

## Genome-wide association scan identifies novel multiple sclerosis susceptibility loci on chromosomes 12 and 20

### The Australia and New Zealand Multiple Sclerosis Genetics Consortium (ANZgene)

**Study design and management committee:** Melanie Bahlo<sup>1</sup>, David R Booth<sup>2</sup>, Simon Broadley<sup>3,4</sup>, Matthew A Brown<sup>5,6</sup>, Simon J Foote<sup>7</sup>, Lyn Griffiths<sup>8</sup>, Trevor J Kilpatrick<sup>9,10,11</sup>, Jeanette Lechner-Scott<sup>12,13</sup>, Pablo Moscato<sup>13,14</sup>, Victoria M Perreau<sup>9</sup>, Justin P Rubio<sup>10</sup>, Rodney J Scott<sup>12,13,14</sup>, Jim Stankovich<sup>7</sup>, Graeme J Stewart<sup>2</sup>, Bruce V Taylor<sup>7</sup>, James Wiley<sup>15</sup> (Chair)

**Sample processing, data handling and genotyping:** Matthew A Brown<sup>5,6</sup>, David R Booth<sup>2</sup>, Glynnis Clarke<sup>16</sup>, Mathew B Cox<sup>14,13</sup>, Peter A Csurhes<sup>17</sup>, Patrick Danoy<sup>5</sup>, Karen Drysdale<sup>7</sup>, Judith Field<sup>10</sup>, Simon J Foote<sup>7</sup>, Judith M Greer<sup>17</sup>, Lyn R Griffiths<sup>8</sup>, Preethi Guru<sup>7</sup>, Johanna Hadler<sup>5</sup>, Ella Hoban<sup>7</sup>, Cathy J Jensen<sup>10</sup>, Laura J Johnson<sup>10</sup>, Ruth McCallum<sup>19</sup>, Marilyn Merriman<sup>19</sup>, Tony Merriman<sup>19</sup>, Karen Pryce<sup>5</sup>, Rodney J Scott<sup>12,13,14</sup>, Graeme J Stewart<sup>2</sup>, Lotfi Tajouri<sup>8</sup>, Ella J Wilkins<sup>10</sup>, Justin P Rubio<sup>10</sup> (Corresponding author)

**Data analysis:** Melanie Bahlo<sup>1</sup>, Matthew A Brown<sup>5,6</sup>, Brian L Browning<sup>20</sup>, Sharon R Browning<sup>20</sup>, Devindri Perera<sup>7</sup>, Justin P Rubio<sup>10</sup>, Jim Stankovich<sup>7</sup> (Chief analyst)

**Phenotyping:** Simon Broadley<sup>3,4</sup>, Helmut Butzkueven<sup>10</sup>, William M Carroll<sup>21,22</sup>, Caron Chapman<sup>23</sup>, Allan G Kermode<sup>21,22</sup>, Mark Marriott<sup>11</sup>, Deborah Mason<sup>24</sup>, Robert N Heard<sup>2</sup>, Michael P Pender<sup>18,25</sup>, Mark Slee<sup>26</sup>, Niall Tubridy<sup>27</sup>, Jeanette Lechner-Scott<sup>12,13</sup>, Bruce V Taylor<sup>7</sup>, Ernest Willoughby<sup>28</sup>, Trevor J Kilpatrick<sup>9,10,11</sup> (Chair)

### AFFILIATIONS FOR AUTHORS

(AUS=Australia, NZ=New Zealand, VIC=Victoria, NSW=New South Wales, QLD=Queensland, South Australia=SA, TAS=Tasmania, WA=Western Australia) .

