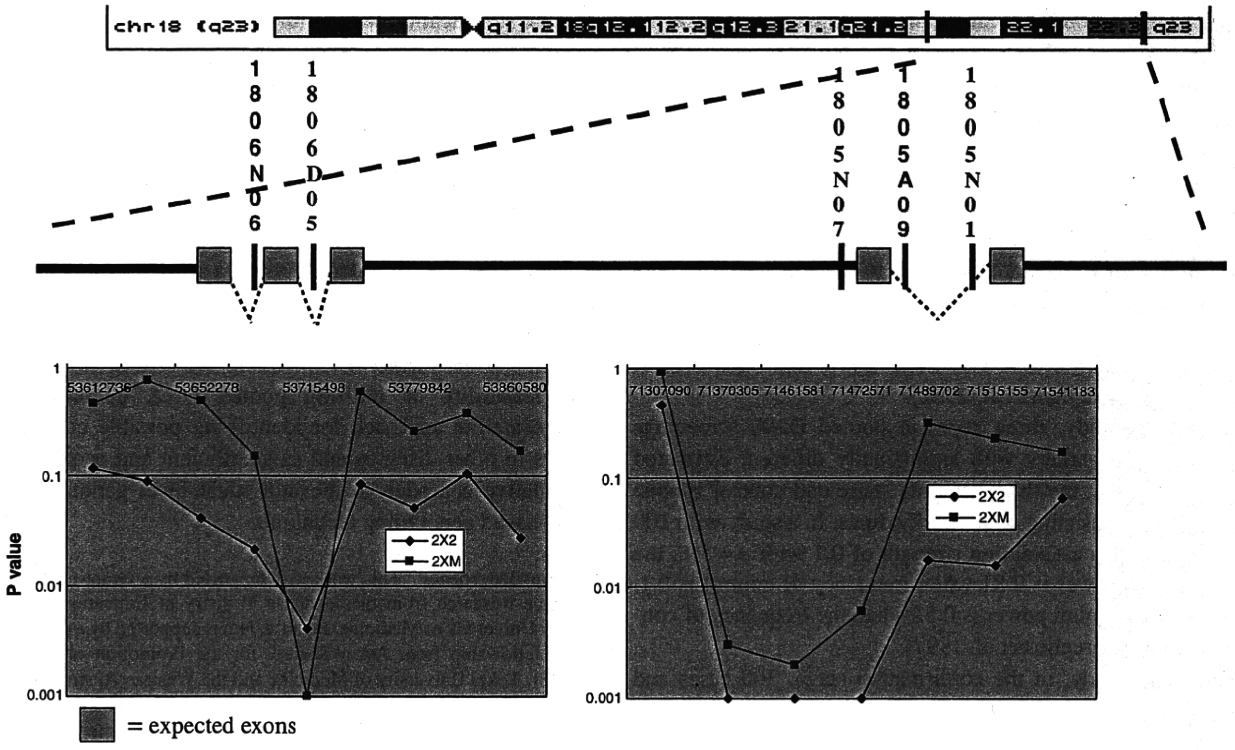

Fig. 3 Fine mapping results showing location and predicted susceptibility gene location on chromosome 18. The predicted exons are shown. Markers with significant *P* values are located within the introns of the predicted genes

tion of genetic factors to disease manifestations. In the present study, we have chosen to use positive response to questions on doctor-diagnosed asthma or ever been treated for asthma/bronchial asthma as the definition of asthma in this study. Since there is no universally accepted 'gold standard' definition of 'asthma', misclassification of case status is possible in our study. However, this would remain a limitation even if other operant definitions of asthma, such as current wheeze plus hyperresponsiveness or current use of asthma medications (Peat et al. 1992), are used. Potential misclassifications will be non-differential and therefore favour the null hypothesis. The associations that we have identified using this conservative approach will therefore remain valid. Other operant definitions of asthma (such as current wheeze plus airway hyperresponsiveness, current use of asthma medications) may alter the prevalence of asthma but will suffer from the same degree of misclassification as using doctor-diagnosed asthma.

Linkage to a non-specific IgE locus was previously reported on chromosome 18 near D18S844 (Mathias et al. 2001). D18S844 is within 1 cM of nuclear factor of activated T cells 1 (NFATC1) and approximately 900 kb from 1805A09 (D18S0325i). Suppression of NFATC1 was associated with highly increased serum IgE levels in patients with severe atopic dermatitis (Wierenga et al. 1999). These results suggest that a gene in this region of chromosome 18 could control IgE levels which may play a role in the pathogenesis of asthma. Further studies are required to determine the expression pattern of the two novel genes within chromosome 18 identified in this GWA study.

The prevalence of asthma has increased over the past 30 years, and its social and economic costs are a major problem in many developed countries (Rachelefsky 2007).

The human genome is now charted by approximately 30,000 highly polymorphic microsatellites (Kawashima et al. 2006; Tamiya et al. 2005). In this study, we have performed a high-density GWA study using microsatellite markers instead of SNPs for several reasons. Firstly, the average length of linkage disequilibrium for 30,000 microsatellites is approximately 100 kb (Tamiya et al. 2005), which is considerably higher than that of SNPs and therefore allows the capture of a larger region of linkage disequilibrium per marker when compared to SNP. Polymorphic microsatellite markers also provide a higher information content when compared with biallelic SNPs. Furthermore, interpolation variability is known to be smaller for microsatellite markers (Jorde et al. 2000; Sawyer et al. 2005). In addition, SNP haplotypes are virtual products, based on statistical hypothesis that depend on other conditions (for example, $r^2 > 0.8$, minor allele frequency > 0.1). Often, recombination events cannot be included in the deduced haplotypes due to the lack of a heterozygous SNP at the

relevant position. This can result in the generation of incorrect haplotype blocks which can, in turn, lead to false negative in GWA studies based on SNPs only. Microsatellites, on the other hand, do not generally suffer from these problems and can therefore be used as a stand-alone source of markers in GWA studies (Bahram and Inoko 2007). Microsatellites may also have important intrinsic functional relevance (Kashi and King 2006). Allelic differences in microsatellite markers are known to cause a wide range of hereditary disorder including Fragile X (Tassone et al. 2007), Huntington's disease (Vassos et al. 2007), spinocerebellar ataxia (Stevanin and Brice 2007) and cleidocranial dysplasia (Vaughan et al. 2002).

In this study, three separate pooled DNA screenings yielded 18 markers with significantly different estimated frequencies in the three separate "case and control" pools: each pool consisting of 120 individuals. In a study with 100 controls at an association strength of 0.1 with $\alpha = 5\%$, the maximum power is 0.81 when frequency of control is 0.1, and the minimum power is 0.52 when the frequency of control is 0.5 (Barcellos et al. 1997).

In this study, in the confirmation stage, 360 cases and 360 controls were individually typed. The maximum power when number of controls is 300 at an association strength of 0.1 with $\alpha = 5\%$ is 1 when frequency of control is 0.1 and power = 0.94 when the control frequency is 0.5. With a three-stage screenings and a final confirmation using a total of 720 individuals (360 controls and 360 cases), we therefore have enough power to detect association even when the control frequency is 0.5.

In a multi-stage GWA study, it is important to ensure that sufficiently large proportion of markers with effect are selected for the subsequent stage. False discoveries can always be eliminated in future stages but markers with effects that have been eliminated cannot be recovered (van den Oord et al. 2007).

Asthma, as it is defined clinically and epidemiologically, is based on symptoms which may result from a number of patho-physiological mechanisms related to airway inflammation and remodelling. Airway pathology may also be present in the absence of symptoms. Therefore, asthma remains phenotypically heterogeneous. This may partly explain why no consistent positive associations for asthma have been identified to date (Bosse and Hudson 2007). Conflicting findings may also be due to genetic heterogeneity and poor study design. Result from a recent GWA study of childhood onset asthma identified markers on chromosome 17q21 to be associated with childhood onset asthma (Moffatt et al. 2007). Further analyses concluded that genetic variants regulating ORMDL3 expression are responsible for childhood onset asthma. We did not detect any consistent significant differences in our adult asthma cases and controls on chromosome 17. One explanation is that no stratification was performed in relation to the age of onset of asthma in the current study. Genetic susceptibility for early onset asthma may also represent a different type of disease. Further analyses of additional phenotypes with consideration of interaction effects as well as characterisation of the two predicted proteins are underway.

In conclusion, this is the first asthma GWA study using in excess of 23,000 microsatellite markers, evenly spaced within the genome to identify asthma susceptibility gene in an unbiased fashion. Our results suggest that chromosome 18 harbours at least two novel asthma susceptibility loci which warrant further investigation. We also demonstrated the feasibility of utilising pooled DNA genome-wide microsatellite approach for identifying possible candidate genes in other diseases and as an efficient and economical alternative approach to the individual level genome-wide chip-based genotyping technique.

Acknowledgments This work was supported by a Grant-in-Aid for Science Research from the Japanese Ministry of Education and by Tokai University of Medicine, Japan. J. Hui is supported by a post-doctoral fellowship from Japan Society for the Promotion of Science (JSPS), Tokai University of Medicine and the Western Australia Institute for Medical Research. The author would like to thank the Busselton Population Medical Research Foundation and the Busselton community for their participation in the study. We also thank Euzebiusz Jamrozik and Mark Divitini for statistical advices and Kaori Yamaguchi, Hiromi Kamura, Erika Matsushita and Hiromi Takase at the Department of Molecular Life Science, Tokai University School of Medicine, for technical assistance in the genome-wide screens.

References

Allen M, Heinzmann A, Noguchi E, Abecasis G, Broxholme J, Ponting CP, Bhattacharyya S, Tinsley J, Zhang Y, Holt R, Jones EY, Lench N, Carey A, Jones H, Dickens NJ, Dimon C, Nicholls R, Baker C, Xue L, Townsend E, Kabesch M, Weiland SK, Carr D, von Mutius E, Adcock IM, Barnes PJ, Lathrop GM, Edwards M, Moffatt MF, Cookson WO (2003) Positional cloning of a novel gene influencing asthma from chromosome 2q14. Nat Genet 35:258–263

Anderson GG, Leaves NI, Bhattacharyya S, Zhang Y, Walshe V, Broxholme J, Abecasis G, Levy E, Zimmer M, Cox R, Cookson WO (2002) Positive association to IgE levels and a physical map of the 13q14 atopy locus. Eur J Hum Genet 10:266–270

Bahram S, Inoko H (2007) Microsatellite markers for genome-wide association studies. Nat Rev Genet 8:164

Barcellos LF, Klitz W, Field LL, Tobias R, Bowcock AM, Wilson R, Nelson MP, Nagatomi J, Thomson G (1997) Association mapping of disease loci, by use of a pooled DNA genomic screen. Am J Hum Genet 61:734–747

Bosse Y, Hudson TJ (2007) Toward a comprehensive set of asthma susceptibility genes. Annu Rev Med 58:171–184

Chanock SJ, Manolio T, Boehnke M, Boerwinkle E, Hunter DJ, Thomas G, Hirschhorn JN, Abecasis G, Altshuler D, Bailey-Wilson JE, Brooks LD, Cardon LR, Daly M, Donnelly P, Fraumeni JF Jr, Freimer NB, Gerhard DS, Gunter C, Guttmacher AE, Guyer MS, Harris EL, Hoh J, Hoover R, Kong CA, Merikangas KR, Morton CC, Palmer LJ, Phimister EG, Rice JP, Roberts J, Rotimi C, Tucker MA, Vogan KJ, Wacholder S, Wijsman EM, Winn DM, Col-

lins FS (2007) Replicating genotype–phenotype associations. Nature 447:655–660

Collins HE, Li H, Inda SE, Anderson J, Laiho K, Tuomilehto J, Seldin MF (2000) A simple and accurate method for determination of microsatellite total allele content differences between DNA pools. Hum Genet 106:218–226

Hersh CP, Raby BA, Soto-Quiros ME, Murphy AJ, Avila L, Lasky-Su J, Sylvia JS, Klanderman BJ, Lange C, Weiss ST, Celedon JC (2007) Comprehensive testing of positionally cloned asthma genes in two populations. Am J Respir Crit Care Med 176(9):849–857

Holloway JW, Koppelman GH (2007) Identifying novel genes contributing to asthma pathogenesis. Curr Opin Allergy Clin Immunol 7:69–74

Iizuka M, Katsuyama Y, Mabuchi T, Umezawa Y, Matsuyama T, Ozawa A, Kawada H, Inoko H, Morita E, Ota M (2002) Genetic association analysis using microsatellite markers in atopic dermatitis. Tokai J Exp Clin Med 27:51–56

Jorde LB, Watkins WS, Bamshad MJ, Dixon ME, Ricker CE, Seielstad MT, Batzer MA (2000) The distribution of human genetic diversity: a comparison of mitochondrial, autosomal, and Y-chromosome data. Am J Hum Genet 66:979–988

Kashi Y, King DG (2006) Simple sequence repeats as advantageous mutators in evolution. Trends Genet 22:253–259

