Human Molecular Genetics, 2011, Vol. 20, No. 14 **2861–2868**
doi:10.1093/hmg/ddr169
Advance Access published on April 19, 2011

A genome-wide association study reveals association between common variants in an intergenic region of 4q25 and high-grade myopia in the Chinese Han population

Zhiqiang Li[1,†], Jia Qu[2,†], Xun Xu[3], Xiangtian Zhou[2], Haidong Zou[3], Ning Wang[3], Tao Li[1], Xiaohan Hu[1], Qian Zhao[1], Peng Chen[1], Wenjin Li[1], Ke Huang[1], Jun Yang[1], Zangdong He[7], Jue Ji[1], Ti Wang[1], Junyan Li[1], You Li[1], Jie Liu[1], Zhen Zeng[1], Guoyin Feng[5], Lin He[1,4,6,*] and Yongyong Shi[1,8,*]

[1]Bio-X Institutes, Key Laboratory for the Genetics of Developmental and Neuropsychiatric Disorders (Ministry of Education), Shanghai Jiao Tong University, Shanghai 200030, P.R. China, [2]Wenzhou Medical College, Wenzhou 325003, Zhejiang Province, P.R. China, [3]Department of Ophthalmology, Shanghai First People's Hospital, Affiliated Shanghai Jiao Tong University, Shanghai 200080, P.R. China, [4]Institute for Nutritional Sciences, Shanghai Institute of Biological Sciences, Chinese Academy of Sciences, Shanghai 200031, P.R. China, [5]Shanghai Institute of Mental Health, Shanghai 200030, P.R. China, [6]Institute of Biomedical Sciences, Fudan University, Shanghai 200032, P.R. China, [7]Division of Biostatistics of School of Medicine, Indiana University, Indianapolis, IN 46202, USA and [8]Institute of Neuropsychiatric Science and Systems Biological Medicine, Changning Mental Health Center, the Bio-X Institutes Hospital, 299 XieHe Road, Shanghai 200042, P.R. China

Received October 17, 2010; Revised and Accepted April 14, 2011

High-grade myopia (HM) is highly heritable, and has a high prevalence in the Han Chinese population. We carried out a genome-wide association study involving 102 HM cases suffering from retinal degeneration, and 335 controls who were free from HM and fundus diseases. Significant single-nucleotide polymorphisms were replicated in two follow-up studies: stage I involved 2628 independent cases and 9485 controls, and stage II involved a further 263 cases and 586 HM-free controls. The results were combined in a meta-analysis. Cases and controls were drawn from the Chinese Han population. A locus in an intergenic region at 4q25, within MYP11 (4q22–q27, OMIM: 609994), was found to be associated with HM (rs10034228, $P_{meta} = 7.70 \times 10^{-13}$, allelic odds ratio = 0.81, 95% confidence interval 0.76–0.86). There are no known genes in the region but a number of expressed sequence tags (ESTs) have been located there, one of which (BI480957) has been reported to express in the native human retinal pigment epithelium. In addition, a predicted gene was identified in this region. The gene's predicted protein sequence is highly similar to tubulin, beta 8 and beta-tubulin 4Q. Several previous studies have shown that tubulin plays an important role in eye development. Our result is compatible with a previous linkage study in the Han Chinese population (mapping in MYP11, 4q22–q27), and provides a more accurate locus for HM. Although there is insufficient evidence to indicate that expressed EST and the predicted gene play an important role in developing HM, this region merits further study as a candidate for the disease.

*To whom correspondence should be addressed at: Bio-X Institutes, Shanghai Jiao Tong University, the Little White House, 1954 Huashan Road, Shanghai 200030, P.R. China. Tel/Fax: +86 2162822491; Email: shiyongyong@gmail.com (Y.S.); Institute of Biomedical Sciences, Fudan University, 130 Dongan Road, Shanghai 200032, People's Republic of China. Email: helinhelin@gmail.com (L.H.)
†These authors contributed equally to this study.

INTRODUCTION

Myopia (nearsightedness) is one of the most common visual disorders affecting both children and adults. It is a refractive error in the eye which causes light rays to focus in front of the retina, making distant objects appear blurred, and is measured in diopters (D) using a corrective lens that focuses distant images on the retina. Myopia up to −6D is usually referred to as high-grade myopia (HM) (1). HM, sometimes, known as pathological myopia or degenerative myopia, is a progressive condition that can create problems because of its association with degenerative changes at the back of the eye, increasing the chance of developing irreversible visual impairment in the retina and the risk of glaucoma and cataracts (2,3).

In the populations of the western world, myopia alone causes ∼25% of uncorrected decreased vision (4). The myopia rates among East Asians, especially in the Chinese and Japanese populations, are much higher than in Europeans (5,6). Around 400 million Chinese (33% of the population, 1.5 times the world average) suffer from myopia, of whom about 20% are in the high-degree category, and the prevalence has increased markedly in recent years (http://news.xinhuanet.com/politics/2010-01/03/content_12745349.htm, Xinhua News Agency).

The etiology of myopia is known to be complex (7). Epidemiological studies have indicated the importance of environmental influences in the development of the disease (8,9). However, a significant number of studies, in particular, familial and twin studies, have identified genetic factors as playing an important role (10,11). Myopia is known to be inherited in different patterns: X-linked (12–15), autosomal dominant (16,17) and autosomal recessive (18,19). A considerable number of loci have been mapped for susceptibility, such as MYP1 (Xq28) (13,15), MYP2 (18p11.31) (20,21), MYP3 (12q) (22), MYP4 (7p15) (23,24), MYP5 (17q) (25), MYP6 (22q12) (26,27), MYP7 (11p13) (28), MYP8–MYP10 (3q26, 4q12 and 8p23) (28), MYP11 (4q22–q27) (29,30), MYP12 (2q37.1) (31), MYP13 (Xq23–q25) (14,32,33), MYP14 (1p36) (33), MYP15 (10q21.1) (34) and MYP16 (5p15) (35). The different patterns of inheritance and the variety of the genetic loci linked to myopia development may reflect different types of myopia.