<sup>1</sup>The Walter and Eliza Hall Institute of Medical Research, Parkville, VIC, AUS, 3050. <sup>2</sup>The Westmead Millenium Institute, Westmead, NSW, AUS 2145. <sup>3</sup>School of Medicine, Griffith University, QLD, AUS, 4222. <sup>4</sup>Department of Neurology, Gold Coast Hospital, QLD, AUS 4215. <sup>5</sup>Diamantina Institute of Cancer, Immunology and Metabolic Medicine, Princess Alexandra Hospital, University of Queensland, Brisbane, QLD, AUS 4102. <sup>6</sup>Botnar Research Centre, Nuffield Department of Orthopaedic Surgery, University of Oxford, Oxford, OX3 7BN, UK. <sup>7</sup>Menzies Research Institute, University of Tasmania, Hobart TAS 7001. <sup>8</sup>Genomics Research Centre, Griffith University, QLD, AUS 4222. <sup>9</sup>Centre for Neuroscience, University of Melbourne, VIC, AUS 3010. <sup>10</sup>The Howard Florey Institute, University of Melbourne, VIC, AUS 3010. <sup>11</sup>Royal Melbourne Hospital, Parkville, VIC, AUS 3050. <sup>12</sup>John Hunter Hospital, Hunter New England Health Service, Newcastle NSW, AUS 2310. <sup>13</sup>Hunter Medical Research Institute, Newcastle, NSW, AUS 2308. <sup>14</sup>Centre for Bioinformatics, Biomarker Discovery and Information-based Medicine, University of Newcastle, NSW, AUS 2308. <sup>15</sup>Department of Medicine, Nepean Hospital, Penrith, NSW, AUS 2751. <sup>16</sup>Christchurch School of Medicine & Health Sciences, University of Otago, New Zealand. <sup>17</sup>UQ Centre for Clinical Research, University of Queensland, QLD, AUS 4029. <sup>18</sup>School of Medicine, University of Queensland, QLD, AUS 4029. <sup>19</sup>Department of Biochemistry, University of Otago, Box 56, Dunedin, NZ. <sup>20</sup>Department of Statistics, The University of Auckland, Auckland, NZ. <sup>21</sup>Sir Charles Gairdner Hospital, Nedlands, WA, AUS 6009. <sup>22</sup>Australian Neuromuscular Research Institute, Nedlands WA, AUS 6009. <sup>23</sup>Barwon Health, Geelong, VIC, AUS 3220. <sup>24</sup>Canterbury District Health

Board, Christchurch, NZ. <sup>25</sup>Department of Neurology, Royal Brisbane and Women's Hospital, QLD, AUS 4029. <sup>26</sup> School of Medicine, Department of Neurology, Flinders University of South Australia, Bedford Park, Adelaide, SA, AUS 5042. <sup>27</sup>Department of Neurology, St. Vincent's University Hospital, Dublin, Republic of Ireland. <sup>28</sup>Auckland District Health board, Auckland, NZ.

Supplementary Table 1: Demographic information for MS patients

	GWAS (n=1618)	Replication study (n=2256)
Gender (n)		
Females	1173	1766
Males	445	490
Female-to-male ratio	2.64:1	3.61:1
Age at disease onset (years)		
average	34.1	34.5
range	7–67	4–73
Clinical course (no.)		
PPMS <sup>a</sup>	407	78
PPMS female-to-male ratio	1.27:1	1.52:1
Relapsing MS <sup>b</sup>	1211	2178
Relapsing MS female-to-male ratio	3.55:1	3.75:1

<sup>a</sup> PPMS (primary progressive MS) is defined as meeting standard clinical criteria for this diagnosis including never having a relapse and having progressive symptoms consistent with MS for at least 12 months

<sup>b</sup> Relapsing MS is defined as either relapsing-remitting or secondary progressive MS

