The *COL1A1* Gene and High Myopia Susceptibility in Japanese

Yumiko Inamori[1], Masao Ota[2], Hidetoshi Inoko[3], Eiichi Okada[4], Ritsuko Nishizaki[3,4], Tomoko Shiota[3,4], Jeewon Mok [3,5], Akira Oka[3], Shigeaki Ohno[6], Nobuhisa Mizuki[1]

1) Department of Ophthalmology, Yokohama City University School of Medicine, Yokohama, Kanagawa, Japan
2) Department of Legal Medicine, Shinsyu University School of Medicine, Matsumoto, Nagano, Japan
3) Department of Basic Medical Science and Molecular Medicine, Tokai University School of Medicine, Isehara, Kanagawa, Japan
4) Okada Eye Clinic, Yokohama, Kanagawa, Japan
5) Laboratory of Ophthalmology & Visual Science, Catholic Research Institutes of Medical Science College of Medicine, The Catholic University of Korea, Seoul, Korea
6) Department of Ophthalmology, Hokkaido University School of Medicine, Sapporo, Hokkaido, Japan

Key words: SNP , *COL1A1*, myopia, Japanese population

Running title: Association of the *COL1A1* gene and high myopia

Correspondence: Masao Ota, Ph.D. Department of Legal Medicine, Shinshu University School of Medicine, 3-1-1 Asahi Matusmoto, 390-8621 Japan

E-mail:otamasao@sch.md.shinshu-u.ac.jp, Tel:+81263373217, Fax:+81263373084

ABSTRACT

The collagen type I alpha I (COL1A1) gene encodes the extracellular matrix component, collagen, and resides in candidate MYP5 for high myopia on the chromosome 17q22-q23.3. This locus has recently been implicated in playing an important role in the pathogenesis of experimental myopia. We investigated the association of disruptions of *COL1A1* gene with high myopia by analyzing the frequency of ten SNPs in a Japanese population of 330 subjects with high myopia of -9.25 D or less and 330 randomized controls without high myopia. Two SNPs (rs2075555 and rs2269336) were significantly associated with high myopia ($P<0.05$, $Pc<0.1$). Two different haplotype blocks in *COL1A1* were observed by the pair-wise linkage disequilibrium between the SNPs. The frequency of GGC/GGC diplotype constructed by three SNPs (rs2075555, rs2269336, rs1107946) was significantly high (OR=1.6) and associated with high myopia ($P=0.028$, $Pc<0.084$). Together our results provide the first evidence for *COL1A1* as a gene associated with high myopia.

INTRODUCTION

Myopia is the most common eye disorder in the world and a significant public health problem. Myopia is far more frequent in Asian populations than in the USA or Europe, and Japan has one of the highest incidences of myopia in the world (Saw et al. 1996). In the United States, the prevalence of myopia in high school students is approximately 25% (Burton 1989; Curtin 1970; Katz et al. 1997; Leibowitz et al. 1980; Ministry of Education 2004), compared to nearly 60% in Japan (Burton 1989). High or pathologic myopia can cause blindness, or severe visual acuity loss due to retinal detachment, submacular hemorrhage, glaucoma, or macular degeneration (Burton 1990). High myopia is especially common in Asia, and in Japan, accounts for 6% to 18% of myopia cases and 1% to 2% of the general population (Curcio et al. 1987). Severe myopia, therefore, has been a major cause of visual affliction.

The refractive power of the eye is correlated with ocular axial length and it is well established that myopia is caused by increased axial eye size (Tokoro and Sato 1982). The development of high myopia in humans is associated with marked thinning of the sclera, the tough outer coat of the eye that facilitates any change in eye size. Scleral thinning is greatest at the posterior pole of the eye, the anatomical region of greatest retinal photoreceptor density and vital to detailed visual discrimination (Curtin 1985). Collagen accounts for 90% of scleral dry weight, and the majority of this collagen is collagen type I (Rada et al. 2006). Mammalian sclera also contains small amounts of other fibrillar and fibril-associated collagens (Marshall et al. 1993; Tamura et al. 1991; Wessel et al. 1997) and studies have shown that scleral fibrils are heterologous, comprised of collagen types I, III, and V (Marshall et al. 1993).

