci for High Myopia on 18p, 12q, and 17q in 51 U.K. Families

Jane E. Farbrother,[1] George Kirov,[2] Michael J. Owen,[2] Ricardo Pong-Wong,[3] Chris S. Haley,[3] and Jeremy A. Guggenheim[1]

Purpose. To determine the extent to which high myopia in a cohort of 51 U.K. families can be attributed to currently identified genetic loci.

Methods. The families comprised 245 subjects with phenotypic information and DNA available, of whom 170 were classified as affected. Subjects were genotyped for microsatellite markers spanning ~40cM regions on 18p (MYP2), 12q (MYP3) and 17q, together with markers flanking *COL2A1*, *COL11A1*, and *FBN1*. Two-point linkage analyses were performed using the same disease gene segregation model as was used in the original publications, followed by nonparametric and multipoint analyses using Genehunter (http://linkage.rockefeller. edu/soft/gh/ provided in the public domain by Rockefeller University, New York, NY), with additional maximization over the parameter α, the proportion of linked families.

Results. Evidence of linkage was found for the *MYP3* locus on 12q (two-point $Z_{max} = 2.54$, $P = 0.0003$ and multipoint hLOD = 1.08 at $\alpha = 0.24$, $P = 0.023$ for marker *D12S332*; nonparametric linkage [NPL] = 1.49, $P = 0.07$ for marker *D12S1607*). For the 17q locus there was weak evidence of excess allele sharing and linkage under a recessive model (NPL = 1.34, $P = 0.09$ for marker *D17S956*; two-point hLOD = 1.24 at $\alpha = 0.30$ for marker *D17S1795*; multipoint hLOD = 1.24 at $\alpha = 0.17$, $P = 0.014$ for marker at 77.68 cM, between markers *D17S956* and *D17S1853*). No significant linkage was found to the *MYP2* locus on 18p, or to the *COL2A1*, *COL11A1*, and *FBN1* genes.

Conclusions. These results suggest that the *MYP3* locus on 12q could be responsible for high myopia in approximately 25% of the U.K. families showing apparent autosomal dominant transmission, but that the loci on 18p and 17q are less common causes. Thus, additional loci for high myopia are likely to be the cause of the majority of cases of high myopia in the United Kingdom. (*Invest Ophthalmol Vis Sci.* 2004;45:2879–2885) DOI:10.1167/iovs.03-1156

The role of genes and environmental factors in the development of refractive errors is poorly understood. Twin and population studies suggest multifactorial inheritance, with additive polygenes exerting the greatest influence.[1-4] However, by analogy with other multifactorial diseases like diabetes and cancer, there may be "rare genes of large effect" segregating in certain families that give rise to extremes of refraction.[5] The strongest evidence that this is the case comes from genetic linkage studies, in which the continuous distribution of refractive error is dichotomized to give *affected* and *unaffected* groups corresponding to high myopes and non–high myopes, respectively.[6-9]

Dichotomizing on the basis of a refractive error of −6.00 D of myopia in the most minus meridian of the right eye (RE MMM), Young et al.[8] identified a locus for fully penetrant high myopia in a large family of Chinese descent (two-point LOD score 3.90 for marker *D18S1138* at 18p11.31). Six of the other seven families described in this report were also consistent with linkage to this locus, which has since been named *MYP2*. A follow-up study in Hong Kong[10] reported evidence of linkage to this locus in 5 of the 15 participating families, and an association study[11] showed that the *MYP2* locus was an important contributor to the inheritance of high myopia in an isolated community in Sardinia. A second locus[9] for fully penetrant nonsyndromic high myopia, *MYP3* on 12q, was found in a single, large, German-Italian family (two-point LOD score 3.85 for markers *D12S1706* and *D12S327* at 12q21-23). More recently, a third locus has been identified on 17q in a single large family of English-Canadian descent (two-point LOD score 3.17 for marker *D17S1604* on 17q21-22, dichotomization criterion −5.00 D RE MMM)[7] and there has been suggestive evidence for a fourth locus on 7q in a study of 23 small families in France.[6] Linkage analysis in 53 nuclear families of juvenile-onset myopes by Mutti et al.[12] found no evidence of involvement of the *MYP2* or *MYP3* loci (dichotomization criterion −0.75 D RE least minus meridian with onset before 17 years of age).

Therefore, it is known from molecular genetic studies that at least three high-penetrance loci exist for autosomal dominant high myopia. However, it is not clear to what extent these genes of large effect are responsible for causing high myopia in the general population. To address this question, we assessed the extent of genetic linkage to the 18p, 12q, and 17q loci in 51 U.K. pedigrees in which at least two siblings had myopia of greater than −6.00 D. These results have been presented in abstract form (Farbrother JE, et al. *IOVS* 2003;44:ARVO E-Abstract 4780).

METHODS

Ethical approval for these studies was granted by the Human Science Ethical Committee, Cardiff University and the local National Health Service Research Ethics Committees for subjects recruited from U.K. hospitals. All participants provided signed, informed consent. The research adhered to the tenets of the Declaration of Helsinki.

Pedigrees containing at least two siblings with myopia of $\leq$ −6.00 D in the least-minus meridian of both eyes were recruited for the study with the help of optometrists distributed across the United Kingdom and ophthalmologists at selected U.K. hospitals. (Seven of the families recruited from optometric practice were ascertained as part of a

Corresponding author: Jeremy A. Guggenheim, School of Optometry and Vision Sciences, Cardiff University, Redwood Building, King Edward VII Avenue, Cardiff CF10 3NB, UK; guggenheim@cardiff.ac.uk.

