

FULL PAPER

Intergenomic consensus in multifactorial inheritance loci: the case of multiple sclerosis

P Serrano-Fernández¹, SM Ibrahim¹, UK Zettl², H-J Thiesen¹, R Göttsche³, JT Epplen³ and S Möller¹

¹Institute of Immunology, University of Rostock, Schillingallee 70, Rostock, Germany; ²Institute of Neurology, University of Rostock, Gehlsheimer Straße 20, Rostock, Germany; ³Department of Human Genetics, University of Bochum, Bochum, Germany

Genetic linkage and association studies define chromosomal regions, quantitative trait loci (QTLs), which influence the phenotype of polygenic diseases. Here, we describe a global approach to determine intergenomic consensus of those regions in order to fine map QTLs and select particularly promising candidate genes for disease susceptibility or other polygenic traits. Exemplarily, human multiple sclerosis (MS) susceptibility regions were compared for sequence similarity with mouse and rat QTLs in its animal model experimental allergic encephalomyelitis (EAE). The number of intergenomic MS/EAE consensus genes (295) is significantly higher than expected if the animal model was unrelated to the human disease. Hence, this approach contributes to the empirical evaluation of animal models for their applicability to the study of human diseases.

Genes and Immunity (2004) 5, 615–620. doi:10.1038/sj.gene.6364134

Keywords: multiple sclerosis (MS); experimental allergic encephalomyelitis (EAE); quantitative trait locus (QTL); intergenomics; comparative genomics

Introduction

Genetic linkage analyses on individuals suffering from a polygenic disease and their relatives help to identify chromosomal regions responsible for susceptibility and for particular traits of the disease (quantitative trait loci, QTLs). Such analyses have proven successful to identify genes and pathways determining the pathogenesis.¹

The process of reducing a previously determined QTL is referred to as fine mapping. Unfortunately, classical approaches for fine-mapping are time and cost intensive and experimentally complex.² In order to save resources, other techniques are used to direct the analysis of larger QTLs, for example, global analysis of single-nucleotide polymorphisms (SNPs),³ large scale sequencing and gene expression profiling.¹ The latter point to genes that are differentially expressed in diseased animals, thus setting preferences among candidate genes within a QTL. This has been performed recently⁴ for the mouse experimental autoimmune encephalomyelitis (EAE). EAE is recognized as the best animal model for multiple sclerosis (MS), one of the most common human neurological diseases. Both are subject of the present study.

Intergenomic approaches have focused mainly on phylogeny,⁵ genetic transfer,⁶ gene functionality⁷ and the prediction of genes or regulatory regions.⁸ Here, we introduce an *in silico* approach for a genome-wide search

of loci and genes that presumably affect the phenotype of polygenic traits shared by multiple species.

QTL mapping efforts have been performed independently for several species for the same traits. If the animal models are driven by the same genetic mechanisms as those for the human diseases, we should expect to find common conserved sequences shared by QTLs and susceptibility regions of all three organisms. Genes in those shared sequences are best candidates to be relevant for the onset and/or further development of the disease. Barton *et al*⁹ evidenced this as they identified a new susceptibility region for rheumatoid arthritis supported by syntenic correspondence to QTLs of an animal model. Xu *et al*¹⁰ identified by the same principle new susceptibility regions associated with MS and syntenic to QTLs of different experimental autoimmune diseases in the rat. Here, we expand that prospective approach to a systematic retrospective comparison of the human disease with combined information from multiple animal models and different species at a genome-wide scale to fine map already known susceptibility regions, and provide an estimation of the validity of those models for the human disease.

Results

Fine mapping

QTLs involved in rat EAE extend over a total size of 200 Mbp. They were scanned for synteny with the mouse genome on the basis of a minimum of 90% sequence similarity. The EnSEMBL 'compara' database yielded 46 000 syntenic DNA segments with sizes ranging

Correspondence: Dr P Serrano-Fernández, Institute of Immunology, University of Rostock, Schillingallee 70, 18055 Rostock, Germany.
E-mail: serrano@pzzr.uni-rostock.de

Received 17 March 2004; revised and accepted 03 August 2004

between 100 bp and 150 kbp. Those segments were merged (see Materials and methods) into 29 000 sequences. Of these sequences, 11 300 were located within EAE QTLs of the mouse and extended over 5.3 Mbp.

