Objective. Genome scans for rheumatoid arthritis (RA) have yielded inconsistent results. The absence of replication of linkage might be due to lack of power of individual studies. We performed a genome scan meta-analysis of published data to increase statistical power and to assess evidence for linkage of RA across genome scan studies.

Methods. Four RA whole-genome scans containing 767 families with 964 sibling pairs were included for the genome scan meta-analysis (GSMA). The GSMA method was applied to pool the results obtained from four genome scans. For each study, 120 genomic bins of ~30 centimorgans were defined and ranked according to maximum evidence for linkage within each bin. Bin ranks were weighted and summed across all studies. The summed rank for each bin was assessed empirically for significance using permutation methods.

Results. A total of nine bins lay above the 95% confidence level \( (P = 0.05) \) and four bins were above the 99% confidence level \( (P = 0.01) \) in the RA GSMA, suggesting that these bins contain RA-linked loci: bins 6.2, 6.4, 8.1, 18.3, 12.3, 12.2, 1.5, 6.3 and 16.2. The strongest evidence for linkage occurred on chromosome 6p22.3-6p21.1 (bin 6.2), containing the HLA region \( (P_{\text{sumrank}} = 0.0000008) \).

Conclusion. This RA GSMA confirmed the evidence for HLA loci as the greatest susceptibility factor to RA and showed evidence for linkage at non-HLA loci, such as chromosomes 1p, 6, 8p, 12, 16 and 18q, across studies. These data may provide a basis to carry out targeted linkage and candidate gene studies, particularly in the regions.

Key words: Rheumatoid arthritis, Genome scan, Linkage, Meta-analysis.

Rheumatoid arthritis (RA) is a chronic inflammatory disease predominantly involving synovial joints and affecting up to 1% of the population worldwide [1]. Although the aetiology of RA remains unsolved, a genetic component of RA susceptibility has been established by data from twin and family studies. Twin studies estimated the heritability of RA liability as up to 60% [2]. Family-based studies have also suggested an increased risk for siblings of RA probands compared with that for the general population \( (\lambda_s = 5.0–7.2) \) [3]. A consistent association between RA and human leucocyte antigen (HLA) loci has been observed in many populations [4]. The HLA class II molecules constitute the most powerful recognized genetic factor for RA. However, results of family studies suggest that this association accounts for only one-third of the genetic susceptibility and non-HLA genes are also involved in disease susceptibility [5].

Several whole-genome scans for RA have been performed to identify the HLA and non-HLA loci involved in RA susceptibility [6–11]. They identified several linkage loci for RA, but the findings of most of the studies have not been replicated. It is not surprising that linkage studies have shown inconsistent results, because they have been limited by small sample size, low statistical power and clinical or genetic heterogeneity [12].

Meta-analysis combines the linkage results from several studies, providing greater statistical power. Meta-analysis may also identify regions where the genetic effect is too small to be detected in an individual study. Recently, the genome search meta-analysis (GSMA) method has been developed to combine genome scan results in an attempt to enhance evidence of linkage [13]. The GSMA is a non-parametric ranking method to identify genomic regions that show consistent linkage evidence based on the linkage scores obtained in each scan. We applied the GSMA method to genome scans of RA to assess evidence for linkage across the studies.

Materials and methods

Selection of genome scans

Genome scans for RA were identified via literature databases. We included RA genome scans using more than 300 microsatellite markers covering the whole genome. When there were genome scans using duplicated data, we chose the study that used larger numbers of markers because genotyping more markers increases the likelihood of detecting true susceptibility loci. The GSMA method assumes a uniform map in each scan, so we did not consider the second stages of genome scans where candidate regions were more densely mapped or samples were modified. For consistency of this meta-analysis, genome scans that did not show all linkage results with \( P < 0.05 \) were excluded from this study. All loci with a \( P \)-value \( < 0.05 \) were obtained from each study for the rank-ordering procedure. The linkage scores were input as \( P \)-values in the bins and linkage scores with \( P \)-values of 0.5 were substituted as \( P \)-values of 0.5 and were considered ties.