Kawashima M, Tamiya G, Oka A, Hohjoh H, Juji T, Ebisawa T, Honda Y, Inoko H, Tokunaga K (2006) Genomewide association analysis of human narcolepsy and a new resistance gene. Am J Hum Genet 79:252–263

Keicho N, Ohashi J, Tamiya G, Nakata K, Taguchi Y, Azuma A, Ohishi N, Emi M, Park MH, Inoko H, Tokunaga K, Kudoh S (2000) Fine localization of a major disease-susceptibility locus for diffuse panbronchiolitis. Am J Hum Genet 66:501–507

Kiley J, Smith R, Noel P (2007) Asthma phenotypes. Curr Opin Pulm Med 13:19–23

Mathias RA, Freidhoff LR, Blumenthal MN, Meyers DA, Lester L, King R, Xu JF, Solway J, Barnes KC, Pierce J, Stine OC, Togias A, Oetting W, Marshik PL, Hetmanski JB, Huang SK, Ehrlich E, Dunston GM, Malveaux F, Banks-Schlegel S, Cox NJ, Bleecker E, Ober C, Beaty TH, Rich SS (2001) Genome-wide linkage analyses of total serum IgE using variance components analysis in asthmatic families. Genet Epidemiol 20:340–355

Matsuzaka Y, Makino S, Okamoto K, Oka A, Tsujimura A, Matsumiya K, Takahara S, Okuyama A, Sada M, Gotoh R, Nakatani T, Ota M, Katsuyama Y, Tamiya G, Inoko H (2002) Susceptibility locus for non-obstructive azoospermia is localized within the HLA-DR/DQ subregion: primary role of DQB1*0604. Tissue Antigens 60:53–63

Moffatt MF, Kabesch M, Liang L, Dixon AL, Strachan D, Heath S, Depner M, von Berg A, Bufe A, Rietschel E, Heinzmann A, Simma B, Frischer T, Willis-Owen SA, Wong KC, Illig T, Vogelberg C, Weiland SK, von Mutius E, Abecasis GR, Farrall M, Gut IG, Lathrop GM, Cookson WO (2007) Genetic variants regulating ORMDL3 expression contribute to the risk of childhood asthma. Nature 448:470–473

Nicolae D, Cox NJ, Lester LA, Schneider D, Tan Z, Billstrand C, Kuldanek S, Donfack J, Kogut P, Patel NM, Goodenbour J, Howard T, Wolf R, Koppelman GH, White SR, Parry R, Postma DS, Meyers D, Bleecker ER, Hunt JS, Solway J, Ober C (2005) Fine mapping and positional candidate studies identify HLA-G as an asthma susceptibility gene on chromosome 6p21. Am J Hum Genet 76:349–357

Noguchi E, Yokouchi Y, Zhang J, Shibuya K, Shibuya A, Bannai M, Tokunaga K, Doi H, Tamari M, Shimizu M, Shirakawa T, Shibasaki M, Ichikawa K, Arinami T (2005) Positional identification of an asthma susceptibility gene on human chromosome 5q33. Am J Respir Crit Care Med 172:183–188

Ohashi J, Tokunaga K (2003) Power of genome-wide linkage disequilibrium testing by using microsatellite markers. J Hum Genet 48:487–491

Oka A, Tamiya G, Tomizawa M, Ota M, Katsuyama Y, Makino S, Shiina T, Yoshitome M, Iizuka M, Sasao Y, Iwashita K, Kawakubo Y, Sugai J, Ozawa A, Ohkido M, Kimura M, Bahram S, Inoko H (1999) Association analysis using refined microsatellite markers localizes a susceptibility locus for psoriasis vulgaris within a 111 kb segment telomeric to the HLA-C gene. Hum Mol Genet 8:2165–2170

Ota M, Mizuki N, Katsuyama Y, Tamiya G, Shiina T, Oka A, Ando H, Kimura M, Goto K, Ohno S, Inoko H (1999) The critical region for Behcet disease in the human major histocompatibility complex is reduced to a 46-kb segment centromeric of HLA-B, by association analysis using refined microsatellite mapping. Am J Hum Genet 64:1406–1410

Peat JK, Haby M, Spijker J, Berry G, Woolcock AJ (1992) Prevalence of asthma in adults in Busselton, Western Australia. BMJ 305:1326–1329

Raby BA, Weiss ST (2004) ADAM33: where are we now? Am J Respir Cell Mol Biol 31:1–2

Rachelefsky GE (2007) Difficult-to-control asthma: underlying factors, clinical implications, and treatment strategies. Curr Med Res Opin

Raymond M, Rousset F (1995) GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. J Hered 86:248–249

Sawyer SL, Mukherjee N, Pakstis AJ, Feuk L, Kidd JR, Brookes AJ, Kidd KK (2005) Linkage disequilibrium patterns vary substantially among populations. Eur J Hum Genet 13:677–86

Stevanin G, Brice A (2007) Spinocerebellar ataxia 17 (SCA17) and Huntington's disease-like 4 (HDL4). Cerebellum 1–9

Tamiya G, Shinya M, Imanishi T, Ikuta T, Makino S, Okamoto K, Furugaki K, Matsumoto T, Mano S, Ando S, Nozaki Y, Yukawa W, Nakashige R, Yamaguchi D, Ishibashi H, Yonekura M, Nakami Y, Takayama S, Endo T, Saruwatari T, Yagura M, Yoshikawa Y, Fujimoto K, Oka A, Chiku S, Linsen SE, Giphart MJ, Kulski JK, Fukazawa T, Hashimoto H, Kimura M, Hoshina Y, Suzuki Y, Hotta T, Mochida J, Minezaki T, Komai K, Shiozawa S, Taniguchi A, Yamanaka H, Kamatani N, Gojobori T, Bahram S, Inoko H (2005) Whole genome association study of rheumatoid arthritis using 27 039 microsatellites. Hum Mol Genet 14:2305–2321

Tassone F, Adams J, Berry-Kravis EM, Cohen SS, Brusco A, Leehey MA, Li L, Hagerman RJ, Hagerman PJ (2007) CGG repeat length correlates with age of onset of motor signs of the fragile X-associated tremor/ataxia syndrome (FXTAS). Am J Med Genet B Neuropsychiatr Genet 144:566–569

van den Oord E, McClay J, York T, Murrelle L, Robles J (2007) Genetics and diagnostic refinement. Behav Genet 37:535–545

Van Eerdewegh P, Little RD, Dupuis J, Del Mastro RG, Falls K, Simon J, Torrey D, Pandit S, McKenny J, Braunschweiger K, Walsh A, Liu Z, Hayward B, Folz C, Manning SP, Bawa A, Saracino L, Thackston M, Benchekroun Y, Capparell N, Wang M, Adair R, Feng Y, Dubois J, FitzGerald MG, Huang H, Gibson R, Allen KM, Pedan A, Danzig MR, Umland SP, Egan RW, Cuss FM, Rorke S, Clough JB, Holloway JW, Holgate ST, Keith TP (2002) Association of the ADAM33 gene with asthma and bronchial hyperresponsiveness. Nature 418:426–430

Vassos E, Panas M, Kladi A, Vassilopoulos D (2007) Effect of CAG repeat length on psychiatric disorders in Huntington's disease. J Psychiatr Res

Vaughan T, Pasco JA, Kotowicz MA, Nicholson GC, Morrison NA (2002) Alleles of RUNX2/CBFA1 gene are associated with differences in bone mineral density and risk of fracture. J Bone Miner Res 17:1527–1534

Wierenga EA, Walchner M, Kick G, Kapsenberg ML, Weiss EH, Messer G (1999) Evidence for suppressed activity of the transcription

factor NFAT1 at its proximal binding element P0 in the IL-4 promoter associated with enhanced IL-4 gene transcription in T cells of atopic patients. Int Immunol 11:297–306

Wilk JB, DeStefano AL, Arnett DK, Rich SS, Djousse L, Crapo RO, Leppert MF, Province MA, Cupples LA, Gottlieb DJ, Myers RH (2003) A genome-wide scan of pulmonary function measures in the National Heart, Lung, and Blood Institute Family Heart Study. Am J Respir Crit Care Med 167:1528–1533

Zhang Y, Leaves NI, Anderson GG, Ponting CP, Broxholme J, Holt R, Edser P, Bhattacharyya S, Dunham A, Adcock IM, Pulleyn L, Barnes PJ, Harper JI, Abecasis G, Cardon L, White M, Burton J, Matthews L, Mott R, Ross M, Cox R, Moffatt MF, Cookson WO (2003) Positional cloning of a quantitative trait locus on chromosome 13q14 that influences immunoglobulin E levels and asthma. Nat Genet 34:181–186

BMC Genomics

BioMed Central

Research article

Open Access

Major histocompatibility complex (Mhc) class Ib gene duplications, organization and expression patterns in mouse strain C57BL/6

Masato Ohtsuka*[1], Hidetoshi Inoko[1], Jerzy K Kulski[2] and Shinichi Yoshimura[1]

Address: [1]Division of Basic Molecular Science and Molecular Medicine, School of Medicine, Tokai University, Bohseidai, Isehara, Kanagawa 259-1193, Japan and [2]Centre for Forensic Science, The University of Western Australia, Nedlands, WA, Australia

Email: Masato Ohtsuka* - masato@is.icc.u-tokai.ac.jp; Hidetoshi Inoko - hinoko@is.icc.u-tokai.ac.jp; Jerzy K Kulski - kulski@mac.com; Shinichi Yoshimura - syoshimu@is.icc.u-tokai.ac.jp

* Corresponding author

Published: 17 April 2008

BMC Genomics 2008, **9**:178 doi:10.1186/1471-2164-9-178

Received: 7 November 2007
Accepted: 17 April 2008

This article is available from: http://www.biomedcentral.com/1471-2164/9/178

Abstract

Background: The mouse has more than 30 *Major histocompatibility complex* (*Mhc*) class Ib genes, most of which exist in the *H2* region of chromosome 17 in distinct gene clusters. Although recent progress in *Mhc* research has revealed the unique roles of several *Mhc* class Ib genes in the immune and non-immune systems, the functions of many class Ib genes have still to be elucidated. To better understand the roles of class Ib molecules, we have characterized their gene duplication, organization and expression patterns within the *H2* region of the mouse strain C57BL/6.

Results: The genomic organization of the *H2-Q*, *-T* and *-M* regions was analyzed and 21 transcribed *Mhc* class Ib genes were identified within these regions. Dot-plot and phylogenetic analyses implied that the genes were generated by monogenic and/or multigenic duplicated events. To investigate the adult tissue, embryonic and placental expressions of these genes, we performed RT-PCR gene expression profiling using gene-specific primers. Both tissue-wide and tissue-specific gene expression patterns were obtained that suggest that the variations in the gene expression may depend on the genomic location of the duplicated genes as well as locus specific mechanisms. The genes located in the *H2-T* region at the centromeric end of the cluster were expressed more widely than those at the telomeric end, which showed tissue-restricted expression in spite of nucleotide sequence similarities among gene paralogs.

Conclusion: Duplicated *Mhc* class Ib genes located in the *H2-Q*, *-T* and *-M* regions are differentially expressed in a variety of developing and adult tissues. Our findings form the basis for further functional validation studies of the *Mhc* class Ib gene expression profiles in specific tissues, such as the brain. The duplicated gene expression results in combination with the genome analysis suggest the possibility of long-range regulation of *H2-T* gene expression and/or important, but as yet unidentified nucleotide changes in the promoter or enhancer regions of the genes. Since the *Mhc* genomic region has diversified among mouse strains, it should be a useful model region for comparative analyses of the relationships between duplicated gene organization, evolution and the regulation of expression patterns.

Background

The *Major Histocompatibility Complex* (*MHC*) genomic region harbors duplicated genes that express protein molecules responsible for the rejection of transplanted tissue, restricted antigen presentation and the recognition of self and non-self [1,2]. The *Mhc* genomic region in the mouse, located on chromosome 17, is named *H2* and the genes within this region are usually classified into three distinct classes (I to III) based on their structure and function [3]. The class I molecules generally elicit immune responses by presenting peptide antigens derived from intracellular proteins to T lymphocytes and their genes can be classified into two groups, the classical *Mhc* class I (class Ia) genes and the non-classical *Mhc* class I (class Ib) genes. The classical *Mhc* class Ia genes, such as *H2-K* and *-D* in the mouse, are highly polymorphic, expressed widely and present antigens to CD8+ cytotoxic T cells. To date, most studies of the *MHC* class I genomic region have been focused on the immunological function of class Ia molecules [4-6].

