International Immunology, Vol. 10, No. 12, pp. 1891–1895

Identification of the gene loci that predispose to rheumatoid arthritis

Shunichi Shiozawa, Sachiko Hayashi, Yasuo Tsukamoto, Hideki Goko, Hiroki Kawasaki, Takahiko Wada, Kazuhiko Shimizu, Norikazu Yasuda[1], Naoyuki Kamatani[2], Kiyoshi Takasugi[3], Yasushi Tanaka[4], Kazuko Shiozawa[4] and **Shigeaki Imura[4]**

Kobe University School of Medicine Faculty of Health Science, 7-10-2 Tomogaoka, Sumaku, Kobe 654-0142, Japan
[1]Division of Information on Genetic Diseases, Research Center for Biological Information and Resources, Association of Genetics Promotion, 1175-195 Yata-Sakuragaoka, Mishimashi, Shizuoka 411-0801, Japan
[2]Tokyo Women's Medical College, Institute of Rheumatology, 10-22 Kawadacho, Shinjikuku, Tokyo 162-0045, Japan
[3]Dogo Onsen Hospital, 21-21 Dogo-Himezuka, Matsuyamashi, Ehime 790-2502, Japan
[4]Center for Rheumatic Diseases, Kakogawa National Hospital, 1545-1 Saijo, Kannocho, Kakogawa 675-8545, Japan

Keywords: gene locus, genetic predisposition, microsatellite marker, rheumatoid arthritis

Abstract

We have searched the human genome for genes that predispose to rheumatoid arthritis (RA) using fluorescence-based microsatellite marker analysis and affected sib-pair linkage study. A panel of 41 Japanese families, each with at least two affected siblings, was typed for genome-wide 358 polymorphic microsatellite marker loci. Markers were amplified by the PCR using fluorescence-tagged primers and sized based on the difference of CA repeats on DNA. Linkage analysis was made using maximum lod score (MLS). The MLS for D1S214 and D8S556 was 3.27 and 3.33, while the MLS for the HLA-DRB1 region was <3.0. According to detailed analysis by single-point analysis using MAPMAKER/SIBS, the MLS for D1S253 and D1S214 was 3.77 and 3.58. The MLS by multi-point analysis was 6.13 for D1S253. The MLS for D8S556 by single-point analysis was 4.20. The MLS for DXS1232 was 2.35 by single-point analysis, whereas the MLS for the region 2 cM right to DXS1232 and the region between DXS1227 and DXS1200 was 3.03 and 2.93 by multi-point analysis. Three principal chromosome regions of linkage, D1S253/214, D8S556 and DXS1232, have been identified which we call RA1, RA2 and RA3 for RA disease loci.

Introduction

Rheumatoid arthritis (RA) is a chronic symmetric polyarthritis of unknown etiology that affects ~1% of the population worldwide, where genetic and environmental factors are suspected to be important in its pathogenesis (1). Previous studies indicate that the risk of the disease in the siblings of affected individuals (λ_s) is much increased in RA (2), suggesting that genetic factors may be important as a cause of familial clustering (2,3). Genetic studies to date have focused primarily on the role of HLA molecules in RA (1,3). However, while the ratio of the risk for siblings of patients with a disease and the population prevalence of that disease (λ_s) (4) was 8 for RA, the λ_s for HLA was significantly low at 1.6 in Caucasian patients with RA (2,5). Thus, the HLA-linked susceptibility locus did not appear to sufficiently explain the observed familial clustering of RA (3). At least one non-HLA-linked disease locus has been proposed by recent genetic epidemiology studies (3,6).

We have searched the human genome for genes that predispose to RA using fluorescence-based microsatellite marker analysis and affected sib-pair linkage study. Since parents were not typed in most of the RA families because of late onset nature of the disease, we have typed one unaffected sibling in addition to the affected sib-pair for each family. Using computer programs, one produced by us and the other by other investigators, the maximum likelihood analysis was performed to detect markers with positive linkage to the disease (7). Both single-point and multi-point sib-pair analyses were performed, the

Correspondence to: S. Shiozawa

Transmitting editor: K. Okumura

Received 21 August 1998, *accepted* 3 September 1998

latter being powerful when both genotype data from multiple linked loci and genetic distances between the loci were available (8).