Genome search meta-analysis

The GSMA was performed as described [13]. In brief, the autosomes were divided into 120 30-centimorgan (cM) bins defined by Genethon markers (CEPH-Genethon Integrated Map web site http://www.ceph.fr/ceph-genethon-map.htm). On the Marshfield map (available at http://www.marshfieldclinic.org/research/genetics), the average bin width was 29.1 cM. Each marker was placed within...
Table 1. Summary of genome searches included in the genome search meta-analysis

<table>
<thead>
<tr>
<th>Study characteristics</th>
<th>European study</th>
<th>US study</th>
<th>US study</th>
<th>UK study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Study population</td>
<td>France, Italy Spain, Belgium</td>
<td>USA (90.1% Caucasian)</td>
<td>USA (92.9% Caucasian)</td>
<td>UK (Caucasian)</td>
</tr>
<tr>
<td>No. of families</td>
<td>97</td>
<td>257</td>
<td>256</td>
<td>157</td>
</tr>
<tr>
<td>No. of siblings pairs</td>
<td>114</td>
<td>301</td>
<td>332</td>
<td>217</td>
</tr>
<tr>
<td>Relative weighting factor</td>
<td>0.70</td>
<td>1.13</td>
<td>1.20</td>
<td>0.97</td>
</tr>
<tr>
<td>No. of autosomal markers</td>
<td>313</td>
<td>379</td>
<td>313</td>
<td>11245</td>
</tr>
<tr>
<td>Analysis program</td>
<td>SIBPALNA</td>
<td>SIBPAL</td>
<td>SIBPAL</td>
<td>MERLIN</td>
</tr>
<tr>
<td>Test statistic output</td>
<td>( P )-value</td>
<td>( P )-value</td>
<td>( P )-value</td>
<td>( P )-value</td>
</tr>
</tbody>
</table>

Table 2. Genome search meta-analysis results, showing chromosomal bins with a summed rank within the highest 10% of all observed summed rank values in weighted analyses

<table>
<thead>
<tr>
<th>Bin</th>
<th>Begin</th>
<th>End</th>
<th>Cytogenetic location</th>
<th>( R_{\text{sum}} )</th>
<th>( P_{\text{sumrnk}} )</th>
<th>( P_{\text{ord}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.2</td>
<td>32.62</td>
<td>65.14</td>
<td>6p22.3-p21.1</td>
<td>480.0</td>
<td>0.0000008</td>
<td>0.0001000</td>
</tr>
<tr>
<td>6.4</td>
<td>99.01</td>
<td>131.07</td>
<td>6q15-q23.2</td>
<td>411.1</td>
<td>0.0007950</td>
<td>0.0037996</td>
</tr>
<tr>
<td>8.1</td>
<td>0</td>
<td>27.4</td>
<td>8pter-p22</td>
<td>378.4</td>
<td>0.0075842</td>
<td>0.0589496</td>
</tr>
<tr>
<td>18.3</td>
<td>62.84</td>
<td>90.48</td>
<td>18q12.3-q22.1</td>
<td>378.0</td>
<td>0.0077142</td>
<td>0.0105989</td>
</tr>
<tr>
<td>12.3</td>
<td>53.28</td>
<td>82.12</td>
<td>12p11.21-q15</td>
<td>347.0</td>
<td>0.0224458</td>
<td>0.1039900</td>
</tr>
<tr>
<td>12.2</td>
<td>24.45</td>
<td>53.28</td>
<td>12p12.1-p11.21</td>
<td>346.9</td>
<td>0.0227133</td>
<td>0.0337966</td>
</tr>
<tr>
<td>1.5</td>
<td>113.69</td>
<td>142.24</td>
<td>1p31.1-p13.3</td>
<td>344.1</td>
<td>0.0254791</td>
<td>0.0143986</td>
</tr>
<tr>
<td>6.3</td>
<td>65.14</td>
<td>99.01</td>
<td>6p21.1-q15</td>
<td>341.3</td>
<td>0.0289908</td>
<td>0.0092991</td>
</tr>
<tr>
<td>16.2</td>
<td>32.07</td>
<td>67.62</td>
<td>16p13-q12.2</td>
<td>328.3</td>
<td>0.0495275</td>
<td>0.0914909</td>
</tr>
<tr>
<td>2.9</td>
<td>206.74</td>
<td>233.62</td>
<td>2q34-q35</td>
<td>307.5</td>
<td>0.0998799</td>
<td>0.8566140</td>
</tr>
</tbody>
</table>

