

Tamao Akesaka · Seong-Gene Lee · Jun Ohashi
Makoto Bannai · Naoyuki Tsuchiya · Yongsook Yoon
Katsushi Tokunaga · Kyuyoung Song

Comparative study of the haplotype structure and linkage disequilibrium of chromosome 1p36.2 region in the Korean and Japanese populations

Received: 2 July 2004 / Accepted: 2 August 2004 / Published online: 6 October 2004
© The Japan Society of Human Genetics and Springer-Verlag 2004

Abstract The patterns of linkage disequilibrium (LD) in the human genome provide important information for disease gene mapping. LDs may vary depending on chromosomal regions and populations. We have compared LD and haplotypes defined by SNPs in the chromosome 1p36.2 region of the Korean and Japanese populations. Fifty-eight SNPs in about 418 kb ranging from tumor necrosis factor receptor 2 (*TNFR2:TNFRSF1B*) to procollagen-lysine, 2-oxoglutarate 5-dioxygenase (*PLOD*) gene were examined in 96 healthy Koreans and Japanese each by direct sequencing and fluorescence correlation spectroscopy combined with the PCR-sequence specific primer method (PCR-SSP-FCS), respectively. Upon pair-wise LD analysis, a total of 25 and 16 out of 58 SNPs greater than MAF 10% were included in LD blocks, encompassing almost 81 kb and 55 kb in total, in Koreans and Japanese, respectively. Both similarities and differences were observed in LD strength and haplotype frequencies between the populations. Considerable similarities were observed in the telomeric region where a long-range block of approximately 80 kb including three genes was found to have strong LDs in both Koreans and Japanese. Significant

difference in LD strength was present near the *TNFR2* region between the Japanese and Korean populations.

Keywords Chromosome 1p36.2 · LD · Haplotype · Korean · Japanese

Introduction

The extent of linkage disequilibrium (LD) in the human genome provides valuable information on the identification of functional polymorphisms predisposing to human diseases. Because there are a limited number of functional polymorphisms across the genome, the positional cloning of genes for disease susceptibility depends on the polymorphisms that are in LD with functional polymorphism. Thus, the information on LD in a population is critical for the design of association mapping studies. Previous studies of LD patterns in the human genome have shown that LD appeared to vary among populations, chromosomal regions, and between pairs of markers in close proximity (Taillon-Miller et al. 2000; Abecasis et al. 2001; Reich et al. 2001).

Recent studies of human haplotype structures have demonstrated that the genome is comprised of regions of strong intermarker LD, or haplotype blocks, interspersed by presumed recombination hot spots (Taillon-Miller et al. 2000; Daly et al. 2001; Jeffreys et al. 2001; Patil et al. 2001; Dawson et al. 2002; Gabriel et al. 2002; Phillips et al. 2003). Haplotype blocks are characterized by low haplotype diversity, strong associations between alleles, and rare recombination. Analyses of the gene-based LD and haplotype patterns showed that the extent of LD and haplotype diversity varied depending on genes (Nickerson et al. 1998; Johnson et al. 2001; Turet et al. 2002). Similar studies in extended regions also showed highly variable patterns of LD depending on a region (Daly et al. 2001; Jeffreys et al. 2001; Gabriel et al. 2002). A few studies have attempted to analyze the

T. Akasaka · J. Ohashi · M. Bannai · N. Tsuchiya · K. Tokunaga
Department of Human Genetics,
University of Tokyo, Tokyo, Japan

S.-G. Lee · K. Song
Asan Institute for Life Science,
University of Ulsan College of Medicine,
Seoul, 138-736, South Korea

Y. Yoon · K. Song (✉)
Department of Biochemistry and Molecular Biology,
University of Ulsan College of Medicine,
Seoul, 138-736, South Korea
E-mail: kysong@amc.seoul.kr
Tel.: +82-2-30104277
Fax: +82-2-30104248

Present address: S.-G. Lee
Department of Biotechnology,
College of Agriculture and Life Sciences,
Chonnam National University, Gwangju, Korea

heterogeneous distribution of LD along a chromosome (Patil et al. 2001; Dawson et al. 2002; Phillips et al. 2003). Patil et al. (2001) analyzed the haplotype patterns along the entire chromosome 21 using rodent–human hybrid cell lines derived from 20 ethnically diverse individuals. They identified haplotype blocks on chromosome 21 for which >80% of chromosomes were represented by a few common haplotypes. In the analysis of human chromosome 22 with a marker density of one SNP per 15 kb, Dawson (2002) reported a highly variable pattern of LD along the chromosome, in which extensive regions of complete LD up to 804 kb in length were interspersed with regions of no detectable LD.