To identify genetic loci associated with HM in the Chinese Han population, we conducted a genome-wide association study (GWAS, stage 1) using Affymetrix Genome-wide Human SNP Array 6.0 in 102 strictly defined HM cases and 335 controls free from myopia and fundus diseases, and subsequently replicated the most significantly associated loci in follow-up studies. The stage 2 follow-up study included an independent sample set of 2628 cases and 9485 randomly selected controls. We also carried out another replication study (stage 3) with 263 independent cases and 586 HM-free controls.

RESULTS

To identify possible stratification between cases and controls, we conducted a multidimensional scaling (MDS) analysis of an identity-by-state sharing matrix of all individuals (stage 1) plus HapMap samples. MDS analysis showed no obvious population substructure in our samples (Supplementary Material, Fig. S1). The genomic inflation factor ($\lambda = 1.065$) indicated a slight inflation of the GWAS results due to population stratification. To correct for potential population stratification, the observed test statistics were corrected using the genomic control method. A Manhattan plot of genetic signals for HM shows no global significance in stage 1, the P-value being $<5 \times 10^{-8}$ (Supplementary Material, Fig. S2). The corresponding Q–Q plot is shown in Supplementary Material, Figure S3.

Only the top three single-nucleotide polymorphisms (SNPs) had P-values $<10^{-6}$ (rs6842731, $P = 7.22 \times 10^{-7}$; rs3850476, $P = 9.16 \times 10^{-7}$; rs6845600, $P = 5.11 \times 10^{-7}$), and they were close to each other at 4q25 (distances range from 3 to 10 kb), which was consistent with a previous linkage study of a Chinese family (29), mapping an autosomal-dominant HM locus to MYP11 (4q22–q27). In the replication study, we therefore genotyped all 10 SNPs (within a size of ∼212 kb region), with P-values $<5 \times 10^{-4}$ in this highly significant region (including rs10034228, $P = 9.41 \times 10^{-5}$; Fig. 1, Table 1). Of the 10 SNPs within 4q25, 5 were validated in the stage 2 replication study, and the most significant SNP was rs10034228 ($P = 1.73 \times 10^{-9}$). In the stage 3 replication study, four were positive with rs10034228 still the most significant ($P = 6.47 \times 10^{-3}$). All these SNPs were in Hardy–Weinberg equilibrium (HWE) in cases and controls (Supplementary Material, Table S1). The genotype distributions for each of the SNPs in cases and controls with P-values for each stage are shown in Supplementary Material, Table S2.

In the combined study, we performed a meta-analysis of the data from the three stages. Heterogeneity across the three stages was evaluated using Cochran's Q test, and the P-values of Cochran's Q test for rs2218817, rs10034228, rs4440293, rs1585471 and rs6837348 were 0.003, 0.015, 0.004, 0.006 and 0.009, respectively. As all the samples were recruited from a relatively small region (the adjacent Shanghai and Zhejiang Province), and the number of studies was small, the meta-analysis was carried out using the Mantel–Haenzel method (36) with a fixed-effect model (37). A previous GWAS adopted the fixed-effect model with a significant heterogeneity, $P = 0.012$ (38). In addition, we produced combined results by directly merging the three-stage studies (Supplementary Material, Table S3). The results produced by the meta-analysis and directly merged analysis were similar. After combining the GWAS and replication data in the meta-analysis, the most significant association was seen at the SNP rs10034228 ($P_{meta} = 7.70 \times 10^{-13}$, allelic odds ratio = 0.81, 95% confidence interval (CI) 0.76–0.86). Nominal associations were also observed for rs2218817 ($P_{meta} = 4.16 \times 10^{-5}$), rs4440293 ($P_{meta} = 5.49 \times 10^{-5}$), rs1585471 ($P_{meta} = 2.14 \times 10^{-6}$) and rs6837348 ($P_{meta} = 3.21 \times 10^{-6}$).

The linkage disequilibrium (LD) structure of the loci (Fig. 1) shows that the five associated SNPs belong to two blocks (rs2218817 and rs10034228 locate in block 1 and rs4440293, rs1585471 and rs6837348 locate in block 2). The SNPs (rs6842731, rs3850476 and rs6845600) that were found to be most significantly associated in stage 1 belong to another block (block 3). The HapMap CHB LD data show that most of the SNPs in strong LD ($r^2 > 0.5$) with