Myopia is a highly prevalent, complex phenotype involving both genetic and environmental factors. Recently, myopia susceptibility genes have been identified in the 14 genomic loci (MYP1on Xq28, MYP2 on chromosome18p, MYP3 on chromosome 12q, MYP4 on chromosome 7q, MYP5 on chromosome 17q, MYP6 on chromosome 22q12, MYP7 on chromosome 11p13, MYP8 on chromosome 3q26, MYP9 on chromosome 4q12, MYP10 on chromosome 8p23, MYP11 on chromosome 4q22-q27, MYP12 on chromosome 2q37.1, MYP13 on Xq23-q25, and MYP14 on chromosome 1p36) (Hammond

et al. 2004; Naiglin et al. 2002; Paluru et al. 2003; Paluru et al. 2005; Schwarts et al. 1990; Stambolian et al. 2004; Wojciechowski et al. 2006; Young et al. 1998a; Young et al. 1998b; Zhang et al. 2006).

The 18 kb *COL1A1* gene, encoded by 52 exons, is located on chromosome 17 (17q21.33) where the MYP5 (17q21-22) susceptibility locus was identified (Paluru et al. 2003). This gene encodes the major component of type I collagen. The scleral tissue contains approximately 90% collagen, consisting predominantly of type I collagen (Zorn et al. 1992). Mutations in *COL1A1* have been described in individuals with type 1 osteogenesis imperfecta, Ehlers-Danlos syndrome type VIIA and VIIB, osteoporosis, and Marfan syndrome, all systemic disorders with scleral thinning and myopia as a clinical component (Dalgleish 1997).

The association between variations in the *COL1A1* gene and high myopia has yet to be investigated. In this study, we screened the *COL1A1* gene for single-nucleotide polymorphisms (SNPs) that could potentially be associated with myopia.

MATERIALS AND METHODS

A total of 330 Japanese subjects with high myopia and ethnically matched unrelated control subjects (*n*=330) attending the Ophthalmology Clinic at the Yokohama City University or the Okada eye clinic were recruited for this study. The myopia phenotypes were classified according to the mean spherical equivalent refractive error below or above the threshold level of –9.25 Dsph. The mean refractive error of myopia cases was –11.55 ± 2.17 Dsph with a mean axial length of 27.78±1.30 mm. The mean age of case subjects was 37.82 ± 11.97 years, and the male–female ratio of the cases was 0.66:1.00. A group of 330 subjects of same sex and age with moderate or no myopia was used as the control group. None of the subjects involved exhibited any systemic connective tissue diseases, keratoconus, or other genetic disorders known to cause myopia. All subjects involved were Japanese from a similar social background and from the same urban area, and informed consent was obtained from all patients. The study was conducted in accordance with the Declaration of Helsinki and subsequent revisions.

SNP-specific polymerase chain reaction (PCR) primers and fluorogenic probes were designed using ABI (Applied Biosystems, Foster City, CA, USA). This technique has been utilized extensively in genotyping other candidate genes with multiple single nucleotide polymorphisms. The fluorogenic probes were labeled with a reporter dye (either FAM or VIC) and specific for one of the two possible bases in the promoter region. A MGB quencher probe was utilized on the 3′ end by a linker arm. The 2X PCR mix (TaqMan Universal PCR Master Mix , Applied Biosystems) was optimized for TaqMan reactions and contained AmpliTaq-Gold DNA polymerase, dNTPs with UTP and a passive reference dye. Primers, probes, and genomic DNA were added to final concentrations of 300 nmol/l, 100 nmol/l, and 0.5–2.5 ng/µl respectively. Control samples (without DNA template) were run as a negative control for contaminating DNA and a reference positive control DNA used to verify the identified polymorphisms. The amplification reactions were carried out in an ABI Prism 7700HT Sequence Detection System (Applied Biosystems GeneAmp® PCR System 9700) with an initial hold step (95°C for 10 min) followed by 40 cycles of a two-step PCR (denaturation at 92°C for 15s, annealing at 60°C for 1 min). The fluorescence intensity of each sample was measured at each temperature change to monitor the amplification of the base pair *COL1A1* promoter region. The targeted nucleotide was determined by the

fluorescence ratio of the two SNP-specific fluorogenic probes. The fluorescence signal increased when the specific probe matched the single stranded template DNA and was digested by the 5′-3′ exonuclease activity of AmpliTaq-Gold DNA polymerase (Applied Biosystems), whereupon the probe released the fluorescent reporter dye (either FAM or VIC) from the quencher dye.