The aforementioned studies demonstrated the complexity of the LD and human haplotype structure. Although differences of LD patterns between populations and ethnic groups have been reported (Abecasis et al. 2001; Reich et al. 2001; Zavattari et al. 2000), little information is available on the haplotype structure in different populations other than the recent study by Gabriel et al. (2002). As a first step toward a better understanding of the haplotype differences between such populations with strong genetic affinities as Koreans and Japanese, we conducted a comparative study of the haplotype structure and LD for a 418-kb region on human chromosome 1p36.2. We chose this region based on the following facts: (1) it has been suggested to be one of the candidate regions for systemic lupus erythematosus (SLE) by genome-wide linkage analyses (Shai et al. 1999; Gaffney et al. 2000). In fact, one of the genes in the region, *TNFR2*, showed significant association with susceptibility to SLE (Komata et al. 1999); (2) *TNFR2* has also been reported to associate with various diseases, including hypertension and hypercholesterolemia (Glenn et al. 2000), familial combined hyperlipidemia (Geurts et al. 2000; van Greevenbroek et al. 2000), Alzheimer's disease (Perry et al. 2001), and rheumatoid arthritis (RA) (Barton et al. 2001; Fabris and Tolusso 2002; Kyogoku et al. 2003).

Materials and methods

Samples

Genomic DNAs from 96 healthy Korean individuals were isolated from peripheral blood leukocytes according to standard procedures with proteinase K–RNase digestion followed by phenol–chloroform extraction. Genomic DNAs from 96 unrelated, healthy Japanese were isolated from peripheral blood leukocytes using the QIAamp blood kit (Qiagen, Hilden, Germany). The present studies were reviewed and approved by the ethics committees of University of Ulsan College of Medicine and University of Tokyo, Graduate School of Medicine.

In order to minimize the required amount of genomic DNA, a whole genome amplification (WGA) method was applied to Japanese samples in the study. Primer extension preamplification (I-PEP)-PCR method by

Dietmaier et al. (1999) was used for the WGA. Briefly, 20 ng of gDNA was amplified in the reaction mixture consisting of 3.6 U of a mixture of Taq polymerase and Pwo polymerase (Expand High Fidelity PCR System, Boehringer Mannheim, Japan), 16 μ M totally degenerated 15-nucleotide-long primer (Sigma Genosys, Japan), 0.1 mM dNTP, and 2.5 mM MgCl₂.

SNP screening

One hundred nineteen SNP sites registered in a 418-kb region of 1p36.2 were selected from dbSNP (<http://www.ncbi.nlm.nih.gov/SNP>) and JSNP (<http://snp.im.s.u-tokyo.ac.jp>) databases. The occurrence of polymorphisms was assessed using 16 unrelated healthy Japanese samples by direct sequencing (ABI 3100 sequencer, Applied Biosystems, Japan). Among the 119 sites selected from the public databases, at least one heterozygous sample was observed for 79 SNP sites, among which 58 SNPs were selected for genotyping. A complete list of the 58 SNPs, their reference sequence numbers, positional information, minor allele frequencies in both populations, and the list of assay primers will be available upon request.

Typing method

Japanese samples were genotyped using the PCR-SSP-FCS method, as described elsewhere (Bannai et al. 2004). Briefly, the first PCR was performed to amplify a fragment including an SNP site using 1 μ l of WGA products as templates. Then a sequence-specific primer (SSP)-PCR was performed with allele-specific, seminested fluorescence- (TAMRA or Cy5) labeled primers using the first PCR products as templates. The SSP-PCR products were then analyzed by fluorescence correlation spectroscopy (FCS) measurement using the single-molecule fluorescence detection system (MF10, Olympus Corporation, Japan). The primers used for the first PCR and SSP-PCR will be available upon request. Korean samples were genotyped by direct sequencing.

Statistical analysis

The genotype frequencies for each SNP were checked for consistency between the observed values and those expected from Hardy–Weinberg equilibrium using a commercial program SNP Alyze V 3.0 (Dynacom Co., Japan) in each population, respectively. Haploview version 2.05, as well as the SNP Alyze software package based on the expectation-maximization (EM) method (Excoffier and Slatkin 1995), were used to estimate the haplotype frequencies, the Lewontin's coefficients D' (Lewontin 1988), and correlation coefficient r (Hill and Robertson 1968). The block structures and their haplotype frequencies were estimated using Haploview version

2.05. Calculation of the gene-average LD was done as described (Tiret et al. 2002). Briefly, the gene-based median D' and r^2 were calculated by averaging the absolute value $k(k-1)/2$ pair-wise disequilibrium coefficients D' and r^2 (k means number of biallelic polymorphisms within a gene). The SNP Alyze program was used to assess the statistical significance of haplotype profile differences and individual haplotype frequency differences between the two populations.