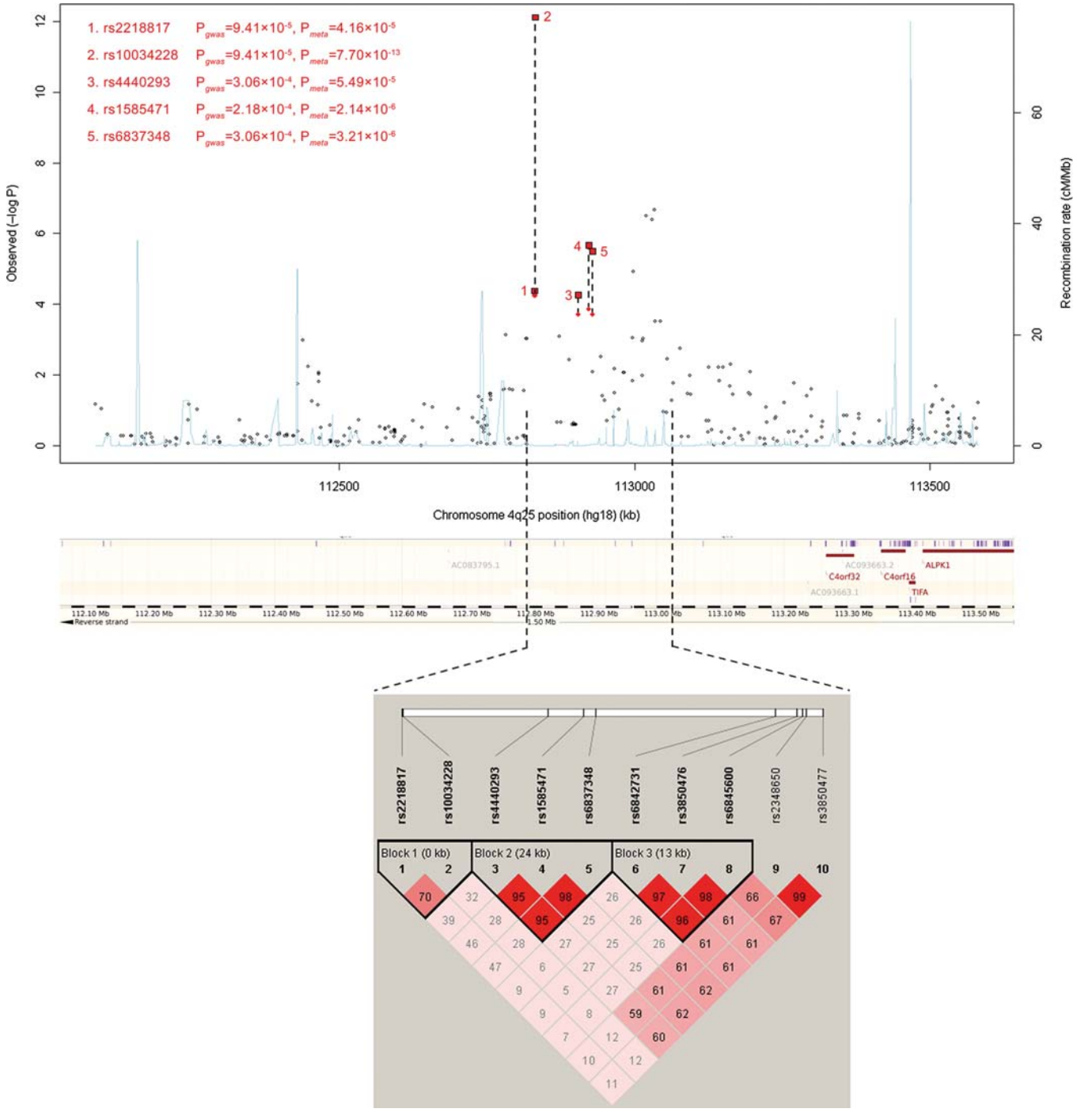

Figure 1. Details of the loci showing most significant association in the 4q25 region (1.5 Mb, centered on rs10034228). (**A**) Single-marker association statistics (as $-\log10\ P$; left *y*-axis) from the GWAS (circle) and from the combined analysis (squares). Recombination rates across each region in HapMap are shown in light blue (right *y*-axis). (**B**) Genomic context of this region, including genes (red or gray) and EST (purple). (**C**) Pairwise LD of the replication SNPs (measured by D′).

rs10034228 lie within 165 kb (15 kb upstream and 150 kb downstream). There are no known genes in the region, although a number of ESTs have been located there, one of which (BI480957) has been reported to be expressed in the native human retinal pigment epithelium (RPE) (39). In addition, the N-SCAN (40–42) program was used to identify a predicted gene (chr4.3.076.a) in this region (Supplementary Material, Fig. S4). The predicted protein sequence of chr4.3.076.a is highly similar to tubulin, beta 8 (bit score = 325, *e*-value = $2e^{-87}$) and beta-tubulin 4Q (bit score = 325, *e*-value = $2e^{-87}$). In addition, we searched for regulatory elements in the candidate region using the UCSC Genome