Methods

Family study

Japanese families including affected sib-pairs who met the diagnostic criteria of the American College of Rheumatology (formerly, the American Rheumatism Association) (9) with clear and identifiable joint destruction of more than stage 2 of Steinbroker's X-ray classification (10) were studied under the consent of patients and their family members on this type of familial genetic search. Peripheral blood (10 ml) was withdrawn using EDTA and gently mixed with 20 ml of buffer I (0.32 M sucrose, 5% v/v Triton X-100, 5 mM $MgCl_2$ and 12 mM Tris–HCl, pH 7.6) to lyse cell membrane. After centrifugation, the precipitate of nuclei was reacted with buffer II (4 M guanidine thiocyanate, 12 mM EDTA, 375 mM NaCl, 0.5% sodium *N*-lauroyl sarcosinate, 0.1 M β-mercaptoethanol and 12 mM Tris–HCl, pH 7.6) at 55°C for 10 min to lyse nuclear membrane. Genomic DNA was extracted by ethanol precipitation.

Microsatellite DNA sizing (11)

Fluorescence-tagged primers for 358 microsatellite markers that define a 10.8 cM resolution human index map were purchased from PE Applied Biosystems-Roche Molecular Systems, (Branchburg, NJ). The markers D1S502 and D6S344 were not used because of technical difficulty in amplifying their DNA. D6S276 was substituted to D6S299, D6S265 and D6S273 to specifically examine the HLA-D region. Other microsatellite markers were newly synthesized according to the published primer sequences (12). PCR was performed in 96-well microtiter plates (6511; Corning Coster, Cambridge, MA) in a 15 μl volume containing 30 ng of DNA, 0.2 μM of primer mixture, 0.2 mM of each dNTP, 1 U of AmpliTaq Gold DNA polymerase (PE Applied Biosystems), 2.5 mM $MgCl_2$ and 1×PCR buffer II (PE Applied Biosystems), and amplification conditions in a MJ Research (Watertown, MD) PTC-100 thermocycler were 94°C for 10 min followed by 27 cycles of 94°C for 30 s, 55°C for 1 min, 72°C for 2 min and 72°C for 5 min after the last cycle. PCR products labeled with 6-FAM, TET or HEX were mixed together with TAMURA-labeled size standards and applied on a 36 cm well-to-read gel (4% acrylamide/6 M urea) plate. Electrophoresis was carried out in a ABI 377 sequencer (PE Applied Biosystems) at 3000 V for 2 h. The peak, height and area of DNA fragments were identified using the GENESCAN (version 2.0.0) computer analysis software (PE Applied Biosystems) with reference to TAMURA-labeled size standards. The DNA fragments were properly sized by lane-to-lane gel comparison and automatic allele calling using the data analysis and transformation software GENOTYPER (version 1.1) (PE Applied Biosystems). Information obtained was transformed into the MS-DOS-based SIB-ADONE computer program (which is available on request by e-mail to shiozawa@ams.kobe-u.ac.jp).

HLA-DRB1 typing

HLA-DRB1 was typed by modifying the PCR restriction fragment length polymorphism (RFLP) method by introducing DRB1 group-specific primers (13). PCR was performed in a 50 μl volume containing 100 ng of DNA, 0.2 mM of each dNTP, 1.25 U of Taq DNA polymerase (PE Applied Biosystems), 1.5 mM $MgCl_2$, 50 mM KCl, 0.01% gelatin, 10 mM Tris–HCl, pH 8.4, and a set of primer (13), and amplification conditions were 94°C for 10 min followed by 30 cycles of 95°C for 30 s, 59–64°C for 30 s, 72°C for 30 s and 72°C for 5 min after the last cycle. After digestion using a battery of restriction enzymes, HLA-DRB1 was typed by RFLP band patterns in a 12% acrylamide gel.

Linkage analysis

Linkage analysis was made using the SIB-ADONE program which was developed by us according to the method of Holmans and Clayton (14). This program is specified for the maximum likelihood analysis of sib-pairs for each of which parents are not typed but a pair of affected sib and an unaffected sib are typed. Probabilities of allele sharing of affected sibs but not those of other types of pairs were considered and this was expected to yield a close approximation to the full likelihood especially when penetrance was low (14). The genotypes of unaffected sibs were used to infer parental genotypes. The contribution of the *i*th pair to the likelihood (L_i) is:

$$L_i = \sum_{P \in \mathrm{P}} \Pr\left\langle \begin{matrix} \textit{parental} \\ \textit{genotypes} \cdot P \end{matrix} \right\rangle \Pr\left\langle \begin{matrix} \text{genotype of} \\ \text{unaffected} \\ \text{sib} \end{matrix} \middle| P \right\rangle \times \sum_{j=0}^{2} \Pr\left\langle \begin{matrix} \textit{genotypes of} \\ \textit{affected pair} \end{matrix} \middle| \begin{matrix} j \cdot \textit{IBD} \\ P \end{matrix} \right\rangle z_j,$$

