

11-1-2003

Identification of three genomic haplotypes 5' to the human CD1D gene and their distribution in four ethnic groups

Q. Y. Chen
Children's Hospital of New Orleans

N. Jackson
Children's Hospital of New Orleans

A. Vargas
LSUHSC School of Medicine

S. Chalew
LSUHSC School of Medicine

J. Rao
LSUHSC School of Medicine

See next page for additional authors

Follow this and additional works at: https://repository.lsu.edu/biosci_pubs

Recommended Citation

Chen, Q., Jackson, N., Vargas, A., Chalew, S., Rao, J., Batzer, M., Lan, M., Chang, Y., Mokhashi, M., & Liu, D. (2003). Identification of three genomic haplotypes 5' to the human CD1D gene and their distribution in four ethnic groups. *Tissue Antigens*, 62 (5), 442-448. <https://doi.org/10.1034/j.1399-0039.2003.00116.x>

This Article is brought to you for free and open access by the Department of Biological Sciences at LSU Scholarly Repository. It has been accepted for inclusion in Faculty Publications by an authorized administrator of LSU Scholarly Repository. For more information, please contact ir@lsu.edu.

Authors

Q. Y. Chen, N. Jackson, A. Vargas, S. Chalew, J. Rao, M. Batzer, M. S. Lan, Y. H. Chang, M. Mokhashi, and D. Liu

Q.-Y. Chen
N. Jackson
A. Vargas
S. Chalew
J. Rao
M. Batzer
M.S. Lan
Y.-H. Chang
M. Mokhashi
D. Liu

Identification of three genomic haplotypes 5' to the human CD1D gene and their distribution in four ethnic groups

Key words:

CD1d; haplotypes; promoter; SNP

Acknowledgments:

The work was sponsored by intramural research grants to QYC from the Children's Hospital of New Orleans and the Department of Pediatrics, LSUHSC. The authors thank Dr Seth Pincus for his many suggestions for the study and his critical reading of the manuscript.

Abstract: CD1d presents lipid antigen to a conserved population of natural killer (NK) T cells, which participate in host immune defense, tumor cell rejection and suppression of autoimmunity. The levels of human CD1d expression vary significantly between individuals. To understand such variation, we sequenced the region up to 1.7 kb 5' upstream of the translation start site and partially through exon 2 in 44 white Americans. We also studied two tagged single nucleotide polymorphisms (SNP) in 112 white Americans, 60 African-Americans, 88 Europeans, and 84 Chinese people from the region. Six SNP present in the region (-836C→T, -773C→T, -764C→G, -713A→T, -365A→G and +363A→G) were found to be in a complete linkage disequilibrium and comprised three haplotypes. Haplotype 1 had -836C, -773C, -764C, -713A, -365A and +363A. Haplotype 2 had -836C, -773T, -764C, -713A, -365A and +363A. Haplotype 3 had -836T, -773C, -764G, -713T, -365G and +363G. -773C→T and -764C→G can serve as the tagged SNP to differentiate the three haplotypes. The frequency of haplotype 1 was significantly higher in African Americans than in the other three ethnic groups, whereas the frequency of haplotype 3 was significantly higher in the Chinese people than those in the other three groups. The finding of the three haplotypes provides a genetic marker for CD1d and facilitates the study of the functional role of the genetic variations in human CD1d expression and regulation.

Author's affiliations:

Q.-Y. Chen^{1,2},
N. Jackson¹,
A. Vargas²,
S. Chalew²,
J. Rao²,
M. Batzer³,
M.S. Lan^{1,2},
Y.-H. Chang⁴,
M. Mokhashi²,
D. Liu⁵

¹Research Institute for Children, Children's Hospital,

²Department of Pediatrics, Louisiana State University Health Science Center, New Orleans,

³Department of Biological Sciences, Louisiana State University, Baton Rouge, LA,

⁴Chung Shan Medical University, Taichung, Taiwan, ROC,

⁵University of Chicago, Chicago, IL

Correspondence to:

Qiao-Yi Chen
Research Institute for Children
200 Henry Clay Ave
New Orleans LA 70118
Tel: +504 896 2709
Fax: +504 896 2722
e-mail: qchen@chnola-research.org