Table 1. Association test results for the replication SNPs

rs number [minor allele in case]	GWAS (102 cases/335 controls)			*P*-value[a]	Replication I (2628 cases/9485 controls)			*P*-value	Replication II (263 cases/586 controls)			*P*-value	Meta-analysis[b]	*P*-value
	Freq in case	Freq in control	Allelic OR (95% CI)		Freq in case	Freq in control	Allelic OR (95% CI)		Freq in case	Freq in control	OR (95% CI)		Allelic OR (95% CI)	
rs2218817 [C]	0.376	0.541	0.51 (0.37–0.71)	$\mathbf{9.41 \times 10^{-5}}$	0.478	0.502	0.91 (0.85–0.97)	$\mathbf{2.51 \times 10^{-3}}$	0.484	0.533	0.82 (0.67–1.01)	0.066	0.88 (0.83–0.94)	$\mathbf{4.16 \times 10^{-5}}$
rs10034228 [A]	0.376	0.541	0.51 (0.37–0.71)	$\mathbf{9.41 \times 10^{-5}}$	0.457	0.504	0.83 (0.78–0.88)	$\mathbf{1.73 \times 10^{-9}}$	0.463	0.537	0.74 (0.60–0.92)	$\mathbf{6.47 \times 10^{-3}}$	0.81 (0.76–0.86)	$\mathbf{7.70 \times 10^{-13}}$
rs4440293 [A]	0.392	0.544	0.54 (0.39–0.75)	$\mathbf{3.06 \times 10^{-4}}$	0.481	0.504	0.91 (0.85–0.97)	$\mathbf{4.45 \times 10^{-3}}$	0.470	0.531	0.78 (0.63–0.97)	**0.022**	0.88 (0.83–0.94)	$\mathbf{5.49 \times 10^{-5}}$
rs1585471 [A]	0.387	0.542	0.53 (0.38–0.74)	$\mathbf{2.18 \times 10^{-4}}$	0.475	0.504	0.89 (0.84–0.95)	$\mathbf{3.17 \times 10^{-4}}$	0.471	0.532	0.78 (0.64–0.96)	**0.021**	0.87 (0.82–0.92)	$\mathbf{2.14 \times 10^{-6}}$
rs6837348 [G]	0.392	0.544	0.54 (0.39–0.75)	$\mathbf{3.06 \times 10^{-4}}$	0.475	0.504	0.89 (0.84–0.95)	$\mathbf{1.27 \times 10^{-4}}$	0.481	0.536	0.80 (0.65–0.98)	**0.035**	0.87 (0.82–0.92)	$\mathbf{3.21 \times 10^{-6}}$
rs6842731 [A]	0.423	0.236	2.37 (1.70–3.33)	$\mathbf{7.22 \times 10^{-7}}$	0.257	0.265	0.95 (0.89–1.03)	0.23	0.243	0.231	1.06 (0.84–1.37)	0.59		
rs3850476 [C]	0.423	0.237	2.35 (1.68–3.30)	$\mathbf{9.16 \times 10^{-7}}$	0.253	0.263	0.95 (0.88–1.02)	0.16	0.238	0.233	1.03 (0.80–1.32)	0.82		
rs6845600 [A]	0.438	0.246	2.39 (1.71–3.33)	$\mathbf{5.11 \times 10^{-7}}$	0.261	0.273	0.94 (0.88–1.01)	0.09	0.238	0.243	0.98 (0.76–1.25)	0.84		
rs2348650 [C]	0.216	0.115	2.13 (1.40–3.23)	$\mathbf{4.63 \times 10^{-4}}$	0.119	0.123	0.97 (0.88–1.06)	0.48	0.140	0.113	1.27 (0.93–1.73)	0.13		
rs3850477 [A]	0.216	0.115	2.13 (1.40–3.23)	$\mathbf{4.63 \times 10^{-4}}$	0.118	0.124	0.95 (0.86–1.05)	0.29	0.138	0.112	1.26 (0.93–1.72)	0.14		

[a]The *P*-values are genomic control adjusted.
[b]Statistical results using the Mantel–Haenzel method as a fixed-effect model. *P*-values less than 0.05 are shown in bold.

Browser and found three enhancers (element_559, element_928 and element_926, Supplementary Material, Fig. S4) based on the experimental data set of conserved non-coding human sequences. The enhancer activity has been tested in transgenic mice (43).

DISCUSSION

We conducted a GWAS of HM in the Chinese Han population. A relatively small number of HM cases, but with extreme phenotypes, were included in the discovery phase (GWAS), and several SNPs within 4q25 were found to be strongly associated with HM, which is consistent with a previous Chinese HM linkage study. The top 10 most significantly associated SNPs within the region were therefore selected for replication in an extended and larger follow-up study, and rs10034228 exhibited genome-wide significance in stage 2 and in the combined stages.

As shown in Supplementary Material, Table S4, the sample age varied in these three stages. For HM patients, fundus changes are associated with a high refractive error, a condition which deteriorates over time. In stage 1, to ensure effective assessment of the severity of myopia, only older patients were selected for the samples. The case consisted of subjects whose myopia was worse than −8D and who exhibited fundus degeneration, and the controls were free from HM and fundus diseases. The *P*-value 4.15×10^{-7} for the older severe cases was quite high, considering only 102 cases were involved; however, they were high severity cases. In stage 2, adolescent cases were principally involved, and the severity of their condition was milder than in the older cases (−6D cut off and no requirement of fundus disease), while the controls were randomly selected samples. East Asian populations tend to have very high rates of adolescent of myopia, compared with the previous generation (44). The relatively weaker power of the larger cohort is probably explained by this recent increase in myopia rates in the younger population who have a milder form of the disease. A 25-fold larger replication cohort achieved a very significant $P = 1.73 \times 10^{-9}$ (rs10034228). In stage 3, the levels of myopia in the cases were somewhere between that of the previous two stages. Due to the limitations of sample size, the statistical power was weak and only nominal significant associations were observed. However, we noticed that the order of odds ratio for these stages was consistent with that of the severity of myopia cases. In general, the age of the samples had no effect on the outcome of the study, adding robustness to our results.

No known genes were found in the adjacent region of rs10034228. However, expressed sequence tags (ESTs) BI480957 within this region has been reported to be expressed in the human RPE, and the N-SCAN program has identified a predicted gene (chr4.3.076.a) in the region whose predicted protein sequence is highly similar to that of the tubulin family. Several previous studies have shown that tubulin plays an important role in eye development. A proteomic analysis of chicken retinas reported that the tubulin alpha-1 chain can act as a 'GO/GROW' signal for myopia (45). Good visual acuity requires an appropriate axial length of the ocular to focus the images of distant objects onto the

retina, but the 'GO/GROW' signals will accelerate axial growth and then induce myopia. A gene expression study identified tubulin alpha 3 (*TUBA3*) as a candidate gene for keratoconus (OMIM: 148300), a non-inflammatory bradytrophic human corneal disease, and found that bilateral distortion induces myopia and astigmatism (46). In addition, beta 3-tubulin is one of the most important neuronal markers, and expression of beta 3-tubulin can be detected in the developing retina (47,48).