We used the NCBI dbSNP database (http://www.ncbi.nlm.nih.gov/SNP/) to extract the available information on the SNPs in *COL1A1* gene. Forward or reverse strand sequences of the SNP sites were selected depending on their GC content, to avoid primer dimerization and hairpin structure, and based on the absence of any other SNP site. The extension reaction was controlled by a mixture of dideoxy-terminated nucleotides, such that one single-base extension product and one double-base extension product was created corresponding to a SNP allele. This scheme generated two peaks in the mass spectrometer separated by approximately 300 Da.

Allele frequencies of sequence alterations in patients and controls were evaluated using Chi−square tests or Fisher exact tests. Allelic frequencies of all detected SNPs were also assessed for Hardy−Weinberg equilibrium. Statistical analyses were performed on computer using the Statview software (ver.5.0 SAS Institute Inc. USA). We corrected these P values (Pc) by the Bonferroni's correction where the coefficient was the total number of the contingency tables tested. P value <0.05 and Pc value <0.1 were considered statistically significant. The R package "haplo.stats" http://www.r-project.org/) was used to evaluate haplotype structure. Based on the haplotype structures the pair-wise linkage disequilibrium between SNPs was measured with LD coefficient (Lewontin's D') obtained from the R package "genetics" in the R Project for Statistical Computing.

RESULTS

We sequenced 10 SNPs of the *COL1A1* gene, chosen using criteria such as population-frequncy validation, multiple submitters and high-profile submitters, using an electronic database 'dsSNP'(http://www.ncbi.nlm.nih.gov/SNP/) for the final selection of ten SNPs for genotyping. A total of 10 SNPs in the *COL1A1* gene were screened in all the myopia cases and control subjects (Figure.1). Each marker was tested to see if it is in Hardy-Weinberg equilibrium and that this hypothesis was not rejected for any of the markers. Of the SNP analyzed, one was non-synonymous, one was a synonymous substitution, three were intronic, two were untranslated and three were unknown (Figure1 and Table1). Two of them, rs2586486 and rs1800211, were monomorphisms (Table 1).

Two SNPs (rs2075555 and rs2269336) were significant under a model assuming dominance of the 1st allele listed for each locus (Table 1). The G allele in intron 11 (rs2075555) was significantly more common in the patient group (P=0.0071; OR=1.36) than in the controls (Table 2). In addition, the frequencies of G allele (rs2269336) and the C allele (rs1107946) in the 5' upstream region of the *COL1A1* gene were significantly increased in the patient group (P=0.0140, OR=1.31 and P=0.0278, OR=1.28, respectively; Table 2).

The pairwise LD mapping confirmed that haplotype structure in the *COL1A1* gene was constructed by two blocks (rs748075 to rs207558 and rs207555 to rs1107946) from LD index values (Table 3). The estimated haplotype (GGC/GGC) was significantly more prevalent in the patient group than the control group (P=0.0084, Pc=0.025, OR=1.60 in Table 4). Frequencies of pairwise haplotypes consisting of 8 polymorphic SNPs are listed in Table 5. No significant difference found between the patient and controls was detected.

DISCUSSION

Myopia is an extremely common ocular condition that affects approximately one billion people worldwide(Marshall et al. 1993). Many studies have shown that the development of human myopia is influenced by multifactorial etiology with underlying complex genetic factors and undefined environment factors (Bear 1991; Goss et al. 1988; Goss 2000). Studies with animal models have indicated that environmental factors such as visual deprivation and the effects of a negative lens may contribute to the development of myopia. Likewise, there is a significant correlation between myopia and the amount of near work such as reading and writing (Bear 1991, Saw et al. 2004). However, evidence for genetic pathways contributing to myopia is strengthened by the observations that myopic parents are much more likely to have myopic children (Mutti et al. 2002), and that myopia is far more frequent in Asian populations than in the USA or Europe, even if the populations examined have performed similar amounts of near work (Feldkamper and Schaeffel. 2003). Thus, in addition to environmental effects, myopia is also likely to result from alterations of multiple genetic factors.