TABLE 1. 18p Linkage Results

Marker	cM	Two-Point										Multipont	
		Total LOD Score (θ)							θ_{max}	Z_{max}	~Max hLOD	hLOD	NPL
		0.00	0.01	0.05	0.1	0.2	0.3	0.4					
D18S59	0.00	$-\infty$	−18.21	−8.39	−4.19	−0.93	0.02	0.12	0.37	0.13	0.29 $\alpha = 0.15$	0.160, $\alpha = 0.10$	0.697
D18S476	2.84	$-\infty$	−13.68	−6.80	−3.67	−1.13	−0.29	−0.07	—	—	—	0.146, $\alpha = 0.09$	0.599
D18S481	6.94	$-\infty$	−17.06	−8.13	−4.25	−1.19	−0.23	−0.03	—	—	—	0.140, $\alpha = 0.09$	0.012
D18S63	8.30	$-\infty$	−23.07	−11.61	−6.60	−2.38	−0.76	−0.16	—	—	—	0.000	−0.334
D18S1132	11.21	$-\infty$	−19.82	−10.46	−6.12	−2.29	−0.75	−0.15	—	—	—	0.000	−0.028
D18S452	18.70	$-\infty$	−25.70	−13.01	−7.36	−2.60	−0.79	−0.13	—	—	—	0.000	0.179
D18S1163	24.08	$-\infty$	−10.35	−5.26	−2.87	−0.87	−0.17	0.01	0.44	0.02	0.02 $\alpha = 0.50$	0.000	0.315
D18S464	31.17	$-\infty$	−13.65	−6.34	−3.29	−0.93	−0.23	−0.07	—	—	—	0.000	0.811
D18S1158	38.92	$-\infty$	−3.45	−0.71	0.39	0.91	0.69	0.32	0.20	0.91	0.91 $\alpha = 1.00$	0.039. $\alpha = 0.07$	0.850

Parametric analysis model autosomal dominant (AD), $f_{DD} = 1.0$, $f_{Dd} = 1.0$, $f_{dd} = 0.0$, $P = 0.0133$. All families consistent with this model ($n = 42$) are included.

contemporaneous investigation into the genetic epidemiology of high myopia[13].) All families meeting these criteria who agreed to participate in the study were included. Six of the 51 families were bilineal (that is, high myopia was present in more than one of the unrelated founder or married-in individuals). These families were not excluded, since simulation studies suggest that bilineality has little effect on the mean maximum LOD score and the estimate of the recombination fraction.[14] The ocular history and current spectacle or contact lens prescription was obtained from each subject's optometrist. For subjects who had undergone cataract removal, presurgery spectacle prescriptions were obtained. The family of any subject with high myopia as part of a known syndrome or associated with a systemic or ocular condition was excluded from the study. The 51 families comprised 306 individuals, with phenotypic information available for 254. According to the criteria used by Young et al.,[8,9] 173 subjects were classed as affected. DNA was available for 245 subjects, of whom 170 were affected. The subjects' phenotype information is presented in Supplemental Table S1 at www.iovs.org/cgi/content/full/45/9/2879/DC1. Pedigree diagrams are presented in Supplemental Figure S1 at www.iovs.org/cgi/content/full/45/9/2879/DC1.

DNA was obtained by postal collection of saline mouthwashes.[15] Microsatellite repeat sequences were selected for genotyping on the basis of their genetic location within and adjacent to the previously linked regions on 18p, 12q, and 17q.[7–9] In addition, markers flanking the genes at which causal mutations have been identified for Stickler syndrome type I (*COL2A1*),[16] Stickler syndrome type II (*COL11A1*),[17] and Marfan syndrome (*FBN1*)[18] were genotyped to exclude linkage to these regions, under the assumption that these syndromes are the most prevalent systemic causes of high myopia.[19] Physical locations were established for all but one of the markers.

For genotyping, PCR reaction volumes of 12 µL comprised 1× PCR buffer (Promega, Madison, WI), 200 µM each dNTP (Amersham-Pharmacia, Piscataway, NJ), 1.5 mM MgCl$_2$, 0.6 or 0.3 µM of each primer (Sigma-Genosys, Woodland, TX), 0.025 U *Taq* DNA polymerase (Promega) and ~14 ng template genomic DNA. One primer was 5′ labeled with a fluorescent dye (FAM, TET, or HEX). Microsatellites were PCR amplified individually, multiplexed along with TAMRA 350 size standards (Perkin Elmer, Boston, MA) and electrophoresed on a 10% denaturing polyacrylamide gel (LongRanger acrylamide; BMA Bioproducts, Lowell, MA) on an automated sequencer and analyzed with a model 373 DNA sequencer running Genescan 3.0 software (Applied Biosystems, Foster City, CA). A sample of DNA pooled from 83 control subjects was included on each plate to standardize the allele size calling between gels. Alleles were sized manually using the peak labeling function of the software program (Genotyper 1.1; ABI).

