

ARTICLE

Association of haplotypes in the β -chemokine locus with multiple sclerosis

Tamara Vyshkina¹, Yin Yao Shugart², Gary Birnbaum³, Thomas P Leist⁴ and Bernadette Kalman^{*,1}

¹Department of Neurology, SLRHC, Columbia University, New York, NY, USA; ²Department of Epidemiology, Johns Hopkins University, Baltimore, MD, USA; ³Multiple Sclerosis Treatment and Research Center, Minneapolis, MN, USA; ⁴Department of Neurology, Thomas Jefferson University, Philadelphia, PA, USA

Linkage studies in multiple sclerosis (MS) identified several susceptibility loci. One of these regions includes chromosome 17q11 where a meta-analysis of data from three genome scans suggested linkage. This region encodes a cluster of genes for β -chemokines or CC chemokine ligands (CCLs), which may be involved in the development of MS lesions. Here we aimed to test if CCL alleles and haplotypes are associated with MS. Using methods of linkage and association, we observed deviations from the expected 50% transmission of haplotypes from unaffected parents to their affected children at CCL2, CCL11–CCL8–CCL13 and CCL3 within the investigated 1.85 MB chromosomal segment. Analyses of the linkage disequilibrium map support that variants with possible relevance to MS can be located within these subregions. Identification of MS associated CCL variants may have direct clinical significance, as it can lead to the design of small competitive antagonists of these molecules with beneficial effects in the treatment of patients with early and active disease.

European Journal of Human Genetics (2005) 13, 240–247. doi:10.1038/sj.ejhg.5201295
Published online 20 October 2004

Keywords: multiple sclerosis; susceptibility; β -chemokines; SNP; haplotype

Introduction

Multiple sclerosis (MS) is a complex trait disorder of unknown etiology. Its phenotypic presentations represent a spectrum with characteristics of autoimmunity and neurodegeneration.^{1–4} The involvement of autoimmunity is supported by epidemiological, clinical and pathological studies.^{4–6} A survey of genome-wide scans reveals that 65% of the positive non-HLA linkage data map into 18 overlapping clusters of MS and other autoimmune diseases.³ The risk of this nonspecific autoimmunity is estimated to be $\lambda_s = 1.65$,^{7–8} and involves a complex mechanism with chemokines playing an important role in it.⁹

The main function of chemokines and their receptors is to attract immune cells to the site of inflammation. In addition, these small molecules are involved in T-cell differentiation, apoptosis, cell cycle regulation, angiogenesis, metastatic processes and in the generation of soluble inflammatory products such as free radicals, nitric oxide, cytokines and matrix metalloproteases.⁹ In humans, approximately 50 chemokine genes have been described, which are divided into four subfamilies based on the characteristic patterns of cysteine residues close to the N-terminal end of molecules. In the CC chemokine ligand family (CCL) (also known as β -chemokines or small cytokine group A – (SCYA)), the two cysteines are adjacent to each other, while in the CXC and CX₃C chemokine families the two cysteines are separated by one or three amino acids, respectively. In the XC family, only one cysteine is present.¹⁰ Chemokine receptors bind multiple chemokines. Currently, 10 CC chemokine receptors (CCRs), six CXCRs, one CX₃CR1 and one XCR1 are known.¹⁰

*Correspondence: Dr B Kalman, Department of Neurology, SLRHC, Columbia University, 432W 58th Street, New York, NY 10019, USA. Tel: +1 212 523 8676; Fax: +1 212 523 8859; E-mail: bk2131@columbia.edu
Received 31 March 2004; revised 5 July 2004; accepted 18 August 2004

A survey of CC chemokines and their receptors by Trebst and Ransohoff¹¹ demonstrated that each of the CC chemokine receptors 1–5 (CCR1–CCR5) is present on infiltrating monocytes, macrophages and lymphocytes in MS lesions and interacts with multiple CCLs. Furthermore, several members of the CCL family (CCL2 = MCP-1, CCL3 = MIP-1 α , CCL4 = MIP-1 β , CCL5 = RANTES, CCL7 = MCP-3, CCL8 = MCP-2) are expressed by astrocytes, microglia and other inflammatory cells within plaques. These observations support the notion the CC chemokines play a role in the development of MS lesions.

In linkage studies, several MS susceptibility loci have been identified.^{8,12–16} A meta-analysis of three MS genome scans detected the highest score of nonparametric linkage (NPL) = 2.58 at 17q11.¹⁷ While numerous candidate genes are located in this region (eg NOS2A, OMG, NF1), a 1.85 MB segment at 17q11.2–q12 encodes a cluster of evolutionarily closely related β -chemokines including CCL2, CCL7, CCL11, CCL8, CCL13, CCL1, CCL5, CCL16, CCL14, CCL23, CCL18, CCL3 and CCL4, respectively.