In this study, we examined whether the *COL1A1* gene is a disease susceptibility gene for high myopia in Japan. We performed SNP analysis on 10 SNPs in 330 patients with high myopic change (refractive error greater than －9.25Dsph) and compared the findings to those in a control group of 330 individuals with no myopia. As myopia is a globally prevalent disease, we selected universally high heterozygosity SNPs from the NCBI SNP database for the association study. However, two of these SNPs were monomorphisms in Japanese.

Two of the 10 SNPs (rs2075555 and rs2269336) showed significantly different frequencies between cases and controls ($P<0.05$, $Pc<0.1$). The SNPs were not missense mutations, and one of them exists in an intron while the others are upstream of the *COL1A1* gene. How these SNPs disrupt the function is unknown. Together this suggests that the *COL1A1* gene is likely associated with high myopia, although the associated variations of this gene don't result in obvious changes in the COL1A1 protein, it is possible that these SNPs are in LD with an, as yet undiscovered, causative mutation in the COL1A1 gene. We investigated the10 SNPs genotypes and allele dominant models (Table2) and found two SNPs to be statistically different, rs2075555(A>C) in allele2(C) dominant model($P<0.05$),

and rs2269336(C>G) in allele1(C) dominant model($P<0.01$).

A Gentle et al reported that collagen type I mRNA expression was reduced and scleral collagen accumulation was decreased in the tree shrews sclera of myopic eyes (Gentle et al. 2003), suggesting that disruption of collagen function in myopia could potentially be regulated at the transcriptional level. Together this suggests the possibility that these three SNPs could affect the promoter and/or trascrptional control of the *COL1A1* gene.

Our susceptibility gene mapping showed that the limited haplotype block (GGC/GGC) constructed by three SNPs in 5'-genomic region of the *COL1A1* gene was significantly associated with the susceptibility to high myopia.

In conclusion, *COL1A1* may be a candidate gene for high myopia based on its mapped location within the MYP5 candidate region as well as the evidence presented in this report linking SNPs in this region to susceptibility to high myopia. The associated SNPs do not appear to be causal mutation in high myopia phenotypes in Japan. Further studies of candidate genes are needed to determine the molecular genetic factors that contribute to genetic influence of myopia.

Acknowledgment

This work was supported by the Health and Labour Sciences Research Grants in Japan and the Johnson & Johnson K.K. Vison Care Company.

REFERENCES

Bear JC (1991) Epidemiology and genetics of refractive anomalies. In Grosvenor, T., Glom, M.C. (eds.), Refractive anomalies : research and clinical applications. Stoneham, MA, Butterworth-Heinemann

Burton TC (1989) The influence of refractive error and lattice degeneration on the incidence of retinal detachment. Trans Am Ophthalmol Soc 87:143–157

Curcio CA, Sloan KR, Packer O, Hendrickson AE, Kalina RE (1987) Distribution of cones in human and monkey retina: individual variability and radial asymmetry. Science 236:579-582

Curtin BJ (1970) Myopia: a review of its etiology, pathogenesis, and treatment. Surv Ophthalmol, 15:1–17

Curtin BJ(1985) The Myopias: Basic Science and Clinical Management, Harper &Row, Philadelphia

Dalgleish R (1997) The human type I collagen mutation database. Nucleic Acids Res 25:181-187

Feldkamper M, Schaeffel F (2003) Interactions of genes and environment in myopia. Dev Ophthalmol 37:34-49

Gentle A, Liu Y, Martin JE, Conti, GL, McBrien, NA (2003) Collagen gene expression and the altered accumulation of scleral collagen during the development of high myopia. J Biol Chem 278:16587-16594

Goss DA, Hampton MJ, Wickham MG (1988) Selected review on genetic factors in myopia. J Am Optom Assoc 59:875-884

Goss DA (2000) Nearwork and myopia. Lancet 356:1456-1457

Hammond CJ, Andrew T, Mak YT, Spector TD (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 75:294-304