Results

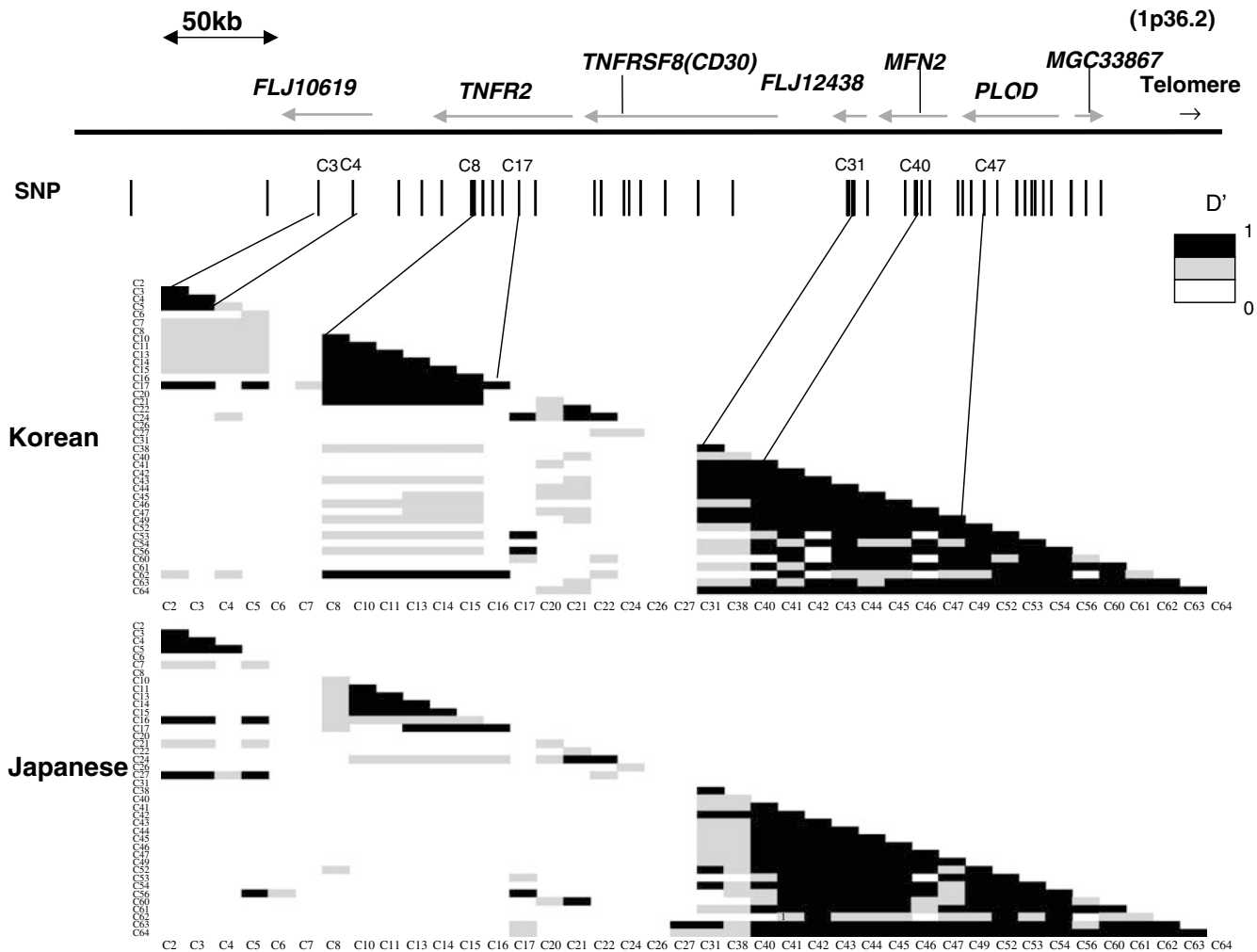
Patterns of linkage disequilibrium

We studied the extent of LD by choosing a region of 418 kb (12.1–11.7 Mb, build 34) on chromosome 1p36.2. Total 119 SNPs were chosen from public databases (dbSNP, JSNP). Initial screening was performed in 16 healthy Japanese individuals. Out of 119 SNPs, 79

were found to be polymorphic in these samples. To reduce the typing cost, 60 SNPs (designated C1–C64) were selected from 79 polymorphic SNPs based on the following criteria: (1) one of the two markers in close proximity; (2) genotype frequency distribution in agreement with Hardy–Weinberg's equilibrium ($P > 0.05$). Then the selected 60 SNPs were genotyped in 96 unrelated Korean and Japanese individuals each. Of 60 SNPs, two were nonpolymorphic in the Korean samples. Of 58 SNPs, 42 and 40 were polymorphic with minor allele frequency $> 10\%$ in the Korean and Japanese samples, respectively, and the minor allele frequencies of SNPs C19 and C23 were less than 10% in the Japanese only. The 42 markers showed an average spacing of 10 kb; however, they were not evenly spaced (Fig. 1). The SNPs with minor allele frequency < 0.10 were omitted from the LD analysis.

To investigate patterns of LD in the region, two measures of LD (D' and r^2) were estimated between all pair-wise combinations of markers in each population. Lewontin's standardized coefficient $D' = 1$ indicates absolute LD when one or two haplotypes are missing (Lewontin 1988) whereas $r = 1$ indicates absolute LD when two of four haplotypes are missing (Hill and

Fig. 1 The location of SNPs examined and pairwise D' between SNPs with minor allele frequency of $> 10\%$



Robertson 1968). The D' value for pairs of 42 SNPs in both population samples were shown in Fig. 1. Upon the pair-wise LD analysis, the LD strength varied considerably depending on the regions.

We have assessed the pattern of LD by calculating average D' and r^2 in successive segments of an average of 74.6 kb (11.9 kb overlap) (Fig. 2). The general patterns of LD appeared similar between the two populations except in the region near *TNFR2* (*TNFRSF1B*). A high LD was observed in both populations across the three genes toward the telomeric region. In contrast, in the

TNFR2 (*TNFRSF1B*) region, a high LD was observed within the gene and in Koreans only.

