Research article

An investigation of polymorphisms in the 17q11.2-12 CC chemokine gene cluster for association with multiple sclerosis in Australians

Matthew J Bugeja1, David Booth1, Bruce Bennetts2,3, Robert Heard1, Justin Rubio4 and Graeme Stewart*1

Address: 1The Institute for Immunology and Allergy Research, Westmead Millennium Institute, Westmead Campus, University of Sydney, Westmead, NSW 2145, Australia, 2Department of Molecular Genetics, The Children's Hospital at Westmead, University of Sydney, Westmead, NSW 2145, Australia, 3Discipline of Paediatrics and Child Health, University of Sydney, NSW 2006, Australia and 4The Howard Florey Institute and the Southern MS Genetics Consortium, University of Melbourne, Parkville, Victoria 3052, Australia

Email: Matthew J Bugeja - mbugeja@med.usyd.edu.au; David Booth - davidb@westgate.wh.usyd.edu.au; Bruce Bennetts - BruceB@chw.edu.au; Robert Heard - heard@idx.com.au; Justin Rubio - j.rubio@hfi.unimelb.edu.au; Graeme Stewart* - stewartg@westgate.wh.usyd.edu.au

* Corresponding author

Abstract

Background: Multiple sclerosis (MS) is a disorder of the central nervous system (CNS) characterised by inflammation and neuronal degeneration. It is believed to result from the complex interaction of a number of genes, each with modest effect. Chemokines are vital to the migration of cells to sites of inflammation, including the CNS, and many are implicated in MS pathogenesis. Most of the CC chemokine genes are encoded in a cluster on chromosome 17q11.2-12, which has been identified in a number of genome wide screens as being potentially associated with MS.

Methods: We conducted a two-stage analysis to investigate the chemokine gene cluster for association with MS. After sequencing the chemokine genes in several DNA pools to identify common polymorphisms, 12 candidate single-nucleotide polymorphisms (SNPs) were genotyped in a cohort of Australian MS trio families.

Results: Marginally significant (uncorrected) transmission distortion was identified for four of the SNPs after stratification for several factors. We also identified marginally significant (uncorrected) transmission distortion for haplotypes encompassing the $CCL2$ and $CCL11$ genes, using two independent cohorts, which was consistent with recent reports from another group.

Conclusion: Our results implicate several chemokines as possibly being associated with MS susceptibility, and given that chemokines and their receptors are suitable targets for therapeutic agents, further investigation is warranted in this region.

Background

Multiple sclerosis (MS) is the most common chronic neurological disease in young adults. It is characterised by inflammation of the central nervous system (CNS), believed to be the result of an autoimmune reaction resulting in demyelination and destruction of neural supporting cells [1]. Epidemiological studies suggest a multifactorial aetiology for MS, implicating a complex interplay...
between environmental and genetic factors [2]. In the past ten years, a large number of genome wide screens have been conducted, including the recent GAMES collaboration (Genetic Analysis of Multiple sclerosis in EuropeanS) [3]. Multiple regions of potential linkage and association with MS have been identified, suggesting that genetic predisposition to MS might result from the modest contribution of many genetic factors, which, if identified, may present important new therapeutic targets [4].

The inflammatory response that is characteristic of MS requires the targetted migration of leukocytes into the CNS, which is under the control of chemokines. Over 40 members of the human chemokine family have been identified, which act upon a variety of leukocytes via interactions with almost 20 seven-transmembrane domain chemokine receptors [5]. Chemokines are small molecules of approximately 8–10 kDa in size, and are primarily classified on the basis of the relative positioning of two conserved cysteines. In the CC chemokines, the two cysteines are adjacent, and in the CXC chemokines, a single amino acid residue separates the cysteines. Two smaller subgroups have also been identified; the CX3C chemokines, and the C chemokines.

Functionally, CC chemokines chemoattract a wide range of cells, including lymphocytes, dendritic cells, monocytes and some granulocytes [6,7], whilst CXC chemokines are chemoattractant for neutrophils and lymphocytes [6,7]. There is increasing evidence for chemokines possessing abilities beyond that of migration, including T-helper cell subset differentiation [8], T cell costimulation [9,10], and macrophage and natural killer cell maturation [11,12].

Substantial evidence supports the involvement of CC chemokines in the pathogenesis of MS. In the mouse model of MS, experimental autoimmune encephalomyelitis (EAE), knockout of CCL2 leads to resistance to disease induction [13], whilst CCL3, CCL4 and CCL5 have all been implicated in EAE development [14,15]. However, CCL3-knockout mice were found to be fully susceptible to myelin oligodendrocyte glycoprotein (MOG)-induced EAE [16]. In MS lesions, expression of CC chemokines, including CCL2, CCL3, CCL4, CCL5, CCL7 and CCL8, and their receptors, have been identified on a wide variety of cells, such as astrocytes, microglia and perivascular T cells [17-22]. In addition, altered levels of CC chemokines and receptors have been identified in the serum and cerebrospinal fluid (CSF) of MS patients; some are elevated (including CCL5), whilst CCL2 is decreased in the CSF [22-26], possibly due to removal by CCR2-positive migrating cells as they cross the blood-brain barrier [27].

Fourteen of the 28 CC chemokine genes are clustered on chromosome 17q11.2-12 (Figure 1) [28]. This cluster spans slightly less than 2 Mb, and is split into two subclusters separated by a gap of 1.5 Mb. The 17q11.2-12 region has been implicated in genome-wide screens for linkage and association with MS [29-34], and in a meta-analysis of three genome screens, the most significant nonparametric linkage score was obtained for this region [35]. The 17q region is also syntenic to an EAE quantitative trait locus on chromosome 10, which includes a chemokine gene cluster [36], and non-synonymous polymorphisms in murine CCL1, CCL2 and CCL12 were identified as candidates for the eae7 quantitative trait locus [37].