Allele frequencies were calculated from the founder and married-in individuals ($n = 87$). Parametric, two-point linkage analysis was performed using MLINK.[20,21] The disease gene segregation models used

by Young et al.[8,9,22] were used for the parametric analyses. For the regions on chromosome 18p and 12q, this consisted of a 100% penetrant, dominant disease gene with frequency 0.0133 and no phenocopies. For the region on 17q and the Marfan and Stickler loci, this consisted of a 90% penetrant, dominant disease gene, with a frequency of 0.0133 and no phenocopies. The presence of heterogeneity in the participating families was considered likely, and hence the A test[23,24] was performed using the Table program to maximize the two-point LOD scores over the additional parameter α, the proportion of linked families. GeneHunter 2.1[25] was used to perform nonparametric linkage analyses and calculate multipoint LOD scores. The data for all markers were checked for Mendelian errors with PedCheck,[26] and unlikely double recombinants were checked for maximum-likelihood haplotypes generated by Genehunter 2.1. Statistical significance for the multipoint heterogeneity LOD scores and power were estimated from analysis of simulated marker data generated for the pedigrees by SLINK[27,28] for an eight-allele marker with equal allele frequencies at $\alpha = 0.0$ and $\theta = 0.00$, respectively.

After the planned linkage analyses had been performed (i.e., using linkage models that replicated those used for the discovery of the 18p, 12q, and 17q loci), a series of exploratory linkage analyses were undertaken to examine the effect of analyzing refractive error as a quantitative trait. Refractive error in the least minus meridian (RE LMM) was chosen as the quantitative trait of interest, rather than mean spherical equivalent (MSE) or RE MMM, to avoid potential bias from cylinder powers.[29] Refractions were transformed using a Box-Cox type function[30] {$RE_{TRANS} = \text{sign}(RE\ LMM) \times [(|RE\ LMM| + 1)^{0.415} - 1] \times 5$} which has the property of removing the kurtosis and most of the skew from the refractive error distribution of the general population. Linkage analysis was performed using Merlin[31] (for standard quantitative trait locus [QTL] or variance components analysis) or Merlin-Regress, which implements an extension of the Haseman-Elston procedure.[32] The correlation in refractive error between eyes was $r_s = 0.95$ ($P < 0.0001$). Quantitative analyses repeated with the average refractive error in the two eyes used as the quantitative trait instead of that in just the right eye produced very similar results.

RESULTS

The linkage analysis results are presented in Tables 1 through 5 and summarized in Figure 1.

MYP2 Locus on 18p

There was no statistically significant evidence of linkage or excess allele sharing for any of the markers to a fully penetrant disease gene at the 18p locus (Table 1). From power simulations, for a fully penetrant dominant disease gene, a heteroge-

TABLE 2. 12q Linkage Results

| | | Two-Point | | | | | | | | | | Multipont | |
| | | Total LOD Score (θ) | | | | | | | | | | | |
Marker	cM	0.00	0.01	0.05	0.1	0.2	0.3	0.4	θ_{max}	Z_{max}	~Max hLOD	hLOD	NPL
D12S80	83.19	$-\infty$	-15.64	-7.11	-3.56	-0.84	-0.04	0.08	0.38	0.08	0.08 $\alpha = 1.00$	0.000	-0.412
D12S326	86.40	$-\infty$	-14.80	-6.69	-3.26	-0.66	0.03	0.08	0.36	0.09	0.13 $\alpha = 0.20$	0.000	-0.197
D12S64	89.42	$-\infty$	-12.83	-6.07	-3.14	-0.78	0.01	0.12	0.38	0.12	0.12 $\alpha = 1.00$	0.018, $\alpha = 0.03$	0.175
D12S316	94.49	$-\infty$	-14.13	-6.37	-3.12	-0.62	0.02	0.03	0.34	0.06	0.20 $\alpha = 0.15$	0.248, $\alpha = 0.13$	0.563
D12S348	100.92	$-\infty$	-13.18	-6.18	-3.24	-0.95	-0.25	-0.08	—	—	—	0.252, $\alpha = 0.13$	1.025
D12S332	105.18	$-\infty$	-2.90	1.00	2.26	2.38	1.51	0.51	0.15	2.54	2.60 $\alpha = 0.45$	1.083, $\alpha = 0.24$	1.438
D12S1607	107.86	$-\infty$	-9.46	-2.12	0.47	1.67	1.24	0.45	0.21	1.67	1.67 $\alpha = 1.00$	1.053, $\alpha = 0.24$	1.487
D12S78	111.87	$-\infty$	-15.52	-5.93	-2.23	0.09	0.36	0.10	0.28	0.37	0.72 $\alpha = 0.20$	0.931, $\alpha = 0.22$	1.290
D12S1605	116.66	$-\infty$	-10.63	-4.33	-1.76	0.07	0.42	0.23	0.30	0.42	0.43 $\alpha = 0.25$	0.695, $\alpha = 0.20$	0.587
D12S354	123.77	$-\infty$	-17.18	-8.01	-4.13	-1.11	-0.20	-0.03	—	—	—	0.361, $\alpha = 0.14$	0.841

Parametric analysis model autosomal dominant (AD), $f_{DD} = 1.0$, $f_{Dd} = 1.0$, $f_{dd} = 0.0$, $P = 0.0133$. All families consistent with this model ($n = 42$) included.

neity LOD score (hLOD) of ≥ 0.5 would be expected to occur 87% of the time if 60% of the families were linked and 68% of the time if 40% of the families were linked.