The subsequent scan of the human genome for syntenic regions with at least 60% similarity with the aforementioned 11 300 chromosomal regions in mouse EAE QTLs yielded 10 900 smaller DNA fragments, which were merged into 4200 larger DNA sequences. A total of 1600 comprised genes that had been previously described as candidates for the human MS or were located in regions predisposing to MS. Those sequences constitute 14 major consensus regions (Table 1) distributed over five chromosomes (1, 5, 6, 17 and 18) and spanning over 1.2 Mbp (Figure 1).

In the first step, the consensuses between rodents reduce the size of QTLs associated with EAE by 97.4% for the rat (from 200 to 5.3 Mbp) and by 99.2% for the mouse (from 630 to 5.3 Mbp). These rodent consensuses are reduced again in the second step to those shared additionally by human MS susceptibility regions. Only 22.6% of the previous rodent consensus (from 5.3 to 1.2 Mbp) passes this filter. The human MS susceptibility regions considered in this study are thus reduced by 99.8% (from 790 to 1.2 Mbp) at the end of the process.

More strictly, if we consider only those regions affected by consensuses, the reduction is 98.8% for human susceptibility loci (from 100 to 1.2 Mbp).

Consensus genes

A total of 242 genes are mapped completely or partially inside the consensus regions. Additionally, we searched for further common genes by means of pair-wise gene homology (see methods). A total of 251 genes were found in this way; 198 of them are shared with the former set. A total of 44 genes are found exclusively by means of synteny search and 53 genes only by homology search. The combination of both sets of genes features a total of 295 particularly interesting candidate genes valid for all three species (Supplementary Table 1). Those genes with evidence for association with MS have been highlighted in Table 1.

The aforementioned reduction of susceptibility regions affected by consensuses can also be expressed in terms of genes. In the human, the 14 consensuses reduce the original number of candidate genes by 76.5% (from 1257 to 295 genes) with local values ranging from 16.4% (first affected susceptibility region in chromosome 17) to 89% (last two affected susceptibility loci in chromosome 17).

Table 1 Consensus regions for EAE/MS

Bands	Start-end (Mbp)	References genetic association (human)	Known susceptibility genes
<i>Chromosome 1</i>			
q21.3	150–150	<i>Genes Immun</i> 2002; 3: 279–285	
q42.13	224.9–225	<i>J Neuroimmunol</i> 2003; 143: 124–128	
<i>Chromosome 5</i>			
q31.1	133.5–134.1	<i>J Neuroimmunol</i> 2003; 143: 65–69 <i>J Neuroimmunol</i> 2003; 143: 116–119	
q33.2–34	152.9–161	<i>Genes Immun</i> 2001; 2: 205–210	IL12B ⁴¹
<i>Chromosome 6</i>			
p21.33–32	29.8–33.2	<i>J Neuroimmunol</i> 2003; 143: 60–64 <i>J Neuroimmunol</i> 2003; 143: 93–96 <i>J Neuroimmunol</i> 2003; 143: 124–128	TAP1, TAP2 ¹² HLA-DMB, DRA ¹⁸ HLA-DQA, DQB ¹⁹
p21.1–12.3	45.9–47.2	<i>J Neuroimmunol</i> 2003; 143: 60–64 <i>J Neuroimmunol</i> 2003; 143: 120–123	
<i>Chromosome 17</i>			
p13.2–13.1	6.2–7.5	<i>J Neuroimmunol</i> 2003; 143: 79–83	
p13.1–12	9.7–12.2	<i>J Neuroimmunol</i> 2003; 143: 65–69 <i>J Neuroimmunol</i> 2003; 143: 124–128	
q11.2	27.1–29.8	<i>J Neuroimmunol</i> 2003; 143: 88–92	
q21.2	41		
q21.31–32	44.1–46	<i>Genes Immun</i> 2003; 4: 559–563	
q23.2	57.7		
q23.3–24.2	60.9–67.6	<i>J Neuroimmunol</i> 2003; 143: 65–69 <i>J Neuroimmunol</i> 2003; 143: 124–128 <i>Genes Immun</i> 2003; 4: 559–563 <i>Hum Mol Genet</i> 2002; 11: 2257–2267	PRKCA ²⁵
q24.3–25.1	70.6–73.5	<i>J Neuroimmunol</i> 2003; 143: 88–92 <i>J Neuroimmunol</i> 2003; 143: 53–55	
<i>Chromosome 18</i>			
q21.2	49.9–50.7	<i>J Neuroimmunol</i> 2003; 143: 84–87	