Candidate gene studies of CC chemokines in MS have not been extensive. Rather, the majority of studies have focussed on other diseases. Polymorphisms from several CC chemokine genes have been variably associated with diseases such as tuberculosis [38], HIV [39], arthritis [40], and asthma and atopy [41,42]. The CCR5Δ32 mutation has been widely studied in MS. Whilst it was found not to be a general susceptibility factor for MS [43], it may have effects on age of onset or disease progression [44,45]. More recently, several moderate single-loci and haplo-

---

Figure 1
Schematic representation of the 17q11.2-12 CC chemokine gene cluster (not to scale).
typic associations with MS were identified for single nucleotide polymorphisms (SNPs) from the CC chemokine gene cluster [46].

It was the aim of our study to conduct candidate gene analyses for the CC chemokines in order to identify polymorphisms and/or haplotypes associated with MS. We undertook a two-stage analysis. The first stage involved the scanning of the genes for published (online) and novel polymorphisms in several DNA pools by DNA sequencing. From these, 12 candidate SNPs were then individually genotyped in a cohort of MS trio families in the second stage. We also sought confirmation of our findings in an independent cohort of Australian MS families. Marginally significant (uncorrected) transmission distortion was identified for four of the SNPs, as well as for haplotypes encompassing the CCL2 and CCL11 genes.

Methods

Subjects

For the main study, all sporadic MS cases and parents (including 373 MS trio families) were recruited by our facility at the Institute for Immunology and Allergy Research, Westmead Hospital (Australia). The familial DNA pool was derived from MS probands from multicycle families obtained from the National Register of Multiple Sclerosis Families (Rex Simmons, Canberra Hospital, Australia). Control individuals were composed of local staff members and spouses of sporadic MS cases. MS patients and controls had the same ethnic composition, and were of similar average age (patients = 50yo; controls = 52.8yo). 90% of MS patients were of northern European origin, while the remainder were of southern European origin. The ratio of female to male patients was 4:1; the control ratio was 1.5:1. Approximately 60% of MS patients were HLA-DRB1*1501 positive; 73% had relapsing-remitting (RR)-MS, 20% secondary-progressive (SP)-MS, 5% primary-progressive (PP)-MS, and 2% progressive relapsing (PR)-MS. An additional cohort of 208 Australian MS trio families was obtained from the Southern MS Genetics Consortium for independent validation of our initial results. All cases were verified as having MS as the primary diagnosis. Between 100–200 ng of DNA was extracted from whole blood using a rapid salting out method [48]. Methods for the construction of the DNA pools and a single individual, included for SNP allele frequency correction purposes and for potential identification of rare SNPs. PCR products for each gene segment were amplified using the ABI 2× PCR Master Mix, with 50 ng of genomic DNA and appropriate amplification primer pairs (Sigma Genosys) to a final concentration of 5 ng/μL. PCR reactions followed a standard PCR program: 1 cycle of 95°C for 10 minutes; 35 cycles of 94°C for 30 seconds, 62°C for 30 seconds, and 72°C for 1 minute and 10 seconds; a final elongation step of 72°C for 10 minutes. PCR products were purified from the PCR cocktail by use of either the QIAQuick® PCR purification kit (QIAGEN), or ExoSAP-IT (USB Corporation). Sequencing reactions were performed using the ABI Big Dye® Terminator v3.1 Cycle Sequencing Kit. Final samples were sequenced using an ABI 3100 sequence analyser. Sequences were viewed using the ABI PRISM™ EditView software.

Genotyping in Trio Families

SNaPshot

An initial set of seven CC chemokine gene SNPs were selected for individual genotyping in a cohort of 204 MS trio families (Table 1): CCL2 -2581, CCL2 -2138, CCL11 -67, CCL5 -471, CCL15 -1284, CCL15 136+88 and CCL23 -289. These SNPs were genotyped using the SNaPshot assay, which allows the genotyping of multiple SNPs simultaneously in a single-tube multiplexed reaction, using primer-extension methodology [51] (see additional file 3: 17qSuppTable3 for SNaPshot primer details for each of the seven SNPs).

SNaPshot genotyping was performed on the pooled PCR products of each individual, amplified using either the ABI 2× PCR master mix, or the Fermentas 2× PCR master mix (Progen). SNaPshot reactions were optimised in individuals of known genotype. Between 100–200 ng of DNA was used from each individual with the appropriate amplification primer pair (Sigma Genosys) to a final concentration of 5 ng/μL. PCR products were amplified using the standard conditions, as described above. The successful amplification of PCR products was verified on 1% agarose gels. The PCR products for each individual were then

DNA sequencing primers were designed to encompass the exons of all genes analysed, as well as segments of the putative promoter region, and often part of the 3’ untranslated region (UTR) (see additional file 1: 17qSuppTable1). Where possible, primers were also designed to maximise coverage of published SNPs, which were identified from online SNP databases.

DNA sequencing of the chemokine genes was performed on each of the DNA pools and a single individual, included for SNP allele frequency correction purposes and for potential identification of rare SNPs. PCR products for each gene segment were amplified using the ABI 2× PCR Master Mix, with 50 ng of genomic DNA and appropriate amplification primer pairs (Sigma Genosys) to a final concentration of 5 ng/μL. PCR reactions followed a standard PCR program: 1 cycle of 95°C for 10 minutes; 35 cycles of 94°C for 30 seconds, 62°C for 30 seconds, and 72°C for 1 minute and 10 seconds; a final elongation step of 72°C for 10 minutes. PCR products were purified from the PCR cocktail by use of either the QIAQuick® PCR purification kit (QIAGEN), or ExoSAP-IT (USB Corporation). Sequencing reactions were performed using the ABI Big Dye® Terminator v3.1 Cycle Sequencing Kit. Final samples were sequenced using an ABI 3100 sequence analyser. Sequences were viewed using the ABI PRISM™ EditView software.