We adapted the criteria of an LD block as any series of two or more markers in a contig for which all pair-wise values of $D' > 0.8$ on average with 95% confidence bound between 0.7 and 0.98 (Gabriel et al. 2002). Using the Haploview program, two common haplotype blocks were identified for a 418-kb region of chromosome 1p36.2 in both population samples and spanned the intervals from 8.3 kb to 47.1 kb in length. As shown in Fig. 3, a block of very strong LD ($D' > 0.8$) was observed in the Korean

Fig. 2 Distribution of linkage disequilibrium on chromosome 1p36.2. The *solid* and *dotted* lines present sliding-window plots of average D' and r^2 , respectively, in mean interval of 74.6 kb with 11.9 kb overlap for markers with minor allele frequency ≥ 0.10

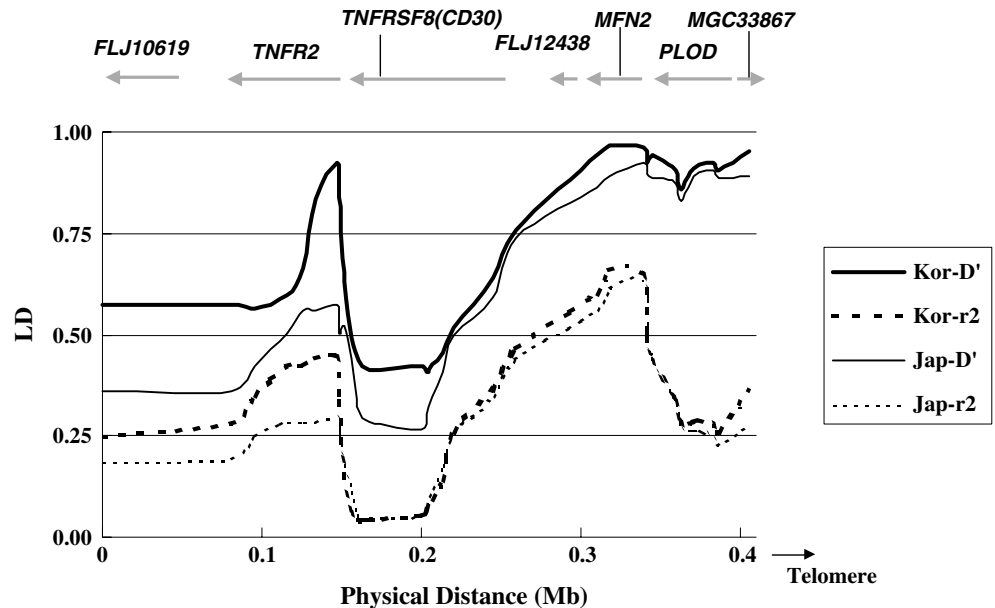


Fig. 3 Haplotype distribution in each LD block. Tag SNPs for Koreans and Japanese were shown in *stars* and *diamonds*, respectively. Difference in the haplotype profile between the populations was assessed through permutation test

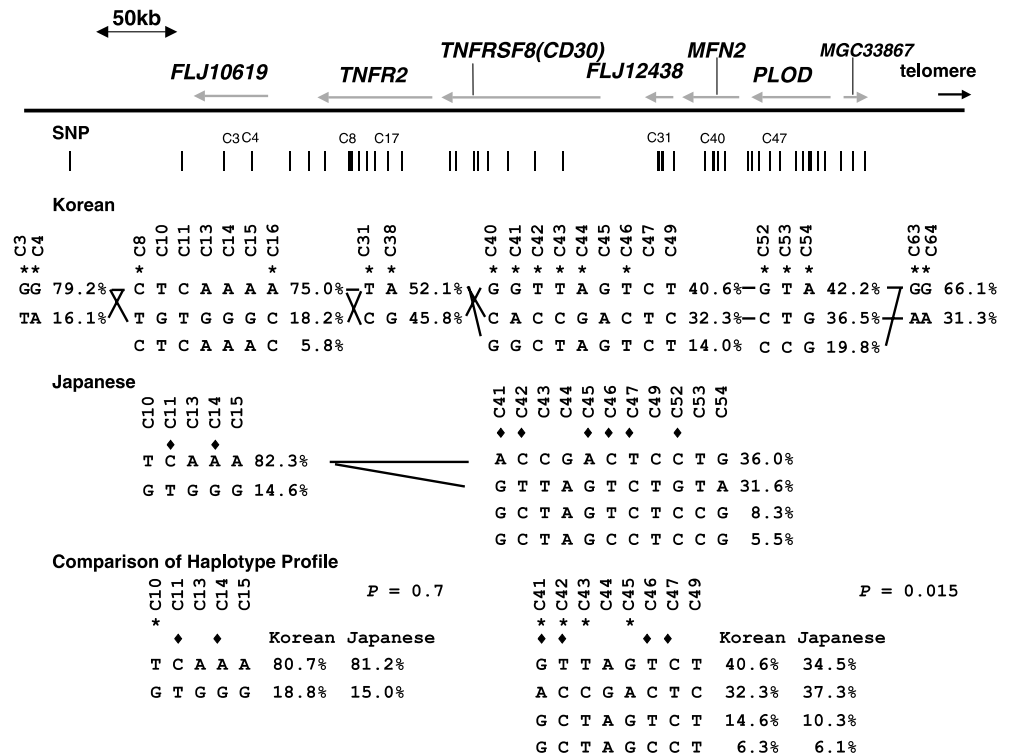


Table 1 Median D' and r^2 values for each gene

Gene	Loci	Korean		Japanese	
		D'	r^2	D'	r^2
TNFR2	10	0.851	0.424	0.711	0.287
CD30	9	0.459	0.035	0.443	0.036
FLJ12438	11	0.946	0.526	0.859	0.400
MFN2	8	0.969	0.513	0.905	0.513
PLOD	12	0.907	0.230	0.792	0.188
MGC33867	2	1.000	0.888	0.895	0.781