CD1d molecules are cell surface glycoproteins that are non-covalently associated with β 2-microglobulin to form a heterodimeric three-dimensional structure that is similar to the molecules of class I major histocompatibility complex (MHC) (1). The gene encoding CD1d molecule or CD1D gene belongs to group II of the CD1 gene family in human. Group I CD1 genes include CD1A, CD1B, CD1C and CD1E based on their sequence homology (2, 3). The CD1 genes are located in chromosome 1 q22-23 in the human (4, 5). CD1D gene has no significant sequence homology in the 5' untranslated region with the group I genes, whereas the latter genes share significant homology in the region (6). Unlike MHC class I and II genes, the polymorphism in

Received 3 April 2003, revised 20 May 2003,
accepted for publication 22 May 2003

Copyright © Blackwell Munksgaard 2003
Tissue Antigens. ISSN 0001-2815

Tissue Antigens 2003 62: 442-448
Printed in Denmark. All rights reserved

the coding region of the CD1 gene family is very limited. Only two alleles have been described for CD1A, -B, -C, and -D and four alleles for CD1E (7–9). Among the CD1 molecules, CD1d is highly conserved across species and is the only group CD1 molecule functionally present in mice and rats (10–13). Both human and mouse CD1d can present α -glactosylceramide (α -GlaCer) to natural killer T (NK T) cells, which express a restricted range of T cell receptors bearing a single invariant V α -chain (V α 14J α 281 in mice and V α 24^{inv} in human) to stimulate specific immune responses (14, 15).

Genetic factors as well as environmental factors play a role in human CD1d expression and regulation. The expression of CD1d molecules in mice is constitutive in the bone-marrow-derived cells including B-lymphocytes, macrophages, and dendritic cells (16–19). Unlike mouse CD1d molecules, the expression of CD1d in human does not appear constitutive (1). The levels of CD1d expression varied among different donors on T lymphocytes (20), monocytes, or monocyte-derived dendritic cells (21). CD1d expression can be up-regulated on intestinal epithelial cells and keratinocytes by gamma interferon (22, 23) or on peripheral blood T cells by mitogen stimulation (20, 24). Patients with psoriasis have an over-expression of CD1d on keratinocytes (23) and patients with allergic reactions to cow's milk have higher levels of CD1d expression in the duodenal lamina propria than healthy controls (25). In addition, patients with primary biliary cirrhosis, an autoimmune liver disease, have an increased expression of CD1d on the epithelial cells of the small bile ducts (26). It remains to be investigated whether the difference in the CD1d expression between individuals is due to their genetic variation or an environmental stimulation or both.

Knowledge of the genetic factors that are involved in CD1d expression would help in understanding how the gene products function in NK T cell-associated immune regulation. Multiple SNP in the upstream region of the human CD1D gene have been noted and listed in the human genome databank. Such SNP may serve as unique genetic markers for CD1d or may be used to study their functional role in CD1d expression and regulation. However, these SNP have not been validated or systematically studied. To confirm the presence of the SNP in the upstream region of the CD1D gene and to look for genetic markers that can be used for a CD1D gene related disease association study, we sequenced the region 1.7 kb upstream of the translation start site of the CD1D gene, exon 1, intron 1 and the 5' portion of exon 2 in white Americans. We found that the SNP in the region are in complete linkage disequilibria and comprise three haplotypes. Two tagged SNP that can be used to differentiate the three haplotypes were also tested in white Americans, African Americans, Europeans and Chinese people. We found that the distribution of the three haplotypes based on the two tagged SNP differed significantly among these four ethnic groups.

The subjects were 112 white Americans, 60 African-Americans, 88 Europeans, and 84 Chinese people. The participants were either healthy volunteers or children being evaluated for non-immune conditions. Subjects were studied after written informed consents were obtained.

Genomic DNA samples were prepared from peripheral blood mononuclear cells, or buccal cells collected using swirl swabs. The isolation of genomic DNA from the peripheral blood mononuclear cells was as previously described (27) using proteinase K digestion, followed by phenol-chloroform extraction. For the isolation of genomic DNA from buccal cells, participants were asked to rinse out their mouths twice with water. Buccal cells were collected with a sample collection swab (Epicentre, Madison, WI) by rolling it firmly on the inside of the cheek approximately 20 times on each side. The cells on the swab were suspended in phosphate buffered saline (PBS), pH 7.4. After the pellet was washed once with PBS, the pellet was suspended in a proteinase K digestion buffer containing 10 mM Tris (pH 7.8), 5 mM ethylenediaminetetraacetic acid (EDTA), and 0.5% sodium dodecyl sulfate. Proteinase K (Sigma, St. Louis, MO) at a final concentration of 30 μ g/mL was added to the solution and incubated at 55°C overnight. The DNA was extracted twice with phenol chloroform. DNA was precipitated with 2-isopropanol, and the pellet was washed twice with 70% ethanol. The isolated DNA was diluted in Tris-EDTA buffer and stored at -20°C.