Supplementary Table 2: SNP associations at novel, suggestively associated loci

SNP	CHR	Position	Locus	Selection criteria for replication	Minor allele	Major allele	GWAS					Replication				Combined			
							Genotyped/ imputed	MAF cases	MAF controls	P- value	OR	MAF cases	MAF controls	P- value	OR	MAF cases	MAF controls	P- value	OR
rs6984045	8	131,161,595	ASAP1/DDEF1	ANZ, IMSGC	C	T	Imp	0.035	0.018	5.0E-06	1.98	0.030	0.024	3.6E-02	1.26	0.032	0.021	1.6E-06	1.59
rs39767	8	90,889,143	RIPK2	ANZ, cand	A	G	Gen	0.414	0.460	1.1E-05	0.83	0.428	0.445	5.0E-02	0.93	0.422	0.454	2.5E-05	0.88
rs9644677	8	9,449,218	Transcript M27826	ANZ	G	A	Gen	0.302	0.264	7.0E-05	1.21	0.290	0.273	3.7E-02	1.09	0.295	0.268	5.1E-05	1.14
rs11685349	2	112,179,596	Regulatory conserved	IMSGC	A	G	Gen	0.437	0.403	1.1E-03	1.15	0.422	0.403	3.4E-02	1.08	0.428	0.403	3.4E-04	1.12
rs754292	19	16,560,813	MED26	IMSGC	G	A	Gen	0.471	0.441	6.1E-03	1.13	0.470	0.447	1.3E-02	1.10	0.470	0.444	4.1E-04	1.11
rs3772054	2	241,327,201	KIF1A	cand	G	A	Gen	0.290	0.320	2.4E-03	0.87	0.295	0.312	3.7E-02	0.92	0.293	0.317	6.6E-04	0.89
rs10983396	9	118,693,400	ASTN1	ANZ	T	A	Imp	0.035	0.051	1.6E-05	0.67	0.054	0.065	1.5E-02	0.83	0.046	0.056	9.2E-04	0.81
rs1609798	4	103,756,488	NFKB1	cand	A	G	Gen	0.344	0.314	2.8E-03	1.15	0.335	0.319	4.9E-02	1.08	0.339	0.316	9.9E-04	1.11
rs11000204	10	73,608,572	ASCC1	cand	A	G	Gen	0.499	0.468	3.8E-03	1.13	0.500	0.482	4.5E-02	1.07	0.500	0.474	1.1E-03	1.10

\*SNP positions are derived from NCBI dbSNP genome build 128 (Oct 2007). Criteria for selection of SNPs for replication genotyping are: ‘ANZ’ denotes SNPs ranked in the top 50 most associated non-MHC SNPs that were genotyped, or the top 100 imputed SNPs in the current study (Methods). ‘IMSGC’ indicates SNPs located in a genomic region showing evidence of association ( $p < 0.05$ ) in a GWAS performed by the International MS Genetics Consortium (IMSGC)<sup>6</sup>. ‘Cand’ refers to SNPs that are located either within or close to a gene that was considered a biologically plausible candidate.

GWAS= genome-wide association scan, SNP=single nucleotide polymorphisms, CHR = chromosome, MAF=minor allele frequency, OR = odds ratio, Gen = SNPs genotyped in the GWAS, Imp = SNPs imputed in the GWAS

Supplementary Table 3. Genotypic odds ratios and DR15-stratified allelic odds ratios at (a) novel loci and (b) known and putative MS susceptibility genes

(a) novel loci								Genotypic ORs*			DR15-stratified allelic ORs		
SNP	CHR	Position	Locus	Minor allele	Major allele	Dataset	Allelic OR	Het	Hom	P-value	DR15 neg	DR15 pos	P-value†
rs703842	12	56,449,006	CYP27B1	G	A	combined	0.82	0.81	0.67	0.92	0.82	0.78	0.55
rs6074022	20	44,173,603	CD40	G	A	combined	1.19	1.17	1.48	0.42	1.27	1.14	0.14
rs6984045	8	131,161,595	ASAP1/DDEF1	C	T	replication	1.26	1.30	0.52	0.35	1.22	1.29	0.86
rs39767	8	90,889,143	RIPK2	A	G	combined	0.88	0.87	0.77	0.77	0.90	0.89	0.85
rs9644677	8	9,449,218	Transcript M27826	G	A	combined	1.15	1.17	1.27	0.51	1.18	1.13	0.51
rs754292	19	16,560,813	MED26	G	A	combined	1.11	1.08	1.25	0.40	1.08	1.14	0.39
rs3772054	2	241,327,201	KIF1A	G	A	combined	0.89	0.90	0.79	0.75	0.90	0.91	0.93
rs11685349	2	112,179,596	Regulatory conserved	A	G	combined	1.11	1.13	1.23	0.71	1.11	1.11	0.99
rs11000204	10	73,608,572	ASCC1	A	G	combined	1.11	1.06	1.23	0.32	1.11	1.12	0.91
rs10983396	9	118,693,400	ASTN1	T	A	replication	0.83	0.86	0.20	0.10	0.87	0.72	0.34
rs1609798	4	103,756,488	NFKB1	A	G	combined	1.11	1.10	1.24	0.78	1.10	1.08	0.78