sample, C8–C10–C11–C13–C14–C15–C16, spanning 13.4 kb from introns 7 to 1 of the *TNFR2* (*TNFRSF1B*) gene. This block was smaller in the Japanese sample, C10–C11–C13–C14–C15, spanning 8.3 kb from exon 6 to intron 1 of the *TNFR2* (*TNFRSF1B*) gene. The second block of strong LD in the Japanese sample was the longest observed in both populations, C41–C42–C43–C44–C45–C46–C47–C49–C52–C53–C54, ranging from intron 11 of the *MFN2* gene to intron 6 of *PLOD* gene, ~47.1 kb. In the Korean sample, this block was divided into two blocks, C40–C41–C42–C43–C44–C45–C46–C47–C49 and C52–C53–C54, 35.2 kb and 3.2 kb, respectively. Three additional blocks consisting of two markers were identified in the Korean sample only and spanned the interval from 5.9 kb to 14.7 kb in length.

On the other hand, slightly different block structures were identified when the 24 and 26 SNPs with minor allele frequency ≥ 0.2 in Koreans and Japanese, respectively, was used to aid identification of LD blocks (data not shown). Of note was a long LD block consisting of 14 markers, C41–C42–C43–C44–C45–C46–C47–C49–C52–C53–C54–C61–C63–C64, observed in the Japanese sample across the 80-kb region from intron 11 of the *MFN2* gene to intron 2 of *MGC33867*, including three genes toward the telomeric region. In the Korean sample, three blocks, C31–C38, C40–C41–C42–C43–C44–C45–C46–C47–C49, C52–C53–C54–C61–C63–C64, of 8.3, 35.2, and 35.8 kb, respectively, were identified in the telomeric region.

Genotype and haplotype diversity

To investigate the haplotype diversity within each block, haplotypes were inferred based on unphased SNP data using the Haploview program (Fig. 3). The same program was used to identify tag SNPs. To avoid problems associated with the differences in block boundaries, we chose to include the markers common to both population samples in the two blocks. In block 1 (C10–C11–C13–C14–C15), of the two (frequency $\geq 5\%$) haplotypes predicted, they accounted for 99.5% and 96.2% of all those seen in the Korean and Japanese samples, respectively (Fig. 3). And C10 and two SNPs (C11 and C14) were sufficient to represent the two haplotypes in the Korean and Japanese samples, respectively. In block 2 (C41–C42–C43–C44–C45–C46–C47–C49), four common haplotypes accounted for 93.8% and 88.2% of the Korean and

Table 2 Distribution of gene-based haplotypes

Gene (Loci)	Korean (%)	Japanese (%)	P value
TNFR2 (C7–C18)			0.002 ^a
CCTCAAAAGT	33.7	26.8	0.428 ^b
TCTCAAAAGT	25.9	30.7	0.553
CCTCAAAAAT	12.1	5.0	0.101
CTGTGGGCGT	11.2	3.9	0.023 ^c
TTGTGGGCGT	7.0	4.3	0.319
TCTCAAACGT	5.0	1.2	0.174
TCTCAAAAAT	2.7	6.8	0.200
CD30 (C19–C27)			0.979
GAGCTACAT	13.7	13.8	0.978
GAGTTGCAT	9.0	13.6	0.487
GAGCTGCAT	8.5	5.7	0.827
GAGTTGCGT	7.3	5.7	0.799
GAAGTGCAT	7.0	1.1	0.131
GAGCCGCAT	6.0	0.0	0.024 ^c
GGACTGCAT	2.2	11.0	0.031 ^c
FLJ12438 (C28–C38)			0.370
ACCTCCGGAAA	49.0	51.1	0.738
ACCCCGGAGG	43.1	41.5	0.771
GTTTTCAATAA	2.0	0.0	0.007 ^c
ACCTCCGGAGG	1.1	2.3	0.194
MFN2 (C39–C46)			0.356
TGGTTAGT	40.6	35.8	0.413
TCACCGAC	32.3	38.0	0.260
TGGCTAGT	14.0	11.5	0.399
CCGCTAGC	4.7	4.2	0.623
TCGCCAGC	3.1	1.0	0.071
PLOD (C47–C62)			0.332
CTGTAAGGGGGA	40.0	36.6	0.511
TCCTGAGGGGAA	29.7	31.7	0.676
CTCCGGGGGGGA	4.7	1.9	0.145
CTCCGATAAAGG	4.7	1.9	0.151
CTCCGGGGGAGG	4.7	4.0	0.840
CCCTGAGGGGAA	3.1	1.2	0.162
CTCCGGGGGAGA	2.6	2.5	0.994
CTCCGATAAAGGA	2.1	3.6	0.437
MGC33867 (C63–C64)			0.339
GG	66.1	63.0	0.532
AA	31.3	31.8	0.835
AG	2.6	2.9	0.751