Pooled DNA sequencing

DNA was extracted from whole blood using a rapid salting out method [48]. Methods for the construction of the DNA pools has been described elsewhere [49]. Four DNA pools were constructed: 217 HLA-DRB1*1501 positive sporadic MS patients; 155 HLA-DRB1*1501 negative sporadic MS patients; 169 MS probands from multicase families; 185 unrelated, unaffected controls. HLA-DRB1*1501 status was genotyped using a SYBR Green assay (Applied Biosystems (ABI)), described elsewhere [50].
pooled, and an aliquot purified using ExoSAP-IT (USB Corporation). To this aliquot was then added 2.5 μL of SNaPshot Multiplex reagent (ABI), and pooled SNaPshot primers as defined in Supplementary Table 3, and the SNaPshot procedure was followed as per the manufacturer's instructions (ABI). Samples were then scanned using an ABI 3100. SNaPshot results were viewed using the GeneScan™ software (ABI) and genotypes determined using the GeneMapper™ software package (ABI).

**SNPlex™**

An additional five CC chemokine gene SNPs were selected for genotyping in the full cohort of 373 MS trio families (Table 1): **CCL11** -488, **CCL8** -205, **CCL16** -595, **CCL14** -649 and **CCL23** 316. These SNPs were genotyped using the SNPlex™ genotyping system (ABI), for which SUPAMAC (University of Sydney, Australia) were contracted. This system allows high-throughput genotyping of up to 48 SNPs simultaneously in a single tube using an oligonucleotide ligation assay. Initially, candidate SNP details were submitted, after which suitability for the assay was determined by running the SNPs through an algorithm. Oligonucleotides were then designed for each SNP, and applied to the DNA samples (500 ng genomic DNA). Samples were scanned using an ABI 3730. Genotyping results were determined using the GeneMapper™ software package (ABI), utilising an allelic discrimination cluster analysis to allocate genotype calls with maximum stringency.

**Additional genotyping of validation cohort**

Genotyping of the additional cohort of MS trio families was contracted to the Australian Genome Research Facility (AGRF; Brisbane, Australia). Only the **CCL2** -2138A>T and **CCL11** -488C>A SNPs were genotyped using the Sequenom Autoflex Mass Spectrometer. After submission of SNP details, PCR oligonucleotides were designed and applied to the DNA samples.

### Table 1: Summary of SNPs identified by sequencing of CC chemokine genes in DNA pools that were selected for individual genotyping. Other detected SNPs are described in Supplementary Table 2.

| Locus | rsID     | Control pool | HLA-DRB1*1501 positive pool | HLA-DRB1*1501 negative pool | Familial pool |
|-------|----------|--------------|-----------------------------|----------------------------|--------------|
| CCL2  | -2581A>G | 1024611      | 0.5                         | 0.9                        | 0.8          |
|       | -2138A>T | 1024610      | 0.3                         | 1.2                        | 1.5          |
| CCL11 | -488C>A  | 17735961     | 0.2                         | 1.2                        | 1.2          |
|       | 67G>A (A>T)* | 3744508     | 0.1                         | -                          | -            |
| CCL8  | -572C>T  | 3138035      | 0.3                         | 1.1                        | 1.2          |
|       | † 205A>C (K>Q)* | 3138038 | 0.2                         | 0.8                        | 0.8          |
| CCL5  | -417C>T  | 2107538      | 0.2                         | 1.2                        | 1.1          |
| CCL16 | -595C>A  | 854880       | 0.1                         | -                          | -            |
| CCL14 | -649T>A  | 854862       | 0.2*                        | 1.1*                       | 1.0*         |
| CCL15 | -1284A>C | 854628       | <0.1                        | -                          | Undetectable |
|       | 136+88C>T | Novel       | 0.3                         | 0.8                        | 0.6          |
| CCL23 | -289A>C  | 854655       | 0.2*                        | 0.7*                       | 0.5*         |
|       | 316T>C (M>V)+ | 1003645 | 0.2*                        | 1.0*                       | 0.7*         |

All SNPs classified relative to translation start site. Relative risks determined relative to control pool. SNPs in bold were individually genotyped in second stage.

(-) Gene encoded in anti-sense direction; SNPs reclassified to account for this

* Coding change in brackets

† The CCL8 205 SNP was originally selected for genotyping by SNPlex™, but failed the SNPlex™ algorithm, and was replaced with the CCL8 -572 SNP

* Heterozygous individual correction applied

rsID: SNP identification number
**Analyses & statistics**

Sequencing results from DNA pools and individuals were compared and scanned for published SNPs and novel polymorphisms. All SNPs were classified according to the nomenclature recommended by den Dunnen and Antonarakis [52], and are relative to the translation start site. For genes transcribed in the anti-sense direction (indicated in Table 1 with (-)), the complementary nucleotide for each allele was used, and is used throughout this manuscript. Minor allele frequencies (MAF) of SNPs were estimated by comparing the relative peak heights of the alleles. Where the individual was heterozygous for a particular SNP, the MAF in the DNA pools could be corrected for SNP-specific variation in peak-height intensity. The significance of differences between the MS patient pools and the control pool was measured by estimated relative risk (ERR) calculations.