MYP3 Locus on 12q

For the MYP3 locus, the maximum two-point LOD score (Z_{max}) was 2.54 at $\theta = 0.15$ for *D12S332* (asymptotic $P = 0.0003$; Table 2). The multipoint analysis results in this region were lower than the two-point results, but remained positive. The maximum multipoint hLOD was 1.08, for $\alpha = 24\%$ for marker *D12S332* (2-*df*, asymptotic $P = 0.041$; empiric estimate $P = 0.023$). The maximum multipoint nonparametric linkage (NPL) was 1.49 for marker *D12S1607* ($P = 0.07$). Haplotype assessment demonstrated that 21 of the 42 families were consistent with linkage, but of these only one pedigree (number 24) had a multipoint LOD exceeding 1.0. Recombination events in the families defined a linked region between the markers *D12S348* and *D12S1605*.

17q Locus

Two-point and multipoint analyses provided no evidence of linkage to a 90% penetrant dominant disease gene on 17q (Table 3). However, because of tentative evidence of excess allele sharing (NPL of 1.34 at marker *D17S956*, $P = 0.09$), and because the dominant model was not consistent with linkage, parametric analyses were repeated under a recessive model, maintaining 90% penetrance and a disease gene frequency of

0.0133 (Table 4). Two-point and multipoint results were higher under this model (maximum hLOD = 1.24, $\alpha = 0.30$ for *D17S1795*, maximum multipoint hLOD = 1.24 between markers *D17S956* and *D17S1853*, $\alpha = 0.17$; 2-*df*, asymptotic $P = 0.029$; empiric estimate $P = 0.014$).

Connective Tissue Disease Genes

Markers adjacent to the Stickler type I (*COL2A1*), type II (*COL11A1*), and Marfan syndrome (*FBN1*) loci produced no evidence for linkage or allele sharing (Table 5). However, the possibility of isolated families in the cohort linked to one of these loci could not be excluded.

Analysis of Myopia as a Quantitative Trait

After assessing the evidence for linkage to the three known myopia loci with linkage models that replicated those used to identify the loci originally, we performed a series of exploratory linkage analyses in which refractive error was examined as a quantitative trait. There was no evidence of linkage to any of the three loci, when using variance components analysis (Merlin-VC) or an extension of the Haseman-Elston regression method (Merlin-Regress). Using the QTL analysis method implemented in Merlin, there was no evidence of linkage at the 18p ($Z_{max} = 0.76$ at 0 cM, $P = 0.20$) or 12q ($Z_{max} = 1.10$ at 111.87 cM, $P = 0.14$) loci, but there was suggestive evidence of linkage at the 17q locus ($Z_{max} = 2.42$ at 68.44 cM, $P = 0.008$). Because of the exploratory nature of these quantitative

TABLE 3. 17q Linkage Results for Dominant Disease Model

| | | Two-Point | | | | | | | | | | Multipont | |
| | | Total LOD Score (θ) | | | | | | | | | | | |
Marker	cM	0.00	0.01	0.05	0.1	0.2	0.3	0.4	θ_{max}	Z_{max}	~Max hLOD	hLOD	NPL
D17S1868	64.16	-13.98	-8.80	-3.45	-1.02	0.58	0.66	0.28	0.25	0.72	0.66 $\alpha = 0.75$	0.047, $\alpha = 0.06$	1.183
D17S1795	68.44	-22.33	-16.06	-8.25	-4.42	-1.24	-0.21	0.01	0.42	0.02	0.02 $\alpha = 0.75$	0.061, $\alpha = 0.07$	1.151
D17S956	73.62	-9.61	-6.44	-2.96	-1.22	0.07	0.25	0.09	0.28	0.26	0.27 $\alpha = 0.55$	0.013, $\alpha = 0.03$	1.342
D17S1853	80.38	-21.63	-15.65	-8.06	-4.24	-1.05	-0.01	0.17	0.39	0.17	0.17 $\alpha = 1.00$	0.000	1.273
D17S1290	82.00	-29.30	-22.46	-12.52	-7.21	-2.48	-0.70	-0.12	—	—	—	0.000	1.282
D17S942	85.94	-20.55	-16.26	-9.85	-6.20	-2.66	-1.05	-0.30	—	—	—	0.004, $\alpha = 0.02$	1.250
D17S2059	93.27	-11.02	-8.56	-4.43	-2.19	-0.32	0.21	0.21	0.34	0.25	0.21 $\alpha = 0.90$	0.001, $\alpha = 0.01$	0.620
D17S1831	97.60	-20.31	-14.82	-8.26	-4.88	-1.81	-0.59	-0.13	—	—	—	0.000	0.155
D17S1817	103.53	-24.27	-18.50	-10.90	-6.65	-2.55	-0.78	-0.11	—	—	—	0.000	0.355

Parametric analysis model autosomal dominant (AD), $f_{DD} = 0.9$, $f_{Dd} = 0.9$, $f_{dd} = 0.0$, $P = 0.0133$. All families included ($n = 51$).