^aGlobal P values between the Korean and Japanese populations based on 1,000 permutations

^bPermutation test significance levels for individual haplotype frequency comparisons between the two populations

^c $P < 0.05$, accepted level of significance

Japanese samples, respectively. The order of the most and second-most frequent haplotypes in the Japanese samples were reversed in the Korean samples. Four SNPs, C41, C42, C43, and C45 for Korean and C41, C42, C46, and C47 for Japanese, would be sufficient enough to define the four most common haplotypes. As shown in Fig. 3, block 2 showed statistically significant differences in the haplotype profile between the two populations based on 1,000 permutations ($P = 0.015$).

Regardless of the criteria of LD blocks, if genes happened to be in the region of poor LD, it becomes necessary to analyze gene-based haplotypes. We tried to analyze gene-based haplotypes of the six genes, including *TNFR2* (*TNFRSF1B*), *CD30* (*TNFRSF8*), *FLJ12438*, *MFN2*, *PLOD*, and *MGC33867*. Median D' and r^2 for each gene were shown in Table 1. As was

expected from the analysis of LD patterns (Figs. 1, 2), CD30 showed poor median LD values. For CD30, seven haplotypes comprised of nine SNPs were identified, with the most frequent one of about 14% (Table 2). On the other hand, for *MFN2*, where all eight SNP markers were in strong LD, the most frequent haplotype showed an average of 38% in the two populations. The median LD values for *TNFR2 (TNFRSF1B)* were the second lowest (Table 1). *TNFR2 (TNFRSF1B)* also showed the largest differences in median LD values between the populations. SNP Alyze software package was used to assess the statistical significance of haplotype profile differences between the samples. As shown in Table 2, *TNFR2 (TNFRSF1B)* alone showed statistically significant differences in the haplotype distribution between the Japanese and Korean samples based on 1,000 permutations ($P=0.002$). A few haplotypes, including one from *TNFR2*, two from *CD30*, and one from *FLJ12438*, showed statistically significant differences in the comparison of individual haplotypes between the two populations (Table 2).

Discussion

Recent studies have shown that the genome is broken into blocks of strong haplotype structure separated by shorter regions of shattered haplotype structure. So far, only a few studies have attempted to analyze haplotype structures in different populations (Gabriel et al. 2002; Shifman et al. 2003; Stenzel et al. 2004). As the International Haplotype Map (HapMap) Project, including the Japanese and Han Chinese is ongoing, it is to our interest whether such a map could be useful for the Korean population as well. One of the potential problems with our study would be ascertainment bias in SNP frequency because all the SNPs typed have been chosen based on 16 Japanese samples. Another problem would be that the accuracy of the haplotype inference method used in this study is not known.

Various analytical approaches were available to define haplotype blocks, and a block definition used in the present study was rather stringent (Gabriel et al. 2002). In the present study encompassing the 400-kb region of chromosome 1p36.2, we were able to identify from two to six blocks, depending on the criteria of a LD block, spanning the interval from 3.2 kb to 47.1 kb in length. The boundaries for the two major blocks obtained with SNPs of a minor allele frequency of 0.10 varied between these closely related populations. Because of the differences in LD strength and block boundaries, different tag SNPs were identified between the two populations for each block. Furthermore, haplotype profiles of the second block showed statistically significant differences between the populations. When the SNPs of a minor allele frequency threshold of 0.20 instead of 0.10 were included in an LD block, we were able to identify only one long block toward the telomeric region in the Japanese samples whereas three blocks were identified in the

Korean samples spanning the same and neighboring regions.

Regardless of the criteria of an LD block, it appeared that most cores of haplotype blocks in the Korean samples coincided with those in the Japanese. Although it still remains to be seen whether the populations used in this study reflect the divergence of both populations, our data of considerable similarities in LD strength and the kinds of haplotype appear to reflect strong genetic affinities between the two populations (Tokunaga et al. 1996).

For the genes in the region where LD block structure cannot be identified, it becomes necessary to analyze gene-based haplotypes. Median LD value was the lowest for *CD30 (TNFRSF8)* in both populations, and the most frequent haplotype showed a population frequency of about 14% compared to 50% for *FLJ12438* with 11 markers. The ethnic difference in median LD values was the largest for *TNFR2 (TNFRSF1B)* (Table 1). Median LD was higher in Koreans than in Japanese. This was the region that showed significant differences in LD strength between the populations (Fig. 2). However, this difference in LD was mainly due to the markers C7, C8, C16, C17, and C18. As shown in Fig. 3, in block 1 (from exon 6 to intron 1 of the *TNFR2 (TNFRSF1B)* gene, C10–C11–C13–C14–C15), it seemed that there was no significant difference in haplotype profiles between the two populations. However, when all the markers of the *TNFR2 (TNFRSF1B)* gene were included, the difference in LD was reflected in the significance level of differences in the *TNFR2* haplotype profile between the populations ($P=0.002$). This data suggested that if a candidate gene such as *TNFR2* were used for an association study in Japanese, it could be hard to obtain reproducible data in Koreans.