Recently, two quantitative trait loci (QTL), *ee6* and *ee7*, were investigated by classic marker-specific linkage and interval mapping in a murine experimental allergic encephalomyelitis (EAE) model. *Eae 6* and *ee7* mapped to a region of mouse chromosome 11 syntenic to human chromosome 17q11. *Eae 6* and *ee7* were implicated in the severity and duration of the clinical presentation of EAE. However, *ee7* also conferred susceptibility to a monophasic remitting/non-relapsing (MRNR) form of the disease.¹⁸ Sequencing of cDNA for all 11 β -chemokines and NOS2 encoded within *ee7* only revealed variations in *Scya1*, *Scya2* and *Scya12* (corresponding to the human *CCL1*, *CCL2* and *CCL12* genes, respectively). A striking divergence in these *Scya1*, *Scya2* and *Scya12* variations was noted when the cDNA molecules of EAE-resistant (eg Balb/cj and B10.S/DvTe) and EAE-susceptible strains (eg SJL/J) were compared. As multiple nonsynonymous sequence polymorphisms within *Scya1* (TCA-3), *Scya2* (MCP-1) and *Scya12* (MCP-5) are now established candidates for *ee7*

(a QTL for the severity and duration, and a susceptibility locus for the MRNR subtype of EAE), we postulate that β -chemokine variants may also confer susceptibility to MS.

Patients, families and DNA specimens

DNA specimens from families with various structures were obtained from the Multiple Sclerosis DNA Bank (MSDB), University of California San Francisco, San Francisco (UCSF), CA, USA, the collection of the Canadian Multiple Sclerosis Collaborative Group (CMSCG) (Dr Ebers) and the Multiple Sclerosis Treatment and Research Center (MSTRC), Minneapolis, MN, USA (Dr Birnbaum) (Table 1). Separated data sets were received from various sources at various times and genotyped accordingly. The diagnosis of MS for patients whose DNA specimens were obtained from existing collections (MSDB, CMSCG) was made by the Poser criteria.¹⁹ Patients whose specimens were prospectively collected for this study were diagnosed using the McDonald's criteria.²⁰ The diagnosis of primary-progressive (PP)-MS was established if sustained progression of disability was observed for at least 1 year in the absence of relapses.²¹ If one or more relapses occurred superimposed on a progressive course from onset, the disease was classified as progressive-relapsing (PR)-MS.

Genotyping

For genotyping, we contracted ACGT Inc. (Northbrook, IL, USA), a biotech company using the 5' Nuclease (or TaqMan) assay for allelic discrimination of single nucleotide polymorphisms (SNPs) with high sensitivity and specificity on an ABI7900HT Instrument (ABI PE Inc., Foster City, CA, USA). PCR amplification of the template DNA sequence was performed with unlabeled forward and reverse primers in the presence of a specifically designed TaqMan probe having a nonfluorescent quencher plus a minor groove binder attached to the 3' end, and a reporter

Table 1 Families studied

Dataset	Total families	Individuals	Trio	ASP	Incomplete	Multiplex	Origin
DS101–105	66	365	7	34	10	15	MSDB, UCSF
DS106–108	66	198	6	0	0	0	MSDB, UCSF
DS109–112	50	300	0	39	0	11	CMSCG
DS113–114	58	174	42	0	16	0	MSDB, UCSF
DS115	17	48	2	0	15	0	MSTRC
Total	257	1085	117	73	41	26	

Specimens of North-American Caucasian MS patients and their families were obtained in five shipments (DS101–105, DS106–108, DS109–112, DS113–114 and DS115) from two existing DNA collections of the UCSF-based MS DNA Bank (MSDB) and the London, ON-based collection of the Canadian MS Collaborative Group (CMSCG), and from a prospective collection of the MS Treatment and Research Center (MSTRC), Minneapolis, MN, USA. DS101, DS102, ... DS115 represent subsets of specimens in separate 96-well plates. All specimens were similarly genotyped and the analyses of combined data from 257 families are presented below. ASP: affected sib pair family (two or more affected and one or more unaffected children with their unaffected parents); trio: an affected child with his/her unaffected parents; incomplete family: an affected individual with one unaffected parent and/or an unaffected sibling; multiplex family: multiple affected and unaffected family members in two or three generations.

fluorescent label (Fam or Vic) attached to the 5' end. A fully hybridized probe remained bound during strand displacement, resulting in an efficient cleavage of the reporter dye by the 5' nuclease activity of TaqGold DNA polymerase. Each reaction mixture contained two probes having a single nucleotide difference and Fam or Vic labels. The release of the reporter dyes correlated with the proportion of the matching alleles, and differed between homozygous and heterozygous states. The genotyping error rate with this method was <0.5%.

Single nucleotide polymorphisms

SNPs were identified in the NCBI database (<http://www.ncbi.nlm.nih.gov/SNP>). A total of 31 assays were successfully developed and validated. All SNPs were

genotyped in three data sets (DS101–105, DS106–108, DS109–112). In DS113–114 and DS115, data from only 27 assays were obtained, as assays of CCL2B, CCL7N, CCL11X and CCL15M were not included. Table 2 shows the list of genes and SNPs assayed, the location of markers and intermarker distances, and the heterozygosity and the minor allele frequency of SNPs calculated from data of unaffected parents in DS101–115. Heterozygosity was calculated by using the Pedigree Statistics in MERLIN (Multipoint Engine for Rapid Likelihood Inference, <http://www.sph.umich.edu/csg/abecasis/Merlin/>).²²

Analysis Files

Genotyping errors and Mendelian inconsistencies were screened by Pedcheck, and manually corrected based on

Table 2 SNPs genotyped (Contig NT010799, chromosome 17q11.2–q12)