TABLE 4. 17q Linkage Results For Recessive Disease Model

| | | Two-Point | | | | | | | | | | Multipont | |
| | | Total LOD Score (θ) | | | | | | | | | | | |
Marker	cM	0.00	0.01	0.05	0.1	0.2	0.3	0.4	θ_{max}	Z_{max}	~Max hLOD	hLOD	NPL
D17S1868	64.16	−38.53	−18.39	−6.90	−2.59	0.11	0.44	0.19	0.28	0.45	0.44 α = 1.00	0.586, α = 0.17	1.183
D17S1795	68.44	−44.90	−19.96	−6.93	−2.17	0.61	0.77	0.30	0.26	0.85	1.24 α = 0.30	0.688, α = 0.17	1.151
D17S956	73.62	−45.94	−22.30	−8.87	−3.63	−0.13	0.47	0.23	0.30	0.47	0.53 α = 0.25	0.555, α = 0.15	1.342
D17S1853	80.38	−51.58	−24.42	−9.27	−3.54	0.10	0.61	0.27	0.29	0.62	0.95 α = 0.15	1.172, α = 0.13	1.273
D17S1290	82.00	−58.20	−30.08	−13.18	−6.37	−1.33	0.01	0.13	0.37	0.15	0.64 α = 0.10	1.086, α = 0.12	1.282
D17S942	85.94	−49.11	−26.42	−12.58	−6.71	−2.03	−0.43	0.02	0.44	0.05	0.02 α = 0.75	0.000	1.250
D17S2059	93.27	−44.47	−25.78	−12.75	−7.07	−2.41	−0.72	−0.13	—	—	—	0.000	0.620
D17S1831	97.60	−52.32	−28.20	−13.09	−6.91	−2.12	−0.51	−0.03	—	—	—	0.000	0.155
D17S1817	103.53	−46.28	−25.73	−11.68	−5.85	−1.52	−0.27	−0.01	—	—	—	0.000	0.355

Parametric analysis model autosomal recessive (AR), f_{DD} = 0.9, f_{Dd} = 0.0, f_{dd} = 0.0, P = 0.0133. All families included (n = 51).

analyses and the multiple testing that this entailed, we consider the absolute values of these Z-scores and their associated probabilities to be less meaningful than their relative levels.

DISCUSSION

To our knowledge, this study is the largest investigation of linkage analysis in highly myopic pedigrees reported to date. In simulations with 500 replicates, the study's power to detect a LOD score ≥ 3 in the absence of locus heterogeneity was 100% in all models. However, the power to detect a locus by using linkage analysis inevitably decreased as the proportion of linked families (α) decreased. For our families, the power to detect an hLOD ≥ 3 exceeded 80% when 80% of families were linked, but decreased to 20% to 25% when only 40% of families were linked, depending on the exact linkage model used. Only five families (pedigrees 24, 30, 32, 33, and 50) were of sufficient size to produce a LOD score ≥ 1 in isolation. Hence, although consistency with linkage can be described in the case of individual families, this is not conclusive evidence of inheritance due to the locus in question.

MYP2 Locus on 18p

Of the 42 families consistent with fully penetrant autosomal dominant transmission, linkage to *MYP2* (telomeric to *D18S52*)[33] could be excluded in 21, in which maximum multipoint LOD scores were negative and haplotype analysis was inconsistent with linkage. Coupled with the absence of significant linkage and excess allele-sharing across the entire data

set, the findings strongly suggest that the *MYP2* locus is not responsible for most cases of fully penetrant autosomal dominant high myopia in the United Kingdom. Because seven of the eight pedigrees examined by Young et al.[8] in their original report were linked to *MYP2*, it could be that the etiology of high myopia in the United Kingdom is different from that in the United States. Alternatively, the high proportion of pedigrees linked to the *MYP2* locus in the cohort studied by Young et al. could have occurred by chance. Consequently, fully penetrant disease allele(s) at the MYP2 locus may be rarer in many population groups than originally seemed the case.

MYP3 Locus on 12q

The multipoint linkage results for the MYP3 locus just reached statistical significance when considered in isolation (P = 0.023). However, once multiple testing is taken into account, the possibility that this result is a false positive becomes more likely. If seeking to replicate linkage to three separate loci is considered as performing three independent statistical tests, the results did not reach statistical significance (empiric estimate P = 0.07). If the weak linkage signal is not a false-positive result, then this study suggests that the *MYP3* locus could be a relatively common cause of familial high myopia in the United Kingdom. Our estimate of the proportion of linked families, α = 0.24 (95% confidence interval; 0.02–0.48), suggests that approximately one quarter of seemingly autosomal dominant families are linked to the *MYP3* locus. The 95% confidence interval suggests that it is unlikely that either none of the families, or more than half of the families are linked to this

TABLE 5. Connective Tissue Gene Linkage Results

| | | Two-Point | | | | | | | | | | Multipont | |
| | | Total LOD Score (θ) | | | | | | | | | | | |
Marker	cM	0.00	0.01	0.05	0.1	0.2	0.3	0.4	θ_{max}	Z_{max}	~Max hLOD	hLOD	NPL
FBN1													
D15S992	45.62	−28.48	−21.06	−12.00	−7.18	−2.73	−0.89	−0.17	—	—	—	0.000	−1.508
D15S1003	47.85	−22.54	−16.85	−9.58	−5.74	−2.23	−0.77	−0.16	—	—	—	0.000	−1.396
COL2A1													
D12S1701	62.54	−25.16	−18.30	−10.03	−5.87	−2.17	−0.69	−0.12	—	—	—	0.000	−0.890
D12S361	64.96	−26.71	−19.62	−10.93	−6.41	−2.29	−0.64	−0.04	—	—	—	0.000	−1.038
COL11A1													
D1S2626	136.34	−24.97	−17.54	−10.04	−6.10	−2.39	−0.86	−0.23	—	—	—	0.000	−0.047
D1S2888	136.88	−31.76	−23.50	−14.12	−9.04	−4.05	−1.71	−0.54	—	—	—	0.000	−0.358