Taken together, the HapMap might facilitate comprehensive genetic association studies of human diseases for some regions where significant similarities in LD and haplotype structure were present between the populations of interest and those used in the construction of HapMap. As reports on the population differences in LD pattern across the human genome demonstrated that chromosome region-specific effects appear more important than population-specific effects in influencing the extent of LD (Zavattari et al. 2000), further studies on various regions across the human genome are needed to address the population differences in block boundaries and haplotype frequencies.

Acknowledgements This work was supported by the grants (to KS) from the Korean HapMap Project of MOST (Ministry of Science and Technology of Korea) and Asan Foundation, and by the grants (to J.O., N.T. and K.T.) from the New Energy and Industrial Technology Development Organization (NEDO) in Japan.

References

- Abecasis G, Noguchi E, Heinzmann A, Traherne J, Bhattacharya S (2001) Extent and distribution of linkage disequilibrium in three genomic regions. *Am J Hum Genet* 68:191–197

- Bannai M, Higuchi K, Akesaka T, Furukawa M, Yamaoka M, Sato K, Tokunaga K (2004) Single-nucleotide-polymorphism genotyping for whole-genome-amplified samples using automated fluorescence correlation spectroscopy. *Anal Biochem* 327:215–221
- Barton A, John S, Ollier WER, Silman A, Worthington J (2001) Association between rheumatoid arthritis and polymorphism of tumor necrosis factor receptor II, but not tumor necrosis factor receptor I, in Caucasians. *Arthritis Rheum* 44:61–65
- Daly MJ, Rioux JD, Schaffner SF, Hudson TJ, Lander ES (2001) High-resolution haplotype structure in the human genome. *Nat Genet* 29:229–232
- Dawson E, Abecasis GR, Bumpstead S, Chen Y, Hunt S, Beare DM, Pabial J, Dibling T, Tinsley E, Kirby S, Carter D, Pappaspyridonos M, Livingstone S, Ganske R, Lohmussaar E, Zernant J, Tonisson N, Remm M, Magi R, Puurand T, Vilo J, Kurg A, Rice K, Deloukas P, Mott R, Metspalu A, Bentley DR, Cardon LR, Dunham I (2002) A first-generation linkage disequilibrium map of human chromosome 22. *Nature* 418:544–548
- Dietmaier W, Hartmann A, Wallinger S, Heinmoller E, Kerner T, Endl E, Jauch KW, Hofstadter F, Ruschoff J (1999) Multiple mutation analyses in single tumor cells with improved whole genome amplification. *Am J Pathol* 154:83–95
- Excoffier L, Slatkin M (1995) Maximum-likelihood estimation of molecular haplotype frequencies in a diploid population. *Mol Biol Evol* 12:921–927
- Fabris M, Tolusso B, Di Poi E, Assaloni R, Sinigaglia L, Ferraccioli G (2002) Tumor necrosis factor- α receptor II polymorphism in patients from southern Europe with mild-moderate and severe rheumatoid arthritis. *J Rheumatol* 29:1847–1850
- Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, Blumenstiel B, Higgins J, DeFelice M, Lochner A, Faggart M, Liu-Cordero SN, Rotimi C, Adeyemo A, Cooper R, Ward R, Lander ES, Daly MJ, Altshuler D (2002) The structure of haplotype blocks in the human genome. *Science* 296:2225–2229
- Gaffney PM, Ortmann WA, Selby SA, Shark KB, Ockenden TC, Rohlf KE, Walgrave NL, Boyum WP, Malmgren ML, Miller ME, Kearns GM, Messner RP, King RA, Rich SS, Behrens TW (2000) Genome screening in human systemic lupus erythematosus: results from a second Minnesota cohort and combined analyses of 187 sib-pair families. *Am J Hum Genet* 66:547–556
- Geurts JM, Janssen RG, van Greevenbroek MM, van der Kallen CJ, Cantor RM, Bu X, Aouizerat BE, Allayee H, Rotter JJ, de Bruin TW (2000) Identification of TNFRSF1B as a novel modifier gene in familial combined hyperlipidemia. *Hum Mol Genet* 9:2067–2074
- Glenn CL, Wang WYS, Benjafield AV, Morris BJ (2000) Linkage and association of tumor necrosis factor receptor 2 locus with hypertension, hypercholesterolemia and plasma shed receptor. *Hum Mol Genet* 9:1943–1949
- van Greevenbroek MM, van der Kallen CJ, Geurts JM, Janssen RG, Buurman WA, de Bruin TW (2000) Soluble receptors for tumor necrosis factor- α (TNF-R p55 and TNF-R p75) in familial combined hyperlipidemia. *Atherosclerosis* 153:1–8