Each SNP was analysed for transmission distortion within the trios using the transmission disequilibrium test (TDT) [53]. Analysis was performed using the GENEHUNTER program [54].

The HelixTree genetics analysis software (Golden Helix Inc, Bozeman, USA) was used to calculate linkage disequilibrium (LD) for the 12 SNPs analysed and to determine haplotypes in parents only. By using the Expectation/Maximisation (EM) algorithm, the software was able to calculate probabilities of each haplotype occurring, based on multi-locus genotypes. It calculated D’ and r² values, which are accepted measures of LD, and performed a χ² comparison for each pair of SNPs. HaploBlockFinder [55] was used to establish the haplotype block structure across the CC chemokine gene cluster.

TRANSMIT [56] was then used to analyse for haplotype transmission disequilibrium, estimating χ² values for individual haplotypes, as well as global χ² values for all haplotypes analysed. Mendelian transmission was checked using MERLIN [57].

**Results**

**Identification and analysis of SNPs in CC chemokine genes by DNA pool sequencing**

The CC chemokine genes were scanned using DNA pool sequencing in order to identify common polymorphisms in the region, including the potential identification of novel SNPs, and to establish whether associations might exist for these SNPs. In all, 50 SNPs were identified across the CC chemokine gene cluster (a list of all variants identified is available in the additional file 2: 17qSuppTable2), 48 of which were known; novel SNPs were detected in CCL8 and CCL15. Five common SNPs within exons were confirmed, four of which cause codon changes. Twenty-nine SNPs had MAF >0.15 in the control pool, designated as ‘common’ here, and for which we had statistical power to detect associations for ERR >1.5.

The greater than 90% DNA sequence homology between CCL3 and CCL3L1, and between CCL4 and CCL4L1, made it impractical to design specific primers for each of these genes. Thus, these four genes were not sequenced.

**Population genetics of individually genotyped markers**

Based on their likely functional significance (codon changing, putative promoter region, reported functional significance) and ERR from the pooled DNA sequencing, 12 SNPs were chosen for genotyping in the MS cohort, either by SNaPshot or SNPlex™ (Table 1). The CCL8 -205 SNP was originally chosen for analysis by SNPlex™, but failed to pass the SNPlex™ algorithm. Thus, it was replaced with the CCL8 -572 SNP, which is in LD with the original SNP (International HapMap Project [58]).

Genotype information retrieval for the seven SNPs genotyped using SNaPshot was 100%. Genotype information available for single-locus analysis for the five SNPs genotyped by SNPlex™ ranged between 84–87%; losses due to failed genotyping, and the exclusion of genotypes called with low stringency. The SNPlex™ genotyping, which is based on a highly multiplex PCR, was quite sensitive to DNA quality. Unambiguous genotypes could be determined using the SNPlex™ calling algorithm, which we have used at the highest stringency. The genotype frequencies for all SNPs conformed to Hardy-Weinberg equilibrium estimates, except for the CCL11 -488 SNP in mothers. The Mendelian error rate for the seven SNPs was ≤2%. Table 2 lists the MAF of the 12 markers as determined by individual genotyping.

**Linkage disequilibrium and haplotypes**

The HelixTree genetics analysis software (Golden Helix Inc, Bozeman, USA) was used to conduct a pair-wise analysis of LD across the 12 CC chemokine gene markers, using the parents only. This software calculated both D’ and r² measures of LD, and the data are represented in Figure 2. These results were confirmed using the Haploviz software package [59]. Results suggest a clear separation of the SNPs into two haplotype blocks of moderate to strong LD, confirmed by HaploBlockFinder [55] (represented in Figure 2). The first block extended from the CCL2 -2581 SNP to the CCL11 67 SNP; a distance of 33.1 kb. The second block extended from the CCL4 -469 SNP to the CCL23 -289 SNP, but did not include the CCL15 136+88 SNP, which interestingly was not in LD with any of its neighbouring SNPs. This block extended across 30.9 kb.

The HelixTree software was then used to construct haplotypes only across the two clear haplotype blocks, in parents only (350–650 individuals; actual numbers varied...
for each SNP); and to estimate haplotype frequencies. Table 2 lists the common haplotypes (frequency >0.05) across the two haplotype blocks and their frequencies within the unaffected parents. Four common haplotypes were identified in the 5' haplotype block, with frequencies between 0.18 and 0.33. In the 3' haplotype block, three common haplotypes were identified, but the TATA haplotype was the most common by a substantial margin (frequency = 0.76).

Analysis for transmission distortion across single markers and haplotypes

Single-locus transmission disequilibrium

Analyses for transmission distortion were undertaken for the 12 chemokine gene markers in all of the trios, as well as subgroups stratified based on HLA-DRB1*1501 status, disease course and gender. Results are listed in Table 3. No significant transmission distortion was found for any of the SNPs when analysed in all individuals. Upon stratification, a trend towards excess transmission of the CCL5 -471T allele was found in the HLA-DRB1*1501 negative group (P(uncorrected) = 0.03). Interestingly, DNA sequencing had also suggested an increased ERR for the T allele in the HLA-DRB1*1501 negative pool. For the CCL14 -649T>A SNP, transmission distortion of the T allele was found in patients with RR-MS (P(uncorrected) = 0.03). In contrast, transmission distortion of the A allele was found in SP-MS patients (P(uncorrected) = 0.02). Transmission distortion in SP-MS patients was also found for the CCL23 316C allele (P(uncorrected) = 0.02). For the CCL23 -289A>C SNP, transmission distortion of the C allele was found in male MS patients (P(uncorrected) = 0.03). This distortion was accentuated in male patients with SP-MS (P(uncorrected) = 0.007), however this result was inconclusive given the small number of individuals in this group (n = 10 informative transmissions). None of these results would survive a conservative Bonferroni correction for multiple comparisons.