Regulatory/enhancer activity has also been found within the candidate region. The activities with lacZ stain of element_559 and element_926 had been detected in midbrain and unidentified abdominal structure of the 11.5 day mouse embryo (E11.5), respectively. Element_928 did not show observable enhancer activities at E11.5 (43). Since an assay captures only a single embryonic time point, this does not imply activity for these enhancers at another time point. Equally, a weak enhancer at the time point may be below the detectable range at the level of resolution. However, we are convinced that some enhancers exist in the candidate region. An enhancer could affect the transcription of a gene which may be far away from it (sometimes several hundred thousand base pairs upstream or downstream of the start site). These elements may therefore lead to illegitimate transcription of the flanking genes (PITX2 and C4orf32). PITX2 encodes a member of the RIEG/PITX homeobox family, which is involved in the development of eyes, teeth and abdominal organs. Mutations in this gene are associated with Axenfeld–Rieger syndrome (49), iridogoniodysgenesis syndrome (50) and sporadic cases of Peters anomaly (51). All these syndromes are characterized by abnormal eye development.

Although there is insufficient evidence to indicate that expressed EST, the predicted gene, or the enhancers play an important role in developing HM, this region merits further study as a candidate for the disease.

Recently, three GWASs of HM have been reported. Nakanishi *et al.* (52) identified 11q24.1 as a susceptible locus for pathological myopia in a Japanese sample, and Solouki *et al.* (53) and Hysi *et al.* (54) reported 15q14 and 15q25 as susceptibility loci for refractive errors and myopia in the Caucasian population. However, none of the SNPs in these regions proved significant in our study with $P < 0.05$ (data not shown). This inconsistency in results across different studies may be due to the following factors: (i) genetic heterogeneity exists among ethnicities; therefore, the risk of a locus can be different in different populations. Obviously, our risk identification within MYP11 was not comparable with that in the sample size studied by Nakanishi *et al.*, Solouki *et al.* and Hysi *et al.*, and the risks of 11q24.1, 15q14 and 15q25 might not be large enough among ethnic Chinese to be detected by the relatively small sample size in our initial study. (ii) In particular, the design of our study was different from previous studies, as our initial GWAS focused on extreme cases suffering from severe myopia, i.e. worse than −8D and with retinal degeneration, and the comparison was with controls completely free from myopia and fundus diseases. Therefore, the exact phenotype we studied was not quite the same as that of previous studies. (iii) In general, HM is recognized as a complex disease. Lots of genes with low penetrance may be involved in the development of HM. Each initial GWAS focuses only on a random sample set, which can only represent part of the whole story thus running the risk of false positives or false negatives. Each individual study might 'touch' only one tip of the iceberg, but a large number of such 'touches' in the future may lead to an understanding of the overall pathogenesis of myopia.

In summary, we identified one HM susceptibility locus, rs10034228, with combined genome-wide significance in the Chinese Han population, located within a region (4q22−q27) mapped in a previous linkage study. Though there are no known genes near this SNP, potential functional evidence can be found in the surrounding genomic region. The identification of this locus should therefore provide new insights into the pathogenesis of HM. Our sample set had a limited possibility of detecting risk loci, but our findings strongly suggest that wider-targeted GWA analyses with much larger samples should be undertaken to identify additional common genetic risk factors and to shed light on the pathophysiology of physiological short-sightedness in humans.

MATERIALS AND METHODS

Participants

The characteristics of cases and controls are summarized in Supplementary Material, Table S4. The myopic spherical refractive power of all cases studied was six diopters or higher. (i) Stage 1 of the study focused on individuals with extreme phenotypes. For the GWAS sample set, all the samples were recruited from local Shanghai residents. The mean age of cases (38 males and 64 females) was 66.3 ± 8.1, and the mean age of controls (148 males and 187 females) was 74.4 ± 6.0. All cases were HM patients with myopia of −8D or less suffering from fundus injury, and with onset age <35. All controls were free of myopia and fundus diseases. Fundus photos were recorded for each sample in this stage (Supplementary Material, Fig. S5). (ii) For the first replication study (stage 2), the mean age of HM cases (1380 males and 1248 females) was 17.8 ± 1.5, with onset age <35 and the mean age of controls (4238 males and 5247 females) was 49.6 ± 11.2. Controls used in this stage were randomly selected from the general population. Both cases and controls were local Shanghai residents. (iii) For the second replication study (stage 3), all samples were recruited from the local Chinese Han population in Zhejiang Province. The mean age of HM cases (115 males and 148 females) was 24.0 ± 7.0 with onset age <35 and the mean age of controls (263 males and 323 females) was 42.5 ± 13.3. All controls for this stage were myopia-free.

Approval was received for the study from the local Ethics Committee of Human Genetic Resources and written informed consent was obtained from all participants in the study.

DNA extraction

All DNA samples were extracted from peripheral blood using a Tiangen DNA extraction kit according to the manufacturer's protocol. Genomic DNA was diluted to working concentrations of 50 ng/μl for genome-wide genotyping and 10 ng/μl for the replication step.

GWAS genotyping and quality control

The study was carried out using a two-tiered design, comprising a GWAS (stage 1) followed by the validation of formally significant association in larger, independent samples (stages 2 and 3). Genotyping for the GWAS was carried out by an Affymetrix service facility (South San Francisco, CA, USA) using the Affymetix Genome-Wide Human SNP Array 6.0. For internal quality control, Affymetrix adopts a preliminary 0.4 or greater Contrast QC value that is used by default to pass arrays for further data analysis. Final genotypes were generated using the birdseed algorithm (55). Those markers which had a missing genotyping rate >5%, a minor allele frequency <1% or deviated significantly from HWE ($P_{HWE} \leq 0.001$) in controls were excluded from subsequent analysis, resulting in a total final number of 681 783 SNPs.

Replication genotyping

Genotyping for the replication study was performed using the ligation detection reaction method in the laboratory, with technical support from the Shanghai Biowing Applied Biotechnology Company.