The non-classical class Ib molecules are structurally similar to the classical class Ia proteins, but in contrast to the classical class Ia proteins, they have limited or no polymorphisms. They are more restricted in their tissue expression and some have functions other than antigen presentation to CD8+ T cells. The non-classical class Ib proteins have shorter cytoplasmic tails and some of them lack consensus residues associated with peptide binding [7]. The mouse is considered to have more than 30 *Mhc* class Ib genes in the genome [3]. Most *Mhc* class Ib genes are located at the telometric end of the 2 Mb-*H2* region within the *H2-Q*, *-T* and *-M* sub-regions, which were originally mapped and defined by recombination analysis. Although the non-classical class Ib genes are involved in immunological functions like the classical class Ia genes, they generally serve a more specialized role in the immune responses. The expression and function of some non-classical class Ib genes, including *H2-T23* (Qa-1), *-M3* and *-T3* (TL antigen), have been analyzed in detail. For example, Qa-1 is involved in the suppression of CD4+ T cell responses via CD94/NKG2A or CD94/NKG2C receptors [8,9]. The peptide presentation by the Qa-1 molecule may also have a role in CD8+ regulatory T cell activity [10]. H2-M3 molecules prime the rapid response of CD8+ T cells by presenting *N*-formylated bacterial peptides [11]. The TL antigen is involved in the formation of memory CD8+ T cells [12] and in the regulation of iIEL responses in the intestine by interaction with homodimeric CD8 alpha receptors [13].

The class Ib molecules are also involved in non-immune functions. For example, the *H2-M1* and *-M10* families of the class Ib genes specifically interact with the V2R class of pheromone receptors presented on the cell surfaces of the vomeronasal organ [14,15]. The Qa-2 proteins encoded by *H2-Q7* and *-Q9* class Ib genes influence the rate of preimplantation embryonic development and subsequent embryonic survival [16]. In addition, the class I molecules have recently been shown to contribute to the development and plasticity of the brain [17,18]. So far, there is little information about which of the non-classical class Ib genes are involved in this function.

The molecular functions of many of the other class Ib molecules are still far from being understood and even the expression patterns for many of the *Mhc* class Ib genes remain to be elucidated. The *Mhc* class Ib genes are members of gene clusters that have been generated by different rounds of duplication and deletion [19]. In the mouse, the telomeric 1 Mb of the *Mhc* including the *H2-M* region was well characterized using the 129/Sv inbred strain [20]. The possible evolutionary fates of duplicated genes are nonfunctionalization, neofunctionalization or subfunctionalization [21]. Genes recently duplicated may even have the same functions by having and using identical or similar expression domain sequences. In order to better understand the role of class Ib molecules expressed by duplicated genes in different tissues, we have undertaken to examine, identify and characterize the *Mhc* class Ib gene duplication, organization and expression patterns within the *H2* region of the mouse strain C57BL/6.

The whole genome of the laboratory mouse strain C57BL/6J has been almost fully sequenced [22]. However, the genomic organization of the *Mhc* class I region of mice varies markedly between different haplotypes and inbred strains [20]. In the present study, we selected *Mhc* class Ib DNA sequences from the mouse genome database (NCBI Entrez Genome Project ID 9559), and characterized the organization of the *Mhc* class Ib genomic region for the mouse C57BL/6 strain (haplotype b). Expression patterns of each of the *Mhc* class Ib genes were examined by RT-PCR using gene-specific primer sets, and we identified *Mhc* class Ib genes with either tissue-restricted expression or tissue-wide expression. We also identified monogenic and multigenic duplicated regions within the *H2-T* region of the mouse inbred-strain, C57BL/6. Based on the results of our comprehensive analysis of the *Mhc* class Ib gene duplication, organization and expression patterns, we discuss the possible relationships and regulatory outcomes between the genomic location and expression patterns of the mouse *Mhc* class Ib duplicated genes.

Results and Discussion

Identification and genomic organization of transcribed Mhc class Ib genes

As the aim of this study was to determine the tissue expression patterns for each of the duplicated *Mhc* class Ib genes, we first needed to identify the location and the number of transcribing *Mhc* class Ib genes in the mouse genomic

sequence [22]. Although a nearly complete mouse genomic sequence of this region was available in the public database, there were many large sequence gaps and incomplete annotations for the sequence when we started this study. Therefore, we predicted the putative *Mhc* class Ib genes from the genomic contig NT_039650.2 by using the GENSCAN program. This analysis identified 19 *Mhc* class Ib-like sequences with coding potential (data not shown). Based on these sequences and the information obtained from the public databases, we designed gene-specific primer sets (Table 1) and confirmed the expression of the predicted genes by RT-PCR against a panel of cDNA tissues as described below. The nucleotide sequences, determined by direct-sequencing of the RT-PCR-amplified fragments, were registered with the GenBank/DDBJ sequence database and given the accession numbers, [GenBank:AB266872, AB266873, AB267092–AB267096]. As a result, a total of 15 expressed genes were identified and mapped onto the current genomic sequence ("GS" number in Figure 1). Although there may be a possibility of misassemblies or misequencing of genomic sequence, most of the assembled sequence, especially the order of genes, is thought to be correct considering the fact that the distributions of restriction sites (such as *Eco*RI, *Bam*HI and *Kpn*I) are consistent with previous reports (data not shown) [23,24], and that the cDNA sequences we examined were perfectly matched with genomic sequence. However, there was no genomic sequence corresponding to the *H2-Q8* and *-Q9* genes that are believed to be present in C57BL/10 (haplotype b). At present, we do not know with certainty whether the assembly of the genome is completely correct in this region. Although the *H2-Q5* locus was annotated as *H2-Q8* in the genome database, we designated this locus as *H2-Q5* for the following reasons. 1) This locus was consistent with the physical map position of *H2-Q5* in the previous report [23], and 2) the DNA sequence of this locus is different from the *H2-Q8* gene of C57BL/10 (U57392). This analysis in combination with a previous report [25] revealed that a total of at least 21 *Mhc* class Ib genes, 7 in the *H2-Q* region, 11 in the *H2-T* region and 3 in the *H2-M* region are definitely transcribed in the C57BL/6 mouse. However, in the present study, we did not consider the *H2-M1, -10* family of *Mhc* class Ib genes that are located outside the *H2-Q* and *-T* genomic regions.

Table 2 presents the *H2* gene numbering system for C57BL/6 mice (haplotype b) that we have used in this paper. We designated each *H2* gene with reference to the genomic locations and designations used by others [23,24,26]. The nucleotide sequences were determined for the full-length cDNAs expressed by the genes *H2-T23, -T22, -T15, -T5* and *-M5* and submitted to the GenBank database [GenBank:AB359227–AB359231]. All genes exhibited a standard class I structure with an alpha 1, alpha 2, alpha 3 and transmembrane (TM) domain.

Monogenic and multigenic duplications

Mhc class I genes tend to diversify between species or strains as a result of local duplications and deletions [27]. As local duplication often generates similar genes with similar expression pattern and functional redundancy, it is important to understand the genomic organization and evolution of the *Mhc* class Ib regions. Hence, dot-plot analysis was conducted by comparing the sequences of the *H2-Q* and *-T* regions to themselves (Figure 1B; 240,000 bp for *H2-Q* cluster, 250,000 bp for *H2-T* cluster). In addition to the short diagonal lines seen in the dot-plots due to the similarity of each *Mhc* class I gene, long diagonal lines that indicate evidence of local duplications are seen in both the *H2-Q* and *-T* regions. Regarding the *H2-Q* region, duplication is evident in approximately a 52-kb region from *H2-Q4* to *-Q10* (Figure 1B left). A long diagonal line is also seen in the *H2-T* region (Figure 1B right) indicating a multigenic duplication event within the *H2-T* region from *H2-T23* to *-T5* (Figure 2). The phylogenetic tree of *H2-T* genes (Figure 1C) supports the occurrence of a multigenic duplication event that produced some gene sets with a high sequence similarity (> 85% in coding region, e.g. *H2-T11* and *-T23, H2-T9* and *-T15, H2-T10* and *-T22*, and *H2-T5, -T7* and *-T13*). Similarities between these genes are seen not only in the coding region, but also in the untranslated region, introns and intergenic regions (Figures 1B and 3), indicating the possibility that these genes have a redundant function and/or expression pattern.

Figure 2 shows a schematic representation of a single multigenic tandem duplication of four ancestral genes that generated eight genes within the genomic D1 and D2 duplication products. The model also shows that before the occurrence of the multigenic duplication event a single monogenic tandem duplication had probably generated a copy of the *H2-T5* gene. This parsimonious model helps to explain the gene organization (Figure 1B), phylogenetic topologies of the gene sequences (Figure 1C) and the sequence similarities (Figure 3) between *H2-T23* and *-T11, H2-T22* and *-T10, H2-T15* and *-T9*, and *H2-T13* and *-T7*. However, the multigenic duplication model presented here for the mouse *H2-T* region has not taken into account the presence of pseudogenes *T1* and *T2* and other evolutionary mechanisms that may have contributed to diversity within this region, such as gene conversions and unequal cross-overs with other haplotypes. Nevertheless, the multigenic duplication model for the mouse *H2-T* region is similar to the multigenic tandem duplication models previously proposed for the *Mhc* class I region of human and non-human primates [28,29].

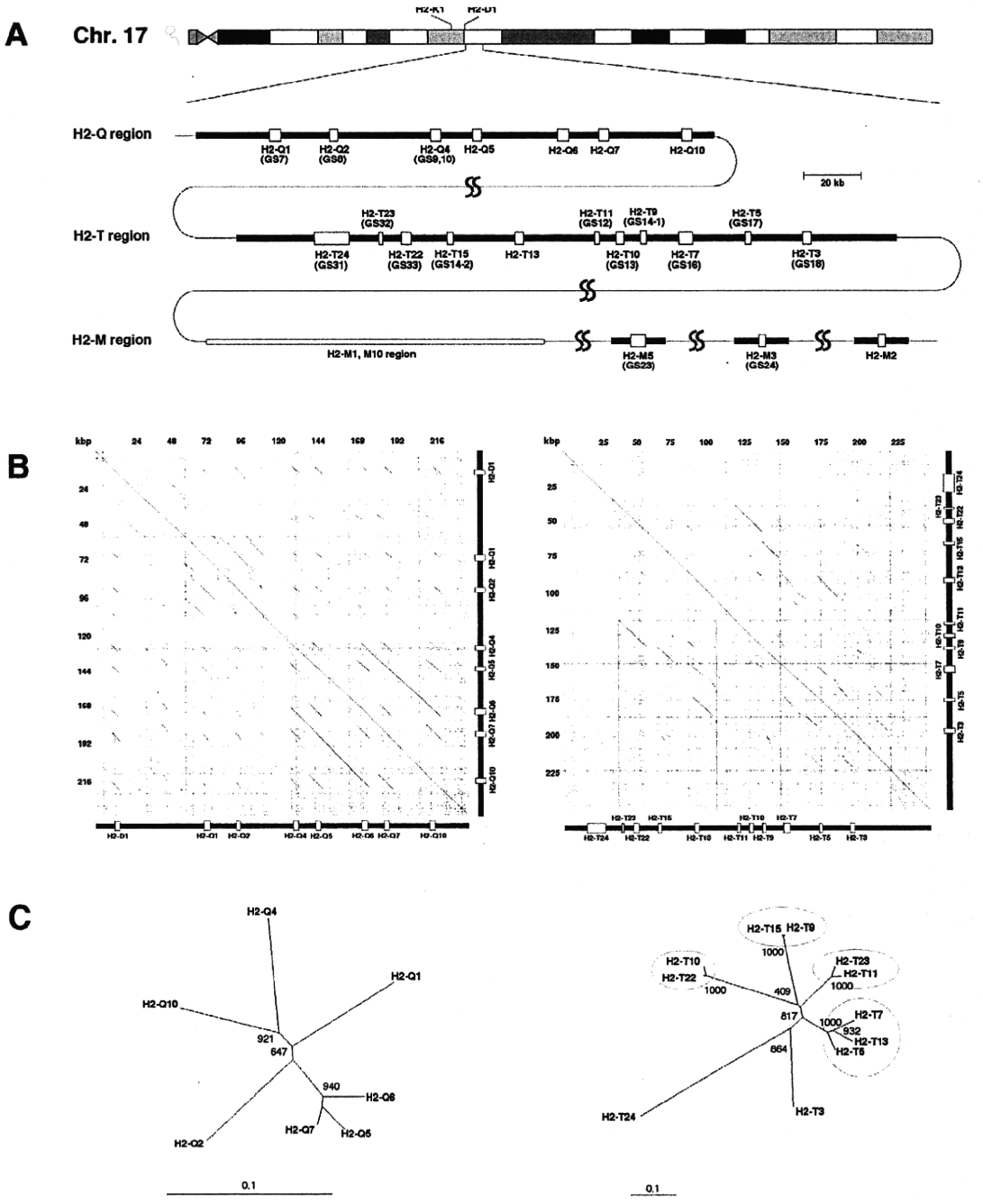

Figure 1

Genomic organization of the *H2-Q*, *-T* and *-M* region. (A) Gene content of the *H2-Q*, *-T* and *-M* regions on chromosome 17 of mouse C57BL/6 strain. The *Mhc* class Ib genes with coding potential are represented by yellow boxes. The genes determined by our initial analysis using GENSCAN program are indicated with a GS number. Gene contents regarding *H2-M1* and *-M10* families were omitted in this figure. The regions indicated by the squiggles are the regions where the non-*Mhc* genes are interspersed. The scale bar indicating 20 kb applies to the *H2-Q* and *-T* regions. (B) Dot-plot comparisons of the mouse *H2-Q* (left) and *-T* (right) regions. Comparison of the sequence to itself reveals the duplicated regions. (C) Phylogenetic tree analysis of *H2-Q* (left) and *-T* (right) genes based on nucleotide sequences of the entire coding region. Gene pairs showing highly similar sequences (>85%) in *H2-T* region are represented by red circles. Bootstrap values (1000 replicates) are indicated. A scale bar of "0.1" represents a branch-length of 0.1 nucleotide substitutions per site.