Gene with SNP	Alternative names	NCBI SNP #	Contig location (bp)	Interval distances (bp)	Heterozygosity	Minor allele frequency
CCL2B	MCP-1	rs2857657	7316912		0.32	0.21
CCL2X	MCP-1	rs13900	7317691	779	0.38	0.29
CCL7X	MCP-3	rs3091237	7331583	13892	0.26	0.15
CCL7N	MCP-3	rs3091322	7333640	2057	0.22	0.13
*CCL11X	Eotaxin	rs3744508	7346674	13034	0.27	0.21
CCL11B	Eotaxin	rs1860184	7346912	238	0.43	0.35
CCL8X	MCP-2	rs3138036	7381324	34412	0.27	0.16
CCL8Y	MCP-2	rs138037	7381524	200	0.32	0.19
CCL13A	MCP-4	rs159313	7418169	36645	0.45	0.42
CCL13B	MCP-4	rs2072069	7418771	602	0.49	0.46
CCL1N	I-309	rs3136682	7421222	2451	0.09	0.04
CCL1A	I-309	rs2282692	7421963	741	0.41	0.33
CCL1X	I-309	rs3138032	7422822	859	0.20	0.11
CCL5B	RANTES	rs1065341	8932373	1509551	0.09	0.05
CCL5A	RANTES	rs2280789	8940783	8410	0.28	0.15
CCL16X	NCC4	rs2063979	9037346	96563	0.39	0.29
CCL16A	NCC4	rs917015	9039305	1959	0.35	0.25
CCL14Y	HCC1	rs2075746	9059312	20007	0.09	0.05
CCL15O	MIP1 δ	rs864104	9059550	238	0.31	0.20
*CCL15M	MIP1 δ	rs854625	9062241	2691	0.08	0.05
CCL15N	MIP1 δ	rs2075746	9062512	271	0.50	0.45
*CCL23M	MPIF-1	rs1003645	9074064	11552	0.30	0.19
CCL23A	MPIF-1	rs1719208	9078021	3957	0.32	0.19
CCL18B	MIP-4	rs2015052	9125793	47772	0.08	0.04
CCL18X	MIP-4	rs712042	9126660	867	0.24	0.14
CCL18A	MIP-4	rs14304	9132275	5615	0.38	0.28
CCL3B	MIP-1 α	rs1719130	9150026	17751	0.33	0.22
CCL3A	MIP-1 α	rs1719134	9150726	700	0.33	0.22
CCL4X	MIP-1 β	rs3744597	9165650	14924	0.18	0.09
*CCL4P	MIP-1 β	rs3744595	9165758	108	0.18	0.09
CCL4B	MIP-1 β	rs1719147	9165905	147	0.33	0.23

The table shows the experimental designation of SNPs in β -chemokine genes using the current nomenclature (CCL2, CCL7, etc) and alternative names (MCP-1, MCP-2, etc), and the NCBI Rs numbers, contig positions, intermarker distances, heterozygosity and minor allele frequency of SNPs. SNPs within a gene are indicated with a letter attached to the gene name (eg CCL2B and CCL2X). In this analysis, we only included nonsynonymous, untranslated region and intronic SNPs within CCL genes, but we did not include SNPs between these genes. We intended to include at least two SNPs per gene and all nonsynonymous SNPs listed in the NCBI SNP database as of November 2002, when the assay development was initiated. The four nonsynonymous SNPs (CCL11X, CCL15M, CCL23M and CCL4P) in the above list are labeled with '*'. Abbreviations for the alternative β -chemokine names: MCP-1: monocyte chemoattractant protein-1; MCP-3: monocyte chemoattractant protein-3; Eotaxin: eosinophil chemoattractant protein; MCP-2: monocyte chemoattractant protein-2; MCP-4: monocyte chemoattractant protein-4; I-309: inflammatory protein homologous to mouse TCA-3; RANTES: Regulated upon Activation, Normally T Expressed, and Presumably Secreted; NCC4: homologous to mouse macrophage inflammatory protein-1; HCC1: enhance proliferation of CD34 stem cells and share homology with the mouse MIP-1 α and MIP-1 β ; MPIF-1: myeloid progenitor inhibitory factor-1; MIP-4: macrophage inflammatory protein-4; MIP-1 α : macrophage inflammatory protein-1 α ; MIP-1 β : macrophage inflammatory protein-1 β .

scatter plots distribution and relative intensity of fluorescent signals. In case of uncertainties, genotypes were deleted. Allele and genotype frequencies were calculated for 356 unrelated parents in DS101–115. Deviation from the Hardy–Weinberg equilibrium was tested by using the χ^2 -statistics in MERLIN (<http://www.sph.umich.edu/csg/abecasis/Merlin/>).²²

Pedigree disequilibrium test and TRANSMIT

Because of the various family structures, we chose to use the pedigree disequilibrium test (PDT), developed from the classical transmission disequilibrium test (TDT), to determine if a marker locus and the hypothetical disease locus are linked or are in linkage disequilibrium (LD).^{23,24} Under Mendelian inheritance, all alleles have a 50% chance of being transmitted to offspring by the parents. An allele may be associated with the disease risk, if it is transmitted more often than 50% of the time, indicating a transmission distortion. PDT can utilize data from trios, nuclear families and discordant sibships within extended pedigrees.