Genomic DNA located between -1738 and +508 of the human CD1D gene was amplified using PCR. The primers used in the amplification were 5'CTA GAG TGT GGT GCA GTA GC3' and 5'CTC TCT GAA GCC TTG GTC C3'. The PCR amplifications were performed on 50 ng genomic DNA in a 25 μ L volume containing 1 unit Platinum *Pfx* DNA polymerase (Invitrogen, Carlsbad, CA) using the buffer provided and the addition of 8% dimethyl sulfoxide (Sigma). The PCR reactions were subjected to a 35-cycle reaction at 60°C annealing temperature using an automated thermal cycler (9700, PE Applied Biosystems, Foster City, CA). The PCR products were separated in agarose gel and the band corresponding to the expected size was removed. The DNA was isolated using a kit (Qiagen, Valencia, CA) following the manufacturer's instructions. The isolated PCR products were processed for direct sequencing. For TA cloning, the PCR products were reacted with 1 unit of *Taq* polymerase at 72°C for 10 min to add a single deoxyadenosine to the 3' ends, then processed for TA cloning (Invitrogen) as per the manufacturer's instructions. Plasmid DNA was isolated using a commercial kit (Promega, Madison WI) and processed for sequencing. Approximately 10 ng of the PCR product or 300 ng of the plasmid DNA was sequenced using the ABI PRISM 377 DNA Sequencer (PE Applied Biosystems). The primers used in the sequencing reaction were M13-F, M13-R, provided by the manufacturer (Invitrogen), and

the gene specific primers including 5'CTA GAG TGT GGT GCA GTA GC3'; 5'CTC TCT GAA GCC TTG GTC C3'; 5'AGT GAG GGA GAG GGA GGT G3'; 5'AAG AGT GCG CAG GTC AGA G3' and 5'AAC CTA CTG AAG TGA GCG GC3'.

A double restriction enzyme digestion was used to screen for haplotype. A PCR product was generated using a primer pair, 5'-AGT GAG GGA GAG GGA GGT G3' and 5'/GCC GCT CAC TTC AGT AGG TT3' to cover the region of -1009 to -240 upstream of CD1D gene. The PCR condition was same as that described above. The PCR product (10 µL) was digested with the restriction enzyme *NgoM* IV or *Avr* II (New England Biolabs, Beverly, MA) using the conditions recommended by the manufacturer and separated in a 2.5% agarose gel and visualized under UV.

For the calculation of allelic or haplotype frequencies, subjects with heterozygous alleles or haplotypes were counted for the presence of each allele or haplotype of interest, while subjects with homozygous alleles or haplotypes were counted twice. The numbers of a specific alleles or haplotypes were then divided by the total numbers of alleles or haplotypes in the group and multiplied by 100 to express the frequencies as percentages. Chi-square tests with Yates' correction were used and *P*-values of less than 0.05 were considered to be statistically significant.

Genomic DNA samples from 44 white Americans were amplified between -1738 and +508 of the CD1D gene and sequenced using multiple primer pairs in the region. The sequence data were manually reviewed, then analyzed using Sequencher 4.1 software (Gene Codes Corporation, Ann Arbor, MI). Six SNP were detected in this region (Table 1). The nomenclature for the identified SNP was based on a recent recommendation, in which the nucleotide +1 is the A of the ATG-translation initiation codon and the nucleotide 5'- to +1 is numbered -1 (28). The sequence of the published cDNA (2) and a genomic DNA sequence obtained from the GenBank (Accession No. AL138899) were used as the reference to analyze the detected SNP. We found that five of the six SNP were located 5' upstream of the CD1D coding region (-836C→T, -773C→T, -764C→G,

-713A→T, -365A→G) and one was within intron 1 (+363A→G) of the CD1D gene (Table 1). We did not find any other SNP or mutations in the region. The reported SNP +482A→T in the exon 2 (7) was not found in our samples. +482A was present in all 44 samples sequenced.