The localization of the consensuses is indicated in the first column as cytogenetic band and in the second column physically in Mbp. The following columns indicate references for the corresponding association analysis studies in human and previously described susceptibility genes included in that region (reference to the literature list in brackets), respectively. Note: for the MHC locus in chromosome 6, only a selection of references with nonfully overlapping susceptibility regions has been included in the table.

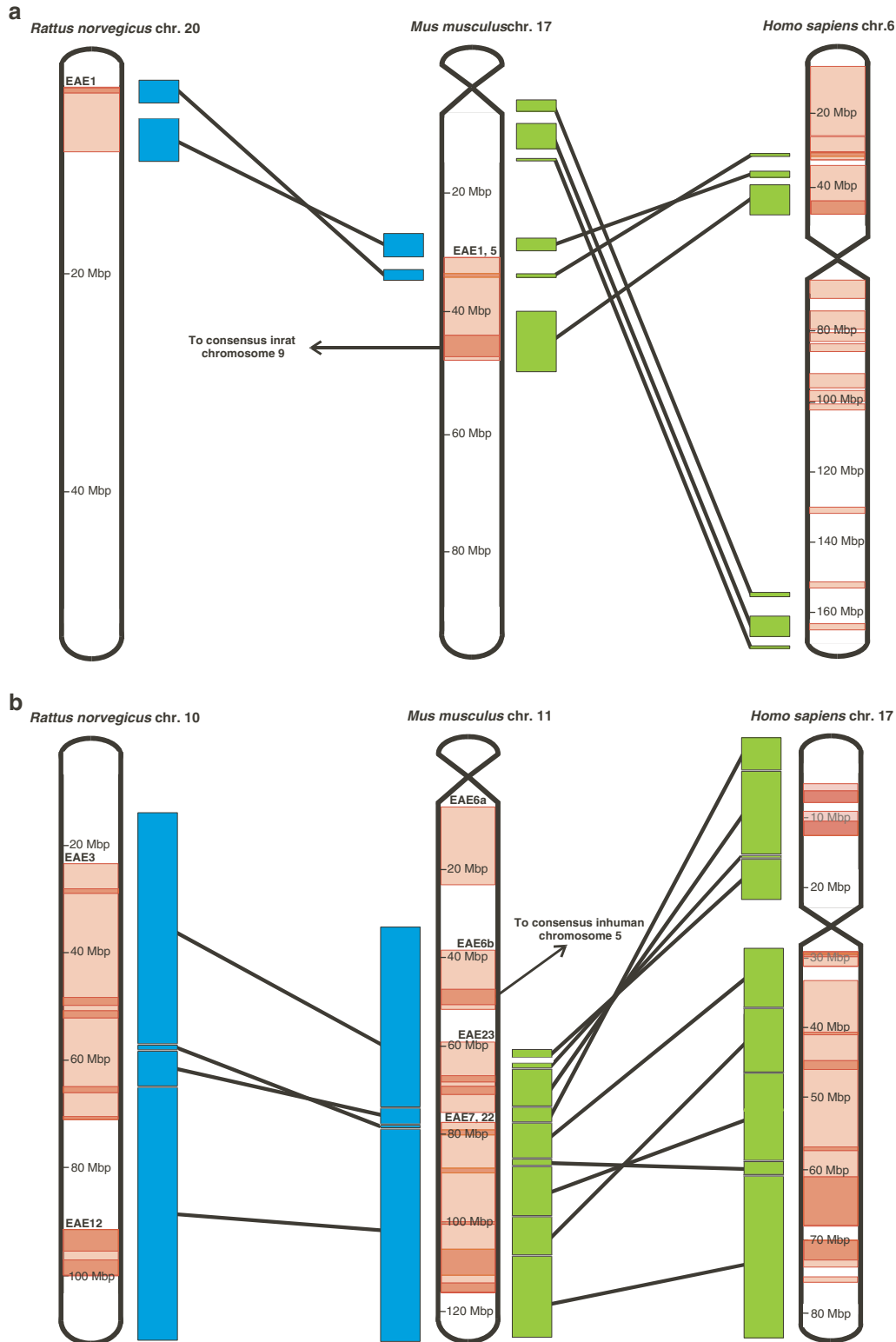


Figure 1 Syntenic regions (blue bars) between rat and mouse chromosomes, and between mouse and human chromosomes (green bars). This comparison indicates a path to reduce the original EAE QTLs and MS susceptibility loci (light red bars) to consensus regions shared by all three species (dark red bars).

Validation

A permutation test on randomized QTL locations showed that the aforementioned correspondence between EAE QTLs and MS susceptibility loci is not the

product of chance. The observed number of genes in consensus surpasses the 97th quantile of the random distribution both for the search based on synteny (242 observed genes >97% threshold: 212 genes) and for the

search based on gene homology (251 observed genes >97% threshold: 244 genes). Thus, the overlap between EAE QTLs and MS susceptibility loci is significantly greater than expected by chance (1000 permutations, $P < 0.03$).

Discussion

This *in silico* approach proved to be valuable for fine mapping large disease-associated regions. Applied to MS/EAE, susceptibility loci were reduced to smaller consensus regions shared among all analysed species. New insights are gathered for the prioritization of candidate genes.

Segregation of susceptibility loci