- Hill WG, Robertson A (1968) Linkage disequilibrium in finite populations. *Theor Appl Genet* 38:226–231
- Jeffreys AJ, Kauppi L, Neumann R (2001) Intensely punctate meiotic recombination in the class II region of the major histocompatibility complex. *Nat Genet* 29:217–222
- Johnson GC, Esposito L, Barratt BJ, Smith AN, Heward J, Di Genova G, Ueda H, Cordell HJ, Eaves IA, Dudbridge F, Twells RC, Payne F, Hughes W, Nutland S, Stevens H, Carr P, Tuomilehto-Wolf E, Tuomilehto J, Gough SC, Clayton DG, Todd JA (2001) Haplotype tagging for the identification of common disease genes. *Nat Genet* 29:233–237
- Komata T, Tsuchiya N, Matsushita M, Hagiwara K, Tokunaga K (1999) Association of tumor necrosis factor receptor 2 (TNFR2) polymorphism with susceptibility to systemic lupus erythematosus. *Tissue Antigens* 53:527–533
- Kyogoku C, Tsuchiya N, Shibue T, Tokunaga K, Matsuta K (2003) TNFR2 position 196 polymorphism in the Japanese patients with rheumatoid arthritis. *Arthritis Rheum* 48:273–274
- Lewontin RC (1988) On measures of gametic disequilibrium. *Genetics* 120:849–852
- Nickerson DA, Taylor SL, Weiss KM, Clark AG, Hutchinson RG, Stengard J, Salomaa V, Vartiainen E, Boerwinkle E, Sing CF (1998) DNA sequence diversity in a 9.7-kb region of the human lipoprotein lipase gene. *Nat Genet* 19:233–240
- Patil N, Berno AJ, Hinds DA, Barrett WA, Doshi JM, Hacker CR, Kautzer CR, Lee DH, Marjoribanks C, McDonough DP, Nguyen BT, Norris MC, Sheehan JB, Shen N, Stern D, Stokowski RP, Thomas DJ, Trulson MO, Vyas KR, Frazer KA, Fodor SP, Cox DR (2001) Blocks of limited haplotype diversity revealed by high-resolution scanning of human chromosome 21. *Science* 294:1719–1723
- Perry RT, Collins JS, Wiener H, Acton R, Go RC (2001) The role of TNF and its receptors in Alzheimer's disease. *Neurobiol Aging* 22:873–883
- Phillips MS, Lawrence R, Sachidanandam R, Morris AP, Balding DJ, Donaldson MA, Studebaker JF, Ankeney WM, Alfisi SV, Kuo FS, Camisa AL, Pazorov V, Scott KE, Carey BJ, Faith J, Katari G, Bhatti HA, Cyr JM, Derohannessian V, Elosua C, Forman AM, Grecco NM, Hock CR, Kuebler JM, Lathrop JA, Mockler MA, Nachtman EP, Restine SL, Varde SA, Hozza MJ, Gelfand CA, Broxholme J, Abecasis GR, Boyce-Jacino MT, Cardon LR (2003) Chromosome-wide distribution of haplotype blocks and the role of recombination hot spots. *Nat Genet* 33:382–387
- Reich DE, Cargill M, Bolk S, Ireland J, Sabeti PC, Richter DJ, Lavery T, Kouyoumjian R, Farhadian SF, Ward R, Lander ES (2001) Linkage disequilibrium in the human genome. *Nature* 411:199–204
- Shai R, Quismorio FP Jr, Li L, Kwon OJ, Morrison J, Wallace DJ, Newwelt CM, Brautbar C, Gauderman WJ, Jacob CO (1999) Genome-wide screen for systemic lupus erythematosus susceptibility genes in multiplex families. *Hum Mol Genet* 8:639–644
- Shifman S, Kuypers J, Kokoris M, Yakir B, Darvasi A (2003) Linkage disequilibrium patterns of the human genome across populations. *Hum Mol Genet* 12:771–776
- Stenzel A, Lu T, Koch WA, Hampe J, Guenther SM, De La Vega FM, Krawczak M, Schreiber S (2004) Patterns of linkage disequilibrium in the human MHC region on human chromosome 6p. *Hum Mol Genet* 13:377–385
- Taillon-Miller P, Bauer-Sardina I, Saccone NL, Putzel J, Laitinen T, Cao A, Kere J, Pilia G, Rice JP, Kwok PY (2000) Juxtaposed regions of extensive and minimal linkage disequilibrium in human Xq25 and Xq28. *Nature Genet* 25:324–328
- Tiret L, Poirier O, Nicaud V, Barbaux S, Herrmann SM, Perret C, Raoux S, Francomme C, Lebard G, Tregouet D, Cambien F (2002) Heterogeneity of linkage disequilibrium in human genes has implications for association studies of common diseases. *Hum Mol Genet* 11:419–429
- Tokunaga K, Imanishi T, Takahashi K, Juji T (1996) On the origin and disposal of East Asian populations as viewed from HLA haplotypes. In: Akazawa T, Szathmary EJ (eds) Prehistoric mongoloid dispersals. Oxford University Press, Oxford, pp 187–197
- Zavattari P, Deidda E, Whalen M, Lampis R, Mulargia A, Lollo M, Eaves I, Mastio G, Todd JA, Cucca F (2000) Major factors influencing linkage disequilibrium by analysis of different chromosome regions in distinct populations: demography, chromosome recombination frequency and selection. *Hum Mol Genet* 9:2947–2957