SNP selection for validation

The following criteria were used for the selection of SNPs for validation: since the most significant association signals in GWAS were observed on 4q25, the top 10 significant SNPs in this region were replicated. The detailed SNP QC information, according to the criteria listed in the GWAS Genotyping and Quality Control section, is given in Supplementary Material, Table S5. The SNP cluster graphs (Supplementary Material, Fig. S6) were created using Genotyping Console v. 3.0.2 (Affymetrix). The mapping between AFFY SNP ID and dbSNP RS ID, which might be needed in Supplementary Material, Figure S5, can be found in Supplementary Material, Table S5.

Statistical analysis

MDS analysis was conducted using the PLINK software package (v 1.03) (56). A representation of the first two dimensions in the MDS analysis is shown in the figure. We estimated a genomic inflation factor based on the median Chi-squared test in the case/control sample.

The GWAS, the basic case/control association test and the HWE analysis were performed using PLINK (v 1.03), and Haploview was used for the genome-wide *P*-value plot (57). The Q–Q plot was used to assess the number and magnitude of observed associations between genotyped SNPs and HM under study, compared with the association statistics expected under the null hypothesis of no association. The Q–Q plot was created using the R qq.plot function (58).

For the replication study, allelic association analysis of the replication sample was conducted using SHEsis (59), and the LD structure of the loci was visualized using Haploview (57).

Summary statistic combination

The meta-analysis was conducted using comprehensive meta analysis V2.0. The heterogeneity across the three stages was evaluated using the Cochran Q test. The meta-analysis was carried out using the Mantel–Haenzel method (36) with a fixed-effect model (37). In addition, we produced combined results by directly merging the three-stage studies.

SUPPLEMENTARY MATERIAL

Supplementary Material is available at *HMG* online.

ACKNOWLEDGEMENTS

We are deeply grateful to all the participants as well as to the doctors working on this project. The authors also thank editors and anonymous reviewers for their valuable comments on the manuscript.

Conflict of Interest statement. None declared.

FUNDING

This work was supported by the 863 Program (2009AA022701, 2006AA02A407), the 973 Program (2010CB529600, 2007CB914703, 2007CB947300), the National Key Technology R&D Program (2006BAI05A09), the Shanghai Municipal Health Bureau program 2008095, the Shanghai Changning Health Bureau program 2008406002, the Shanghai Municipal Commission of Science and Technology Program 09DJ1400601, the National Key Project for Investigational New Drugs 2008ZX09312-003 and the Shanghai Leading Academic Discipline Project (B205).

REFERENCES

1. Zejmo, M., Formińska-Kapu cik, M., Pieczara, E., Filipek, E., Mrukwa-Kominek, E., Samochowiec-Donocik, E., Leszczyński, R. and Smuzyńska, M. (2009) Etiopathogenesis and management of high-degree myopia. Part I. *Med. Sci. Monitor*, **15**, RA199–RA202.
2. Mitchell, P., Hourihan, F., Sandbach, J. and Jin Wang, J. (1999) The relationship between glaucoma and myopia: the Blue Mountains Eye Study. *Ophthalmology*, **106**, 2010–2015.
3. Wong, T.Y., Klein, B.E.K., Klein, R., Tomany, S.C. and Lee, K.E. (2001) Refractive errors and incident cataracts: the Beaver Dam Eye Study. *Invest. Ophthalmol. Vis. Sci.*, **42**, 1449–1454.
4. Be, R. (2004) The prevalence of refractive errors among adults in the United States, Western Europe, and Australia. *Arch. Ophthalmol.*, **122**, 495–505.
5. Hosaka, A. (1988) Population studies—myopia experience in Japan. *Acta Ophthalmol. Suppl.*, **185**, 37–40.
6. He, M., Zheng, Y. and Xiang, F. (2009) Prevalence of myopia in urban and rural children in mainland China. *Optom. Vis. Sci.*, **86**, 40–44.
7. Morgan, I. and Rose, K. (2005) How genetic is school myopia? *Prog. Retin. Eye Res.*, **24**, 1–38.
8. Saw, S.M., Chua, W.H., Hong, C.Y., Wu, H.M., Chan, W.Y., Chia, K.S., Stone, R.A. and Tan, D. (2002) Nearwork in early-onset myopia. *Invest. Ophthalmol. Vis. Sci.*, **43**, 332–339.
9. Wong, T.Y., Foster, P.J., Johnson, G.J. and Seah, S.K.L. (2002) Education, socioeconomic status, and ocular dimensions in Chinese adults: the Tanjong Pagar Survey. *Br. J. Ophthalmol.*, **86**, 963–968.