Different studies in human, mouse and rat have repeatedly identified the MHC locus as associated with MS/EAE. QTLs here are therefore mostly overlapping and it is difficult to segregate them into independent regions. Reading Figure 1a from right to left, two MS susceptibility loci overlap partially with the MHC locus in the shorter arm of chromosome 6 in the human and chromosome 17 of the mouse. However, the consensus regions within are segregated to different QTLs of the rat, even in different chromosomes (9 and 20). That result confirms that, despite the close neighbourhood, the mentioned susceptibility loci harbour at least to independent regions. Similarly, the large rat QTL EAE3 in chromosome 10 (Figure 1b) is very likely to contain several independent loci. The consensus regions in this particular QTL segregate into three nonoverlapping mouse QTLs and five independent human susceptibility loci located even in different chromosomes (5 and 17).

In fact, all described consensus regions are torn apart into distinct QTLs or susceptibility loci in at least one of the three species. This strongly suggests that those consensus regions are independent from each other; each of them is probably including at least one gene influencing the disease.

MHC consensus

The linkage between the MHC locus and MS or EAE has been extensively discussed in literature¹¹ and indeed here it stands out with 81 consensus genes.

Although controversial, TAP2 is proposed by some authors as a contributor to MS susceptibility.¹² The NOTCH pathway is most probably involved in a limited remyelination process taking place during MS.¹³ The complement 4 (C4A/B) is known to be involved in the demyelination process of MS.¹⁴ A dysfunction of NEU1 gene may lead to neurological disorders.¹⁵ PSMB8 and PSMB9 play a role in the aforementioned ubiquitin-proteasome system. HSD17B8 dehydrogenates oestradiol, which has an EAE-protective influence.¹⁶ RXRA, highly related to the RXRB identified here, contributes to the inhibition of IL12 production (see below).¹⁷ Specific HLA-DMB,¹⁸ HLA-DRA,¹⁸ HLA-DQA¹⁹ and HLA-DQB (HLA-DRB)¹⁹ haplotypes are strongly linked to MS.