Table 1: List of gene specific primer sets used for expression pattern analysis

Gene	Primer set		Size (bp)
	Forward	Reverse	
H2-K1	GGGAGCCCCGGTACATGGAA	GGTGACTTTATCTTCAGGTCTGCT	548
H2-D1	TCGGCTATGTGGACAACAAGG	GGCCATAGCTCCAAGGACAC	818
H2-Q1	CTGCGGTATTTCGAGACCTCG	GGTATCTGTGGAGCCACATCAG	502/686
H2-Q2	ACACACAGGTCTCCAAGGAA	TGGATCTTGAGCGTAGTCTCTTA	785
H2-Q4	CTTGCTGAGTTATTTCTACACCT	ACCGTCAGATCTGTGGTGACAT	583
H2-Q5	GGGAGCCCCGGTTCATCATC	CAGGGTGACAGCATCATAAGATA	539
H2-Q6	GTATTTCCACACTGCTGTGTCCT	AAGGACAACCAGAATAGCTACGT	871
H2-Q7	CGGGCCAACACTCGCTGCAA	GTATCTGCGGAGCGACTGCAT	515
H2-Q10	CACACTCCATGAGGTATTTCGAA	CAGATCAGCAATGTGTGACATGATA	590/866
H2-T24	ATGCACAGTACTTCACTCATG	CCCCTAGCATATACTCCTGTCG	736/839
H2-T23	AGTATTGGGAGCGGGAGACTT	AGCACCTCAGGGTGACTTCAT	438
H2-T22	CTGGAGCAGGAGGAAGCAGATA	CAAATGATGAACAAAATGAAAACCA	698
H2-T15	ACCGCCCTGGCCCCGACCCAA	CATCCGTGCATATCCTGGATT	332
H2-T13	GCCCTGACTATGATCGAGACT	CACCTCAGGGTGACATCACCTG	635
H2-T11	CGGTATTTCCACACCGTCGTA	TAGAGATATGCGAGGCTAAGTTG	415/628
H2-T10	CCCTTTGGGTTCACACTCGCTT	CCTGGTCTCCACAAACTCCACTTCT	661
H2-T9	ACCGCCCTGGCCCCGACCCGA	CATCCGTGCATATCCTGGATA	332
H2-T7	CTTCACACGTTCCAGCTGTTGTT	AGGCCTGGTCTCCACAAGCTCT	432
H2-T5	GGTGGTGTTGCAGAGACGCT	CTGCTCTTCAACACAAAAGG	482
H2-T3	TTCAACAGCTCAGGGGAGACTG	AAGCTCCGTGTCCTGAATCAAT	585
H2-M3	CAGCGCTGTGATAGCATTGA	ACAACAATAGTGATCACACCT	806
H2-M2	GAGGAGACCCACTACATGACTGTT	GAAAATGAAAGACTGAGGAGGTCTAC	798
*b2m**	ATGGCTCGCTCGGTGACCCTG	ATTGCTCAGCTATCTAGGATA	546
*GAPDH**	TGAAGGTCGGTGTGAACGGATTTG	GGCCTTCTCCATGGTGGTGAAGAC	314

*Sequences of these primers were obtained from a previous report [25].

Regarding the *H2-Q* region, the genes *H2-Q5*, *-Q6* and *-Q7*, which form a tandem array in the *H2-Q* region (Figure 1B), also grouped relatively closely together in the phylogenetic tree analysis (Figure 1C). Assuming the current genome assembly is correct, then these three genes were probably generated by two separate monogenic tandem duplications much more recently than the duplications previously involved with the generation of the *H2-Q1*, *-Q2*, *-Q4* and *-Q10* genes, which are more distantly related in sequence in the phylogenetic analysis. However, the duplication structure of the *H2-Q* region in C57BL/6 (Figure 1B left) appears to be different to the mouse strain 129/SvJ [30].

Expression of Mhc class Ib genes in adult tissues

To clarify the tissue expression patterns for each of the *Mhc* class Ib genes, we conducted RT-PCR analysis of the cDNAs isolated from various tissues of the adult mouse. Although it is difficult to analyze *Mhc* class I expression due to the sequence similarity of the *Mhc* genes (showing 60 – 95% identities in coding region; data not shown), we circumvented this disadvantage by designing the gene-specific primer sets that are listed in Table 1. Transcription of each *Mhc* class I gene was detected as shown in Figure 4. The gene identities of the amplified cDNAs were confirmed by direct sequencing of the RT-PCR-amplified fragments (indicated by yellow asterisks in Figure 4). Using the specific primer sets, we successfully amplify most of the identified *Mhc* class Ib genes, except for *H2-M5*, which may be expressed at very low levels and below the limit of detection of our RT-PCR assays. We obtained amplified fragments of the *H2-M5* gene from the brain and thymus, but we were unable to detect amplified products in the other tissues (data not shown).

The gene expression patterns were classified into two types: tissue-wide or tissue-specific expression. *H2-Q4*, *-Q7*, *-T24*, *-T23*, *-T22* and *-M3* as well as the class Ia genes (*H2-K1* and *-D1*) exhibited tissue-wide expression. In contrast, *H2-Q1*, *-Q2*, *-Q5*, *-Q6*, *-Q10*, *-T15*, *-T13*, *-T11*, *-T10*, *-T9*, *-T7*, *-T5*, *-T3* and *-M2* genes were expressed in a tissue-specific manner. Regardless of the tissue-wide or tissue-specific expression patterns, most of the class I genes were expressed in the thymus and intestine, both of which are critical organs for immune defense.

The tissue expression patterns of the genes *H2-T11* and *-T10* located within the duplicated D2 region (Figure 2) are more tissue-restricted than those of the respective paralogous genes *H2-T23* and *-T22* (Figures 4 and 5) that are located within the duplicated D1 region (Figure 2), confirming that major changes do occur in the expression

Table 2: List of mouse MHC class Ib genes analysed in this study

Gene name		mRNA sequences referred		
Used in this study	Others	NCBI accession	Ensembl transcript ID	NCBI accessions determined in this study
Q1 (GS7)		U96752	ENSMUST00000073208	-
Q2 (GS8)		AY989880	ENSMUST00000074806	AB266872
Q4 (GS9,10)	Qb-1	XR_034205	ENSMUST00000113887	AB267092, AB266873
Q5		-	ENSMUST00000040240	-
Q6		NM_207648	ENSMUST00000091611	-
Q7	Qa-2, Ped	NM_010394	ENSMUST00000071951	-
Q10		AK131620	ENSMUST00000068291	-
T24 (GS31)		NM_008207	ENSMUST00000066488	-
T23 (GS32)	Qa-1		ENSMUST00000102678	AB359230
T22 (GS33)		AK133985	ENSMUST00000058801	AB359229
T15 (GS14-2)		-	ENSMUST00000113742	AB359227
T13	Bl, blastocyst MHC, T25	AY989821	ENSMUST00000025333	-
T11 (GS12)		XM_975970	ENSMUST00000079918	-
T10 (GS13)		NM_010395	ENSMUST00000074201	-
T9 (GS14-1)		-	-	AB267093
T7 (GS16)		NM_001025208	ENSMUST00000064686	-
T5 (GS17)		NM_001081032	ENSMUST00000040467	AB359231
T3 (GS18)	TL	AK033602	ENSMUST00000025312	-
M5 (GS23)	CRW2	XM_903477	ENSMUST00000113667	AB359228
M3 (GS24)	Hmt	NM_013819	ENSMUST00000038580	AB267096
M2	Thy19.4	AY302212	ENSMUST00000077662	-

Haplotype b (C57BL/6) was used. Nomenclature of each gene was based on the previous reports [23, 26], except for GS number. The H2 prefixes were omitted. The GS numbers in parenthesis represent the gene sequence numbers used in our laboratory.

profiles and functions of recently duplicated genes. Of particular note is the loss of expression in the liver, heart, muscle and testis by *H2-T11*, as previously reported for the liver [31], in comparison to its paralogous gene, *H2-T23*; and the loss of expression of *H2-T10* in all tissues except the thymus, spleen, ovary and placenta in comparison to the tissue-wide expression by its paralogous gene *H2-T22*. The gene paralogs, *H2-T13*, *-T7* and *-T5*, all showed tissue specific expression in the small intestine, except that the brains of adults also expressed the *H2-T13* gene, the thymus and placenta expressed the *H2-T5* gene and the thymus, ovary and placenta expressed *H2-T7* (Figure 4).

The tissue expression patterns of the two flanking genes, *H2-T24* and *-T3*, in the *H2-T* region are markedly different and may be among the oldest of the genes in this region. The centromeric *H2-T24* gene was expressed widely, whereas the telomeric *H2-T3* gene expression was restricted to the thymus and the small intestine (Figure 4) as previously reported [12,13].

As described above, the genes *H2-Q5*, *-Q6* and *-Q7* were probably generated by monogenic tandem duplications. In this regard, *H2-Q7* showed the widest tissue expression, followed by *H2-Q6* and then *H2-Q5*. This suggests that there might have been a gain or loss of tissue specificity with each gene duplication event. Of the other *H2-Q* genes, the most tissue-wide expression was by *H2-Q4*.

The *Mhc* gene expression in the brain is of particular interest because such genes could have a specific function in brain development and plasticity [17]. In this study, we identified 12 class Ib genes, *H2-Q1*, *-Q2*, *-Q4*, *-Q7*, *-T24*, *-T23*, *-T22*, *-T15*, *-T13*, *-T11*, *-M3* and *-M5*, expressed in the brain. The *Mhc* gene expression in the brain warrants further investigation particularly to determine in what cells (neurons and/or various glial cells) and at what stages of brain development these genes are expressed.

Expression of Mhc class Ib genes in embryos and placentas

Some *Mhc* genes are known to express and function during development in the embryo [32,33] and/or in the placenta [34]. Therefore, we determined which of the 20 class Ib genes were expressed in the embryo and placenta (Figure 4 and 5). The expression of some of the class Ib genes gradually increased (e.g. *H2-Q10* and *-T7*) or decreased (e.g. *H2-Q6* and *-M3*) during the course of development. The class Ib genes that were expressed widely in the adult tissues (*H2-Q4*, *-Q7*, *-T24*, *-T23*, *-T22* and *-M3*) also tended to be expressed throughout the developmental stages. This observation suggests that the regions in which these class Ib genes are located may have an open or accessible chromatin configuration from the time of the first

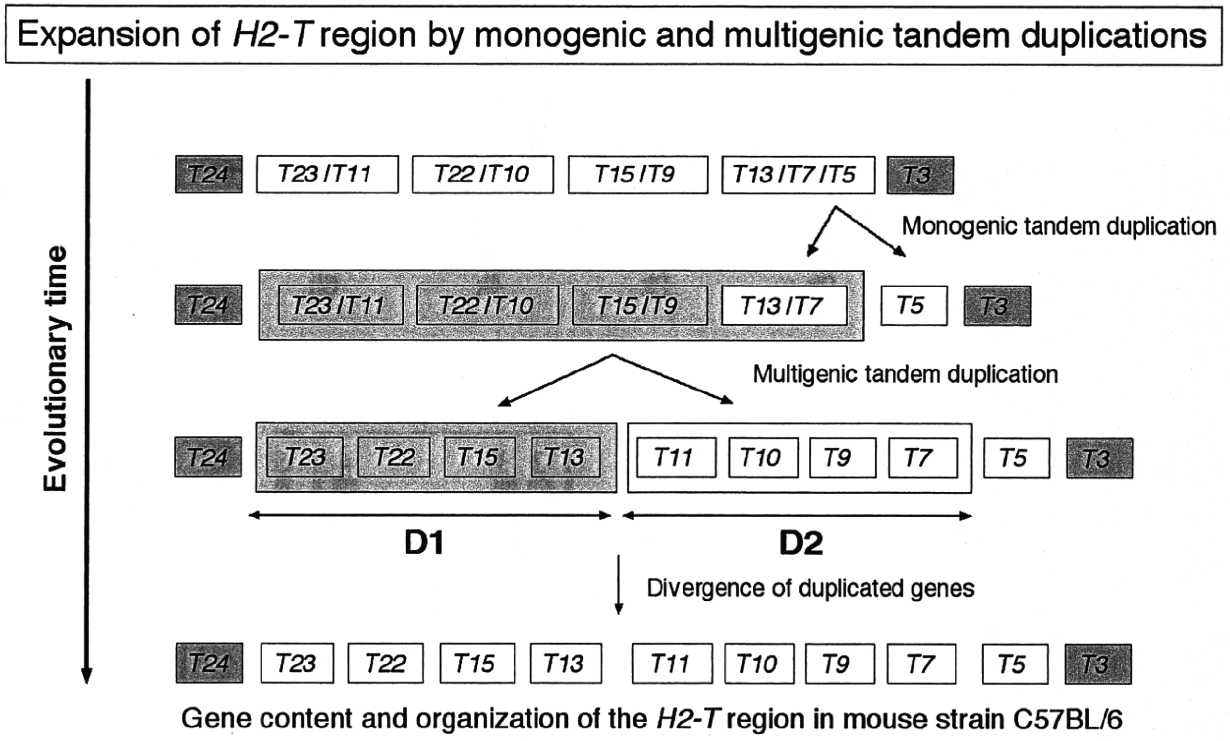

Figure 2
A schematic model of the expansion of the *H2-T* region by monogenic and multigenic tandem duplications. This model represents monogenic and multigenic tandem duplications originating from a hypothetical ancestral *H2-T* genomic sequence consisting of six *H2-T* genes. Each labeled box represents a *H2-T* gene in a linear array (horizontal) at different evolutionary times along the vertical axis. The horizontal double arrows labeled D1 and D2 represent the genomic products of the multigenic tandem duplication, with each product consisting of four genes.