Haplotype TDT

Whilst single-locus analyses did not suggest that any of the SNPs were significantly associated with MS susceptibility, it is conceivable that they may define MS-associated haplotypes. Using the TRANSMIT [56] software, haplotype transmission was evaluated for common haplotypes (>5% frequency) of decreasing size from either block, as well as pairwise analyses across all twelve SNPs. Analyses were conducted only where genotyping information was available for all individuals across all 12 SNPs (n = 162.
Table 4 lists common haplotypes for which transmission distortion was identified. All of these findings were identified within the 5' haplotype block (CCL2-2581.CCL2-2138.CCL11-488.CCL11-67). Marginally significant transmission distortion was found for two four-marker haplotypes, and several three-marker haplotypes (P ≤ 0.05). Marginally significant transmission distortion was also identified for pairwise analyses for CCL2-2581.CCL2-2138 (P(uncorrected) = 0.05) and CCL2-2138.CCL11-488 haplotypes (P(uncorrected) = 0.04). No significant transmission distortion was found for haplotypes across the 3' haplotype block.

Verification of haplotype results in an independent cohort
Independent validation of the results discussed above was sought in an independent cohort of 208 Australian MS trio families obtained from the Southern MS Genetics Consortium. It was determined that genotyping the CCL2-2138A>T and CCL11-488C>A SNPs was sufficient for information extraction across the four markers within the 5' haplotype block.

Genotype and minor allele frequencies for both SNPs were equivalent to those obtained for our cohort, and were in Hardy-Weinberg equilibrium. LD properties between the two markers were also in agreement with our results. No evidence for single-locus transmission distortion was identified for either SNP, supporting our original findings. Full haplotype transmission data could be obtained for 169 trios; losses were due to genotyping failure for either SNP. Two-marker haplotype analysis revealed slight transmission distortion for the A.C and T.C haplotypes in the same direction as our original findings, but the distortion was not significant (Table 4(b)). However, combining the total datasets of the two cohorts restored the original trends for transmission distortion for the A.C and T.C haplotypes (P = 0.05) (Table 4(c)).
Discussion

In this study, we have analysed the members of the CC chemokine gene cluster for association with MS. Variant chemokine expression could diminish or enhance the inflammatory response characteristic of MS. The majority of CC chemokine genes are located in a chromosomal region (17q11.2-12) that has shown suggestive linkage and association with MS in a number of genome wide screens [29-34]. Most recently, in a large genome wide linkage screen, no genome-wide significant results could be identified beyond the MHC [60]. However, suggestive linkage was identified on chromosome 17q23. An important deduction from this work was that linkage studies are under-powered to detect modest associations even in large cohorts.

We utilised a two-stage approach for this study, the first of which was sequencing of four DNA pools, three composed of MS patients and one control DNA pool across the 17q11.2-12 CC chemokine gene cluster. This approach allowed estimation of MAF for all common (MAF >0.15) SNPs across the CC chemokine genes in our Australian Caucasian cohort, and identification of two novel SNPs. This technique has been validated by our group [61]. Twelve SNPs were chosen for further analysis by individual genotyping, based on their likely functional significance (codon changing, location in the putative promoter region), prevalence (MAF >0.1), and/or uneven representation between the MS and control pools.

Overall, a large amount of genetic data has been gathered from pooled DNA sequencing and individual genotyping. The pooled sequencing provided a close estimate of MAF for SNPs across the 17q chemokine gene region, and verified the existence of common SNPs published in online databases. The individual genotyping data from 350–650 unaffected parents allowed the determination of MAF for the 12 SNPs chosen for further study, establishment of LD structure across the region, and calculation of haplotype frequencies in our unaffected Australian cohort. The MAF data facilitate power calculations for additional disease-susceptibility studies, and it and the haplotype data should prove useful in studies of population variation for these immunologically important genes. From individual genotyping, MAF were established for the 12 markers in a large unaffected cohort (up to 600 unaffected parents), whilst LD analysis revealed the haplotype block structure across the cluster, and haplotype frequencies were also established.

Measurement of transmission distortion for each of the 12 SNPs in MS trios provided some evidence for over-transmission of several of the SNPs after stratification, but these trends would not survive a conservative correction for multiple comparisons.

The CCL5 -471C>T SNP, which we found might be associated with MS in HLA-DRB1*1501 negative patients, is of potential functional relevance, as it creates a new trans-
scription factor binding site [41], and has shown associations with atopic dermatitis, atopy and asthma [41,42]. CCL5 is an important inflammatory chemokine, with a range of activities upon eosinophils [62], monocytes and activated T cells [63,64], and has been identified repeatedly in the CNS and CSF of MS patients [19,21,22]. Our data also identify CCL14 and CCL23 as warranting further investigation, with marginally significant (uncorrected) trends towards transmission distortion found in RR-MS and SP-MS patients for \( CCL14 \)-649T>A, and in males for \( CCL23 \)-289A>C. Little is known for either of these chemokines, beyond basic functionality, and neither has been studied with regard to MS pathogenesis. The \( CCL23 \)-289A>C SNP was found to affect a potential SP-1 binding site using the TRANSFAC database [65].

Upon analysing for haplotype transmission distortion across the CC chemokine gene cluster, we found suggestive evidence for association of haplotypes encompassing \( CCL2 \) and \( CCL11 \) (Table 4), which lie within a haplotype block spanning 33.1 kb. CCL2 has been implicated in MS pathogenesis. It is chemotactic for T cells and monocytes [66,67], is important in the induction of inflammation in EAE [13], has been identified in MS lesions [17-19], and in contrast to EAE studies has been associated with remission of MS [68].