Different methods have been used to estimate the required sample size (N) in the TDT/PDT, taking into consideration varying figures of the genotypic risk ratio ($\lambda_s = 1.5–4$), allele frequency ($P = 0.01–0.80$) and proportion of heterozygous parents ($0.025–0.500$), and different modes of inheritance (multiplicative, additive, recessive or dominant).^{25–27} Based on four methods of power assessment in the TDT test of one candidate SNP, a sample size of 53–55 is required for $\lambda_s = 2$ and a sample size of 192–195 is required for $\lambda_s = 1.5$ to achieve 80% power, when assuming a disease-predisposing allele frequency of 0.1 and a multiplicative model.^{26,27} Therefore we considered that $N = 257$ gives a reasonable sample size in a TDT- or PDT-based analysis of SNP variants of MS candidate genes with an expected range of $\lambda_s = 1.5–2$.⁸

Haplotype-based analyses were conducted using the TRANSMIT version 2.5 program (<http://www.gene.cimr.cam.ac.uk/clayton/software/>).²⁸ This program tests for association between genetic markers and disease by examining the transmission of markers and haplotypes from parents to affected offspring. TRANSMIT can analyze multilocus haplotypes, even if phase is unknown and parental haplotypes are missing. The tests are based on a score vector, which is averaged over all possible configurations of parental haplotypes and transmissions consistent with the observed data. Data from unaffected siblings may be used to restrict the number of possible parental genotypes to be considered. When transmission is fully observed, this test is similar to the Pearson's χ^2 -test. The program calculates the following χ^2 -statistics: (1) for each haplotype or allele, a test for excess transmission of that haplotype and (2) a global test for association ($H-1$ df, where H is the number of haplotypes for which transmission data are available). In brief, TRANSMIT v. 2.5 estimates the maximum log-likelihood of haplotype probabilities of paired marker

alleles. The haplotype-based score test evaluates the difference in the observed and expected transmission of each possible haplotype from parents to affected children by using the χ^2 -statistics. We also performed a bootstrap significance test using 10 000 bootstrap samples of haplotypes.

Assessment of linkage disequilibrium by Ldmax

Ldmax in the GOLD program provides maximum likelihood estimates of pairwise disequilibrium (<http://www.sph.umich.edu/csg/abecasis/GOLD/docs/stats.html>) by using the expectation-maximization algorithm of Slatkin and Excoffier.²⁹ SNP alleles were taken from 356 unrelated parents in DS101–115 to assess haplotype frequencies. For details of computing the δ^2 , D and D' values, we refer to the web site. The usual contingency table of χ^2 -test was also generated to calculate significance from an asymptotic distribution with $(r-1)(c-1)$ degrees of freedom, where r and c are alleles of A and B markers. The output file provides the usual χ^2 , P , δ^2 and D' values, where D' ranges between 0 and 1 (greater values indicating stronger LD).

Stratification

To test if the inclusion of patients with PP-MS caused any bias in the outcome, we removed all PP-MS nuclear families and trios from the total DS101–115 data set. The majority (over 94%) of patients in DS101–104, DS106–108, DS109–112 and DS113–114 have relapsing-remitting (RR)/secondary progressive (SP)-MS. DS105 and DS115 are exclusively composed of PP-MS trios and PP-MS incomplete families. Based on natural history studies,³⁰ patients (trios) with PR-MS are also classified here as PP-MS. We omitted seven trios with PP-MS from DS101–104, two trios with PR-MS from DS106–108 and four trios with PP- or PR-MS from DS113–114. Therefore, there were 47 PP-MS trios and incomplete families, and 217 RR/SP-MS families (trios, affected sib pair and multiplex families). PP-MS and SP-MS were both designated as chronic progressive disease in DS109–112 at the time of specimen collection and no clinical update is as yet available for phenotypic stratification. The estimated proportion of RR/SP-MS is over 90–95% in this data set.

Computation

PDT, TRANSMIT version 2.5, MERLIN and the GOLD programs were run in Unix- and Linux-based computational environments at the AMDc Bioinformatics Core Facility operated by the Columbia Genome Center, Columbia University.

Results

Allele and genotype frequencies

Genotyping of DNA was similarly performed in DS101–105, DS106–108, DS109–112, DS113–114 and DS105. SNP allele and genotype frequencies were compared among the data sets, and also calculated for the unrelated parents in

the combined DS101–115 data set. No deviations from the Hardy–Weinberg equilibrium were observed.

Markers showing allelic associations with multiple sclerosis

Analysis of data was carried out separately in each data set, and then in the combined DS101–115 set. As all data sets contain North-American families with European background and with similar allele frequency distributions, we present here the results from the combined DS101–115 families (Table 3). The analysis reveals transmission distortion for SNPs CCL2B, CCL11B, CCL5A, CCL18B, CCL18X and CCL3B (Table 3). Considering that 94% of the DS101–115 families include patients with RR/SP-MS, it is not surprising to detect similar trends of findings in the unstratified and the RR/SP-MS data sets. Although the PP-MS group is small, allelic associations are significant at CCL18B and CCL18X. However, all the significant findings

are modest and disappear after using the conservative Bonferroni correction for multiple comparisons.