observable developmental stage. We could not detect *H2-T13*, *-T10*, *-T9*, *-T3* and *-M2* in the embryo or placenta, although *H2-T13* (*Blastocyst MHC*) was previously shown to express in the placenta of B6 mice [34]. This negative result may be due to the developmental stage examined. Tajima et al. (2003) examined *Blastocyst MHC* gene expression at E3.5, E7.5 and E13.5 and expression at E13.5 was difficult to detect [34], while we analyzed gene expression at the developmental stages from E9.5 – E14.5.

We also examined the expression of class Ib genes in the brains of the E14.5 embryos (Figure 5). Nine genes (*H2-Q1*, *-Q2*, *-Q4*, *-Q7*, *-T24*, *-T23*, *-T22*, *-T11* and *-M3*) were transcribed in the brains of the E14.5 embryos. All of them were also expressed in the adult brain (Figure 4), indicating that these gene products may have a functional role in both adult and embryonic brains.

From the RT-PCR analyses in Figure 4 and 5, we identified alternative splicing variants in the *H2-Q1*, *-Q10*, *-T24*, *-T11*, *-T9* and *-M5* (for *M5* gene, see GenBank:AB378579) genes. The splicing patterns can be classified into four types: A) a common splicing pattern for class I gene, B) a loss of alpha2 domain, C) an unspliced second intron and D) an unspliced fourth intron. The type B variant was seen for *H2-Q10* and *-M5* expression, whereas type C was observed in *H2-Q1*, *-T11* and *-T9* expression. *H2-T24* showed type D variant. It is of interest in future to determine whether these splicing variants have distinct or common functions. The type A and type B variants were previously reported for the *H2-T13* (*Blastocyst MHC*) gene, and the RMA-S cell expressing the type B variant was protected from NK cell-mediated rejection via loading of its signal peptide onto the Qa-I molecules [34].

Expression patterns between duplicated class Ib genes

Since local duplication in the *H2-T* region (Figure 2) have produced gene sets with high sequence similarity (Figures 1B,C and 2) even in the upstream promoter region (Figure 3), a redundant expression pattern was expected between the similar genes. However, as described above, the expression patterns between similar genes were mostly

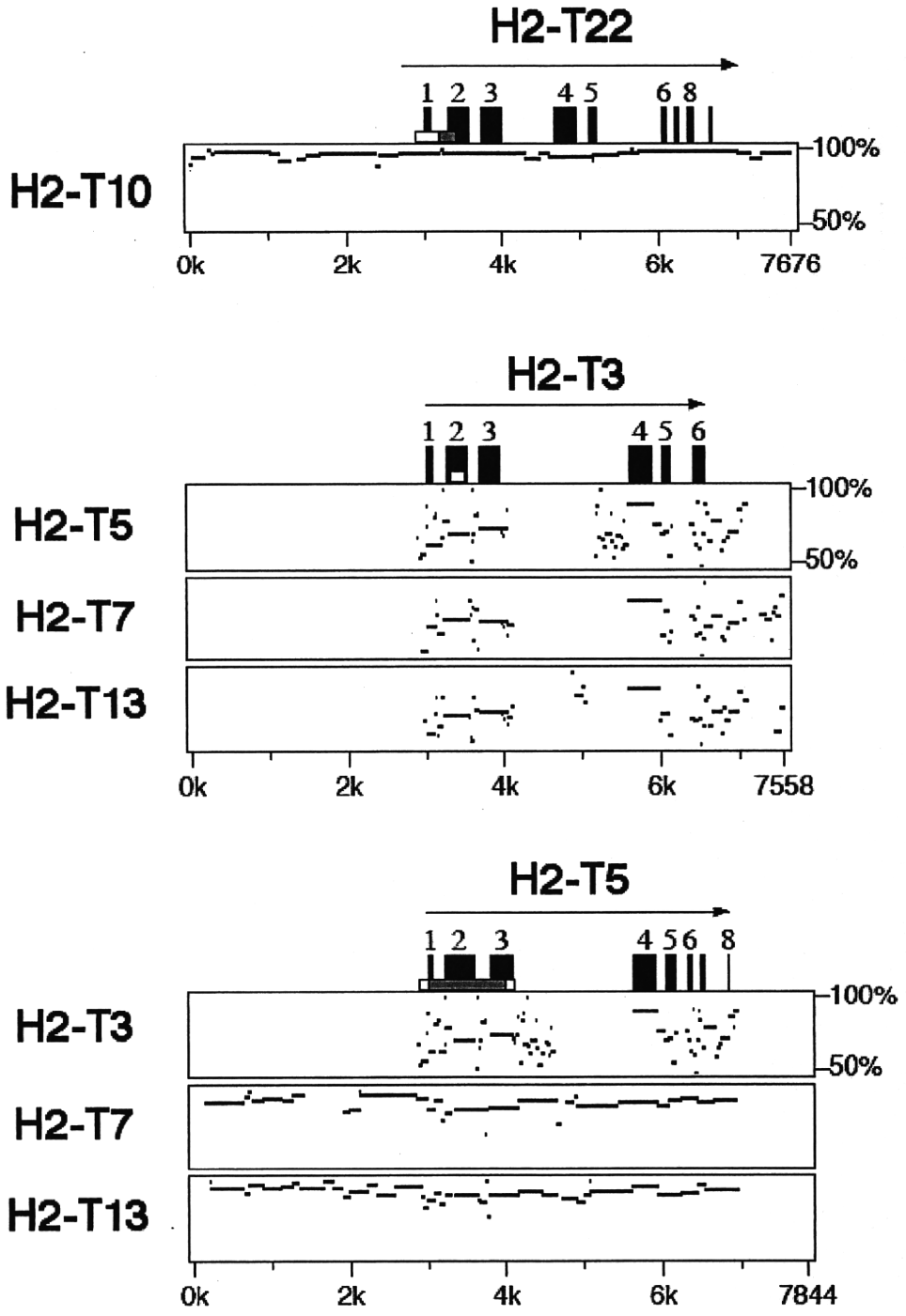

Figure 3
PipMaker analyses of genomic sequences of mouse *H2-T* genes. PipMaker analyses were performed to detect similarity within the promoter region. Sequences used for comparison include 3 kb of 5' upstream region and 1 kb of 3' downstream region of coding sequence for each gene. Exons are indicated by black boxes above the plot.

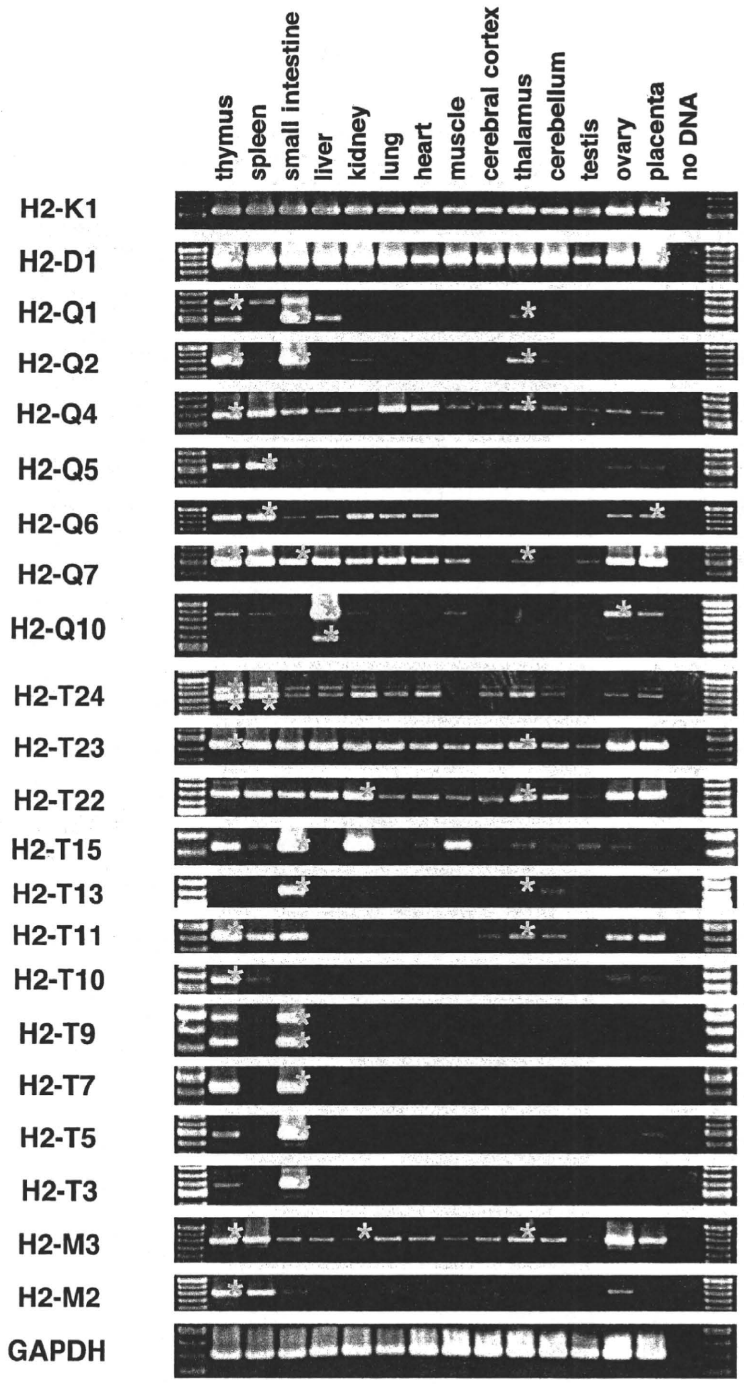

Figure 4
Expression of *Mhc* class Ib genes in adult tissues. RT-PCR was performed on total RNA isolated from tissues of C57BL/6J mouse. Identities of bands were confirmed by amplified sizes and by sequencing (indicated by yellow asterisk). The same reaction conditions were used for PCR.