BTNL2 is closely related to MOG; a highly immunogenic myelin protein, which can elicit chronic EAE. MOG is also a potent candidate autoantigen in MS, since it has several properties typical for autoantigens like (a) molecular mimicry, that is, sharing of epitopes with common infectious agents,²⁰ (b) determinant spreading²¹ and (c) complement binding.²²

Non-MHC consensus

From 295 consensus genes, 214 are located outside of the MHC locus. Many of these fit known disease hypotheses.

In the consensus regions of chromosome 17 we find OMG, a potent inhibitor of neurite outgrowth *in vitro*²³ and NF1, which is associated to neurofibromatosis, a disease that appears to be linked to MS.²⁴ There is also strong evidence for PRKCA to be linked with MS.²⁵ MAP2K6 activates MEF2, which plays a decisive role in the process of neurogenesis and mediates calcium-dependent neuronal survival.²⁶ The calcium channel subunits CACNG 1, 4 and 5 may also be directly associated with the calcium-dependent neuronal survival, and are known to be additionally involved in several inherited human neurological disorders.²⁷

The ubiquitin-proteasome system has been shown to play an important role in axon degeneration.²⁸ PSMC5 and PSMD12, subunits of the proteasome (the latter is a regulatory one) and NM_022739, an ubiquitination regulatory factor, arise therefore as good candidates to mediate in that process. The expression of the ICAM2 receptor is enhanced in early brain lesions in MS,²⁹ which could be implicating ICAM2 as a candidate gene, analogously to ICAM1.¹¹ A further candidate is ACE, which is known to be upregulated in serum of MS patients³⁰ and believed to play a role in EAE inflammation.³¹ Analogously, FALZ has also been found to be upregulated in neurodegenerative diseases.³²

GAS7 and TRAF4 play an important role in neurite outgrowth³³ and neurogenesis,^{34,35} respectively. XAF1 and SSTR2 mediate in apoptosis processes.^{36,37} BLMH, SCO1 and SLC6A4 dysfunctions are associated with several neurological^{38,39} and psychiatric disorders,⁴⁰ respectively.

Among the genes in the consensus region of chromosome 5, IL12 outstands as a proinflammatory cytokine already described as relevant to MS⁴¹ and TCF7 and ITK also mediate in T-cell activity.^{42,43} RAB27 is an interesting candidate gene located in the small consensus of chromosome 18. It is involved in the Griscelli syndrome, which is characterized both by a neurological impairment and by an immunological dysfunction.⁴⁴

Conclusions

Research on the animal model EAE is based on the assumption that at least some of the pathways responsible for susceptibility and/or the pathogenesis are shared by human MS. The analysis presented here statistically tested that assumption. Chromosomal regions predisposing to or modifying the phenotype of the pathology in the animal models and the susceptibility loci of the human disease share 295 genes. That is significantly more than expected in a context where similarities would have arisen simply by chance.

The search for intergenic consensus regions within disease-associated loci offers a systematic approach to objectively test the degree of similarity between a human disease and its animal model. It is directly applicable to fine map susceptibility loci in other diseases for which animal models are available and genetic association analyses have been performed (eg rheumatoid arthritis/CIA, obesity). The potential of this *in silico* fine-mapping

technique will further increase with the inclusion of additional species in the future.

Materials and methods

Data sources

A comprehensive QTL database for rodent EAE was created with data collected from public databases and some recent genetic linkage studies on EAE. The database was complemented with human MS predisposition loci assembled from recent genetic association studies. References to the original publications are shown for each marker in a database [<http://rack2.pzr.uni-rostock.de/ctl/ctlview.php>]. Genetic positions were standardized to physical positions with the help of a map conversion tool.⁴⁵ The boundaries of the susceptibility regions were defined by the original publications or to 1 Mbp on either side of the disease-associated peak marker.