where z_j is the probability of sharing *j* IBD between an affected sib-pair and *P* is the set of genotypes consistent with the observed parental data. The likelihood of the whole sample was maximized with respect to all markers as well as IBD sharing probabilities z_0, z_1 and z_2 under the restriction of $z_1 \leqslant 1/2$, $z_0 \leqslant 1/2\, z_1$ and $z_0 + z_1 + z_2 = 1$.

The maximum lod score (MLS) was calculated as:

$$\text{MLS} = \log10(\text{likelihood under } z/\text{likelihood under null hypothesis})$$

where *z* is the maximum likelihood proportions. The MAPMAKER/SIBS program was used to confirm and extend the result obtained by the SIB-ADONE, in which both single point and multi-point sib-pair analyses were performed (15). Analysis was made within the possible triangle for autosomal loci or under the genetic constraints as described by Cordell *et al.* (16) for X-linked loci but without the assumption of no dominance variance.

Results

A panel of 41 Japanese families, including at least two affected siblings with full-blown clinical RA and clear and identifiable joint destruction under X-ray films (Fig. 1), was typed for genome-wide 358 polymorphic microsatellite marker loci, with a mean heterozygosity of 0.79. The sex-averaged genetic distance between two consecutive marker loci was 10.8 cM.

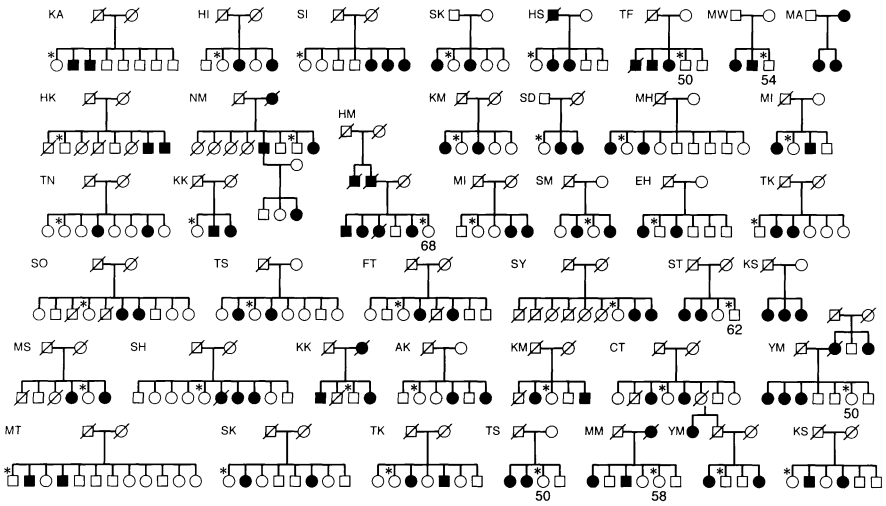

Fig. 1. Family trees of affected sib-pairs with RA. One unaffected elderly sibling (indicated by an asterisk with age) was typed in addition to the corresponding affected sib-pair for each family.