\*Genotypic odds ratios are calculated for heterozygotes (Het) and minor allele homozygotes (Hom) relative to major allele homozygotes (reference group). P-values denote the significance of a dominance term (i.e. a test for departure from multiplicativity of odds ratios)

†P-values in this column denote the significance of a product term, with DR15 genotype coded as a binary variable and genotype at the SNP in question coded as a quantitative variable (0, 1, 2).

Supplementary Table 3 (continued)

(b) known and putative MS susceptibility genes

								Genotypic ORs			DR15-stratified allelic ORs		
SNP	CHR	Position	Locus	Minor allele	Major allele	Dataset	Allelic OR	Het	Hom	P-value	DR15 neg	DR15 pos	P-value
rs9271366	6	32,694,832	HLA-DRB1	G	A	combined	2.78	3.18	6.91	0.003	—	—	—
rs1335532	1	116,902,480	CD58	G	A	combined	0.79	0.76	0.82	0.07	0.81	0.73	0.32
rs6604026	1	93,076,191	EVI5/RPL5	G	A	combined	1.17	1.16	1.38	0.81	1.29	1.03	0.001
rs2104286	10	6,139,051	IL2RA	G	A	replication	0.85	0.85	0.71	0.94	0.87	0.82	0.54
rs6897932	5	35,910,332	IL7R	A	G	combined	0.88	0.92	0.73	0.18	0.90	0.86	0.51
rs8112449	19	10,381,064	TYK2	A	G	combined	0.92	0.90	0.87	0.42	0.92	0.92	0.92
rs2051322	18	65,701,566	CD226	A	G	combined	0.99	0.99	1.00	0.90	1.00	0.99	0.92
rs10492972	1	10,275,699	KIF1B	C	T	replication	0.96	1.02	0.88	0.22	0.93	1.04	0.22

Supplementary Table 4. GWAS p-values at significantly associated SNPs on chromosomes 12 and 20 using various tests of association.

Test	Inflation factor, $\lambda$	rs703842 (chr 12)	rs12368653 (chr12)	rs6074022 (chr 20)	rs1569723 (chr20)
Main test (as in Table 3)	1.100	$4.1 \times 10^{-6}$	$4.6 \times 10^{-5}$	$2.5 \times 10^{-5}$	$3.3 \times 10^{-5}$
(1) No imputation of missing genotypes	1.097	Not tested†	$4.4 \times 10^{-5}$	Not tested†	$3.0 \times 10^{-5}$
(2) Remove outliers > 3SD from mean	1.087	$2.3 \times 10^{-5}$	$5.5 \times 10^{-4}$	$7.6 \times 10^{-5}$	$9.2 \times 10^{-5}$
(3a) Ancestry adjustment: continuous	1.053	$1.3 \times 10^{-5}$	$1.6 \times 10^{-4}$	$1.8 \times 10^{-5}$	$2.2 \times 10^{-5}$
(3b) Ancestry adjustment: categorical	1.057	$1.7 \times 10^{-5}$	$2.6 \times 10^{-4}$	$2.2 \times 10^{-5}$	$2.8 \times 10^{-5}$

To assess the validity of the imputation procedure in the GWAS and the likelihood of false positives due to population stratification, we performed several subsidiary tests of association:

- (1) without using imputed values for missing genotypes (this test was not run on the 46,827 SNPs showing significantly different call rates between the datasets, and it was the main test for SNPs on the X chromosome)
- (2) discarding samples more than 3 standard deviations from the mean along the first 3 principal components in the Eigenstrat analysis leaving 1500 cases and 3322 controls.
- (3) discarding samples as in (2) and adjusting for positions along the first 3 principal components, by treating the principal components as a) continuous or b) categorical covariates in a logistic regression analysis.