Results

Individual genome scans

Four RA genome scans were used for this GSMA based on the inclusion criteria [6–9]. Two genome scan studies of RA were excluded due to duplicated data [11] or absence of all linkage data with \( P < 0.05 \) [10]. This meta-analysis included 767 families with 964 sibling pairs. The main characteristics of genome scans included in the meta-analysis are summarized in Table 1.

Two genome scans were performed in Europe [6, 9] and the other scans were conducted in the USA [7, 8]. The number of sibling pairs ranged from 114 to 332 and individual weighting factors were 0.70, 1.13, 1.20 and 0.97, respectively. The results of four genome scans were inconsistent except for HLA loci. The strongest linkage findings were shown at HLA (6p21.3) in all of four RA genome scans. Major linkage findings except for chromosome 6 in the individual studies were 13q32-qter (D13S1315), 17q25.1 (D17S1301), 10q21.1 (D10S1221) and 1q (152 cM), respectively [6–9].

Discussion

To date, whole-genome linkage studies for RA have revealed a number of regions of the genome that are likely to harbour genes predisposing to RA [6–11]. Four linkage studies for RA included in the meta-analysis have shown the HLA region as containing the most significant linkage loci, but they have revealed inconsistent non-HLA linkage loci [6–9]. Susceptibility to RA is likely to involve several genes of weak effect, and consequently individual studies may have insufficient power to detect linkage. Here we
have conducted a meta-analysis of RA genome scans using the GSMA method in an attempt to increase statistical power and to enhance evidence of linkage.

The highest evidence for linkage was observed in bin 6.2, which includes the HLA loci, with genome-wide significance \( P_{\text{sumrnk}} = 0.0000008 \). The nominally significant \( P_{\text{sumrnk}} \) and \( P_{\text{adj}} \) of the adjacent bins, 6.3 and 6.4, provided additional evidence for linkage. This might have been because, in simulated data, bins adjacent to those containing disease loci often also achieve nominal significance [14], or it may suggest the possibility that genes on non-HLA chromosome 6 regions may also play an important role in susceptibility to RA. Previous analyses of HLA associations have shown that the majority of RA-associated HLA-DRB1 alleles share a common structural feature at positions 70–74 of the DRB1 chain. This finding has given rise to the shared epitope hypothesis [15], and there is a hypothesis suggesting a role for DQ alleles or protective effects of DRB1 alleles [16, 17]. Another candidate gene, tumour necrosis factor (TNF), located in the HLA class III region, has been reported to play a role in susceptibility to RA [18].

Besides confirming linkage of the HLA region, the GSMA found evidence for linkage at non-HLA loci, such as chromosome 1p, 6, 8p, 12, 16 and 18q, across studies. Although all genome scans of RA included in this analysis have shown the HLA region to contain the strongest linkage loci, one or two studies did not reveal significant linkage in the non-HLA loci defined by this meta-analysis. This suggests that a gene with low frequency or weak penetrance may be present in this region, and an extensive collection of RA families will be required to provide sufficient power to detect linkage. Notably, the major linkage findings at non-HLA loci in individual studies were not replicated in this meta-analysis, and several other loci, although ranked high in individual studies, did not achieve a high summed rank. This suggests that some linkage loci observed in individual genome scans may be false-positive results. Also, there was no new locus identified by this meta-analysis that was not previously identified individually in any of the four genome-wide scans.