Haplotypes associated with multiple sclerosis

Although no SNP listed in Table 3 has direct relevance to MS, they may define disease-associated haplotypes. We used the TRANSMIT program to evaluate haplotype transmissions (Table 4a and b). A total of 27 two-marker haplotypes generated in a pairwise manner from the list in Table 2 (but skipping subregions with a gap >30 kb) and five three-marker haplotypes in the subregions of two-marker haplotype associations were tested by TRANSMIT. We detected significant *P*-values for several two-marker haplotypes including CCL2bx, CCL11b–CCL8x, CCL8y–CCL13a, CCL13ab and CCL15om in both DS101–115 and the RR/SP-MS groups and for CCL3ba in the RR/SP-MS group. Only one three-marker haplotype, CCL15omn, showed significant distortion in the two groups. However, the number of transmissions and the variance of observed and expected

Table 3 Allelic associations as determined by PDT

	Alleles	DS101–115, χ^2	DS101–115, <i>P</i> -value	RR/SP-MS, χ^2	RR/SP-MS, <i>P</i> -value
CCL2B	C/G	4.900	0.0269	6.366	0.0116
CCL2X	C/T	0.643	0.4227	0.021	0.8841
CCL7X	C/G	1.584	0.2082	1.780	0.1821
CCL7N	A/G	0.736	0.3910	0.083	0.7731
CCL11X	A/G	0.540	0.4626	0.128	0.7208
CCL11B	A/T	7.134	0.0076	3.777	0.0520
CCL8X	A/G	0.705	0.4010	0.332	0.5642
CCL8Y	A/C	0.120	0.7293	0.229	0.6320
CCL13A	C/T	1.615	0.2038	2.661	0.1029
CCL13B	A/G	0.019	0.8913	0.024	0.8759
CCL1N	A/G	0.119	0.7299	0.701	0.4025
CCL1A	G/T	0.004	0.9469	0.086	0.7697
CCL1X	C/T	0.326	0.5680	0.835	0.3608
CCL5B	A/G	0.582	0.4455	0.129	0.7193
CCL5A	C/T	3.968	0.0464	3.234	0.0721
CCL16X	A/G	1.421	0.2333	1.245	0.2645
CCL16A	A/G	0.439	0.5077	0.447	0.5039
CCL14Y	C/T	1.506	0.2197	1.582	0.2085
CCL15O	A/G	0.003	0.9551	0.527	0.4679
CCL15M	C/T	3.712	0.0540	1.997	0.1576
CCL15N	A/G	0.804	0.3700	1.839	0.1751
CCL23M	A/G	0.137	0.7108	0.145	0.7037
CCL23A	C/T	0.164	0.6854	0.137	0.7115
CCL18B	A/C	6.132	0.0133	2.133	0.1442
CCL18X	A/G	4.195	0.0405	1.292	0.2556
CCL18A	A/G	0.353	0.5523	1.412	0.2347
CCL3B	C/T	3.201	0.0736	5.614	0.0178
CCL3A	C/T	1.460	0.2269	2.905	0.0883
CCL4X	A/C	0.345	0.5570	0.562	0.4536
CCL4P	A/G	0.000	0.9986	0.010	0.9223
CCL4B	A/G	0.618	0.4317	0.001	0.9712

PDT used 253 individual families, 243 trios and 446 discordant sib pairs in 27 assays (DS101–115), and 178 individual families, 199 trios and 411 discordant sib pairs in four assays (CCL2B, CCL7N, CCL11X, CCL15M, see single nucleotide polymorphisms section) (DS101–112). In the RR/SP-MS group, 213 individual families, 228 trios and 402 discordant sib pairs or 159 individual families, 188 trios and 387 discordant sib pairs were used, respectively. PP-MS families (42 individual families, 16 trios and 32 discordant sib pairs) did not show any association, except for CCL18B ($\chi^2 = 6.00$, *P*-value = 0.0143) and for CCL18X ($\chi^2 = 6.756$, *P*-value = 0.0093). Bonferroni correction for multiple comparisons would require a $P \leq 0.00165$ to bring back the overall α level to 0.05. Thus, after Bonferroni correction, none of the above *P*-values indicates significant deviation from the expected distribution of marker allele transmission.

Table 4 Results of two- and three-marker haplotype analysis by TRANSMIT

Gene Haplotypes of marker alleles	DS101–115					RR/SP-MS				
	Observed	Expected	Var (O–E)	χ^2 (1 df)	P-value	Observed	Expected	Var (O–E)	χ^2 (1 df)	P-value
<i>(a) Two-marker haplotype</i>										
CCL2bx										
Haplotype G C	137	153	45.38	5.68	0.0121	116	133	41.94	6.71	0.0086
CCL11b–CCL8x										
Haplotype T A	422	448	85.47	8.22	0.0039	370	390	74.00	5.62	0.0181
CCL8y–CCL13a										
Haplotype A C	237	257	70.40	5.72	0.0158	208	228	61.84	6.96	0.0065
CCL13ab										
Haplotype T G	39	27	12.29	10.26	0.0001	38	25	11.58	12.70	0
CCL15om										
Haplotype G T	0.5	2.5	1.02	4.12	0.0062	0.4	2.4	1.02	4.12	0.0005
CCL3ba										
Haplotype T C	572	558	65.56	2.89	0.0701	499	481	59.32	5.16	0.013
<i>(b) Three-marker haplotype</i>										
CCL15omn										
Haplotype G T G	0.5	2.5	0.99	4.13	0.0007	0.4	2.4	0.99	4.12	0.0011

Significant distortion of two-marker (a) and tri-marker (b) haplotype transmissions is shown in the DS101–115 and RR/SP-MS data set, when using a cutoff *P*-value ≥ 0.02 . The number of nuclear families with transmission to affected offspring was 283 in DS101–115 and 242 in the RR/SP-MS subgroup.

transmissions (<2.5) of the CCL15om and CCL15omn haplotypes are too low to make reliable conclusion.²⁸ As indicated in Table 4a and b, removal of the PP-MS families from DS101–115 did not markedly change the outcome of the analyses, and the PP-MS subgroup alone was too small to define reliably a pattern of phenotype-specific associations.