The SNP were analyzed for the presence of haplotypes in the region. Six SNP combinations were present in the 44 individuals tested (Table 1). The combination in the individuals with the homozygous alleles for the six SNP indicates the presence of three haplotypes as seen in the SNP combination 1-3 or referred to as haplotype 1-3 in Table 1. Haplotype 1 and haplotype 2 differ only in SNP -773C→T. The former is -773C, and the latter is -773T. However, haplotype 3 has diverse SNP compared to haplotype 1 and 2. Haplotype 1 and haplotype 3 differ in -836C→T, -764C→G, -713A→T, -365A→G, and +363A→G. Haplotype 2 and haplotype 3 differ in all of the six detected SNP. In addition, genomic DNA sequenced data derived from two African-Americans, four Hispanic-Americans, two Chinese and one Indian showed the six SNP and gave the same pattern of the three haplotypes (data not shown).

To confirm the presence of the three haplotypes, PCR products covering the region between -1738 and +508 of CD1D gene from 20 additional individuals were cloned into a TA cloning vector and sequenced. The sequences from each of the individuals were aligned for SNP presentation and haplotype formation. We found that all six SNP were present in each of the three haplotypes in individuals with either the heterozygous or homozygous haplotype as described above (data not shown).

The identified SNP were compared with the SNP in the region available in the DNA database of the National Center for Biotechnology Information (NCBI). There are eight listed SNP in the region. These eight SNP were detected primarily by looking at overlaps between clones that cover the same region of the genome and validated either by multiple, independent submissions to the ref SNP cluster, or by frequency or genotype data according to NCBI. Most of these SNP lack validation information, and all of them do not have

SNP and haplotypes in 44 white Americans

SNP combination	-836C→T	-773C→T	-764C→G	-713A→T	-365A→G	+363A→G	<i>n</i> (%)
1	C	C	C	A	A	A	10 (23)
2	C	T	C	A	A	A	5 (11)
3	T	C	G	T	G	G	1 (2)
4	C	C/T	C	A	A	A	12 (27)
5	C/T	C	C/G	A/T	A/G	A/G	10 (23)
6	C/T	T/C	C/G	A/T	A/G	A/G	6 (14)

Table 1

the information on allelic frequency according to NCBI and PubMed line searches. All of the six SNP identified in our study were among the eight SNP. These six SNP were rs859008 (-836C→T), rs3754471(-773C→T), rs859009 (-764A→T), rs85910 (-713A→T), rs859011 (-365A→G) and rs859013 (+363A→G). Two of the listed SNP in the region, rs859007 (-1340G→A) and rs859012 (-45G→A) are not present in our data set. Our data were -1340A and -45G.

We then examined whether the three haplotypes were in concordance with the sequences available in the Genebank database. We identified three separate genomic DNA sequences (Genebank Accession Nos. AL138899.23, X14974.1 and AB045357.1), which covered the entire region that had been sequenced. Interestingly, these three genomic DNA sequences represented each of the three identified haplotypes. The sequence of AL138899.23 had the SNP of haplotype 1. X14974.1 had the SNP of haplotype 2 and AB045357.1 had the SNP of haplotype 3. For the two listed SNP that were not showed to be polymorphic in our data, AL138899.23 had -1340G (rs859007) and -45G (rs859012). AB045357.1 had -1340A and -45A. X14974.1 had -1340A and -45G. Accordingly, at least three haplotypes may be present upstream of the CD1D coding region.

We developed a double restriction enzyme digestion strategy to screen a large number of samples for the three haplotypes. The strategy was based on the finding that the SNP were in complete linkage disequilibria. Haplotypes 1 and 3 had the SNP -773C, whereas haplotype 2 had -773T; haplotypes 1 and 2 had -764C, whereas haplotype 3 had -764G. The detection of these two SNP with restriction endonuclease digestion of PCR amplicons could identify each of the three haplotypes (Fig.1). The results derived

from the double restriction enzyme digestion were in concordance with those derived from the direct sequence of the genomic DNA (data not shown). It was thus possible to assign individuals to each of the three haplotypes based upon this simplified method of restriction enzyme digestion.

The DNA samples derived from the 112 white Americans, 60 African-Americans, 88 Europeans and 84 Chinese were tested for the two tagged SNP and analyzed for their allelic frequencies and haplotype distributions (Fig.2). The African-Americans had a significantly higher frequency for SNP-773C than the white Americans (77.5% (93/120) vs 63.5% (140/224), $\chi^2=8.04$, $P=0.0046$), than the Europeans (77.5% (93/120) vs 61.4% (108/176), $\chi^2=8.52$, $P=0.0035$), and the Chinese (77.5% (93/120) vs 66.7% (112/168), $\chi^2=4.00$, $P=0.0454$). There was no significant difference in the frequency for the SNP-773C between white Americans and Europeans or Chinese. However, the frequency of SNP -764G was significantly higher in Chinese than in white Americans (42.3% (71/168) vs 16.5% (37/224), $\chi^2=31.88$, $P<0.0001$), Europeans (42.3% (71/168) vs 10.8% (19/176), $\chi^2=44.06$, $P<0.0001$), or African-Americans (42.3% (71/168) vs 8.3% (10/110), $\chi^2=39.86$, $P<0.0001$). There was no significant difference in the allelic frequency for SNP -764C or -764G among the white Americans, Europeans or Chinese after Yates corrections.