Fine mapping

A second database, the Ensembl 'compara' version 16,⁴⁶ was used, which comprises information about syntenic chromosomal regions of different species. Starting from QTL regions for EAE in the rat genome, any sequence in the mouse with a similarity above a selectable percentage of its nucleotides is regarded as a syntenic region. That similarity threshold determines the restrictiveness of the reduction process. Overlapping sequences and sequences separated by gaps smaller than 1 kbp are merged in order to avoid redundancy and to reduce complexity, respectively. In the next step, the result is checked in the same way for consensuses in the human [<http://rack2.pzr.uni-rostock.de/ctl/ctlview.php>]. The number of genes in consensus regions is determined additionally on the basis of the Ensembl pair-wise gene homology, where similarity is based on the gene products.

Statistical methods

Consensuses may occur by chance with a certain probability. In order to determine that probability, we analyse the performance of a permutation test of the intergenomics QTL synteny tool on randomly located QTLs of the same size as the original ones. With an increasing number of iterations, the calculated distribution tends to fit the real random distribution. That distribution yields the probability of finding by chance a certain number of genes in consensuses.⁴⁷ Whenever the observed number of genes in consensuses is above the 95th percentile, it can be concluded that such a consensus size is reached or surpassed by chance in only five from 100 cases (ie $P < 0.05$).

Acknowledgements

We thank Michael O Glocker, Michael Kreutzer, Gertrud Fischer, Brigitte Müller-Hilke and Claudia Voigt for their support. The developers of Ensembl are thanked for their extraordinary efforts. This work was supported by the BMBF projects FKZ 01GG9831, 01GG9841 and NBL3 (FKZ 01ZZ0108).

References

- Doerge RW. Mapping and analysis of quantitative trait loci in experimental populations. *Nat Rev Genet* 2002; **3**: 43–52.
- Rogner UC, Alvner P. Congenic mice. Cutting tools for complex immune disorders. *Nat Rev Immunol* 2003; **3**: 243–252.
- Lemon WJ, Swinton CH, Wang M, Berbari N, Wang Y, You M. Single nucleotide polymorphism (SNP) analysis of mouse pulmonary adenoma susceptibility loci 1–4 for identification of candidate genes. *J Med Genet* 2003; **40**: e36.
- Ibrahim SM, Mix E, Bottcher T et al. Gene expression profiling of the nervous system in murine experimental autoimmune encephalomyelitis. *Brain* 2001; **124**: 1897–1899.
- Kirkness EF, Bafna V, Halpern AL et al. The dog genome. Survey sequencing and comparative analysis. *Science* 2003; **301**: 1898–1903.
- Lu M, Kato M, Kakihara F. Destiny of a transgene escape from *Brassica napus* into *Brassica rapa*. *Theor Appl Genet* 2002; **105**: 78–84.
- Chien S, Reiter LT, Bier E, Gribskov M. Homophila: human disease gene cognates in *Drosophila*. *Nucl Acids Res* 2002; **30**: 149–151.
- Ureta-Vidal A, Etmüller L, Birney E. Comparative genomics: genome-wide analysis in metazoan eukaryotes. *Nat Rev Genet* 2003; **4**: 252–262.
- Barton A, Eyre S, Myerscough A et al. High resolution linkage and association mapping identifies a novel rheumatoid arthritis susceptibility locus homologous to one linked to two rat models of inflammatory arthritis. *Hum Mol Genet* 2001; **10**: 1901–1906.
- Xu C, Dai Y, Lorentzen JC, Dahlman I, Olsson T, Hillert J. Linkage analysis in multiple sclerosis of chromosomal regions syntenic to experimental autoimmune disease loci. *Eur J Hum Genet* 2001; **9**: 458–463.
- Dyment DA, Ebers GC, Sadovnick AD. Genetics of multiple sclerosis. *Lancet Neurol* 2004; **3**: 104–110.
- Moins-Teisserenc H, Semana G, Alizadeh M et al. TAP2 gene polymorphism contributes to genetic susceptibility to multiple sclerosis. *Hum Immunol* 1995; **42**: 195–202.
- John GR, Shankar SL, Shafit-Zagardo B et al. Multiple sclerosis. Re-expression of a developmental pathway that restricts oligodendrocyte maturation. *Nat Med* 2002; **8**: 1115–1121.
- Schwab C, McGeer PL. Complement activated C4d immunoreactive oligodendrocytes delineate small cortical plaques in multiple sclerosis. *Exp Neurol* 2002; **174**: 81–88.
- Bonten EJ, Arts WF, Beck M et al. Novel mutations in lysosomal neuraminidase identify functional domains and determine clinical severity in sialidosis. *Hum Molec Genet* 2000; **9**: 2715–2725.
- Matejuk A, Dwyer J, Zamora A, Vandenbark AA, Offner H. Evaluation of the effects of 17[β]-estradiol (17[β]-E2) on gene expression in experimental autoimmune encephalomyelitis using DNA microarray. *Endocrinology* 2002; **143**: 313–319.
- D'Ambrosio D, Cippitelli M, Cocciolo MG et al. Inhibition of IL-12 production by 1,25-dihydroxyvitamin D₃. Involvement of NF- κ B downregulation in transcriptional repression of the p40 gene. *J Clin Invest* 1998; **101**: 252–262.
- Bennetts BH, Teutsch SM, Buhler MM, Heard RN, Stewart GJ. HLA-DMB gene and HLA-DRA promoter region polymorphisms in Australian multiple sclerosis patients. *Hum Immunol* 1999; **60**: 886–893.
- Marrosu MG, Muntoni F, Murru MR et al. Role of predisposing and protective HLA-DQA and HLA-DQB alleles in Sardinian multiple sclerosis. *Arch Neurol* 1993; **50**: 256–260.
- Mokhtarian F, Zhang Z, Shi Y, Gonzales E, Sobel RA. Molecular mimicry between a viral peptide and a myelin oligodendrocyte glycoprotein peptide induces autoimmune demyelinating disease in mice. *J Neuroimmunol* 1999; **95**: 43–54.