Table 5 shows the distribution of the MS-associated haplotypes in correlation with the distribution of inter-marker LD in the chr17q11.2–q12 region. Pairwise distributions of *D'* values indicate that strong LD extends from CCL2B to CCL11B, then it drops significantly and there is no LD between CCL2A and the CCL13A markers. The *D'* values generally correlate with the intermarker distances, and the highest *D'* values (0.7–1.0) can be observed up to 20 kb. As haplotypes of CCL2bx, CCL13ab, CCL15omn and CCL3ba are defined by markers in strong LD with each other, further investigation of these haplotypes may reveal disease relevant mutations. The varying strength and discontinuous nature of LD observed at CCL11b–CCL8x and CCL8y–CCL13a suggest that more than one haplotype may carry disease relevant variants and fine mapping is necessary in this subregion.

Summary of findings

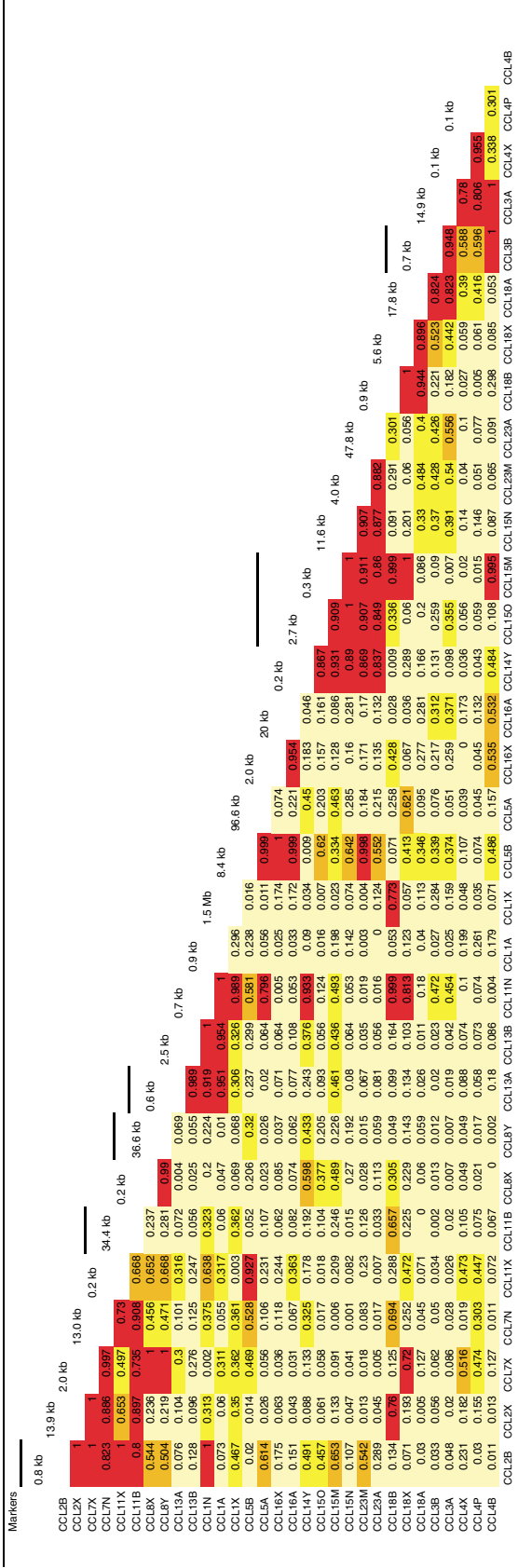
These analyses suggest that variants of CCL genes are associated with MS. Four subregions of haplotypic associations are detected within the 1.85 MB CCL coding region. We note that candidate genes showing association with all MS phenotypes overlap with those detected in the RR/SP subgroup. The inclusion of PP-MS in the combined analyses did not noticeably modify the outcome, and as a subgroup alone, did not have enough power to draw conclusions regarding differential associations.

Discussion

Linkage studies have defined a number of susceptibility loci in MS. These loci, however, still encompass 2–20 cM chromosomal segments encoding large numbers of genes. Comparing allele sharing as a method of linkage and TDT as a method of association, Risch and Merikangas²⁵ demonstrated that TDT is more powerful in detecting genes of modest effect in complex disorders. Subsequently, TDT and related methods assisted in identifying a significant number of candidate genes and alleles within regions of linkage in complex diseases.³¹ Based on these considerations, we genotyped and analyzed SNPs within candidate genes located on chromosome 17q11, a previously defined susceptibility locus of MS.¹⁷ Although none of the marker alleles showed direct association with MS after adjusting for multiple comparisons, disease-causing mutations are likely to be within or near haplotypes located in genes of CCL2 and CCL3 and in the subregion encompassing CCL11–CCL13 (Table 3–5). Because of the low number of transmitted haplotypes, the CCL15 gene region remains to be further investigated. The *D'* values summarized in Table 5 indicate that LD extends up to 20 kb between paired markers in this region. The *D'* value estimates are compatible with those obtained by simulation and experimental assessment of haplotype blocks in several chromosomal regions of Caucasians.³² The LD map (Table 5) facilitates the design of a high-density SNP mapping of disease relevant mutations in or close to the identified haplotypes.