The haplotypes comprised by the two-tagged SNP were also analyzed based on the method illustrated in Fig.1. The distribution of the haplotypes comprised by the two-tagged SNP differed significantly between the four ethnic groups (Table 2). The frequency of the homozygous haplotype of -773C/-764C (haplotype-1/1) was significantly higher in the African-Americans than that in the other three ethnic groups ($P<0.001$). The frequency of the homozygous haplotype of -773T/-764C (haplotype-2/2) was significantly higher in the white Americans than in the African-Americans or the Europeans ($P<0.05$). However, the frequency of the homozygous haplotype of -773C/-764G (haplotype-3/3) was significantly higher in the Chinese than in the white Americans or the Europeans ($P<0.0001$), indicating that haplotype 3 was predominantly present in the Chinese. This was also supported by the data, which were analyzed for the distribution of the three individual haplotypes (Fig. 2). The frequency of the heterozygous haplotype of -773CT/-764C (haplotype-1/2) was significantly higher in the Europeans than those derived from the three other ethnic groups ($P<0.01$). The Chinese had the lowest frequency for this haplotype. However, the frequency of the heterozygous haplotype of -773CT/-764CG (haplotype-2/3) was significantly higher in the Chinese than those derived from the three other ethnic groups ($P<0.01$).

We found that five SNP in the 1.7 kb upstream of CD1D coding region and one SNP in intron 1 of the gene were in complete linkage

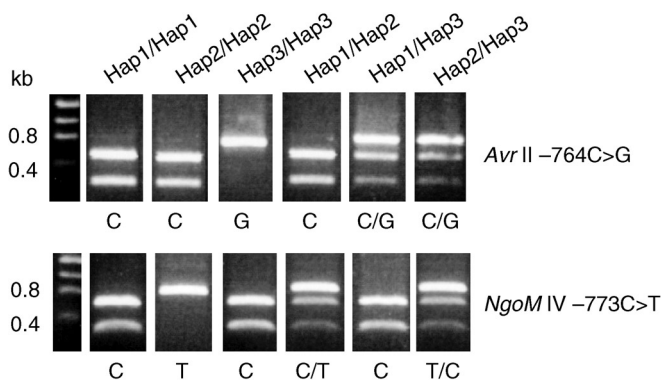


Fig. 1. Identification of three haplotypes using a double restriction enzyme digestion method. Top panel: PCR product (-1009 to -240) was digested with *Avr* II to reveal homozygous -764C (band 246 bp and 524 bp), homozygous -764G (770 bp) or heterozygous -764C/G. Bottom panel: The same PCR product was digested with *NgoM* IV to reveal homozygous -773C (236 bp and 534 bp), homozygous -773T (770 bp) or heterozygous -773C/T. The combination of the two revealed the presence of haplotypes for each individual tested.