10. Hammond, C.J., Snieder, H., Gilbert, C.E. and Spector, T.D. (2001) Genes and environment in refractive error: the twin eye study. *Invest. Ophthalmol. Vis. Sci.*, **42**, 1232–1236.
11. Lyhne, N., Sjølie, A.K., Kyvik, K.O. and Green, A. (2001) 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.*, **85**, 1470–1476.
12. Wold, K.C. (1949) Hereditary myopia. *Arch. Ophthalmol.*, **42**, 225–237.
13. Bartsocas, C.S. and Kastrantas, A.D. (1981) X-linked form of myopia. *Hum. Hered.*, **31**, 199–200.
14. Zhang, Q., Guo, X., Xiao, X., Jia, X., Li, S. and Hejtmancik, J.F. (2006) Novel locus for X linked recessive high myopia maps to Xq23-q25 but outside MYP1. *Br. Med. J.*, **43**, e20.
15. Haim, M. and Fledelius, H.C. (1988) X-linked myopia in Danish family. *Acta Ophthalmol. (Copenh.)*, **66**, 450–456.
16. Franceschetti, A. (1953) Haute myopie avec decollement retinien hereditaire. *J. Hum. Genet.*, **2**, 283–284.
17. Francois, J. (1961) *Heredity in Ophthalmology*. Mosby.
18. Karlsson, J.L. (1975) Influence of the myopia gene on brain development. *Clin. Genet.*, **8**, 314–318.
19. Edwards, M. and Lewis, W.H.P. (1991) Autosomal recessive inheritance of myopia in Hong Kong Chinese infants. *Ophthalmic Physiol. Opt.*, **11**, 227–231.
20. Young, T.L., Ronan, S.M., Drahozal, L.A., Wildenberg, S.C., Alvear, A.B., Oetting, W.S., Atwood, L.D., Wilkin, D.J. and King, R.A. (1998) Evidence that a locus for familial high myopia maps to chromosome 18p. *Am. J. Hum. Genet.*, **63**, 109–119.
21. Young, T.L., Atwood, L.D., Ronan, S.M., Dewan, A.T., Alvear, A.B., Peterson, J., Holleschau, A. and King, R.A. (2001) Further refinement of the MYP2 locus for autosomal dominant high myopia by linkage disequilibrium analysis. *Ophthalmic Genet.*, **22**, 69–75.
22. Young, T.L., Ronan, S.M., Alvear, A.B., Wildenberg, S.C., Oetting, W.S., Atwood, L.D., Wilkin, D.J. and King, R.A. (1998) A second locus for familial high myopia maps to chromosome 12q. *Am. J. Hum. Genet.*, **63**, 1419–1424.
23. Paget, S., Julia, S., Vitezica, Z.G., Soler, V., Malecaze, F. and Calvas, P. (2008) Linkage analysis of high myopia susceptibility locus in 26 families. *Mol. Vis.*, **14**, 2566–2574.
24. Ciner, E., Wojciechowski, R., Ibay, G., Bailey-Wilson, J.E. and Stambolian, D. (2008) Genomewide scan of ocular refraction in African-American families shows significant linkage to chromosome 7p15. *Genet. Epidemiol.*, **32**, 454–463.
25. Paluru, P., Ronan, S.M., Heon, E., Devoto, M., Wildenberg, S.C., Scavello, G., Holleschau, A., Mäkitie, O., Cole, W.G. and King, R.A. (2003) New locus for autosomal dominant high myopia maps to the long arm of chromosome 17. *Invest. Ophthalmol. Vis. Sci.*, **44**, 1830–1836.
26. Stambolian, D., Ibay, G., Reider, L., Dana, D., Moy, C., Schlifka, M., Holmes, T., Ciner, E. and Bailey-Wilson, J.E. (2004) Genomewide linkage scan for myopia susceptibility loci among Ashkenazi Jewish families shows evidence of linkage on chromosome 22q12. *Am. J. Hum. Genet.*, **75**, 448–459.
27. Klein, A.P., Duggal, P., Lee, K.E., Klein, R., Bailey-Wilson, J.E. and Klein, B.E.K. (2007) Confirmation of linkage to ocular refraction on chromosome 22q and identification of a novel linkage region on 1q. *Arch. Ophthalmol.*, **125**, 80–85.
28. Hammond, C.J., Andrew, T., Tat Mak, Y. and Spector, T.D. (2004) A susceptibility locus for myopia in the normal population is linked to the PAX6 gene region on chromosome 11: a genomewide scan of dizygotic twins. *Am. J. Hum. Genet.*, **75**, 294–303.
29. Zhang, Q., Guo, X., Xiao, X., Jia, X., Li, S. and Hejtmancik, J.F. (2005) A new locus for autosomal dominant high myopia maps to 4q22-q27 between D4S1578 and D4S1612. *Mol. Vis.*, **11**, 554–560.
30. Li, Y.J., Guggenheim, J.A., Bulusu, A., Metlapally, R., Abbott, D., Malecaze, F., Calvas, P., Rosenberg, T., Paget, S. and Creer, R.C. (2009) An international collaborative family-based whole-genome linkage scan for high-grade myopia. *Invest. Ophthalmol. Vis. Sci.*, **50**, 3116–3127.
31. Paluru, P.C., Nallasamy, S., Devoto, M., Rappaport, E.F. and Young, T.L. (2005) Identification of a novel locus on 2q for autosomal dominant high-grade myopia. *Invest. Ophthalmol. Vis. Sci.*, **46**, 2300–2307.
32. Zhang, Q., Li, S., Xiao, X., Jia, X. and Guo, X. (2007) Confirmation of a genetic locus for X-linked recessive high myopia outside MYP1. *J. Hum. Genet.*, **52**, 469–472.
33. Wojciechowski, R., Moy, C., Ciner, E., Ibay, G., Reider, L., Bailey-Wilson, J.E. and Stambolian, D. (2006) Genomewide scan in Ashkenazi Jewish families demonstrates evidence of linkage of ocular refraction to a QTL on chromosome 1p36. *Hum. Genet.*, **119**, 389–399.
34. Nallasamy, S., Paluru, P.C., Devoto, M., Wasserman, N.F., Zhou, J. and Young, T.L. (2007) Genetic linkage study of high-grade myopia in a Hutterite population from South Dakota. *Mol. Vis.*, **13**, 229–236.
35. Lam, C.Y., Tam, P.O.S., Fan, D.S.P., Fan, B.J., Wang, D.Y., Lee, C.W.S., Pang, C.P. and Lam, D.S.C. (2008) A genome-wide scan maps a novel high myopia locus to 5p15. *Invest. Ophthalmol. Vis. Sci.*, **49**, 3768–3778.