- 21 McFarland HI, Lobito AA, Johnson MM *et al*. Determinant spreading associated with demyelination in a nonhuman primate model of multiple sclerosis. *J Immunol* 1999; **162**: 2384–2390.
- 22 Johns TG, Bernard CC. Binding of complement component C1q to myelin oligodendrocyte glycoprotein. a novel mechanism for regulating CNS inflammation. *Mol Immunol* 1997; **34**: 33–38.
- 23 Wang KC, Koprivica V, Kim JA *et al*. Oligodendrocyte-myelin glycoprotein is a Nogo receptor ligand that inhibits neurite outgrowth. *Nature* 2002; **417**: 941–944.
- 24 Perini P, Gallo P. The range of multiple sclerosis associated with neurofibromatosis type 1. *J Neurol Neurosurg Psychiatry* 2001; **71**: 679–681.
- 25 Barton A, Woolmore JA, Ward D *et al*. Association of protein kinase C alpha (PRKCA) gene with multiple sclerosis in a UK population. *Brain* 2004; **127**: 1717–1722.
- 26 Mao Z, Bonni A, Xia F, Nadal-Vicens M, Greenberg ME. Neuronal activity-dependent cell survival mediated by transcription factor MEF2. *Science* 1999; **286**: 785–790.
- 27 Pietrobon D. Calcium channels and channelopathies of the central nervous system. *Mol Neurobiol* 2002; **25**: 31–50.
- 28 Zhai Q, Wang J, Kim A *et al*. Involvement of the ubiquitin-proteasome system in the early stages of wallerian degeneration. *Neuron* 2003; **39**: 217–225.
- 29 Washington R, Burton J, Todd RF, Newman W, Dragovic L, Dore-Duffy P. Expression of immunologically relevant endothelial cell activation antigens on isolated central nervous system microvessels from patients with multiple sclerosis. *Ann Neurol* 1994; **35**: 89–97.
- 30 Constantinescu CS, Goodman DB, Grossman RI, Mannon LJ, Cohen JA. Serum angiotensin-converting enzyme in multiple sclerosis. *Arch Neurol* 1997; **54**: 1012–1015.
- 31 Constantinescu CS, Goodman DB, Hilliard B, Wysocka M, Cohen JA. Murine macrophages stimulated with central and peripheral nervous system myelin or purified myelin proteins release inflammatory products. *Neurosci Lett* 2000; **287**: 171–174.
- 32 Styren SD, Bowser R, Dekosky ST. Expression of fetal ALZ-50 reactive clone 1 (FAC1) in dentate gyrus following entorhinal cortex lesion. *J Comp Neurol* 1997; **386**: 555–561.
- 33 Ju Y-T, Chang ACY, She BR *et al*. Gas7: a gene expressed preferentially in growth-arrested fibroblasts and terminally differentiated Purkinje neurons affects neurite formation. *Proc Nat Acad Sci USA* 1998; **95**: 11423–11428.
- 34 Stathakis DG, Hoover KB, You Z, Bryant PJ. Human postsynaptic density-95 (PSD95): location of the gene (DLG4) and possible function in nonneural as well as in neural tissues. *Genomics* 1997; **44**: 71–82.