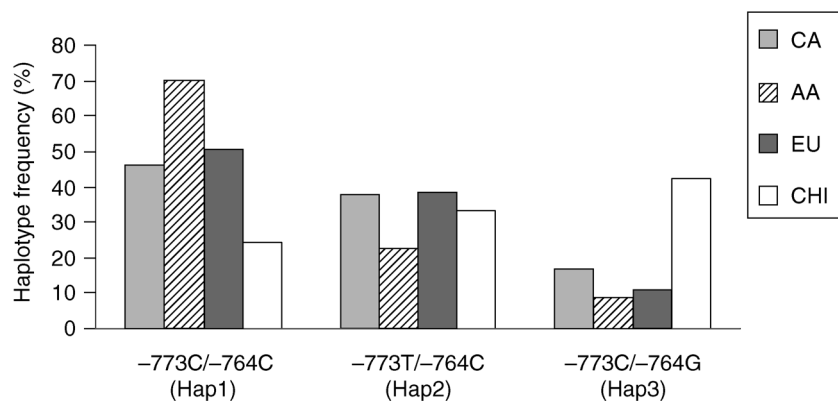


Fig. 2. Distribution of the three haplotypes comprised by the two tagged SNP (-773C→G -764C→G) in 112 white Americans (CA), 60 African-Americans (AA), 88 Europeans (EU), and 84 Chinese (CHI). The frequency of haplotype 1 is significantly higher in the AA than that in the CA ($\chi^2 = 16.91, P < 0.0001$), that in the EU ($\chi^2 = 10.14, P = 0.0015$) or in the CHI ($\chi^2 = 57.21, P < 0.0001$). The frequency of haplotype 1 in CA or EU was also significantly higher than that in the CHI ($\chi^2 = 25.03, P < 0.0001$; $\chi^2 = 19.23, P < 0.0001$, respectively). However, the frequency of haplotype 2 was significantly lower in the AA than that in the CA ($\chi^2 = 8.04, P < 0.0046$), than in the EU ($\chi^2 = 8.52, P = 0.035$), or the CHI ($\chi^2 = 4.00, P = 0.0454$). The frequency of haplotype 3 is significantly higher in the CHI than in the CA ($\chi^2 = 31.88, P < 0.0001$), than in the AA ($\chi^2 = 39.86, P < 0.0001$), or the EU ($\chi^2 = 44.06, P < 0.0001$).

disequilibrium and comprised three haplotypes. We identified two tagged SNP, -773C→T and -764C→G, which could be typed using a method of restriction enzyme digestion to differentiate the three haplotypes. These two tagged SNP were commonly present in the white Americans, African-Americans, Europeans and Chinese, suggesting that the three haplotypes were also commonly present in these ethnic groups. However, application of the two tagged SNP to differentiate the three haplotypes in non-white Americans may be limited as the two tagged SNP were based on the sequence data primarily derived from the White Americans.

The three haplotypes could serve as genetic markers for CD1D. The polymorphism in the coding region of the human CD1D gene was very limited. Only two alleles (+482A and +482T) in exon 2 of the gene had been reported, although the allelic frequency for +482T

was only 1% (7). +482T was not found in our study or in two other studies (29, 30). It would be interesting to learn the types of SNP in 5' upstream of the coding region in the individuals who had +482T. The SNP in the coding region of CD1 genes have been used to establish an association between CD1 genes and Mycobacterium malmoeense pulmonary disease, although negative results were reported (28). The use of the rare alleles of the CD1 genes certainly limits the statistical power in the association studies. It remains to be established for a linkage disequilibrium between the SNP in the 5' upstream of the coding region and the SNP in the coding region of the CD1 genes. However, the common polymorphism and the haplotypes in the 5' upstream of the coding region of the genes should be the candidate genetic markers for the genes, as polymorphism as well as haplotypes in the 5' upstream region have been successfully used in

The distribution of the haplotypes comprised by -773C→T and -764C→G in four ethnic groups

Haplotype	CA, n = 112 n (%)	AA, n = 60 n (%)	EU, n = 88 n (%)	CHI, n = 84 n (%)
-773C/-764C (Hap1/1)*	25 (22.3)	30 (50.0)	16 (18.2)	6 (7.2)
-773T/-764C (Hap2/2)**	20 (17.9)	4 (6.7)	7(8.0)	12 (14.3)
-773C/-764G (Hap3/3)***	6 (5.4)	0 (0.0)	1 (1.1)	18 (21.4)
-773C/T/-764C (Hap1/2)†	36 (32.1)	16 (26.7)	47 (53.4)	13 (15.5)
-773C/-764C/G (Hap1/3)	17 (15.2)	7 (11.7)	10 (11.4)	16(19.0)
-773C/T/-764C/G (Hap2/3)††	8 (7.1)	3 (5.0)	7(8.0)	19 (22.6)

CA, white Americans; AA, African-Americans; EU, Europeans; CHI, Chinese

*AA vs CA: $\chi^2 = 13.76, P = 0.0002$. AA vs EU: $\chi^2 = 16.86, P < 0.00001$. AA vs CHI: $\chi^2 = 34.29, P < 0.00001$. CA vs CHI: $\chi^2 = 8.31, P = 0.0040$. EU vs CHI: $\chi^2 = 4.69, P = 0.0303$

**CA vs AA, $\chi^2 = 4.07, P = 0.0435$. CA vs EU: $\chi^2 = 4.14, P = 0.0419$

***CHI vs CA: $\chi^2 = 11.54, P = 0.007$. CHI vs AA, $\chi^2 = 14.69, P = 0.0001$. CHI vs EU: $\chi^2 = 18.01, P < 0.00001$