The linkage loci shown by this meta-analysis may provide a basis for the location of RA susceptibility genes. It is interesting to examine these regions for candidate genes. A possible candidate gene on chromosome 18q21 is the TNFRSF11A gene, which encodes receptor activator of nuclear factor \( \kappa B \) (RANK). RANK is critically involved in the differentiation of osteoclasts in inflamed synovium and is important for the development of bone resorption in inflammatory arthritis [19]. A risk allele of a haematopoietic-specific protein tyrosine phosphatase, PTPN22, located at 1p13.3-1p13.1, changes the function of the protein, which functions as a negative regulator of T-cell activation, leading to T cells with a lower threshold for T-cell activation. The risk allele was significantly increased in RA patients compared with controls in Caucasians (28 vs 17%) [20] and the minor allele of this SNP recently was implicated in type 1 diabetes [21], suggesting that the variant of PTPN22 may increase susceptibility to autoimmune diseases.

Many candidate genes in other chromosome regions have been reported to be associated with susceptibility to RA, such as TNFR2 on 1p36.3-p36.2 [23], PADI4 on 1p36.13 [24], FcGR3A on 1q23 [25], IL10 on 1q32.1 [26], PARP on 1q41-1q42 [27], CTLA4 on 2q33 [28], SLC22A4 on 5q31.1 [29], IFNG on 12q21.1 [30] and MIF on 22q11.23 [31]. This might reflect the limited power of linkage studies compared with association analyses.

There were limitations to our meta-analysis. First, we looked at published genome scans for meta-analysis. Our analysis was a little different from the RA GSMA method performed by Fisher et al. [32]. They analysed genome data obtained from linkage graphs, whereas we used linkage scores with \( P < 0.05 \). Fisher et al. found significant results on chromosomes 6, 16, 12, 1q, 14q, 8p, 9q, 4q and 3q based on weighted \( P \)-values \( <0.05 \). There was consistency in the four bins 6.2, 16.2, 6.4 and 8.1 between the previous study and the present meta-analysis. The differences in results between the two studies could also have been due primarily to the use of different sets of scan analyses. Fisher et al. included two data sets that were used in this analysis [7, 8], a different data set for one scan [11] and data from one scan that we excluded because of absence of all linkage data with \( P < 0.05 \) [10].

**FIG. 1.** GSMA results of RA whole-genome linkage scans. Individual chromosomes were sub-divided into \( \sim 30 \) cM bins (represented by a dot), and bins were ranked by the significance after summing weighted data across the studies. Significance levels corresponding to 99\% \( (P_{\text{sumrnk}} < 0.01) \) and 95\% \( (P_{\text{sumrnk}} < 0.05) \) were shown by the horizontal lines.
there might be a bias in our analysis, we believe that our meta-
analysis using published data is not invalid for the assessment of evidence
for linkage of RA across genome scan studies. For example, a previous meta-analysis using published data instead of whole-genome data has shown that the analysis is more powerful than one of several genome scans in detecting significant
linkage without increasing false positives and also robust to a considerable amount of heterogeneity [33]. Secondly, although meta-analysis significantly increases the sample size, it also
may introduce heterogeneity arising from phenotypic differences between different populations and ethnic and geographic differ-
ces. Although most populations included in the meta-analysis
were Caucasian, it was possible that the Caucasian population in the USA was admixed with other populations, such as populations with African or Hispanic ancestry, and there was a difference in admixture of Caucasians between Europe and the USA. These
might confound the analysis. There is a need for the judicious
application of phenotypic stratification, including ethnicity, as
a means of reducing heterogeneity in gene mapping studies.
Thirdly, this RA GSMA did not consider the X or Y chromosome
data, so no conclusions can be reached on possible linkage on
these chromosomes.