Recently (and subsequent to the completion of our study), a similar study of polymorphisms across the CC chemokine gene cluster was conducted by Vyshkina et al [46]. In this study, 31 SNPs derived from online databases were selected for genotyping in a variety of individuals. The basis of selection of SNPs was not discussed. This is in contrast to our approach, in which an informed decision for SNP selection was based on several criteria, discussed above, with a particular focus on SNPs with potential functional consequences. The SNPs assayed by Vyshkina et al [46] were predominantly non-coding, with no SNPs from putative promoter regions. Fifteen SNPs overlapped with those identified in our pooled DNA sequencing. Two SNPs, both exonic, were individually genotyped in our study; the remainder did not pass the first stage. One exonic SNP was excluded as it was not common, whilst an exonic SNP in \( CCL4 \) was not analysed as this gene was excluded from our study. Similar to our study, Vyshkina et al [46] found no strong evidence for association with any

### Table 4: Haplotype TDT of CC chemokine gene cluster SNPs.

| Haplotype | Frequency Transmitted | Non-transmitted | \( \chi^2 (1 \text{df}) \) | \( P \) (uncorrected) |
|-----------|-----------------------|----------------|--------------------------|----------------------|
| (a) Local cohort | | | | |
| S’ haplotype block | | | | |
| 4-marker haplotypes | | | | |
| \( CCL2 \)-2581.G.T.C.G \| n = 162 trio families | | | | |
| \( A.A.C.G \) | 0.33 | 57.0 | 85.0 | 5.50 | 0.02 |
| \( A.T.C.G \) | 0.18 | 61.0 | 40.0 | 4.36 | 0.04 |
| 3-marker haplotypes | | | | |
| \( CCL2 \)-2581.G.T.C.G \| n = 162 trio families | | | | |
| \( A.A.C \) | 0.33 | 56.0 | 86.0 | 6.33 | 0.01 |
| \( A.T.C \) | 0.18 | 61.0 | 40.0 | 4.36 | 0.04 |
| \( CCL2 \)-2138.G.T.C.G \| n = 162 trio families | | | | |
| \( A.C.G \) | 0.62 | 70.1 | 92.0 | 2.96 | 0.09 |
| \( T.C.G \) | 0.18 | 60.9 | 40.0 | 4.34 | 0.04 |
| 2-marker haplotypes | | | | |
| \( CCL2 \)-2581.G.T.C.G \| n = 204 trio families | | | | |
| \( A.A \) | 0.52 | 89.1 | 117.1 | 3.80 | 0.05 |
| \( CCL2 \)-2138.G.T.C.G \| n = 162 trio families | | | | |
| \( A.C \) | 0.62 | 69.2 | 93.1 | 3.50 | 0.06 |
| \( T.C \) | 0.18 | 60.8 | 39.9 | 4.32 | 0.04 |
| (b) Independent cohort (n = 169 trio families) \( CCL2 \)-2138.G.T.C.G | | | | |
| \( A.C \) | 0.63 | 75.1 | 86.0 | 0.74 | 0.4 |
| \( T.C \) | 0.20 | 59.9 | 52.0 | 0.56 | 0.5 |
| (c) Combined dataset (n = 331 trio families) \( CCL2 \)-2138.G.T.C.G | | | | |
| \( A.C \) | 0.62 | 144.4 | 179.1 | 3.73 | 0.05 |
| \( T.C \) | 0.19 | 120.6 | 91.9 | 3.89 | 0.05 |
Discussion
We have conducted a two-stage analysis of polymorphisms across the CC chemokine gene cluster. We identified novel SNPs in this region, and added further information to the data available on LD structure and haplotypes across the cluster. Interestingly, our best single-locus findings were for promoter SNPs (CCL5 -649T>A, CCL14 -649T>A, CCL23 -289A>C), which might affect the relative expression of these chemokines and lead to downstream effects on leukocyte migration to the CNS, and therefore influence MS pathogenesis. We also confirmed the potential presence of a haplotypic association across CCL2 and CCL11. Further validation of the association of these SNPs in independent cohorts, and confirmation of their functional significance would support therapeutic targeting of these chemokines and their receptors.

Competing interests
The author(s) declare that they have no competing interests.

Authors’ contributions
MJB was responsible for the conception of the project, acquisition and analysis of the data, interpretation of results and writing of the manuscript. DB, BB and GS were responsible for the conception and design of the project, intellectual input, interpretation of results, and helped to draft the manuscript. RH was responsible for the clinical aspects of this project, including the assessment of MS patients. All authors read and approved the final manuscript.

Additional material

Additional File 1
Primers for the amplification of CC chemokine genes. This table presents the list of all primers used for pooled DNA sequencing.
Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2350-7-64-S1.doc]

Additional File 2
Summary of SNPs identified and analysed by sequencing of CC chemokine genes in DNA pools. This table is an extended version of Table 1, which describes the minor allele frequency and estimated relative risk data for all SNPs identified by pooled DNA sequencing.
Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2350-7-64-S2.doc]

Acknowledgements
The authors would like to thank the members of the Southern MS Genetics Consortium: Justin Rubio, Trevor Kilpatrick, Helmut Burtkueven, Niall Tubridy, Mark Mariott, Caron Chapman, John Carey, Jo Baker, Laura Johnson, Rachel Tan, Simon Foote, Stewart Huxtable, Melanie Bahlo, Jim Stankovich, and Terry Speed, Najwa Marmash for technical assistance, Susan Adams for recruitment of MS trio patients, Rex Simmons for providing familial MS cases, Vanessa Perich, Maria Ban, Natalie Hartley and Suzy Teutsch for preparation of the DNA samples, Mark Wheeler and Ilya Henner for running samples for sequencing and SNaPshot analysis, and Mike Payne for overseeing the genotyping of our samples by SNPlex™. This work was supported by a grant from the National Health and Medical Research Council of Australia (grant no. 153990). M.B. is the recipient of an Australian Postgraduate Award Scholarship and additional support was provided by Biogen Australia.