Our results are supported by functional observations. CCL5 (RANTES) and CCL3 (MIP-1 α) are likely to be involved in the recruitment of mononuclear cells (MNCs) into the central nervous system (CNS) and in the

Table 5 Distribution of intermarker *D'* values



Red color designates $D' > 0.70$; orange designates $D' < 0.70$ and $D' > 0.50$; dark yellow designates $D' < 0.50$ and $D' > 0.30$; light yellow designates $D' < 0.30$. Intermarker distances are indicated above the LD map. Horizontal bars above distance estimates indicate regions where associations with MS were detected by TRANSMIT.

pathogenesis of both MS and EAE. Blockade of CCL3 prevents the development of both acute and relapsing forms of EAE, and the immigration of MNCs into the CNS.³³ CCR1 and CCR5, the main receptors for these chemokines, are expressed on MNCs in blood, cerebrospinal fluid (CSF) and plaques. Evidence supports that monocytes competent to enter the CNS are derived from a minor pool of CCR1 + /CCR5 + MNCs in the peripheral circulation. In the presence of their CCL3 and CCL5 ligands, these cells will be retained in the CNS.³³

CCL2 (MCP-1) and its receptor CCR2 have also been implicated in the pathogenesis of both MS and EAE. A review of available data in blood, CSF and CNS of MS patients and the results of descriptive, transgenic, knock-out and neutralizing antibody studies in EAE suggest that CCL2 and CCR2 play important roles in the development of inflammatory lesions in the CNS.^{9,34–35} MCP chemokines are generally chemoattractants for monocytes. Among them, CCL2 is also a chemoattractant for memory T cells, dendritic cells, natural killer cells and microglia. The expression of CCL2 can be induced in various cell types, including glial cells in the CNS.⁹ Consistent with the above observations concerning CCL2 and CCL3 in MS, our SNP haplotype analyses revealed that variants within these genes may be associated with the disease.

In conclusion, this study reveals that haplotypes within genes of CCL2, CCL11–CCL8–CCL13 and CCL3 are associated with MS. Further investigation of these haplotypes with dense SNP markers in families and in case–control cohorts of various ethnic origins will help to determine the variants conferring the highest risks to the development of the disease. If CCL variants produced by activated residential immune cells play an important role in the recruitment, retention and further activation of peripheral immune cells in the CNS, then genetically engineered small molecules with chemokine receptor antagonist properties may beneficially influence the natural history of the disease with autoimmune features and overt inflammatory activity.

Acknowledgements

We are very grateful to Dr George Ebers for the DNA specimens generously donated from the Collection of the Canadian MS Collaborative Group, London, ON, and to Dr Jorge Oksenberg for the DNA specimens obtained from the Multiple Sclerosis DNA Bank, UCSF, San Francisco. The assistance of Holly Armstrong, Robin Lincoln and Laura Walz in the collection of specimens is greatly appreciated. This study was supported by the RG3334A4/T research grant from the National Multiple Sclerosis Society, USA and by a grant from the Wadsworth Foundation, USA.

References

1 Kalman B, Lublin FD: Genetics of multiple sclerosis. Dossier: multiple sclerosis. *Biomed Pharmacother* 1999; 53: 358–370.