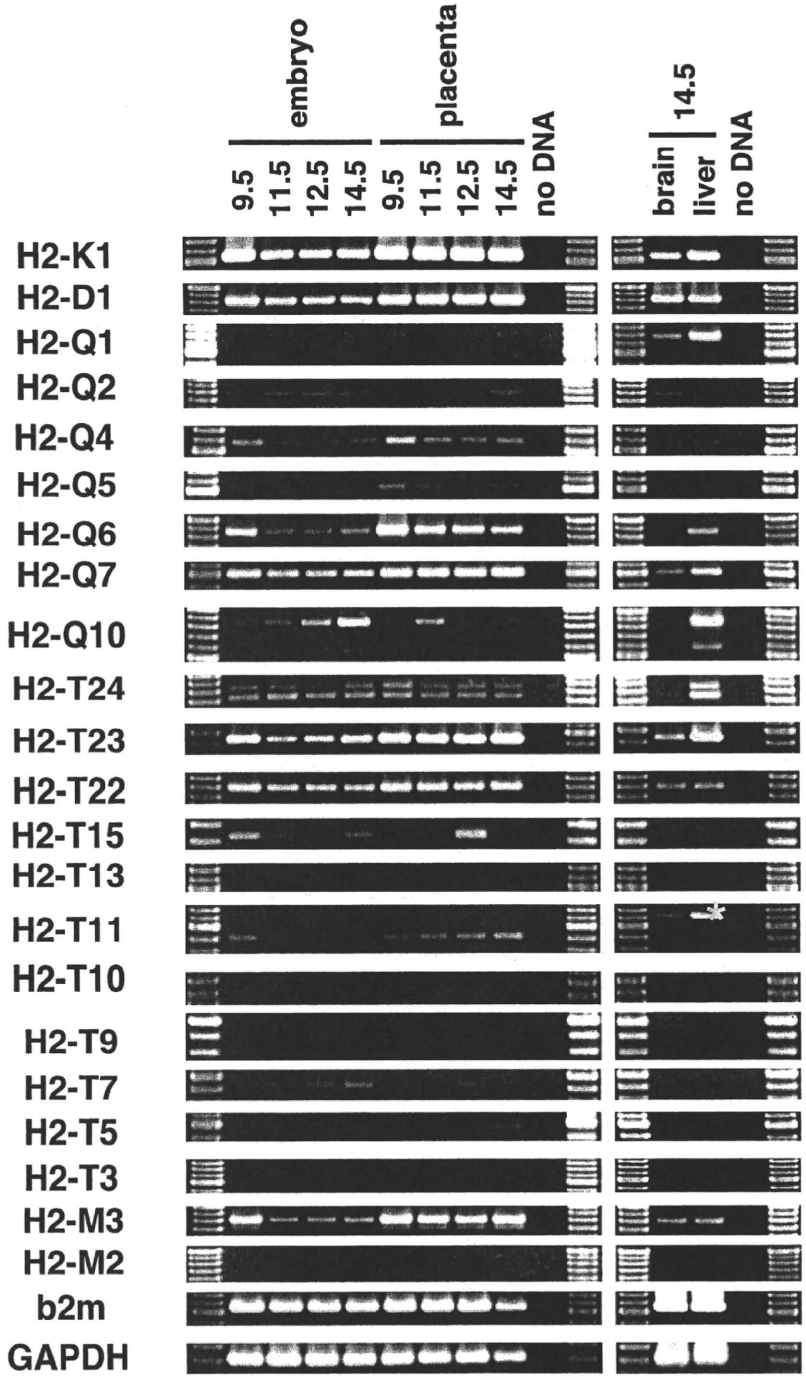

Figure 5
Embryonic and placental expression of *Mhc* class Ib genes. RT-PCR analysis was performed on total RNA isolated from E9.5–14.5 of C57BL/6J embryos and placentas. As for the E14.5 embryo, RT-PCR was conducted using cDNA as templates, derived from total RNA isolated from the brain and liver. Identities of bands were confirmed by amplified sizes and by sequencing (indicated by yellow asterisk). The same reaction conditions were used for PCR.

different. For example, *H2-T23* was expressed widely, whereas the *H2-T11* gene paralog showed a much more restricted expression pattern. This difference in expression between duplicated genes was especially remarkable for *H2-T22* and *-T10* expression (Figure 4) because the sequences of the upstream promoter regions of *H2-T22* and *-T10* are almost identical (Figure 3). In contrast, the *H2-T13*, *-T5* and *-T7* duplicated genes have similar nucleotide sequences, including within their promoter region, and similar expression patterns (predominantly in small intestine). This expression pattern, especially for *H2-T5* and *-T7*, was almost the same as for *H2-T3* that flanks these genes (Figure 2), but exhibited no similarity in the promoter sequence (Figure 3).

The co-expression of neighboring genes, such as *H2-T24* to *-T22* or *H2-T15* to *-T3* (Figure 2), may be regulated by 1) independent cis-acting regulatory elements for each gene that produce similar expression patterns, or by 2) a shared long-range regulatory element that operates over several genes (i.e. a long-range enhancer and/or a chromatin level regulation). Model 1 is appropriate for duplicated regions in which control regions are duplicated together with the coding sequence [35]. This is the most likely explanation for co-expression of *H2-T5* and *-T7* (Figure 3). The possibility that different promoter sequences produce a similar expression pattern might also be explained by model 1. The 2.8 kb promoter region of *H2-T3* was shown previously to direct transgene expression in the epithelial cells of the small and large intestine [36]. Therefore, it will be of interest in future to examine whether the upstream regions of *H2-T5* and *-T7* have the same activity as that of *H2-T3* (Figure 3). We think, however, it is unlikely that all the genes located between *H2-T24* to *-T22* or *H2-T15* to *T3* contain their own cis-regulatory element with similar function. Considering the order of the *H2-T* genes that show tissue-wide or tissue-specific expression, we rather favor model 2. The *H2-T* genes with tissue-wide expression are located within the same 40 kb centromeric portion of the *H2-T* region (*H2-T24* to *-T22*), whereas the genes *H2-T15* to *-T3* located at the telomeric-end exhibited a tissue-specific expression pattern with most of them predominantly expressed in the small intestine (Figure 4). The region containing the genes from *H2-T15* to *-T3* with the restricted tissue expression spans as much as 150 kb, which is consistent with the possibility of a long-range regulation. The long-range regulation may provide a simple explanation of different expression patterns of similar genes (e.g. *H2-T22* and *-T10*) and similar expression pattern of genes with distinct promoter regions (e.g. *H2-T5*, *-T7* and *-T3*) over long distances. This model is supported by recent papers that reported that a special AT-rich binding protein 1 (SATB1), the most characterized matrix attachment regions (MARs)-binding protein (MBP), is involved in the tissue-specific chromatin organization of the human *MHC* class I locus and its expression profile [37,38].

The mouse is known to have strain-specific gene duplications in the *H2-T* region with a number of duplicated *H2-T* gene differences between strains producing considerable variability between haplotypes [39,40]. The genomic features, organization and the expression patterns of the *H2-T* genes in other mouse strains warrant a comparative analysis. The expression pattern analysis of rat *Mhc* class Ib genes [41] may also provide clues for our hypothesis for the long-range regulation of duplicated class Ib gene expression. In addition, an investigation of gene duplications in genetically modified mice may help to distinguish between the different models involved in the regulation of duplicated gene expression. We are currently generating chromosomally engineered mice towards these ends.

Conclusion

We have identified 21 transcribed *Mhc* class Ib genes in the *H2-Q*, *-T* and *-M* regions and examined their expression patterns within a wide array of developmental and adult mouse tissues. Some of the class Ib gene products were expressed tissue-wide, while others were expressed in a tissue-restricted manner. These results provide a basis to select important candidate *Mhc* class Ib genes for future functional validation studies. For example, we found 12 brain-expressed class Ib genes that could have neuronal and other functions in brain development and plasticity. We also found that genes expressed tissue-wide are located in the centromeric region, whereas the tissue-specifically expressed genes are located towards the telomeric end of the *H2-T* region where the number of genes has been increased by local duplication. In this region, there are genes that showed distinct expression patterns in spite of their similar nucleotide sequences, and there is a gene pair that has a similar expression pattern, but dissimilar promoter sequence regions. From these results, the presence of a long-range regulation of *H2-T* genes is suggested, although we cannot dismiss the possibility that nucleotide changes in the promoter and enhancer regions have contributed to the loss or gain of tissue-wide expression. Since this region has diversified not only between rodent species, but also between mouse strains, it should be a good model region to address the relationship between genomic organization and expression patterns.

Methods

Sequence analysis

The genomic sequences of the *H2* region used in this study were obtained from the public databases at the NCBI Entrez Genome Project ID 9559 [42] and the Ensembl Mouse Genome Project [43]. Although we first analyzed the NT_039650.2 genomic contig by using a GENSCAN program [44] to identify the *Mhc* class Ib

genes, we finally utilized the NCBI Mouse Build 36 containing the nearly completely annotated sequence of this region, which was released on June 20th, 2006. Dot matrix analysis was performed on these genomic sequences to detect duplicated regions by using Harrplot Ver. 2.0 as part of the computer software GENETYX package. Complete or partial coding sequences of each *Mhc* class I gene was first predicted by GENSCAN, referred to the annotation, and finally confirmed by the sequencing of RT-PCR products. These coding sequences (nucleic acids) were aligned by the ClustalW program version 1.83 at DDBJ [45] using the default setting and Kimura's two-parameter method to estimate the evolutionary distances. The final outputs as radial phylogenetic trees were generated with the TreeView drawing software. The sequences used for the phylogenetic tree analyses are listed in Table 2 (shown in "Ensembl transcript ID" column for *H2-Q1*, *-Q2*, *-Q5*, *-Q6*, *-Q7*, *-Q10*, *-T24*, *-T13*, *-T11*, *-T10*, *-T7* and *-T3*, in "NCBI accession" column for *H2-Q4*, and in the "Determined in this study" column for *H2-T23*, *-T22*, *-T15*, *-T9*, *-T5*). PipMaker analyses were performed on selected *Mhc* class Ib gene sequences to visualize the DNA sequence similarities [46]. The genomic sequences analyzed by PipMaker contained the regulatory region 3 kb upstream from ATG start codon and the untranslated downstream region 1 kb from the stop codon in addition to the exon and intron sequences.

Reverse transcriptase-polymerase chain reaction (RT-PCR)

The mRNA expression of *Mhc* class Ib genes was determined by RT-PCR analysis. Total cellular RNA was isolated from the thymus, spleen, small intestine, liver, kidney, lung, heart, skeletal muscle, cerebral cortex, thalamus, cerebellum, testis and ovary of adult C57BL/6J mice, and the embryo (E9.5 – E14.5, where embryonic day 0.5 [E0.5] was defined as midday (noon) of day 1 when a vaginal plug was detected after overnight mating.), placenta (E9.5 – E14.5), and embryonic (E14.5) brains and livers of C57BL/6J mice using the guanidine isothiocyanate/CsCl ultracentrifugation method. Complementary DNA (cDNA) was synthesized from isolated RNA using the Gene Amp RNA-PCR core kit (Applied Biosystem) with the oligo-dT primer and 2 μg RNA as template in a 40 μl volume according to the manufacture's protocol. An aliquot of 0.5 μl from the 40 μl of the cDNA was used for RT-PCR reactions of all cDNA samples. The PCR was performed in 20 μl of a total reaction volume under the following conditions: cDNA was denatured at 95°C for 5 min, followed by 35 cycles of amplification (95°C for 45 s, 58°C for 30 s and 72°C for 1 min) and 5 min at 72°C. The PCR primers used for the amplifications are listed in Table 1 (see also additional file 1). The primer sets were manually designed to amplify specific *Mhc* class Ib genes by locating the gene specific polymorphisms within 5-bp of the 3' end as much as possible. All primers were designed within putative cDNA to flank or cross at least one exon-intron border. Resultant RT-PCR products were directly sequenced to verify their identity.

3' Rapid amplification of cDNA end (RACE) and cloning of class Ib cDNAs

To determine the complete cDNA sequences of *H2-T5*, *-T15*, *-T22*, *-T23* and *-M5*, 3'RACE was performed using thymus or duodenum RNAs as template, the oligo-dT-primer with adapter ($GGCCACGCGTCGACTAGTACT_{17}$.), and the forward primers listed in Table 1. The 3'RACE products were cloned into pBSII plasmid (STRATAGENE). RT-PCR covering the translation start site was done using the following forward primers designed from the genomic sequences around the translation start codon (ATG) as predicted by GENSCAN program:

H2-T5; TCTCCTGTATCATCATTCCCAGAT,

H2-T15; ACTGTACTGAGCTCTCTCTATCCCA,

H2-T22; AGTTTATAAAGCTGTCCAAGATCT,

H2-T23; GATTCAGGTTCCTCACAGACCCAG,

H2-M5; TGTATGAGAAGCCCTGCGCTCT, and the reverse primer listed in Table 1. The products were also cloned into pBSII plasmid. The nucleotide sequences of the 3'RACE and RT-PCR products were combined and analyzed.

List of abbreviations

Mhc: major histocompatibility complex; RT-PCR: Reverse transcriptase-polymerase chain reaction.

Authors' contributions

MO designed and performed the experiments, conducted genome analyses, prepared the manuscript, and is responsible for this study. HI is the director of the laboratory and gave suggestions for this study. JKK helped in editing the manuscript and in interpreting the genome analysis. SY carried out the preparation, cloning and sequencing of cDNA, and participated in the design of the study. All authors have read and approved the final manuscript.

Additional material

Additional file 1

Primer positions. *Positions of primers were indicated in alignment of* Mhc *class I sequences. Forward and reverse primers were shown in red and blue, respectively.*

Click here for file

[http://www.biomedcentral.com/content/supplementary/1471-2164-9-178-S1.doc]

Acknowledgements

We thank A. Shigenari, H. Miura, M. Koshimizu and M. Ayabe for technical assistance. We also thank reviewer 3 for his suggestion regarding nomenclature of *H2* genes. This work was supported in part by the Research and Study Program of the Tokai University Educational System General Research Organization (2005), and by 2006 Tokai University School of Medicine Research Aid to MO.