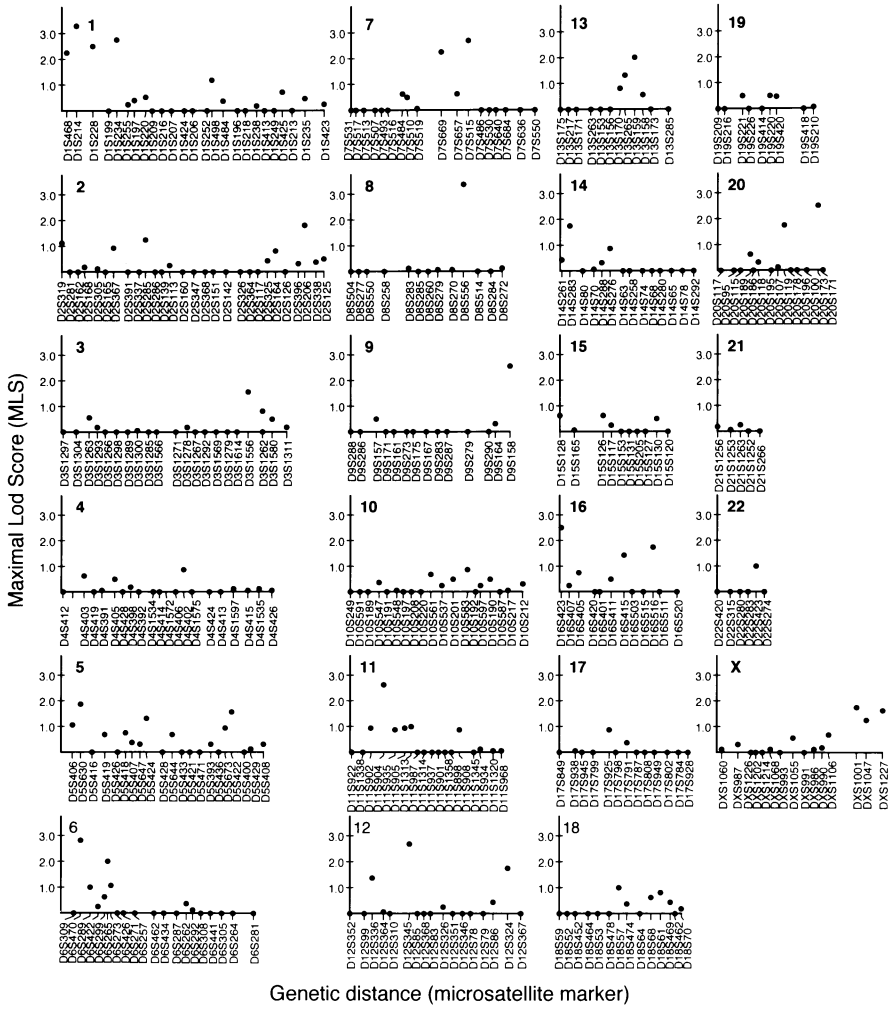

Fig. 2. MLS of the genome-wide microsatellite marker study is shown in relation to the genetic distance expressed using microsatellite marker loci.

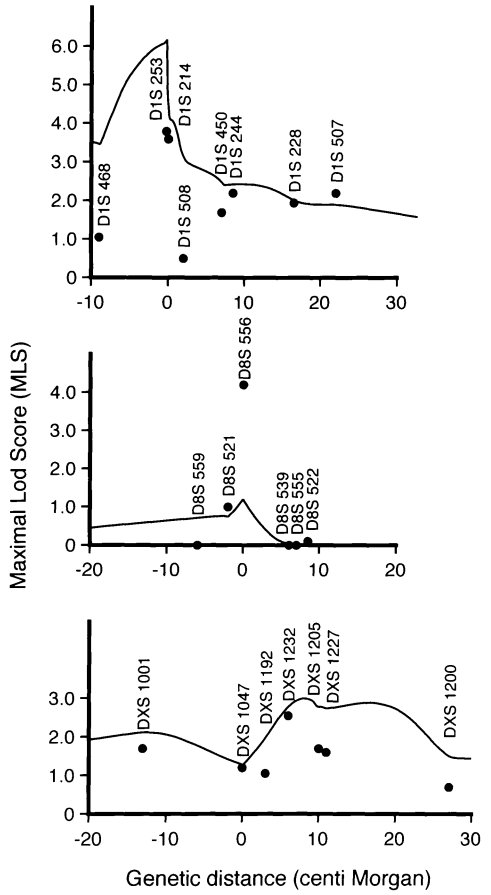

Fig. 3. Detailed MLS mapping of the gene loci including D1S214, D8S556, and the region between DXS1047 and DXS1227. Dots indicate the MLS values from single-point analysis, whereas curves indicate those from multi-point analysis.

Microsatellite markers were amplified by PCR using fluorescence-tagged primers and sized based on the difference of CA repeats on DNA fragments.

Two principal regions of linkage, D1S214 and D8S556, that exceeded the MLS of 3.0 for autosomal chromosomes were identified (Fig. 2): The MLS for D1S214 and D8S556 was 3.27 and 3.33. Estimated proportions of siblings sharing 0, 1 and 2 alleles IBD were $z_0 = 0.065$, $z_1 = 0.130$ and $z_2 = 0.805$ for D1S214 and $z_0 = 0.075$, $z_1 = 0.150$ and $z_2 = 0.775$ for D8S556. Since the linkage study is highly sensitive to the allele frequency in the population, we have typed allele frequencies of 358 marker loci of Japanese population (Hayashi *et al.*, in preparation) and this was referred to adjust the allele frequencies of unaffected siblings of rheumatoid families used for calculation in the present linkage study if necessary.