- 2 Kalman B, Albert RH, Leist TP: Genetics of multiple sclerosis: determinants of autoimmunity and neurodegeneration. *Autoimmunity* 2002; **35**: 225–234.
- 3 Becker KG, Simon RM, Bailey-Wilson JE *et al*: Clustering of non-major histocompatibility complex susceptibility candidate loci in human autoimmune diseases. *Proc Natl Acad Sci USA* 1998; **95**: 9979–9984.
- 4 Lucchinetti C, Bruck W, Parisi J, Scheithauer B, Rodriguez M, Lassmann H: Heterogeneity of multiple sclerosis lesions: implication for the pathogenesis of demyelination. *Ann Neurol* 2000; **47**: 707–717.
- 5 McCombe PA, Chalk JB, Pender MP: Familial occurrence of multiple sclerosis with thyroid disease and systemic lupus erythematosus. *J Neurol Sci* 1990; **97**: 163–167.
- 6 Reveille JD, Wilson RW, Provost T *et al*: Primary Sjogren's syndrome and other autoimmune diseases in families. *Ann Intern Med* 1984; **101**: 748–756.
- 7 Broadley SA, Deans J, Sawcer SJ *et al*: Autoimmune disease in first-degree relatives of patients with multiple sclerosis. A UK survey. *Brain* 2000; **123**: 1102–1111.
- 8 Sawcer S, Jones HB, Feakes R *et al*: A genome screen in multiple sclerosis reveals susceptibility loci on chromosome 6p21 and 17q22. *Nat Genet* 1996; **13**: 464–468.
- 9 Mahad DJ, Ransohoff RM: The role of MCP-1 (CCL2) and CCR2 in multiple sclerosis and experimental autoimmune encephalomyelitis (EAE). *Semin Immunol* 2003; **15**: 23–32.
- 10 Luttichau HR, Cark-Lewis I, Jensen PO, Moser C, Gerstoft J, Schwarz TW: A highly selective CCR2 chemokine agonist encoded by human herpesvirus 6. *J Biol Chem* 2003; **278**: 10928–10933.
- 11 Trebst C, Ransohoff RM: Investigating chemokines and chemokine receptors in patients with multiple sclerosis. *Arch Neurol* 2001; **58**: 1975–1980.
- 12 Ebers GC, Kukay K, Bulman DE *et al*: A full genome search in multiple sclerosis. *Nat Genet* 1996; **13**: 472–476.
- 13 The Multiple Sclerosis Genetics Group: A complete genomic screen for multiple sclerosis underscores a role for the major histocompatibility complex. *Nat Genet* 1996; **13**: 469–471.
- 14 Kuokkanen S, Sundvall M, Terwilinger JD *et al*: A putative vulnerability locus to multiple sclerosis maps to 5p14–p12 in a region syntenic to the murine locus Eae2. *Nat Genet* 1996; **13**: 477–480.
- 15 Kuokkanen S, Gschwend M, Rioux JD *et al*: Genomewide scan of multiple sclerosis in Finnish multiplex families. *Am J Hum Genet* 1997; **61**: 1379–1387.
- 16 Green AJ, Barcellos LF, Rimmler JB *et al*: Sequence variation in the transforming growth factor-beta1 (TFGB1) gene and multiple sclerosis susceptibility. *J Neuroimmunol* 2001; **116**: 116–124.
- 17 The Transatlantic Multiple Sclerosis Genetics Cooperative: A meta-analysis of genomic screens in multiple sclerosis. *Multiple Sclerosis* 2001; **7**: 3–11.
- 18 Teuscher C, Butterfield RJ, Ma RZ *et al*: Sequence polymorphisms in the chemokines Scya1 (TCA3), Scya2 (monocyte chemoattractant protein – MCP-1) and Scya12 (MCP-5) are candidates for eae7, a locus controlling susceptibility to monophasic remitting/non-relapsing experimental allergic encephalomyelitis. *J Immunol* 1999; **163**: 2262–2266.
- 19 Poser CM, Paty DW, Scheinberg L *et al*: New diagnostic criteria for multiple sclerosis: guidelines for research protocols. *Ann Neurol* 1983; **13**: 227–231.
- 20 McDonald IW, Compston A, Edan G *et al*: Recommended diagnostic criteria for multiple sclerosis: guidelines from the international panel on the diagnosis of multiple sclerosis. *Ann Neurol* 2001; **50**: 121–127.
- 21 Thompson AJ, Montalban X, Barkhof F *et al*: Diagnostic criteria for primary progressive multiple sclerosis: a position paper. *Ann Neurol* 2000; **47**: 831–835.
- 22 Abecasis GR, Cherny SS, Cookson WO, Cardon LR: MERLIN – rapid analysis of dense genetic maps using sparse gene flow trees. *Nat Genet* 2002; **30**: 97–101.
- 23 Spielman RS, McGinnis RE, Ewens WJ: Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM). *Am J Hum Genet* 1993; **52**: 506–516.
- 24 Martin ER, Monks SA, Warren LL, Kaplan NL: A test for linkage and association in general pedigrees: the pedigree disequilibrium test. *Am J Hum Genet* 2000; **67**: 146–154.
- 25 Risch N, Merikangas K: The future of genetic studies of complex human diseases. *Science* 1996; **273**: 1516–1517.
- 26 Knapp M: A note on power approximation for the transmission/disequilibrium test. *Am J Hum Genet* 1999; **64**: 1177–1185.
- 27 McGinnis R: General equations for P_t , P_s , and the power of the TDT and the affected-sib-pair test. *Am J Hum Genet* 2000; **67**: 1340–1347.
- 28 Clayton D: A generalization of the transmission/disequilibrium test for uncertain-haplotype transmission. *Am J Hum Genet* 1999; **65**: 1170–1177.
- 29 Slatkin M, Excoffier L: Testing for linkage disequilibrium in genotypic data using the expectation-maximization algorithm. *Heredity* 1996; **76**: 377–383.
- 30 Kremenutzky M, Cottrell D, Rice G *et al*: The natural history of multiple sclerosis: a geographically based study. 7. Progressive-relapsing and relapsing-progressive multiple sclerosis: a re-evaluation. *Brain* 1999; **122**: 1941–1949.
- 31 Croutcher PJP, Mascheretti S, Hampe J *et al*: Haplotype structure and association to Crohn's disease of CARD15 mutations in two ethnically divergent populations. *Eur J Hum Genet* 2003; **11**: 6–16.
- 32 Gabriel SB, Schaffner SF, Nguyen H *et al*: The structure of haplotype blocks in the human genome. *Science* 2002; **296**: 2225–2229.
- 33 Trebst C, Sorensen TL, Kivisakk P *et al*: CCR1+/CCR5+ mononuclear phagocytes accumulate in the central nervous system of patients with multiple sclerosis. *Am J Pathol* 2001; **159**: 1701–1710.
- 34 Mahad DJ, Howell SJL, Woodroffe MN: Expression of chemokines in the CSF and correlation with clinical disease activity in patients with multiple sclerosis. *J Neurol Neurosurg Psychiatry* 2002; **72**: 498–502.
- 35 Sorensen TL, Sellebjerg F, Jensen CV, Strieter RM, Ransohoff RM: Chemokines CXCL10 and CCL2: differential involvement in intrathecal inflammation in multiple sclerosis. *Eur J Neurol* 2001; **8**: 665–672.