Ancestry-adjusted GWAS p-values (Test 3a) for all genotyped SNPs included in the replication phase are listed in Supplementary Table 4.

The genomic inflation factors  $\lambda$  in the table are all for association testing of genotyped SNPs (excluding SNPs in the MHC). The inflation factor for testing of imputed HapMap SNPs outside the MHC in 1500 cases and 3322 controls was 1.087.

†This test was not performed for these SNPs as they showed significantly different call rates between the 3 datasets (ANZ, UK, US).

Supplementary Table 5: Association tests for additional SNP markers genotyped in the replication phase

#	SNP	CHR	Position*	Selection criteria for replication	Minor allele	Major allele	GWAS					Replication				Combined			
							Genotyped/ imputed	MAF cases	MAF controls	P-value	OR	MAF cases	MAF controls	P-value	OR	MAF cases	MAF controls	P-value	OR
1	rs10492972	1	10,275,699	prev	C	T	Imp	0.323	0.315	4.4E-01	1.04	0.316	0.324	7.9E-01	0.96	0.319	0.319	1.0E+00	1.00
2	rs4074733	1	19,003,414	ANZ	C	T	Imp	0.131	0.158	1.7E-06	0.80	0.147	0.142	7.4E-01	1.04	0.140	0.152	1.7E-02	0.91
3	rs3748855	1	26,358,330	ANZ, IMSGC	C	T	Imp	0.077	0.105	4.8E-07	0.70	0.106	0.105	5.5E-01	1.01	0.094	0.105	9.5E-03	0.88
4	rs3819033	1	26,380,880	ANZ, IMSGC	T	C	Imp	0.079	0.109	2.1E-06	0.71	0.099	0.099	5.1E-01	1.00	0.091	0.104	2.1E-03	0.86
5	rs11247868	1	26,392,551	ANZ, IMSGC	A	G	Gen	0.081	0.106	4.4E-05	0.74	0.101	0.100	5.9E-01	1.02	0.093	0.104	8.1E-03	0.87
6	rs10794532	1	26,566,847	ANZ, IMSGC	G	A	Gen	0.176	0.208	1.1E-04	0.81	0.196	0.191	7.2E-01	1.03	0.187	0.201	2.3E-02	0.92
7	rs7514477	1	26,569,213	ANZ, IMSGC	A	G	Gen	0.264	0.303	6.1E-05	0.83	0.287	0.288	4.7E-01	1.00	0.277	0.297	4.0E-03	0.91