Parametric analysis model autosomal dominant (AD), f_{DD} = 0.9, f_{Dd} = 0.9, f_{dd} = 0.0, P = 0.0133. All families included (n = 51).
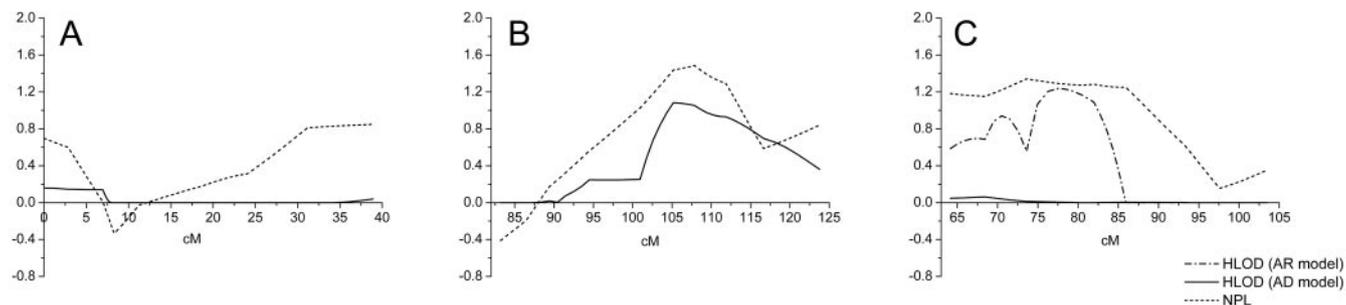
FIGURE 1. Multipoint linkage results for regions (**A**) 18p11.32-11.22, (**B**) 12q21.1-24.21, and (**C**) 17q21.32-25.1. Genetic map distances from the Marshfield database are shown on the *y*-axes.

locus. However, even in view of the wide confidence interval, our estimate of the proportion of linked families must be considered tentative, not least since the accuracy of the maximum likelihood estimate of α is dependent on the accuracy of the linkage model chosen (and this is subject to inevitable uncertainty regarding the criteria used to define affectation status, penetrance, and disease gene frequency) as well as the characteristics of the alternative unlinked loci.[34,35] Assessment of haplotypes demonstrated that 21 of the 42 families were consistent with linkage, but of these only one pedigree (number 24) had a multipoint LOD exceeding 1.0. Recombination events in these families suggest the disease gene is harbored within the narrowed interval (~15 cM) between markers *D12S348* and *D12S1605*.

The difference between the two-point and multipoint results for *MYP3* is explained in part by the extra identity-by-descent (IBD) information available in the analysis of multiple markers. In pedigree 33 in particular, additional data for the centromeric markers reduced the likelihood that all the affected individuals shared the region between *D12S348* and *D12S1605*. Lower multipoint results could also have occurred as a consequence of the constraint of θ, in combination with heterogeneity in the data set.[36] The data were rigorously checked for genotyping and marker map distance errors to exclude these as a potential cause of the two-point versus multipoint LOD score disparity.

None of the myopes among the 80 unaffected subjects included in the linkage analysis were younger than 18 years at the time of the eye examination on which their classification was based, and, as such, their status could be determined with relative confidence[37] (it should be noted that 5 of these 80 unaffected subjects were not genotyped). However, because of the potential for the measurement of spectacle refractions to vary from optometrist to optometrist, the effect of changing the affectation threshold was tested. Increasing the stringency of the threshold by 0.50 D increased the multipoint hLOD by 49% and the NPL statistic by 32%, whereas reducing the stringency by 0.50 D reduced the multipoint hLOD by 15% but increased the NPL statistic by 5%. However, the maxima remained close to marker *D12S332* and the proportion of linked families stayed between 20% and 30%.

17q Locus

The suggestion of excess allele sharing found using nonparametric analysis, even though linkage was absent in the original, dominant parametric model,[22] prompted us to investigate linkage under a recessive model for this locus. When considered in isolation, the evidence for linkage with this recessive model reached statistical significance ($P = 0.014$) and explained the excess allele sharing. After accounting for multiple testing, the result would be of only borderline statistical significance. Yet the fact that, under a quantitative trait model, the linkage signal

for this locus was the highest of the three suggests that it could either be the cause of the high myopia in a small number of the families studied or could act as a susceptibility gene in a larger proportion of them. It is conceivable that allelic variants in the *same* gene at the locus on 17q reported by Paluru et al.[7] and Young et al.[22] could give rise to either dominant or recessive inheritance in different families. Alternatively, our results would be consistent with the presence of a recessive or additive locus for high myopia close to the previously identified high penetrance dominant locus.