References
1. Ewing C, Bernard CC: Insights into the aetiology and pathogenesis of multiple sclerosis. *ImmunoL Cell Biol* 1998, 76(1):47-54.
2. Dyment DA, Sadovnick AD, Ebers GC, Sadnovich AD: Genetics of multiple sclerosis. *Hum Mol Genet* 1997, 6(10):1693-1698.
3. Sawcer S, Compston A: The Genetic Analysis of Multiple sclerosis in EuropeanS: concepts and design. *J Neuroimmunol* 2003, 143(1-2):13-16.
4. Stewart GJ: Multiple sclerosis, multiple genes. *Medical Journal of Australia* 1997, 166:573-574.
5. Brown KD, Zurawski SM, Mosmann TR, Zurawski G: A family of small inducible proteins secreted by leukocytes are members of a new superfamily that includes leukocyte and fibroblast-derived inflammatory agents, growth factors, and indicators of various activation processes. *J Immunol* 1989, 142(2):679-687.
6. Baggioili M: Chemokines and leukocyte traffic. *Nature* 1998, 392(6676):565-568.
7. Gale LM, McColl SR: Chemokines: extracellular messengers for all occasions? *Bioessays* 1999, 21(1):17-28.
8. Karpus WJ, Lukacs NW, Kennedy JX, Smith WS, Hurst SD, Barrett TA: Differential CC chemokine-induced enhancement of T helper type 2 cytokine production. J Immunol 1996, 158(9):4129-4136.

9. Taub DD, Turcovich-Corralles SM, Key ML, Longo DL, Murphy WJ: Chemokines and T lymphocyte activation: I. Beta chemokines costimulate human T lymphocyte activation in vitro. J Immunol 1994, 153(6):2095-2103.

10. Molon B, Giri T, Betteila C, Lanzavecchia A, Martinez AC, Manes S, Viola A: T cell costimulation by chemokine receptors. Nat Immunol 2005, 6(5):465-471.

11. Cappello P, Cao S, Bosticardo M, De Angelis S, Novelli F, Forni G, Giovarelli M: CCL16/LEC powerfully triggers effector functions and antigen-presenting functions of macrophages and enhances T cell cytotoxicity. J Leukoc Biol 2004, 75(1):135-142.

12. Taub DD, Sayers TJ, Carter CR, Ortolano RC, Jr. Alpha and beta chemokines induce NK cell migration and enhance NK-mediated cytotoxicity. J Exp Med 2001, 193(6):713-726.

13. Huang DR, Wang J, Kivisakk P, Rollins BJ, Ransohoff RM: Absence of monocyte chemoattractant protein 1 in mice leads to decreased local macrophage recruitment and antigen-specific T helper type 1 immune response in experimental autoimmune encephalomyelitis. J Exp Med 2001, 193(6):713-726.

14. Karpus WJ, Lukacs NW, McRae BL, Strier RM, Kunkel SL, Miller SD: An important role for the chemokine macrophage inflammatory protein-1 alpha in the pathogenesis of the T cell-mediated autoimmune disease, experimental autoimmune encephalomyelitis. J Immunol 1995, 155(10):5003-5010.

15. Miyagishi R, Kikuchi S, Fukazawa T, Tashiro K: Alpha and beta chemokines induce NK cell migration and enhance NK-mediated cytolysis. J Immunol 1995, 155(10):3877-3888.

16. Miyagishi R, Kikuchi S, Fukazawa T, Tashiro K: Differential CC chemokine-induced enhancement of T cell costimulation by chemokine receptors. J Immunol 1998, 160(1):21-29.

17. McNamara C, Berman JW, Brett FM, Staunton H, Farrell M, Brousseau CF: MCP-1, MCP-2 and MCP-3 expression in multiple sclerosis lesions: an immunohistochemical and in situ hybridization study. J Neuroimmunol 1993, 56(1):21-37.

18. Van Der Voorn P, Tekstra J, Beelen RH, Tensen CP, Van Der Valk P, De Groot CJ: Expression of MCP-1 by reactive astrocytes in demyelinating multiple sclerosis lesions. Am J Pathol 1999, 154(1):45-51.

19. Simpson JF, Newcombe J, Cuzner ML, Woodroffe MN: Expression of monocyte chemoattractant protein-1 and other beta- chemokines by resident glia and inflammatory cells in multiple sclerosis lesions. J Neuroimmunol 1998, 84(2):238-249.

20. Simpson J, Rezaie P, Newcombe J, Cuzner ML, Male D, Woodroffe MN: Expression of the beta-chemokine receptors CCR2, CCR3 and CCR5 in multiple sclerosis central nervous system tissue. J Neuroimmunol 2000, 108(1-2):192-200.
43. Bennetts BH, Teutsch SM, Buhrer MM, Heard RN, Stewart GJ: The CCR5 deletion mutation fails to protect against multiple sclerosis. Nature Immunol 1997, 58(1):52-59.

44. Barcellos LF, Schito AM, Rimmber JF, Vittinghoff E, Shihi A, Lincoln R, Callier S, Etinics MK, Goodkin DE, Haines JL, et al.: CC-chemokine receptor 5 polymorphism and age of onset in familial multiple sclerosis. Multiple Sclerosis Genetics Group. Immunogenetics 2000, 51(4-5):281-288.