In conclusion, this RA GSMA confirmed the evidence that the
HLA loci is the greatest susceptibility factor for RA and revealed
evidence for chromosomes 1p, 6, 8p, 12, 16, and 18q being non-
HLA susceptibility loci across studies. We believe that these
data may provide a basis for carrying out targeted linkage and
candidate gene studies in these chromosomal regions.

The authors have declared no conflicts of interest.

References

1. Harris ED Jr. Rheumatoid arthritis. Pathophysiology and implica-
2. MacGregor AJ, Snieder H, Rigby AS et al. Characterizing the
quantitative genetic contribution to rheumatoid arthritis using data
5. Deighton CM, Walker DJ, Griffiths ID, Roberts DF. The contribu-
for rheumatoid arthritis suggested by a genome-wide linkage study.
7. Jawaheer D, Seldin MF, Amos CI et al. A genomewide screen in
multiplex rheumatoid arthritis families suggests genetic overlap with
8. Jawaheer D, Seldin MF, Amos CI et al. Screening the genome for
rheumatoid arthritis susceptibility genes: a replication study and
combined analysis of 512 multicase families. Arthritis Rheum 2003;
48:906–16.
in a complex disease, using 11,245 single-nucleotide polymor-
phisms: comparison with microsatellites. Am J Hum Genet
2004;75:54–64.
10. Shiozawa S, Hayashi S, Tsukamoto Y et al. Identification of the
gene loci that predispose to rheumatoid arthritis. Int Immunol
analysis of rheumatoid arthritis susceptibility loci in 252 affected
scans of complex human diseases: true linkage is hard to find. Am J
13. Levinson DF, Levinson MD, Segurado R, Lewis CM. Genome scan
meta-analysis of schizophrenia and bipolar disorder, part I: methods
14. Cordell HJ. Sample size requirements to control for stochastic
variation in magnitude and location of allele-sharing linkage statistics
15. Gregersen PK, Silver J, Winchester RJ. The shared epitope hypo-
thesis. An approach to understanding the molecular genetics of
susceptibility to rheumatoid arthritis. Arthritis Rheum 1987;30:
1205–13.
16. Zanelli E, Gonzalez-Gay MA, David CS. Could HLA-DRB1 be the
17. Zanelli E, Krco CJ, Baisch JM, Cheng S, David CS. Immune response
of HLA-DQ8 transgenic mice to peptides from the third hypervariable
region of HLA-DRB1 correlates with predisposition to rheumatoid
variability in the tumor necrosis factor-lymphotoxin region influences
susceptibility to rheumatoid arthritis. Am J Hum Genet 1996;
59:676–83.
19. Li J, Sarosi I, Yan XQ et al. RANK is the intrinsic hematopoietic
cell surface receptor that controls osteoclastogenesis and regulation
of bone mass and calcium metabolism. Proc Natl Acad Sci USA
20. Begovich AB, Carlton VE, Honigberg LA et al. A missense single-
nucleotide polymorphism in a gene encoding a protein tyrosine
phosphatase (PTPN22) is associated with rheumatoid arthritis. Am J
lymphoid tyrosine phosphatase is associated with type I diabetes.
killer cell lectin-like receptors: common occurrence of NKG2-C
tumor necrosis factor receptor II and familial, but not sporadic,
rheumatoid arthritis: evidence for genetic heterogeneity. Arthritis
PADI4, encoding citrullinating enzyme peptidylarginine deiminase 4,
25. Kyogoku C, Tsuchiya N, Matsuta K, Tokunaga K. Studies on
the association of Fc gamma receptor IIA, IIB, IIIA and IIIB
polymorphisms with rheumatoid arthritis in the Japanese: evidence
for a genetic interaction between HLA-DRB1 and FCGR3A. Genes
Polymorphic haplotypes of the interleukin-10 5′ flanking region
determine variable interleukin-10 transcription and are associated
with particular phenotypes of juvenile rheumatoid arthritis. Arthritis
27. Pascual M, Lopez-Navot MA, Caliz R et al. A poly(ADP-ribose)
polymerase haplotype spanning the promoter region confers suscep-