Analysis of Myopia as a Quantitative Trait

The highly selective ascertainment scheme adopted in this study meant that the degree of phenotypic variability in refractive error in these pedigrees was limited. This meant that only one of the three QTL analysis methods tested, the nonparametric QTL analysis option of Merlin, was strictly appropriate for assessing linkage. (Both the variance components and the regression-based analysis methods are designed for traits with a normal distribution and may be sensitive to deviations from normality. In addition, variance components analysis is better suited to nonselectively ascertained cohorts or cohorts ascertained using selection from both tails of the distribution, rather than the one-sided selection scheme used in the present study.[38] Finally, the regression analysis approach is highly sensitive to misspecification of trait parameters.[32]) Using Merlin-QTL, all three loci showed at least a weak linkage signal, with that from the 17q locus being the strongest. We interpret these quantitative analyses as backing up the results of the main linkage analyses, in suggesting that whereas the three loci examined may contribute to high myopia in these U.K. pedigrees, they must either act only in a minority of families, or harbor susceptibility genes, rather than high-penetrance myopia genes.

Implications for Future Linkage Studies

The 51 U.K. pedigrees examined herein were ascertained on the basis of an affected sibling pair, which would be expected to enrich the sample for high penetrance autosomal dominant disease genes. Yet the three high myopia loci for which convincing evidence of linkage has so far been obtained, all of which show apparent high-penetrance autosomal dominant transmission, appeared to be the cause of high myopia in only a minority of cases. This suggests either that additional high myopia loci remain to be discovered and that together these loci account for much of the high myopia in the general population or that most cases of high myopia are not genetic in origin. Our genetic epidemiology investigation of high myopia in the United Kingdom[13] reached the conclusion that high-penetrance dominant inheritance of high myopia is the exception rather than the rule. Thus, finding additional loci may be

difficult, requiring either the fortuitous discovery of further large families suitable for parametric linkage analysis (especially consanguineous pedigrees for mapping recessive high myopia genes) or very large collections of small pedigrees suitable for mapping "common genes of small effect." The results from a recent linkage study in a clinically unselected sample of dizygotic twins in which refractive error was analyzed as a quantitative trait were very encouraging (Hammond, et al. *IOVS* 2003;44:ARVO E-Abstract 2015). Such approaches should be complementary to those using extreme proband selection schemes like the affected sib-pair design, to determine whether the same genes or pathways underlie both high and low myopia.

Conclusions

The *MYP3* locus on 12q could be the cause of approximately 25% of apparent autosomal dominant high myopia in the United Kingdom, with the disease gene apparently situated within the narrowed interval between markers *D12S348* and *D12S1605*. The *MYP2* locus on 18p is likely to account for fewer cases of high myopia than previous studies have suggested and probably accounts for fewer cases than the *MYP3* locus. The recently reported locus on 17q also appears to be an infrequent cause of autosomal dominant high myopia, but may contribute to the risk of high myopia through a recessive, or perhaps an additive, model in some cases. This study strengthens the idea that high myopia should be considered a genetically complex disease.

Electronic Database Information

The following databases were used in the present study. All are available in the public domain:

GeneHunter: http://linkage.rockefeller.edu/soft/gh/ provided by Rockefeller University, New York, NY

Human Genome Browser: http://genome.ucsc.edu/; provided by the University of Santa Cruz, CA

MLINK program of the Human Genome Mapping Project: http://www.hgmp.mrc.ac.uk/ Human Genome Mapping Project Resources Centre, Cambridge, UK

Marshfield Clinic: http://research.marshfieldclinic.org/genetics/; provided by Marshfield Clinic Center for Medical genetics, Marshfield, WI

Merlin: http://csg.sph.umich.edu/pn/index.php?furl= /abecasis/Merlin/index.html; provided by the University of Michigan Center for Statistical Genetics, MI

National Center for Biotechnology Information http://www.ncbi.nlm.nih.gov/

Online Mendelian Inheritance in Man (OMIM) http://www.ncbi.nlm.nih.gov/Omin/; provided by the National Center for Biotechnology Information, Bethesda, MD

Pedcheck: http://watson.hgen.pitt.edu/register/docs/pedcheck.html; provided by Human Genetics, the University of Pittsburgh, PA

Note Added in Proof

The genetic locus for high myopia on 7q36 (OMIM 608367) identified by Naiglin et al.[6] has recently been designated MYP4, and the locus on 17q21-q22 (OMIM 608474) identified by Palaru et al.[7] designated MYP5.