The MLS for the microsatellite markers in the vicinity of the HLA-DRB1 region was 1.98 for D6S299, 0.66 for D6S265 and 1.03 for D6S273. This was in line with the calculation of MLS based on the classical HLA-DRB1 DNA typing using PCR-RFLP. The MLS for HLA-DRB1 using this method was found to be 2.10.

Microsatellite markers in the vicinity of D1S214 and D8S556 were analyzed in detail by electrophoresis using the new primers distributed more closely around D1S214 and D8S556, and the data were analyzed by MAPMAKER/SIBS (Fig. 3). According to single-point analysis of MAPMAKER/SIBS, the MLS for D1S253 and D1S214 was 3.77 and 3.58 respectively. Multi-point analysis detected a peak of MLS of 6.13 at D1S253 (Fig. 3), where maximum likelihood allele sharing proportions were $z_0 = 0.095$, $z_1 = 0.189$ and $z_2 = 0.716$.

The MLS for D8S556 according to single-point analysis was 4.20, in which estimated proportions of siblings sharing 0, 1 and 2 alleles IBD were $z_0 = 0.094$, $z_1 = 0.188$ and $z_2 = 0.717$. The MLS according to multi-point analysis was 1.14 (Fig. 3).

Microsatellite markers in chromosome X were analyzed by MAPMAPER/SIBS. The MLS for DXS1232 was 2.35 by single-point analysis. An MLS peak of 3.03 was detected at the position 2 cM right to DXS1232 and the other of 2.93 at the region between DXS1227 and DXS1200 by multi-point analysis (Fig. 3).

Discussion

We have selected 41 families with affected sib-pairs by paying attention to the fact that affected sib-pairs expressed full-blown clinical RA with clear and identifiable joint destruction under X-ray films. This was especially important because one of the affected sib-pairs often presented with transient arthritis without identifiable joint destruction, which makes analysis difficult. We also paid attention to the fact that unaffected siblings in a family should be older enough than at least one of the affected sib-pair within the family. If an unaffected sib was younger than both of the affected sibs, such a family was excluded unless the age of the unaffected sib was older than either of the ages of the disease onset of the affected sibs. Seven out of 41 families were of this type and the youngest unaffected sibling in these families was 50 years old.

Based on detailed microsatellite marker analyses, linkage has been assigned to the chromosome regions D1S253/214, D8S556 and DXS1232. The MLS for D1S253 and D1S214 was 3.77 and 3.58 respectively by single-point analysis and the MLS for D1S253 was 6.13 by multi-point analysis of MAPMAKER/SIBS. The radiation hybrid map shows the location of D1S253 at 23.56 cR and that of D1S214 at 21.24 cR, and thus they are very close. We may therefore assign the disease gene which we call RA1 in the region of D1S253 and D1S214.

With regard to D8S556, although the MLS by single-point analysis was 4.20, this was not supported by multi-point analysis. This region may not truly be linked to a disease-associated locus; however, we think that this region must be searched further as RA2 for RA disease gene, because the method is basically sensitive to the information of the map and allele frequencies, and that increased mutation has been pointed out in an enzyme localized D8S556 in our preliminary mutation study so far.

With regards to the X chromosome, the MLS for DXS1232 was 2.35 by single-point analysis and the locus 2 cM right to DXS1232 was 3.03 by multi-point analysis. Two peaks were detected near this region; however, since the region is very close, there may be only one disease-associated locus. Being evaluated by the prescribed level 2 for X chromosome, we may assign the RA disease gene RA3 at the locus around DXS1232,

possibly within the region around DXS1232 and DXS1227 (Fig. 3).

Under the single locus and multiplicative models, the relative risks of RA for a sib of an affected proband versus population prevalence were estimated from maximum likelihood allele sharing proportions at autosomal loci to be 2.6 for D1S253/214 linked to RA1 and 2.7 for D8S556 linked to RA2 (4). The contribution of these loci, in addition to HLA-linked loci and RA3, may explain the value of 8 (5) previously estimated for the relative risk of sibs of RA.

The present study did not show significant MLS for the HLA-DR region. This was in line with the calculation of MLS based on the classical HLA-DRB1 DNA typing using PCR-RFLP. The result may be consistent with the previous studies based on genetic epidemiology that the HLA-linked susceptibility locus accounted for <20% of RA cases in the general population (3), and the ratio of the risk for siblings of patients with a disease and the population prevalence of that disease (λ_s) (4) was 8 for RA, whereas λ_s for HLA was less significant at 1.6 (5).