Katz J, Tielsch JM, Sommer A (1997) Prevalence and risk factors for refractive errors in an adult inner city population. Invest Ophthalmol Vis Sci 38:334–340

Leibowitz HM, Krueger DE, Maunder LR, Milton RC, Kini MM, Kahn HA, Nickerson RJ, Pool J, Colton TL, Ganley JP, et al. (1980) The Framingham Eye Study monograph: An ophthalmological and epidemiological study of cataract, glaucoma, diabetic retinopathy,

macular degeneration, and visual acuity in a general population of 2631 adults, 1973-1975. Surv Ophthalmol 24:472–479

Marshall GE, Konstas AG, Lee WR (1993) Collagens in the aged human macular sclera. Curr Eye Res 12:143-153

Ministry of Education, Culture, Sports, Science and Technology(ed.)(2004) Disease Rate among Students, STATISTICAL ABSTRACT.

Mutti DO, Mitchell GL, Moeschberger ML, Jones LA, Zadnik K (2002) Parental myopia, near work, school achievement, and children's refractive error. Invest. Ophthalmol Vis Sci 43:3633-3640

Naiglin L, Gazagne C, Dallongeville F, Thalamas C, Idder A, Rascol O, Malecaze F, Calvas P (2002) A genome wide scan for familial high myopia suggests a novel locus on chromosome 7q36. J Med Genet 39:118-124

Paluru PC, Ronan SM, Heon E, Devoto M, Wildenberg SC, Scavello G, Holleschau A, Makitie O, Cole WG, King RA. et al. (2003) New locus for autosomal dominant high myopia maps to the long arm of chromosome 17. Invest Ophthalmol Vis Sci 44:1830-1836

Paluru PC, Nallasamy S, Devoto M, Rappaport EF, Young TL (2005) Identification of a novel locus on 2q for autosomal dominant high-grade myopia. Invest Ophthalmol Vis Sci 46:2300-2307

Rada JA, Shelton S, Norton TT (2006) The sclera and myopia. Exp Eye Res 82:185-200

Saw SM, Katz J, Schein OD, Chew SJ, Chan TK (1996) Epidemiology of myopia. Epidemiol Rev 18: 175-187

Saw SM, Chua WH, Hong CY, Wu HM, Chan WY, Chia KS, Stone RA, Tan D (2002) Nearwork in early-onset myopia. Invest. Ophthalmol Vis Sci 43:332-339

Schwartz M, Haim M, Skarsholm D (1990) X-linked myopia: Bornholm eye disease. Linkage to DNA markers on the distal part of Xq. Clin Genet 38:281-286

Stambolian D, Ibay G, Reider L, Dana D, Moy C, Schlifka M, Holmes T, Ciner E, Bailey-Wilson JE (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

Tamura Y, Konomi H, Sawada H, Takashima S, Nakajima A (1991) Tissue distribution of

type VIII collagen in human adult and fetal eyes. Invest Ophthalmol Vis Sci 32:2636-2644.

Tokoro T, Sato A. (1982) Results of Investigation of Pathologic Myopia in Japan: Report of Myopic Chorioretinal Atrophy 1982, 32-35 Ministry of Health and Welfare, Tokyo.

Wessel H, Anderson S, Fite D, Halvas E, Hempel J, SundarRaj N (1997) Type XII collagen contributes to diversities in human corneal and limbal extracellular matrices. Invest Ophthalmol Vis Sci, 38:2408-2422

Wojciechowski R, Moy C, Ciner E, Ibay G, Reider L, Bailey-Wilson JE, Stambolian D (2006) Genomewide scan in Ashkenazi Jewish families demonstrates evidence of ocular refraction to a QTL on chromosome 1p36, Hum Genet 119:389-399

Young TL, Ronan SM, Drahozal LA, Wildenberg SC, Alvear AB, Oetting WS, Atwood LD, Wilkin DJ, King RA (1998a) Evidence that a locus for familial high myopia maps to chromosome 18p. Am J Hum Genet 63:109-119

Young TL, Ronan SM, Alvear AB, Wildenberg SC, Oetting WS, Atwood LD, Wilkin DJ, King RA (1998b) A second locus for familial high myopia maps to chromosome 12q. Am J Hum Genet 63:1419-1424