References

1. Goss DA, Hampton MJ, Wickham MG. Selected review on genetic factors in myopia. *J Am Optom Assoc.* 1988;59:875–884.
2. Guggenheim JA, Kirov G, Hodson SA. The heritability of high myopia: a re-analysis of Goldschmidt's data. *J Med Genet.* 2000; 37:227–231.
3. Hammond CJ, Snieder H, Gilbert CE, Spector TD. Genes and environment in refractive error: the twin eye study. *Invest Ophthalmol Vis Sci.* 2001;42:1232–1236.
4. Lyhne N, Sjølie AK, Kyvik KO, Green A. 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.
5. Terwilliger JD, Göring HH. Gene mapping in the 20th and 21st centuries: statistical methods, data analysis, and experimental design. *Hum Biol.* 2000;72:63–132.
6. Naiglin L, Gazagne C, Dallongeville F, et al. A genome wide scan for familial high myopia suggests a novel locus on chromosome 7q36. *J Med Genet.* 2002;39:118–124.
7. Paluru P, Ronan SM, Heon E, et al. New locus for autosomal dominant high myopia maps to the long arm of chromosome 17. *Invest Ophthalmol Vis Sci.* 2003;44:1830–1836.
8. 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.
9. Young TL, Ronan SM, Alvear AB, et al. A second locus for familial high myopia maps to chromosome 12q. *Am J Hum Genet.* 1998; 63:1419–1424.
10. Lam DSC, Tam POS, Fan DSP, et al. Familial high myopia linkage to chromosome 18p. *Ophthalmologica.* 2003;217:115–118.
11. Heath S, Robledo 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.
12. Mutti DO, Semina E, Marazita M, et al. Genetic loci for pathological myopia are not associated with juvenile myopia. *Am J Hum Genet.* 2002;112:355–360.
13. Farbrother JE, Kirov G, Owen MJ, Guggenheim JA. Familial aggregation of high myopia: estimation of the sibling recurrence risk ratio. *Invest Ophthalmol Vis Sci.* 2004;45:2873–2878.
14. Durner M, Greenberg DA. Effect of heterogeneity and assumed mode of inheritance on LOD scores. *Am J Med Genet.* 1992;42: 271–275.
15. Garcia-Closas M, Egan KM, Abruzzo J, et al. Collection of genomic DNA from adults in epidemiological studies by buccal cytobrush and mouthwash. *Cancer Epidemiol Biomark Prev.* 2001;10:687–696.
16. Richards AJ, Baguley DM, Yates JRW, et al. Variation in the vitreous phenotype of Stickler syndrome can be caused by different amino acid substitutions in the position of the type II collagen Gly-X-Y triple helix. *Am J Hum Genet.* 2000;67:1083–1094.
17. Snead MP, Yates JRW. Clinical and molecular genetics of Stickler syndrome. *J Med Genet.* 1999;36:353–359.
18. Collod-Beroud G, Beroud C, Ades L, et al. Marfan database (third edition): new mutations and new routines for software. *Nucleic Acids Res.* 1998;26:229–233.
19. Marr JE, Halliwell-Ewen J, Fisher B, Soler L, Ainsworth JR. Associations of high myopia in childhood. *Eye.* 2001;15:70–74.
20. Lanthrop GM, Lalouel JM, Julier C, Ott J. Strategies for multilocus linkage analysis in humans. *Proc Natl Acad Sci USA.* 1984;81: 3443–3446.

21. Lanthrop GM, Lalouel JM, Julier C, Ott J. Multilocus linkage analysis in humans: detection of linkage and estimation of recombination. *Am J Hum Genet*. 1985;37:482–498.

22. Young T, Paluru P, Heon E, et al. A new locus for autosomal dominant high myopia maps to chromosome 17q21-23 (abstract). *Am J Hum Genet*. 2001;69:526.

23. Smith CAB. Homogeneity test for linkage data. *Proc Sec Int Cong Hum Genet*. 1961;1:212–213.

24. Smith CAB. Testing for heterogeneity of recombination fraction values in human genetics. *Ann Hum Genet*. 1963;27:175–182.

25. Kruglyak L, Daly M, Reeve-Daly M, Lander E. Parametric and nonparametric linkage analysis: a unified multipoint approach. *Am J Hum Genet*. 1996;58:1347–1363.

26. O'Connell JR, Weeks DA. PedCheck: a program for identification of genotype incompatibilities in linkage analysis. *Am J Hum Genet*. 1998;63:259–266.

27. Ott J. Computer-simulation methods in human linkage analysis. *Proc Natl Acad Sci USA*. 1989;86:4175–4178.

28. Weeks DE, Ott J, Lanthrop GM. SLINK: a general simulation program for linkage analysis (Abstract). *Am J Hum Genet*. 1990;47:A204.

29. Guggenheim JA, Farbrother JE. The association between spherical and cylindrical component powers: a statistical note. *Optom Vis Sci*. 2004;81:62–63.

30. Blackie CA, Harris WF. Refraction and keratometry: departures from and transformations toward multivariate normality. *Optom Vis Sci*. 1997;74:452–458.

31. Abecasis GR, Cherny SS, Cookson WO, Cardon LR. Merlin-rapid analysis of dense genetic maps using sparse gene flow trees. *Nature Genet*. 2002;30:97–101.

32. Sham PC, Purcell S, Cherny SS, Abecasis GR. Powerful regression-based quantitative-trait linkage analysis of general pedigrees. *Am J Hum Genet*. 2002;71:238–253.

33. Young TL, Atwood LD, Ronan SM, et al. Further refinement of the MYP2 locus for autosomal dominant high myopia by linkage disequilibrium analysis. *Ophthalmic Genet*. 2001;22:69–75.

34. Logue M, Vieland VJ. The heterogeneity lod cannot be used to estimate the population proportion of linked families (Abstract). *Genet Epidemiol*. 2000;19:259.

35. Vieland VJ, Logue M. HLODs, trait models, and ascertainment: implications of admixture for parameter estimation and linkage detection. *Hum Hered*. 2002;53:23–35.

36. Risch N, Guiffra L. Model misspecification and multipoint linkage analysis. *Hum Hered*. 1992;42:77–92.

37. Goss DA, Winkler RL. Progression of myopia in youth: age of cessation. *Am J Optom Physiol Opt*. 1983;60:651–658.

38. Lynch M, Walsh B. *Genetics and Analysis of Quantitative Traits*. Sunderland, MA: Sinauer; 1997.