References

1. Snell GD: **The Nobel Lectures in Immunology. Lecture for the Nobel Prize for Physiology or Medicine, 1980: Studies in histocompatibility.** *Scand J Immunol* 1992, **36(4):**513-526.
2. Zinkernagel RM, Doherty PC: **The discovery of MHC restriction.** *Immunol Today* 1997, **18(1):**14-17.
3. Kumanovics A, Fischer Lindahl K: **Good copy, bad copy: choosing animal models for HLA-linked diseases.** *Curr Opin Genet Dev* 2004, **14(3):**258-263.
4. Cresswell P, Ackerman AL, Giodini A, Peaper DR, Wearsch PA: **Mechanisms of MHC class I-restricted antigen processing and cross-presentation.** *Immunol Rev* 2005, **207:**145-157.
5. Shen L, Rock KL: **Priming of T cells by exogenous antigen cross-presented on MHC class I molecules.** *Curr Opin Immunol* 2006, **18(1):**85-91.
6. Smyth LA, Afzali B, Tsang J, Lombardi G, Lechler RI: **Intercellular transfer of MHC and immunological molecules: molecular mechanisms and biological significance.** *Am J Transplant* 2007, **7(6):**1442-1449.
7. Stroynowski I, Lindahl KF: **Antigen presentation by non-classical class I molecules.** *Curr Opin Immunol* 1994, **6(1):**38-44.
8. Hu D, Ikizawa K, Lu L, Sanchirico ME, Shinohara ML, Cantor H: **Analysis of regulatory CD8 T cells in Qa-1-deficient mice.** *Nat Immunol* 2004, **5(5):**516-523.
9. Sarantopoulos S, Lu L, Cantor H: **Qa-1 restriction of CD8+ suppressor T cells.** *J Clin Invest* 2004, **114(9):**1218-1221.
10. Lu L, Werneck MB, Cantor H: **The immunoregulatory effects of Qa-1.** *Immunol Rev* 2006, **212:**51-59.
11. Xu H, Chun T, Choi HJ, Wang B, Wang CR: **Impaired response to Listeria in H2-M3-deficient mice reveals a nonredundant role of MHC class Ib-specific T cells in host defense.** *J Exp Med* 2006, **203(2):**449-459.
12. Madakamutil LT, Christen U, Lena CJ, Wang-Zhu Y, Attinger A, Sundarrajan M, Ellmeier W, von Herrath MG, Jensen P, Littman DR, Cheroutre H: **CD8alphaalpha-mediated survival and differentiation of CD8 memory T cell precursors.** *Science* 2004, **304(5670):**590-593.
13. Leishman AJ, Naidenko OV, Attinger A, Koning F, Lena CJ, Xiong Y, Chang HC, Reinherz E, Kronenberg M, Cheroutre H: **T cell responses modulated through interaction between CD8alphaalpha and the nonclassical MHC class I molecule, TL.** *Science* 2001, **294(5548):**1936-1939.
14. Loconto J, Papes F, Chang E, Stowers L, Jones EP, Takada T, Kumanovics A, Fischer Lindahl K, Dulac C: **Functional expression of murine V2R pheromone receptors involves selective association with the M10 and M1 families of MHC class Ib molecules.** *Cell* 2003, **112(5):**607-618.
15. Ishii T, Hirota J, Mombaerts P: **Combinatorial coexpression of neural and immune multigene families in mouse vomeronasal sensory neurons.** *Curr Biol* 2003, **13(5):**394-400.
16. Wu L, Feng H, Warner CM: **Identification of two major histocompatibility complex class Ib genes, Q7 and Q9, as the Ped gene in the mouse.** *Biol Reprod* 1999, **60(5):**1114-1119.
17. Huh GS, Boulanger LM, Du H, Riquelme PA, Brotz TM, Shatz CJ: **Functional requirement for class I MHC in CNS development and plasticity.** *Science* 2000, **290(5499):**2155-2159.
18. Syken J, Grandpre T, Kanold PO, Shatz CJ: **PirB restricts ocular-dominance plasticity in visual cortex.** *Science* 2006, **313(5794):**1795-1800.
19. Amadou C, Younger RM, Sims S, Matthews LH, Rogers J, Kumanovics A, Ziegler A, Beck S, Lindahl KF: **Co-duplication of olfactory receptor and MHC class I genes in the mouse major histocompatibility complex.** *Hum Mol Genet* 2003, **12(22):**3025-3040.
20. Takada T, Kumanovics A, Amadou C, Yoshino M, Jones EP, Athanasiou M, Evans GA, Fischer Lindahl K: **Species-specific class I gene expansions formed the telomeric 1 mb of the mouse major histocompatibility complex.** *Genome Res* 2003, **13(4):**589-600.
21. Postlethwait J, Amores A, Cresko W, Singer A, Yan YL: **Subfunction partitioning, the teleost radiation and the annotation of the human genome.** *Trends Genet* 2004, **20(10):**481-490.
22. Waterston RH, Lindblad-Toh K, Birney E, Rogers J, Abril JF, Agarwal P, Agarwala R, Ainscough R, Alexandersson M, An P, Antonarakis SE, Attwood J, Baertsch R, Bailey J, Barlow K, Beck S, Berry E, Birren B, Bloom T, Bork P, Botcherby M, Bray N, Brent MR, Brown DG, Brown SD, Bult C, Burton J, Butler J, Campbell RD, Carninci P, Cawley S, Chiaromonte F, Chinwalla AT, Church DM, Clamp M, Clee C, Collins FS, Cook LL, Copley RR, Coulson A, Couronne O, Cuff J, Curwen V, Cutts T, Daly M, David R, Davies J, Delehaunty KD, Deri J, Dermitzakis ET, Dewey C, Dickens NJ, Diekhans M, Dodge S, Dubchak I, Dunn DM, Eddy SR, Elnitski L, Emes RD, Eswara P, Eyras E, Felsenfeld A, Fewell GA, Flicek P, Foley K, Frankel WN, Fulton LA, Fulton RS, Furey TS, Gage D, Gibbs RA, Glusman G, Gnerre S, Goldman N, Goodstadt L, Grafham D, Graves TA, Green ED, Gregory S, Guigo R, Guyer M, Hardison RC, Haussler D, Hayashizaki Y, Hillier LW, Hinrichs A, Hlavina W, Holzer T, Hsu F, Hua A, Hubbard T, Hunt A, Jackson I, Jaffe DB, Johnson LS, Jones M, Jones TA, Joy A, Kamal M, Karlsson EK, Karolchik D, Kasprzyk A, Kawai J, Keibler E, Kells C, Kent WJ, Kirby A, Kolbe DL, Korf I, Kucherlapati RS, Kulbokas EJ, Kulp D, Landers T, Leger JP, Leonard S, Letunic I, Levine R, Li J, Li M, Lloyd C, Lucas S, Ma B, Maglott DR, Mardis ER, Matthews L, Mauceli E, Mayer JH, McCarthy M, McCombie WR, McLaren S, McLay K, McPherson JD, Meldrim J, Meredith B, Mesirov JP, Miller W, Miner TL, Mongin E, Montgomery KT, Morgan M, Mott R, Mullikin JC, Muzny DM, Nash WE, Nelson JO, Nhan MN, Nicol R, Ning Z, Nusbaum C, O'Connor MJ, Okazaki Y, Oliver K, Overton-Larty E, Pachter L, Parra G, Pepin KH, Peterson J, Pevzner P, Plumb R, Pohl CS, Poliakov A, Ponce TC, Ponting CP, Potter S, Quail M, Reymond A, Roe BA, Roskin KM, Rubin EM, Rust AG, Santos R, Sapojnikov V, Schultz B, Schultz J, Schwartz MS, Schwartz S, Scott C, Seaman S, Searle S, Sharpe T, Sheridan A, Shownkeen R, Sims S, Singer JB, Slater G, Smit A, Smith DR, Spencer B, Stabenau A, Stange-Thomann N, Sugnet C, Suyama M, Tesler G, Thompson J, Torrents D, Trevaskis E, Tromp J, Ucla C, Ureta-Vidal A, Vinson JP, Von Niederhausern AC, Wade CM, Wall M, Weber RJ, Weiss RB, Wendl MC, West AP, Wetterstrand K, Wheeler R, Whelan S, Wierzbowski J, Willey D, Williams S, Wilson RK, Winter E, Worley KC, Wyman D, Yang S, Yang SP, Zdobnov EM, Zody MC, Lander ES: **Initial sequencing and comparative analysis of the mouse genome.** *Nature* 2002, **420(6915):**520-562.
23. Weiss EH, Golden L, Fahrner K, Mellor AL, Devlin JJ, Bullman H, Tiddens H, Bud H, Flavell RA: **Organization and evolution of the class I gene family in the major histocompatibility complex of the C57BL/10 mouse.** *Nature* 1984, **310(5979):**650-655.
24. Brown GD, Choi Y, Egan G, Meruelo D: **Extension of the H-2 TLb molecular map. Isolation and characterization of T13, T14, and T15 from the C57BL/6 mouse.** *Immunogenetics* 1988, **27(4):**239-251.
25. Guidry PA, Stroynowski I: **The murine family of gut-restricted class Ib MHC includes alternatively spliced isoforms of the proposed HLA-G homolog, "blastocyst MHC".** *J Immunol* 2005, **175(8):**5248-5259.
26. Klein J, Benoist C, David CS, Demant P, Lindahl KF, Flaherty L, Flavell RA, Hammerling U, Hood LE, Hunt SW 3rd, *et al.*: **Revised nomenclature of mouse H-2 genes.** *Immunogenetics* 1990, **32(3):**147-149.
27. Shiina T, Tamiya G, Oka A, Takishima N, Yamagata T, Kikkawa E, Iwata K, Tomizawa M, Okuaki N, Kuwano Y, Watanabe K, Fukuzumi Y, Itakura S, Sugawara C, Ono A, Yamazaki M, Tashiro H, Ando A, Ikemura T, Soeda E, Kimura M, Bahram S, Inoko H: **Molecular dynamics of MHC genesis unraveled by sequence analysis of the 1,796,938-bp HLA class I region.** *Proc Natl Acad Sci U S A* 1999, **96(23):**13282-13287.
28. Kulski JK, Anzai T, Shiina T, Inoko H: **Rhesus macaque class I duplicon structures, organization, and evolution within the alpha block of the major histocompatibility complex.** *Mol Biol Evol* 2004, **21(11):**2079-2091.
29. Kulski JK, Gaudieri S, Martin A, Dawkins RL: **Coevolution of PERB11 (MIC) and HLA class I genes with HERV-16 and retroelements by extended genomic duplication.** *J Mol Evol* 1999, **49(1):**84-97.
30. Kumanovics A, Madan A, Qin S, Rowen L, Hood L, Fischer Lindahl K: **QUOD ERAT FACIENDUM: sequence analysis of the H2-D**

and H2-Q regions of 129/SvJ mice. *Immunogenetics* 2002, **54(7)**:479-489.
31. Oudshoorn-Snoek M, Demant P: **Identification and expression of the Tla region gene T11b and its Qa-like product.** *J Immunol* 1990, **145(4)**:1270-1277.
32. Comiskey M, Goldstein CY, De Fazio SR, Mammolenti M, Newmark JA, Warner CM: **Evidence that HLA-G is the functional homolog of mouse Qa-2, the Ped gene product.** *Hum Immunol* 2003, **64(11)**:999-1004.
33. Wu L, Exley GE, Warner CM: **Differential expression of Ped gene candidates in preimplantation mouse embryos.** *Biol Reprod* 1998, **59(4)**:941-952.
34. Tajima A, Tanaka T, Ebata T, Takeda K, Kawasaki A, Kelly JM, Darcy PK, Vance RE, Raulet DH, Kinoshita K, Okumura K, Smyth MJ, Yagita H: **Blastocyst MHC, a putative murine homologue of HLA-G, protects TAP-deficient tumor cells from natural killer cell-mediated rejection in vivo.** *J Immunol* 2003, **171(4)**:1715-1721.
35. Lercher MJ, Blumenthal T, Hurst LD: **Coexpression of neighboring genes in Caenorhabditis elegans is mostly due to operons and duplicate genes.** *Genome Res* 2003, **13(2)**:238-243.
36. Aihara H, Hiwatashi N, Kumagai S, Obata Y, Shimosegawa T, Toyota T, Miyazaki J: **The T3(b) gene promoter directs intestinal epithelial cell-specific expression in transgenic mice.** *FEBS Lett* 1999, **463(1-2)**:185-188.
37. Kumar PP, Bischof O, Purbey PK, Notani D, Urlaub H, Dejean A, Galande S: **Functional interaction between PML and SATB1 regulates chromatin-loop architecture and transcription of the MHC class I locus.** *Nat Cell Biol* 2007, **9(1)**:45-56.
38. Galande S, Purbey PK, Notani D, Kumar PP: **The third dimension of gene regulation: organization of dynamic chromatin loopscape by SATB1.** *Curr Opin Genet Dev* 2007, **17(5)**:408-414.
39. Teitell M, Cheroutre H, Panwala C, Holcombe H, Eghtesady P, Kronenberg M: **Structure and function of H-2 T (Tla) region class I MHC molecules.** *Crit Rev Immunol* 1994, **14(1)**:1-27.
40. Fischer Lindahl K: **On naming H2 haplotypes: functional significance of MHC class Ib alleles.** *Immunogenetics* 1997, **46(1)**:53-62.
41. Hurt P, Walter L, Sudbrak R, Klages S, Muller I, Shiina T, Inoko H, Lehrach H, Gunther E, Reinhardt R, Himmelbauer H: **The genomic sequence and comparative analysis of the rat major histocompatibility complex.** *Genome Res* 2004, **14(4)**:631-639.
42. **NCBI Entrez Genome Project** [http://www.ncbi.nlm.nih.gov/sites/entrez]
43. **Ensembl Mouse Genome Project** [http://analysis1.lab.nig.ac.jp/Mus_musculus/index.html]
44. Burge C, Karlin S: **Prediction of complete gene structures in human genomic DNA.** *J Mol Biol* 1997, **268(1)**:78-94.
45. **DNA Data Bank of Japan (DDBJ) CLUSTALW Analysis** [http://clustalw.ddbj.nig.ac.jp/top-j.html]
46. Schwartz S, Zhang Z, Frazer KA, Smit A, Riemer C, Bouck J, Gibbs R, Hardison R, Miller W: **PipMaker--a web server for aligning two genomic DNA sequences.** *Genome Res* 2000, **10(4)**:577-586.