†EU vs CA: $\chi^2 = 9.18, P = 0.0024$. EU vs AA: $\chi^2 = 10.44, P = 0.0012$. EU vs CHI: $\chi^2 = 27.22, P < 0.00001$. CA vs CHI: $\chi^2 = 7.11, P = 0.0077$

††CHI vs CA: $\chi^2 = 9.68, P = 0.0019$. CHI vs AA: $\chi^2 = 8.39, P = 0.0038$. CHI vs EU: $\chi^2 = 7.2, P = 0.0073$

Table 2

disease association studies (31). If the three haplotypes had functional effects on CD1d expression or regulation, the haplotypes could serve as functional markers for CD1d molecules as well as the genetic markers for the locus. Our unpublished observation suggests that the region between -985 and +24, which contains the above five SNP, has promoter activity. In addition, presence of the nucleotide '-836C' in the SNP -836C→T introduces two putative *cis*-regulatory elements (TCF-1 and gamma-IRE_CS), and the presence of '-836T' in the position removes both of the elements based on the program of SIGSCAN (32).

It would be interesting to learn whether the three identified haplotypes could be associated with a disease condition. Like MHC class I molecules, CD1d molecules are involved in the antigen presentation and immune regulation. In contrast to the MHC class I molecules, the ligand-binding groove of the CD1d is hydrophobic with a narrower

and deeper shape (33). The polymorphism in the coding region of the MHC class I genes is much higher than that of CD1D gene. This may lead to the fact that class I molecules have diverse ligand peptides for antigen presentation whereas CD1d has only one known lipid antigen, α -GlaCer. Accordingly, the effects of the individual genetic variation on CD1D may be associated with levels of the gene expression, CD1d tissue distribution and the response of the gene to environmental stimulation, rather than with the ligand diversity in the antigen presentation. The reported individual variations in the levels of CD1d expression (20, 21) are possibly due to a genetic variation associated with the transcriptional regulation of the gene. It invites investigation whether the three haplotypes are associated with those diseases that have an abnormal NK T cell function as an abnormal CD1d expression in the thymus or periphery may lead to an abnormal functional development of the CD1d restricted NK T cells.