Zorn N, Hernandez MR, Norton TT, Yang J, Ye HO (1992) Collagen gene expression in the developing tree shrew sclera, Invest Ophthalmol Vis Sci 33:S1053 (ARVO Abstracts)

Zhang Q, Guo X, Xiao X, Jia X, Li S, Hejtmancik JF (2006) Novel locus for X linked recessive high myopia maps to Xq23-q25 but outside MYP1. J Med Genet 43:e20

Table 1 Association of 10 SNPs of COL1A1 gene with high myopia

SNP rs	public position	function	AA change*	allele	case	control	χ2	P	OR	95%CI
rs748075	45615428	3'downstream		C	269(82.8)	259(79.7)		N.S.		
				G	208(64.0)	218(97.8)				
rs1061947	45617118	3'UTR		A	30(9.1)	27(8.2)		N.S.		
				G	328(99.7)	328(99.7)				
rs1061237	45617774	3'UTR		A	276(84.1)	267(81.4)		N.S.		
				G	212(64.6)	212(65.6)				
rs2586486	45618215	exon	Lys→Gln	G	326(100)	329(100)		N.S.		
				T	0(0)	0(0)				
rs2277632	45618902	intron		A	299(91.4)	291(89.8)		N.S.		
				G	170(52.0)	169(52.2)				
rs2075558	45622584	intron		A	124(37.8)	111(33.9)		N.S.		
				C	311(94.8)	313(95.7)				
rs1800211	45626379	exon		C	329(100)	330(100)		N.S.		
				T	0(0)	0(0)				
rs2075555	45629290	intron		G	294(89.6)	274(84.0)	4.47	0.035*1	1.64	1.03−2.61
				T	200(61.0)	228(69.9)				
rs2269336	45635355	5'upstream		C	212(64.4)	246(74.5)	7.94	0.0048*2	1.62	1.15−2.26
				G	280(85.1)	270(81.8)				
rs1107946	45635989	5'upstream		A	191(58.1)	214(65.4)		N.S.		
				C	296(90.0)	280(85.6)				

SNPrs : public reference SNP number from the dbSNP database.
AA change, amino acid change
Numbers in parentheses indicate the percentage.
P value was calculated by χ2 test 2x2 contingency table (df=1).
NS: not significant by χ2 test 2x2 contingency table (df=1).
OR: odds ratio.
95%CI, 95% confidence interval.
*1: Pc (corrected P value) =0.07
*2: Pc=0.0096

Table 2 Two SNPs (rs2075555 and rs2269336) in Japanese patients with high myopia and healthy individuals

SNPs rs	Genotype frequency	case(%)	control (%)	P	Allele frequency	case (%)	control (%)	P	OR	95%CI
rs2075555	G/G	128(39.0)	98(30.1)	0.018*1	G	422(64.3)	372(57.1)	0.0071*3	1.36	1.09−1.70
	G/T	166(50.6)	176(54.0)		T	234(35.7)	280(42.9)			
	T/T	34(10.4)	52(16.0)							
rs2269336	C/C	49(14.9)	60(18.2)	0.018*2	C	261(39.7)	306(46.4)	0.014*4	1.31	1.06−1.64
	C/G	163(49.5)	186(56.4)		G	397(60.3)	354(53.6)			
	G/G	117(35.6)	84(25.4)							

Numbers in parentheses indicate the percentage.
P value was calculated by $\chi 2$ test 3x2 contingency table (df=2), or $\chi 2$ test 2x2 contingency table (df=1).
OR: odds ratio.
95%CI, 95% confidence interval.
*1: Pc (corrected P value) =0.054
*2: Pc=0.054
*3: Pc =0.014
*2: Pc=0.028

Table 3 Pairwise linkage disequilibrium (LD) between 8 SNPs of the *COL1A1* gene for high myopia patients and healthy control groups