8	rs640044	1	78,672,437	ANZ	G	A	Imp	0.034	0.049	2.9E-06	0.68	0.058	0.055	7.0E-01	1.05	0.048	0.052	2.2E-01	0.92
9	rs11808092	1	92,845,816	prev	A	C	Gen	0.292	0.256	9.6E-05	1.20	0.281	0.260	1.3E-02	1.11	0.286	0.258	1.5E-05	1.16
10	rs6604026	1	93,076,191	prev	G	A	Gen	0.310	0.273	1.3E-04	1.20	0.305	0.278	2.1E-03	1.14	0.307	0.275	2.5E-06	1.17
11	rs1335532	1	116,902,480	prev	G	A	Gen	0.101	0.124	1.1E-03	0.80	0.106	0.134	1.1E-05	0.76	0.104	0.128	9.6E-08	0.78
12	rs4657016	1	159,539,150	cand	A	G	Gen	0.492	0.453	2.4E-04	1.17	0.470	0.470	5.0E-01	1.00	0.479	0.460	1.1E-02	1.08
13	rs6711791	2	65,432,732	ANZ	C	A	Gen	0.518	0.466	1.2E-06	1.23	0.482	0.481	4.5E-01	1.01	0.497	0.472	5.3E-04	1.11
14	rs11685349	2	112,179,596	IMSGC	A	G	Gen	0.437	0.403	1.1E-03	1.15	0.422	0.403	3.4E-02	1.08	0.428	0.403	3.4E-04	1.12
15	rs12623828	2	112,180,830	IMSGC	C	T	Imp	0.487	0.448	1.4E-04	1.17	0.467	0.452	7.8E-02	1.06	0.475	0.450	4.0E-04	1.11
16	rs10497702	2	190,053,411	ANZ, IMSGC	G	A	Gen	0.190	0.158	9.0E-05	1.25	0.171	0.174	6.5E-01	0.98	0.179	0.165	1.4E-02	1.10
17	rs10170389	2	190,194,313	ANZ, IMSGC	A	G	Gen	0.273	0.238	2.0E-04	1.20	0.238	0.242	6.8E-01	0.98	0.253	0.240	2.0E-02	1.08
18	rs867294	2	234,937,428	ANZ	A	G	Gen	0.452	0.404	2.9E-06	1.22	0.428	0.411	5.7E-02	1.07	0.438	0.407	1.3E-05	1.14
19	rs3900638	2	234,985,938	ANZ	A	G	Gen	0.419	0.462	3.9E-05	0.84	0.441	0.449	2.3E-01	0.97	0.432	0.457	7.5E-04	0.90
20	rs4663407	2	234,992,080	ANZ	A	T	Imp	0.359	0.395	7.9E-07	0.86	0.451	0.453	4.3E-01	0.99	0.413	0.418	4.3E-01	0.98
21	rs3772054	2	241,327,201	cand	G	A	Gen	0.290	0.320	2.4E-03	0.87	0.295	0.312	3.7E-02	0.92	0.293	0.317	6.6E-04	0.89
22	rs7648424	3	151,906,081	ANZ	A	C	Gen	0.398	0.359	1.3E-04	1.19	0.376	0.392	9.3E-01	0.94	0.385	0.372	1.1E-01	1.05