Acknowledgements

We thank Drs M. Itakura, The University of Tokushima, H. Tateishi, Hyogo Medical College and Y. Kawaguchi, Osaka University, for useful advice and help. We also thank Toyama Chemical Co., Ltd (Tokyo) and Hayashibara Biochemical Co., Ltd (Okayama) for help in promoting this genetic study program. An abstract of this study has been presented at the 61th National Meeting of American College of Rheumatology (*Arthritis Rheum.* 40: Suppl. S329, 1997).

Abbreviations

RA	rheumatoid arthritis
RFLP	restriction fragment length polymorphism
MLS	maximum lod score

References

1 Harris, E. D., Jr. 1990. Rheumatoid arthritis. Pathophysiology and implications for the therapy. *N. Engl. J. Med.* 322:1277.
2 Vyse, T. J. and Todd, J. A. 1996. Genetic analysis of autoimmune disease. *Cell* 85:311.
3 Lynn, A. H., Kwoh, K. C., Venglish, C. M., Aston, C. E. and Chakravarti, A. 1995. Genetic epidemiology of rheumatoid arthritis. *Am. J. Hum. Genet.* 57:150.
4 Risch, N. 1987. Assessing the role of HLA-linked and unlinked determinations of disease. *Am. J. Hum. Genet.* 40:1.
5 Wordsworth, P. 1995. Genes and arthritis. *Br. Med. Bull.* 51:249.
6 Rigby, A. S., Voelm, L. and Silman, A. J. 1993. Epistatic modeling in rheumatoid arthritis: an application of the Risch theory. *Genet. Epidemiol.* 10:311.
7 Ott, J. 1989. Computer-simulation methods in human linkage analysis. *Proc. Natl Acad. Sci. USA* 86:4175.
8 Kruglyak, L., Daly, M. J. and Lander, E. S. 1995. Rapid multipoint linkage analysis of recessive tranits in nuclear families, including homozygosity mapping. *Am. J. Hum. Genet.* 56:519.
9 Arnett, F. C., Edworthy, S. M., Bloch, D. A., McShane, D. J., Fries, J. F., Cooper, N. S., Healey, L. A., Kaplan, S. R., Liang, M. H., Luthra, H. S., Medsger, T. A., Jr, Mitchell, D. M., Neustadt, D. H., Pinals, R. S., Schaller, J. G., Sharp, J. T., Wilder, R. L. and Hunder, G. G. 1988. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. *Arthritis Rheum.* 31:315.
10 Steinbroker, O., Traeger, G. H. and Batterman, R. C. 1949. Therapeutic criteria in rheumatoid arthritis. *J. Am. Med. Ass.* 140:659.
11 Reed, P. W., Davies, J. L., Copeman, J. B., Bennett, S. T., Palmer, S. M., Pritchard, L. E., Gough, S. C. L., Kawaguchi, Y., Cordell, H. J., Balfour, K. M., Jenkins, S. C., Powell, E. E., Vignal, A. and Todd, J. A. 1994. Chromosome-specific microsatellite sets for fluorescence-based, semiautomated genome mapping. *Nature Genet.* 7:390.
12 Gyapay, G., Morissette, J., Vignal, A., Dib, C., Fizames, C., Millasseau, P., Marc, S., Bernardi, G., Lathrop, M. and Weissenbach, J. 1994. The 1993–94 Genethon human genetic linkage map. *Nature Genet.* 7:246.
13 Kaneshige, T., Hashimoto, M., Matsumoto, Y., Kinoshita, T., Hirasawa, T., Uchida, K. and Inoko, H. 1994. Serologic and nucleotide sequencing analyses of a novel DR52-associated DRB1 allele with the DR 'NJ25' specificity, designated DRB1*1307. *Hum. Immunol.* 41:151.
14 Holmans, P. and Clayton, D. 1995. Efficiency of typing unaffected relatives in an affected-sib-pair linkage study with single-locus and multiple tightly linked markers. *Am. J. Hum. Genet.* 57: 1221.
15 Kruglyak, L. and Lander, E. S. 1995. Complete multipoint sib-pair analysis of quantitative and qualitative traits. *Am. J. Hum. Genet.* 57:439.
16 Cordell, H. J., Kawaguchi, Y., Todd, J. A. and Farrall, M. 1995. An extension of the maximum lod score method to X-linked loci. *Ann. Hum. Genet.* 59:435.