45. Kantor R, Balkhanashvili M, Achiron A: A mutated CCR5 gene may have favorable prognostic implications in MS. Neurology 2003, 61(2):238-240.

46. Vyshkina T, Shugart YY, Birnbaum G, Leist TP, Kalman B: Association of haplotypes in the beta-chemokine locus with multiple sclerosis. Eur J Hum Genet 2005, 13(2):240-247.

47. Poser CM, Paty DW, Scheinberg L, McDonald WI, Davis FA, Ebers GC, Johnson KP, Sibley WA, Silberberg DH, Tourtellotte WW: New diagnostic criteria for multiple sclerosis: guidelines for research protocols. Ann Neurol 1983, 13(3):227-231.

48. Lahiri DK, Nurnberger JJ Jr: A rapid non-enzymatic method for the preparation of HMW DNA from blood for RFLP studies. Nucleic Acids Res 1991, 19(19):5444.

49. Teutsch SM, Booth DR, Bennetts BH, Heard RN, Stewart GJ: Identification of 1 intron and common single nucleotide polymorphisms in the interleukin-7 receptor-alpha gene and their associations with multiple sclerosis. Eur J Hum Genet 2003, 11(7):509-515.

50. Ban M, Sawcer S, Heard RN, Bennetts BH, Adams S, Booth D, Perich V, Setakis E, Compston A, Stewart GJ: A genome-wide screen for linkage disequilibrium in Australian HLA-DRB1*1501 positive multiple sclerosis patients. J Neuroimmunol 2003, 143(1-2):60-64.

51. Norton N, Williams NM, Williams HJ, Spurlock G, Kirov G, Morris DW, Hoogendoorn B, Owen MJ, O’Donovan MC: Universal, robust, highly quantitative SNP allele frequency measurement in DNA pools. Hum Genet 2002, 110(5):471-478.

52. den Dunnen JT, Antonarakis E: Nomenclature for the description of human sequence variations. Hum Genet 2001, 109:121-124.

53. 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(3):506-516.

54. Kruglyak L, Daly MJ, Reeve-Daly MP, Lander ES: Parametric and nonparametric linkage analysis: a unified multipoint approach. Am J Hum Genet 1996, 58(6):1347-1363.

55. Zhang K, Jin L: HaploBlockFinder: haplotype block analyses. Bioinformatics 2003, 19(10):1300-1301.

56. Clayton D: A generalization of the transmission/disequilibrium test for uncertain-haplotype transmission. Am J Hum Genet 1999, 65(4):1170-1177.

57. Abecasis GR, Chen X, Heckl L, Leid Jr: Merlin – rapid analysis of dense genetic maps using sparse gene flow trees. Nat Genet 2002, 30(1):97-101.

58. Altmuguer D, Brooks LD, Chakravarti A, Collins FS, Daly MJ, Donnelly P: A haplotype map of the human genome. Nature 2005, 437(7063):1299-1302.

59. Barnett JC, Fry B, Mallor J, Daly MJ: Haplotview: analysis and visualization of LD and haplotype maps. Bioinformatics 2005, 21(2):263-265.

60. Sawcer S, Ban M, Marian M, Yeo TW, Compston A, Kirby A, Daly MJ, De Jager PL, Walsh E, Lander ES, et al.: A high-density screen for linkage in multiple sclerosis. Am J Hum Genet 2005, 77(3):454-467.

61. Bugajs M, Booth DR, Bennetts BH, Heard RN, Stewart GJ: An investigation of polymorphisms in the 4q13.3-21.1 CXC-chemokine gene cluster for association with multiple sclerosis in Australians. Mult Scler 2006 in press.

62. Kameyoshi J, Dorscher A, Maltes AI, Christophers E, Schroder JM: Cytokine RANTES released by thrombin-stimulated platelets is a potent attractant for human eosinophils. J Exp Med 1992, 176(2):587-592.

63. Schall TJ: Biology of the RANTES/SIS cytokine family. Cytokine 1991, 3(3):165-183.

64. Schall TJ, Bacon K, Toy KJ, Gooddel DV: Selective attraction of monocytes and T lymphocytes of the memory phenotype by cytokine RANTES. Nature 1990, 347(6294):669-671.

65. Wingender E, Chen X, Hehl R, Karas H, Liebsch I, Matys V, Meinhardt T, Pruss M, Reuter I, Schacherer F: TRANSFAC: an integrated system for gene expression regulation. Nucleic Acids Res 2000, 28(1):31-31.

66. Taub DD, Proost P, Murphy WJ, Anver M, Longo DL, van Danne J, Oppenheim JJ: Monocyte chemotactic protein-1 (MCP-1), -2, and -3 are chemotactic for human T lymphocytes. J Clin Invest 1995, 95(3):1370-1376.

Pre-publication history

The pre-publication history for this paper can be accessed here:

http://www.biomedcentral.com/1471-2350/7/64/prepub
Author/s:
Bugeja, MJ; Booth, D; Bennetts, B; Heard, R; Rubio, J; Stewart, G

Title:
An investigation of polymorphisms in the 17q11.2-12 CC chemokine gene cluster for association with multiple sclerosis in Australians

Date:
2006-07-26

Citation:
Bugeja, M. J., Booth, D., Bennetts, B., Heard, R., Rubio, J. & Stewart, G. (2006). An investigation of polymorphisms in the 17q11.2-12 CC chemokine gene cluster for association with multiple sclerosis in Australians. BMC MEDICAL GENETICS, 7 (1), https://doi.org/10.1186/1471-2350-7-64.

Persistent Link:
http://hdl.handle.net/11343/262187

File Description:
Published version

License:
CC BY