		Myopia D'							
		rs748075	rs1061947	rs1061237	rs2277632	rs2075558	rs2075555	rs2269336	rs1107946
Control	rs748075	-	**1.00**	0.78	**0.94**	0.01	0.43	0.35	0.34
	rs1061947	**1.00**	-	**1.00**	**1.00**	**1.00**	0.77	0.64	0.67
	rs1061237	**0.80**	**1.00**	-	**0.95**	0.33	0.51	0.36	0.39
D'	rs2277632	**1.00**	**1.00**	**0.96**	-	**0.80**	0.29	0.23	0.20
	rs2075558	0.20	**0.90**	0.28	**0.90**	-	0.33	0.28	0.16
	rs2075555	0.30	0.63	0.36	0.15	0.53	-	**0.90**	**0.83**
	rs2269336	0.29	0.44	0.23	0.07	0.40	**0.88**	-	**0.96**
	rs1107946	0.30	0.51	0.25	0.11	0.08	**0.81**	**0.96**	-

The degree of LD is shown as the LD index of Lewontin correlation (D'). Bold number indicates the strong LD: D'>0.8.

Table 4 Frequency analysis of the haplotype of *COL1A1* gene in Japanese high myopia patients and control groups

Haplotype	Case(%)	Control(%)	P	$\chi 2$	OR	*
GGC/-	171(52.1)	187(57.7)	N.S.			
GGC/GGC	104(31.7)	73(22.5)	0.0084	6.94	1.60	P=0.028
-/-	53(16.2)	64(19.8)	N.S.			$\chi 2$=7.15

Haplotype consists of three SNPs (rs2075555, rs2269336, rs1107946) in the *COL1A1* gene.
Numbers in parentheses indicate the percentage.
P value was calculated by 2x2 contingency $\chi 2$ test, and * 3x2 contingency table (df=2).
OR: odds ratio.
95%CI, 95% confidence interval.

Table 5 Estimated haplotype frequencies of the COL1A1 gene between control groups and Japanese high myopia patients groups

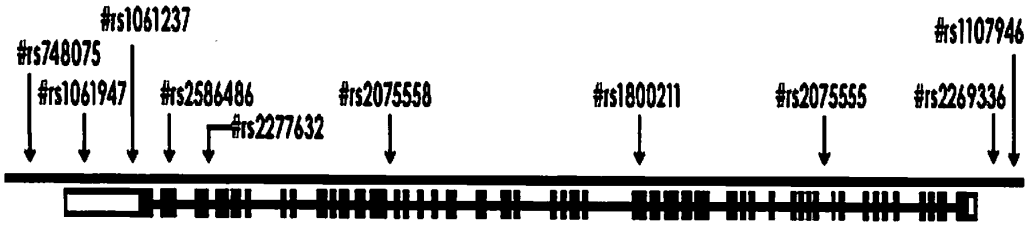
	Haplotype	748075	1061947	1061237	2277632	2075558	2075555	2269336	1107946	frequency (n=310)
Control	1	C	G	A	A	C	G	G	C	0.318
	2	G	G	G	G	C	T	C	A	0.150
	3	G	G	G	G	C	G	G	C	0.132
	4	C	G	A	A	C	T	C	A	0.081
	5	C	G	A	A	A	G	G	C	0.045
	6	G	G	G	A	A	T	C	A	0.045
	7	G	G	A	A	C	T	C	A	0.041
	8	C	A	G	A	A	T	C	A	0.035
	Haplotype	748075	1061947	1061237	2277632	2075558	2075555	2269336	1107946	frequency(n=323)
Case	1	C	G	A	A	C	G	G	C	0.342
	3	G	G	G	G	C	G	G	C	0.136
	2	G	G	G	G	C	T	C	A	0.121
	5	C	G	A	A	A	G	G	C	0.071
	4	C	G	A	A	C	T	C	A	0.060
	8	C	A	G	A	A	T	C	A	0.037
	6	G	G	G	A	A	T	C	A	0.036
	7	G	G	A	A	C	T	C	A	0.026

Figure legends

Figure 1. Human *COL1A1* gene structure and 10 SNPs.

Vertical arrows show 10 SNPs in this study with the #rs numbers. Left of the figure is 5'UTR. White boxes show the untranslated regions and black boxes show the translated regions of this gene. The size of each exon and intron is not to scale.

Figure 1


#rs1061237
#rs1107946
#rs748075
#rs1061947
#rs2586486
#rs2277632
#rs2075558
#rs1800211
#rs2075555
#rs2269336