36. Mantel, N. and Haenszel, W. (1959) Statistical aspects of the analysis of data from retrospective studies. *J. Natl Cancer Inst.*, **22**, 719–748.
37. Borenstein, M., Hedges, L. and Rothstein, H. (2007) *Meta analysis: Fixed effect vs. random effects*, Biostat, Englewood, NJ.
38. Purcell, S.M., Wray, N.R., Stone, J.L., Visscher, P.M., O'Donovan, M.C., Sullivan, P.F., Sklar, P., Ruderfer, D.M., McQuillin, A. and Morris, D.W. (2009) Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature*, **460**, 748–752.
39. Buraczynska, M., Mears, A.J., Zareparsi, S., Farjo, R., Filippova, E., Yuan, Y., MacNee, S.P., Hughes, B. and Swaroop, A. (2002) Gene expression profile of native human retinal pigment epithelium. *Invest. Ophthalmol. Vis. Sci.*, **43**, 603–607.
40. van Baren, M.J. and Brent, M.R. (2006) Iterative gene prediction and pseudogene removal improves genome annotation. *Genome Res.*, **16**, 678–685.
41. Gross, S.S. and Brent, M.R. (2006) Using multiple alignments to improve gene prediction. *J. Comput. Biol.*, **13**, 379–393.
42. Korf, I., Flicek, P., Duan, D. and Brent, M.R. (2001) Integrating genomic homology into gene structure prediction. *Bioinformatics*, **17**, 140–148.
43. Pennacchio, L.A., Ahituv, N., Moses, A.M., Prabhakar, S., Nobrega, M.A., Shoukry, M., Minovitsky, S., Dubchak, I., Holt, A. and Lewis, K.D. (2006) In vivo enhancer analysis of human conserved non-coding sequences. *Nature*, **444**, 499–502.
44. Saw, S.M. (2003) A synopsis of the prevalence rates and environmental risk factors for myopia. *Clin. Exp. Optom.*, **86**, 289–294.
45. Bertrand, E., Fritsch, C., Diether, S., Lambrou, G., Muller, D., Schaeffel, F., Schindler, P., Schmid, K.L., van Oostrum, J. and Voshol, H. (2006) Identification of apolipoprotein AI as a "STOP" signal for myopia. *Mol. Cell Proteomics*, **5**, 2158–2166.
46. Nielsen, K., Birkenkamp-Demtroder, K., Ehlers, N. and Orntoft, T.F. (2003) Identification of differentially expressed genes in keratoconus epithelium analyzed on microarrays. *Invest. Ophthalmol. Vis. Sci.*, **44**, 2466–2476.
47. McKerracher, L., Essagian, C. and Aguayo, A.J. (1993) Marked increase in beta-tubulin mRNA expression during regeneration of axotomized retinal ganglion cells in adult mammals. *J. Neurosci.*, **13**, 5294–5300.
48. Sharma, R.K., O'Leary, T.E., Fields, C.M. and Johnson, D.A. (2003) Development of the outer retina in the mouse. *Dev. Brain Res.*, **145**, 93–105.
49. Alward, W.L.M., Semina, E.V., Kalenak, J.W., Héon, E., Sheth, B.P., Stone, E.M. and Murray, J.C. (1998) Autosomal dominant iris hypoplasia is caused by a mutation in the rieger syndrome (rieg/pitx2) gene. *Am. J. Ophthalmol.*, **125**, 98–100.
50. Kulak, S.C., Kozlowski, K., Semina, E.V., Pearce, W.G. and Walter, M.A. (1998) Mutation in the RIEG1 gene in patients with iridogoniodysgenesis syndrome. *Hum. Mol. Genet.*, **7**, 1113.
51. Doward, W., Perveen, R., Lloyd, I.C., Ridgway, A.E.A., Wilson, L. and Black, G.C.M. (1999) A mutation in the RIEG1 gene associated with Peters' anomaly. *J. Med. Genet.*, **36**, 152.
52. Nakanishi, H., Yamada, R., Gotoh, N., Hayashi, H., Yamashiro, K., Shimada, N., Ohno-Matsui, K., Mochizuki, M., Saito, M. and Iida, T. (2009) A genome-wide association analysis identified a novel susceptible locus for pathological myopia at 11q24.1. *PLoS Genet.*, **5**, e1000660.
53. Solouki, A.M., Verhoeven, V.J.M., van Duijn, C.M., Verkerk, A., Ikram, M.K., Hysi, P.G., Despriet, D.D.G., van Koolwijk, L.M., Ho, L. and Ramdas, W.D. (2010) A genome-wide association study identifies a susceptibility locus for refractive errors and myopia at 15q14. *Nat. Genet.*, **42**, 897–901.
54. Hysi, P.G., Young, T.L., Mackey, D.A., Andrew, T., Fernández-Medarde, A., Solouki, A.M., Hewitt, A.W., Macgregor, S., Vingerling, J.R. and Li, Y.J. (2010) A genome-wide association study for myopia and refractive error identifies a susceptibility locus at 15q25. *Nat. Genet.*, **42**, 902–905.

55. Korn, J.M., Kuruvilla, F.G., McCarroll, S.A., Wysoker, A., Nemesh, J., Cawley, S., Hubbell, E., Veitch, J., Collins, P.J. and Darvishi, K. (2008) Integrated genotype calling and association analysis of SNPs, common copy number polymorphisms and rare CNVs. *Nat. Genet.*, **40**, 1253–1260.
56. Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A.R., Bender, D., Maller, J., Sklar, P., de Bakker, P.I.W. and Daly, M.J. (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.*, **81**, 559–575.
57. Barrett, J.C., Fry, B., Maller, J. and Daly, M.J. (2005) Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics*, **21**, 263–265.
58. Saxena, R., Voight, B.F., Lyssenko, V., Burtt, N.P., de Bakker, P.I., Chen, H., Roix, J.J., Kathiresan, S., Hirschhorn, J.N. and Daly, M.J. (2007) Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. *Science*, **316**, 1331–1336.
59. Shi, Y.Y. and He, L. (2005) SHEsis, a powerful software platform for analyses of linkage disequilibrium, haplotype construction, and genetic association at polymorphism loci. *Cell Res.*, **15**, 97–98.