References

- Porcelli SA, Modlin RL. The CD1 system. antigen-presenting molecules for T cell recognition of lipids and glycolipids. *Annu Rev Immunol* 1999; **17**: 297-329.
- Calabi F, Jarvis JM, Martin L, Milstein C. Two classes of CD1 genes. *Eur J Immunol* 1989; **19**: 285-92.
- Hughes AL. Evolutionary origin and diversification of the mammalian CD1 antigen genes. *Mol Biol Evol* 1991; **8**: 185-201.
- YuCY, Milstein C. A physical map linking the five CD1 human thymocyte differentiation antigen genes. *EMBO J* 1989; **8**: 3727-32.
- Albertson DG, Fishpool R, Sherrington P, Nacheva E, Milstein C. Sensitive and high resolution in situ hybridization to human chromosomes using biotin labelled probes: assignment of the human thymocyte CD1 antigen genes to chromosome 1. *EMBO J* 1988; **7**: 2801-5.
- Balk SP, Bleicher PA, Terhorst C. Isolation and characterization of a cDNA and gene coding for a fourth CD1 molecule. *Proc Natl Acad Sci U S A* 1989; **86**: 252-6.
- Han M, Hannick LI, Dibrino M, Robinson MA. Polymorphism of human CD1 genes. *Tissue Antigens* 1999; **54**: 122-7.
- Brookes AJ. 4th International Meeting on Single Nucleotide Polymorphism and Complex Genome Analysis. Various uses for DNA variations. *Eur J Hum Genet* 2002; **10**: 153-5.
- Mirones I, Oteo M, Parra-Cuadrado JF, Martinez-Naves E. Identification of two novel human CD1E alleles. *Tissue Antigens* 2000 August; **56**: 56: 159-61.
- Bradbury A, Belt KT, Neri TM, Milstein C, Calabi F. Mouse CD1 is distinct from and co-exists with TL in the same thymus. *EMBO J* 1988; **7**: 3081-6.
- Bradbury A, Calabi F, Milstein C. Expression of CD1 in the mouse thymus. *Eur J Immunol* 1990; **20**: 1831-6.
- Balk SP, Bleicher PA, Terhorst C. Isolation and expression of cDNA encoding the murine homologues of CD1. *J Immunol* 1991; **146**: 768-74.
- Porcelli SA. The CD1 family: a third lineage of antigen-presenting molecules. *Adv Immunol* 1995; **59**: 1-98, 1-98.
- Kawano T, Cui J, Koezuka Y et al. CD1d-restricted and TCR-mediated activation of valpha14 NKT cells by glycosylceramides. *Science* 1997; **278**: 1626-9.
- Joyce S. CD1d and natural T cells: how their properties jump-start the immune system. *Cell Mol Life Sci* 2001; **58**: 442-69.
- Brossay L, Jullien D, Cardell S et al. Mouse CD1 is mainly expressed on hemopoietic-derived cells. *J Immunol* 1997; **159**: 1216-24.
- Amano M, Baumgarth N, Dick MD et al. CD1 expression defines subsets of follicular and marginal zone B cells in the spleen: beta 2-microglobulin-dependent and independent forms. *J Immunol* 1998; **161**: 1710-7.
- Mandal M, Chen XR, Alegre ML et al. Tissue distribution, regulation and intracellular localization of murine CD1 molecules. *Mol Immunol* 1998; **35**: 525-36.
- Roark JH, Park SH, Jayawardena J, Kavita U, Shannon M, Bendelac A. CD1.1 expression by mouse antigen-presenting cells and marginal zone B cells. *J Immunol* 1998; **160**: 3121-7.
- Exley M, Garcia J, Wilson SB et al. CD1d structure and regulation on human thymocytes, peripheral blood T cells, B cells and monocytes. *Immunology* 2000; **100**: 37-47.
- Spada FM, Borriello F, Sugita M, Watts GF, Koezuka Y, Porcelli SA. Low expression level but potent antigen presenting function of CD1d on monocyte lineage cells. *Eur J Immunol* 2000; **30**: 3468-77.
- Colgan SP, Morales VM, Madara JL, Polischuk JE, Balk SP, Blumberg RS. IFN-gamma modulates CD1d surface expression on intestinal epithelia. *Am J Physiol* 1996; **271**: C276-C283.
- Bonish B, Jullien D, Dutronc Y et al. Overexpression of CD1d by keratinocytes in psoriasis and CD1d-dependent IFN-gamma production by NK-T cells. *J Immunol* 2000; **165**: 4076-85.
- Salamone MC, Rabinovich GA, Mendiguren AK, Salamone GV, Fainboim L. Activation-induced expression of CD1d antigen on mature T cells. *J Leukoc Biol* 2001; **69**: 207-14.
- Ulanova M, Torebring M, Porcelli SA et al. Expression of CD1d in the duodenum of patients with cow's milk hypersensitivity. *Scand J Immunol* 2000; **52**: 609-17.

26. Tsuneyama K, Yasoshima M, Harada K, Hiramatsu K, Gershwin ME, Nakanuma Y. Increased CD1d expression on small bile duct epithelium and epithelioid granuloma in livers in primary biliary cirrhosis. *Hepatology* 1998; **28**: 620–3.
27. Chen QYD, Nadell, Zheng XY et al. The HLA DRB3*0202/DQA1*0501 haplotype is associated with Graves' disease in African American patients. *J Clin Endocrinol Metab* 2000; **80**: 1545–9.
28. den Dunnen JT, Antonarakis SE. Nomenclature for the description of human sequence variations. *Hum Genet* 2001; **109**: 121–4.
29. Jones DC, Gelder CM, Ahmad T et al. CD1 genotyping of patients with Mycobacterium malmoense pulmonary disease. *Tissue Antigens* 2001; **58**: 19–23.
30. Kukreja A, Cost G, Marker J et al. Multiple immuno-regulatory defects in type-1 diabetes. *J Clin Invest* 2002; **109**: 131–40.
31. Prokunina L, Castillejo-Lopez C, Oberg F et al. A regulatory polymorphism in PDCD1 is associated with susceptibility to systemic lupus erythematosus in humans. *Nat Genet* 2002; **32**: 666–9.
32. Prestridge DS, Signal SCAN. a computer program that scans DNA sequences for eukaryotic transcriptional elements. *Comput Appl Biosci* 1991; **7**: 203–6.
33. Zeng Z, Castano AR, Segelke BW, Stura EA, Peterson PA, Wilson IA. Crystal structure of mouse CD1: An MHC-like fold with a large hydrophobic binding groove. *Science* 1997; **277**: 339–45.