Microsatellite analysis of the GLC1B locus on chromosome 2 points to NCK2 as a new candidate gene for normal tension glaucoma

M Akiyama,[1] K Yatsu,[2,3] M Ota,[4] Y Katsuyama,[5] K Kashiwagi,[6] F Mabuchi,[6] H Iijima,[6] K Kawase,[7] T Yamamoto,[7] M Nakamura,[8] A Negi,[8] T Sagara,[9] N Kumagai,[9] T Nishida,[9] M Inatani,[10] H Tanihara,[10] S Ohno,[11] H Inoko,[3] N Mizuki[1]

► Additional data are published online only at http://bjo.bmj.com/content/vol92/issue9

[1] Department of Ophthalmology, Yokohama City University School of Medicine, Yokohama, Japan; [2] Department of Nephrology, Yokohama City University Medical Center, Yokohama, Japan; [3] Department of Molecular Life Science, Course of Basic Medical Science and Molecular Medicine, Tokai University School of Medicine, Isehara, Japan; [4] Department of Legal Medicine, Shinshu University School of Medicine, Matsumoto, Japan; [5] Department of Pharmacy, Shinshu University Hospital, Matsumoto, Japan; [6] Department of Ophthalmology, University of Yamanashi, Faculty of Medicine, Yamanashi, Japan; [7] Department of Ophthalmology, Gifu University Graduate School of Medicine, Gifu, Japan; [8] Division of Ophthalmology, Department of Surgery, Kobe University Graduate School of Medicine, Kobe, Japan; [9] Department of Ophthalmology, Yamaguchi University Graduate School of Medicine, Ube, Yamaguchi, Japan; [10] Department of Ophthalmology and Visual Science, Graduate School of Medical Sciences, Kumamoto University, Kumamoto, Japan; [11] Department of Ophthalmology and Visual Sciences, Hokkaido University Graduate School of Medicine, Sapporo, Japan

Correspondence to: Professor N Mizuki, Department of Ophthalmology, Yokohama City University School of Medicine, Fukuura 3–9, Kanazawa-ku, Yokohama 236-0004, Japan; mizunobu@med.yokohama-cu.ac.jp

Accepted 13 June 2008

ABSTRACT

Aims: The aim of this study was to investigate the association between normal tension glaucoma and the candidate disease locus glaucoma 1, open angle, B (GLC1B) on chromosome 2. There are many reports describing the results of association or linkage studies for primary open angle glaucoma (POAG), with GLC1B as one of the loci associated with normal or moderately elevated intraocular pressure. However, there are few reports about the association of genes or defined genomic regions with normal tension glaucoma, which is the leading type of glaucoma in Japan. The GLC1B locus is hypothesized to be a causative region for normal tension glaucoma.

Methods: Genomic DNA was extracted from whole blood of normal tension glaucoma (n = 143) and healthy controls (n = 103) of Japanese origin.

Results: Fifteen microsatellite markers within and/or near to the GLC1B locus were genotyped, and their association with normal tension glaucoma was analysed. Two markers D2S2264 and D2S176 had significant positive associations.

Conclusion: The D2S176 marker had the strongest significant association and it is located 24 kb from the nearest gene NCK2, which now becomes an important new candidate gene for future studies of its association with normal tension glaucoma.

The cause of glaucoma has been attributed largely to primary open angle glaucoma (POAG) associated with elevated IOP, but in Japan normal tension glaucoma (NTG) is the leading type of glaucoma. The Tajimi study group reported that the prevalence rate of NTG was 3.60% in Japan.[1] These studies suggested that more emphasis should be placed on the prevention, detection and treatment at an early stage of the disease in order to prevent irreversible blindness.

Elucidation of the genetic aetiology of glaucoma has been increasingly emphasised as an important step for a better understanding of the pathogenesis of this disease, and for ultimately improving the prevention strategies, diagnostic tools, and therapy in the new millennium.[2] A few reports describing the results of association or linkage studies for POAG[3] are available, and they have linked the disease to numerous chromosomal regions.[4–12] The application of linkage analysis to glaucoma with the exception of obvious Mendelian inheritance, has achieved only limited success thus far.

One of the candidate loci that has been identified for POAG is Glaucoma 1, open angle, B (GLC1B, MIM:606689) located on chromosome 2.[5] The identification of this locus was based on a linkage study of six Caucasian families in the UK and the GLC1B locus for adult-onset POAG was identified within a region of 11.2 cM flanked by markers at chromosome 2cen-q13. The POAG patients in these families had clinical characteristics of low to moderate IOP, disease onset in their late 40s, and a good response to medical therapy, indicating that the GLC1B locus might encode a POAG gene that is associated with normal or moderately elevated IOP. Interestingly, eight additional families, in which members had variable clinical presentations, did not show any linkage to this region. Another study on American families also excluded 2cen-q13.[13] These studies suggest that there is genetic heterogeneity for POAG and the possibility of a specific NTG gene.

The aim of our study was to determine the association between 15 selected microsatellite (MS) markers within or near to the GLC1B locus and NTG in Japanese subjects. The 15 MS markers are distributed across the GLC1B region with each neighbouring MS locus less than 100 kb apart and consequently in linkage disequilibrium (LD) with each other. We considered that the number and distribution of the MS markers were enough to identify disease-predisposing genomic variants in terms of LD. The MS markers chosen for this study are highly polymorphic and have a high degree of heterozygosity (on average, approximately 70%) and LD lengths in the 100–200 kb range.[14–19]

MATERIALS AND METHODS

Subjects for microsatellite typing

One hundred and forty-three patients with NTG for patients and 103 healthy individuals participated for controls in this study. The subjects were of Japanese origin from Yokohama City University, Yamanashi University, Gifu University, Kobe University, Yamaguchi University, Kumamoto University and Tokai University in Japan. The diagnosis of NTG was made according to the guidelines of Genetic Variation Project Team in the Japanese Society of Ophthalmology (declared in 2000), which defined NTG as glaucoma with an intraocular pressure (IOP) of less than 21 mm Hg by Goldmann applanation tonometry without medication, including during the 24 h diurnal

curve, open irido-corneal angle, glaucomatous optic disc appearance, and the corresponding glaucomatous visual field defect. The subjects had no systemic abnormalities. Glaucomatous optic disc change was diagnosed when the vertical cup/disc ratio of optic nerve head was more than or equal to 0.7, the rim width at superior portion (11–1 h) or inferior portion (5–7 h) was less than or equal to 0.1 of the disc diameter, the difference of the vertical cup/disc ratio was more than or equal to 0.2 between both eyes, or a nerve fibre layer defect was found. We have excluded individuals diagnosed at an age under 20 or over 60 years. Glaucomatous visual field defect was defined in a hemifield basis using reliable field data examined by the Humphrey static visual field analyser (Carl Zeiss-Meditec) C30-2 program according to the Anderson and Patella's criteria: the hemifield was judged abnormal when the pattern deviation probability plot showed a cluster of three or more non-edge contiguous points having sensitivity with a probability of less than 5% in the upper or lower hemifield and in one of these with a probability of less than 1%. The selection criteria were modified depending on the subjects' age in order to minimise the effects of damage to the retinal ganglion that occurs due to ageing; that is, the mean deviation measured by HFA C-30-2, (i) no requirement if the patient was diagnosed at the age under 50 years (ii) −10.00 dB or worse in at least one eye if the patient was diagnosed at the age between 50 and 55 (iii) −15.00 dB or worse in at least one eye if the patient was diagnosed at the age above 55.

During diagnosis, patients whose refraction values were myopic over −8.0 D of spherical equivalent and had changed due to cataract surgery, refractive surgery, etc were excluded from the study. In cases where a glaucomatous visual field defect was present in only one eye, the refraction value and glaucomatous visual field defect of the affected eye were adopted. Further, in cases where a glaucomatous visual field defect was present in both eyes, the refraction value and glaucomatous visual field defect of the more severely affected eye were adopted.

The control subjects were all healthy volunteers having regular medical check-ups. They are all under 50 years old. All personal identities associated with medical information and blood samples were carefully eliminated and replaced with anonymous identities in each recruiting institution.

This study was performed in accordance with the Declaration of Helsinki, and we obtained informed consent from all patients and healthy individuals whose DNA samples were used in the analyses. Our experimental procedures were conducted in accordance with the Declaration of Helsinki and approved by the relevant ethical committee in each participating university and centre.

DNA genotyping

DNA was extracted using a QIAamp DNA blood kit (QIAGEN, Hilden, Germany) under standardised conditions to prevent any variation in DNA quality. This was followed by 0.8% agarose gel electrophoresis to check for DNA degradation and RNA contamination. Following measurement of the optical density to check for protein contamination, the DNA concentration was determined through three successive measurements using the PicoGreen fluorescence assay (Molecular Probes, Eugene, OR). PCR reactions were performed in a total volume of 12.5 μl containing PCR buffer, genomic DNA, 0.2 mM dinucleotide triphosphates (dNTPs), 0.5 μM primers, and 0.35 U Taq polymerase. The reaction mixture was subjected to 5 min at 94°C, then 30 cycles of 30 s for denaturing at 94°C, 30 s for annealing at 56°C, 1 min for extension at 72°C and then 10 min for final elongation at 72°C using a polymerase chain reaction (PCR) thermal cycler, GeneAmp System 9700 (Applied Biosystems, Foster City, CA). To determine the number of microsatellite repeats, PCR-amplified products were denatured for 2 min at 95°C, mixed with formamide and electrophoresed using an ABI3130 Genetic Analyzer (Applied Biosystems). The number of microsatellite repeats was estimated automatically using the GeneScan 672 software (Applied Biosystems) by the local Southern method with a size marker of GS500 TAMRA (Applied Biosystems).

Marker information

Microsatellite (MS) sequences were computationally detected from all the chromosomes. In this study, we utilised 15 MS markers within or near the GLC1B locus as shown in fig 1.

The PCR primer pairs of the 15 MS markers for the association test were designed in order to improve the efficiency of PCR (table S1, at http://bjo.bmj.com/content/vol92/issue9). The forward primer was labelled at the 5′ end with 6-FAM, NED, PED or VIC (ABI, Tokyo).

Statistical analysis

The measurements of the heights of individual peaks in the DNA were applied to association analysis. To calculate p values, we used χ^2 tests and made corrections using Bonferroni correction. The Hardy–Weinberg test for allele frequency distributions at the microsatellite loci was performed using a probability test for differentiation, as determined by GenePop 3.4. Other basic analyses were carried out using Microsoft Excel.

RESULTS

We genotyped 15 MS markers, and their locations are shown within the GLC1B locus in fig 1. The observed and expected frequencies of each genotype for the 15 markers in the case and control subjects were in Hardy–Weinberg equilibrium (data not shown). Two markers D2S2264 and D2S176 were significantly positive, as shown in tables 1, 2. In considering the LD range, the susceptibility genes for NTG were estimated to reside within a 100–150 kb region flanking the two positive markers D2S2264 and D2S176. The corrected p value for the 245 allele of D2S176 was $\leqslant 0.05$. There was no significant association for the other 13 markers.

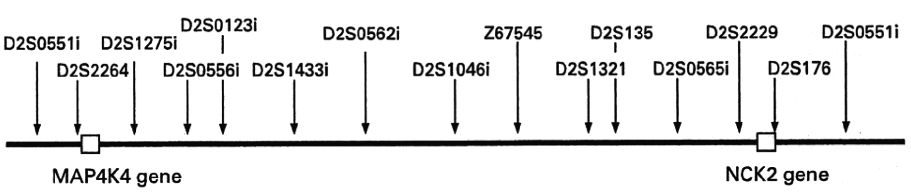

Figure 1 Locations of 15 microsatellite markers. To the left of the figure is 5′UTR. White boxes show the NCK2 gene (42 kb) and MAP4K4 gene (29 kb).