- 35 Regnier CH, Masson R, Kedinger V *et al*. Impaired neural tube closure, axial skeleton malformations, and tracheal ring disruption in TRAF4-deficient mice. *Proc Nat Acad Sci USA* 2002; **99**: 5585–5590.
- 36 Liston P, Fong WG, Kelly NL *et al*. Identification of XAF1 as an antagonist of XIAP anti-caspase activity. *Nat Cell Biol* 2001; **3**: 128–133.
- 37 Guillermet J, Saint-Laurent N, Rochaix P *et al*. Somatostatin receptor subtype 2 sensitizes human pancreatic cancer cells to death ligand-induced apoptosis. *Proc Nat Acad Sci USA* 2003; **100**: 155–160.
- 38 Nivet-Antoine V, Coulhon MP, Le Denmat C *et al*. Apolipoprotein E and bleomycin hydrolase. Polymorphisms: association with neurodegenerative diseases. *Ann Biol Clin* 2003; **61**: 61–67.
- 39 Valnot I, Osmond S, Gigarel N *et al*. Mutations of the SCO1 gene in mitochondrial cytochrome *c* oxidase deficiency with neonatal-onset hepatic failure and encephalopathy. *Am J Hum Genet* 2000; **67**: 1104–1109.
- 40 Dimitrova A, Georgieva L, Nikolov I *et al*. Major psychiatric disorders and the serotonin transporter gene (SLC6A4): family-based association studies. *Psychiatr Genet* 2002; **12**: 137–141.
- 41 Fillion LG, Matusevicius D, Graziani-Bowering GM, Kumar A, Freedman MS. Monocyte-derived IL12, CD86 (B7-2) and CD40L expression in relapsing and progressive multiple sclerosis. *Clin Immunol* 2003; **106**: 127–138.
- 42 Verbeek S, Izon D, Hofhuis F *et al*. An HMG-box-containing T-cell factor required for thymocyte differentiation. *Nature* 1995; **374**: 70–74.
- 43 Woods ML, Kivens WJ, Adelman MA, Qiu Y, August A, Shimizu Y. A novel function for the Tec family tyrosine kinase Itk in activation of beta-1 integrins by the T-cell receptor. *EMBO J* 2001; **20**: 1232–1244.
- 44 Menasche G, Pastural E, Feldmann J *et al*. Mutations in RAB27A cause Griscelli syndrome associated with haemophagocytic syndrome. *Nat Genet* 2000; **25**: 173–176.
- 45 Voigt C, Möller S, Ibrahim SM, Serrano-Fernández P. Non-linear conversion between genetic and physical chromosomal distances. *Bioinformatics* 2004; **20**: 1966–1967.
- 46 Clamp M, Andrews D, Barker D *et al*. Ensembl 2002: accommodating comparative genomics. *Nucleic Acids Res* 2003; **31**: 38–42.
- 47 Mooney CZ. *Monte Carlo Simulation*. Sage University Papers Series, Quantitative Applications in the Social Sciences No 116, Sage Publications: Thousand Oaks, London, 1997.

Supplementary Information accompanies the paper on Genes and Immunity website (<http://www.nature.com/gene>).