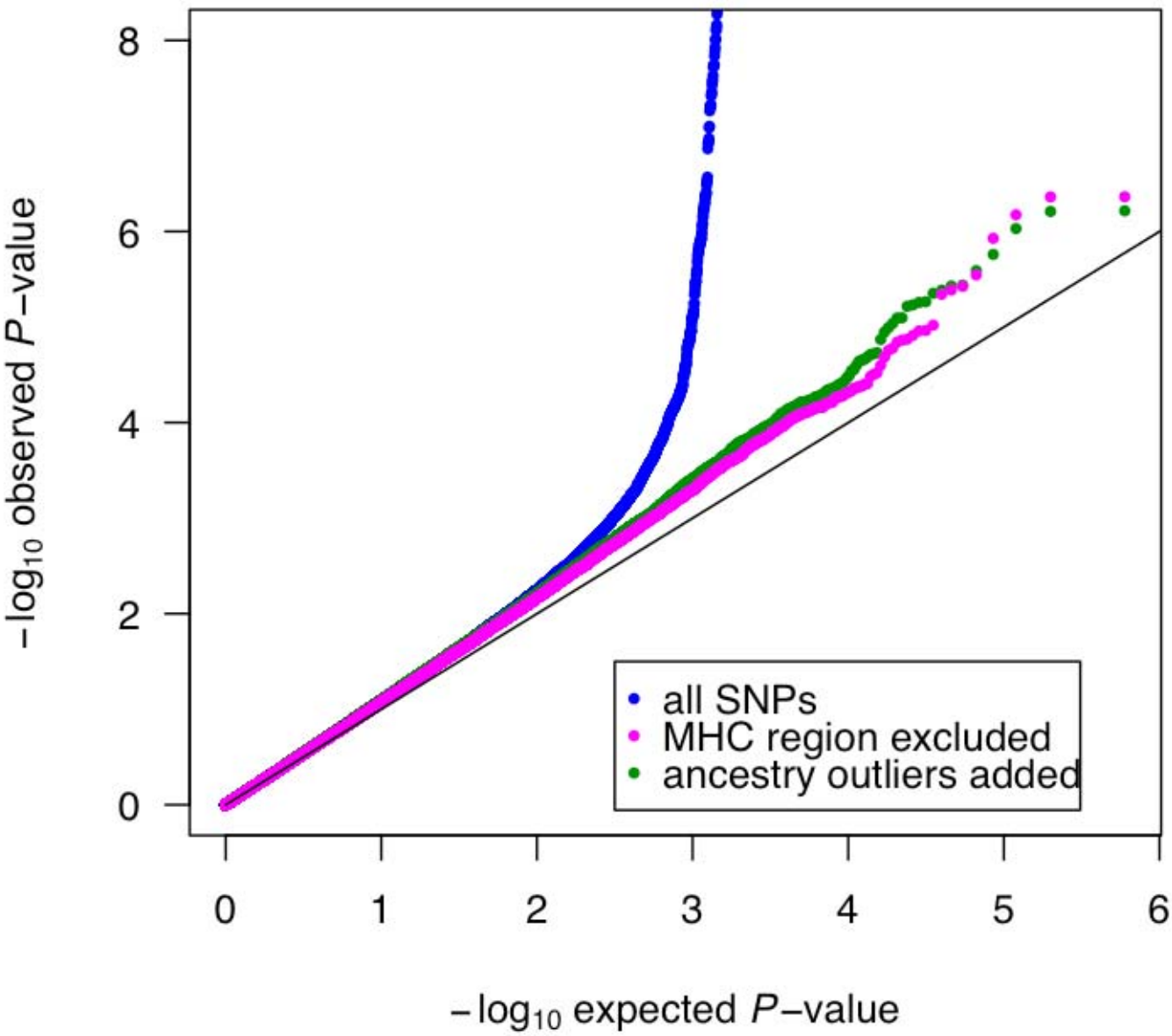
23	rs13074670	3	151,911,821	ANZ	A	C	Gen	0.445	0.491	1.7E-05	0.83	0.463	0.468	3.0E-01	0.98	0.455	0.482	6.7E-04	0.90
24	rs902417	3	185,434,125	IMSGC	A	G	Gen	0.270	0.238	7.2E-04	1.18	0.252	0.245	2.3E-01	1.04	0.259	0.241	3.7E-03	1.11
25	rs1488293	4	26,844,024	ANZ	A	G	Imp	0.032	0.050	1.1E-05	0.63	0.051	0.051	5.3E-01	1.01	0.043	0.050	2.5E-02	0.85
26	rs1609798	4	103,756,488	cand	A	G	Gen	0.344	0.314	2.8E-03	1.15	0.335	0.319	4.9E-02	1.08	0.339	0.316	9.9E-04	1.11
27	rs6838639	4	123,118,615	cand	A	G	Gen	0.228	0.258	9.8E-04	0.85	0.246	0.231	9.5E-01	1.08	0.238	0.248	2.4E-01	0.96
28	rs6897932	5	35,910,332	prev	A	G	Gen	0.253	0.267	1.3E-01	0.93	0.231	0.258	1.3E-03	0.86	0.240	0.264	1.3E-03	0.89
29	rs3194051	5	35,912,031	prev	G	A	Gen	0.255	0.267	1.8E-01	0.94	0.260	0.264	3.5E-01	0.97	0.258	0.266	2.1E-01	0.96
30	rs10515058	5	67,245,631	ANZ	A	T	Imp	0.505	0.459	9.7E-06	1.20	0.471	0.470	4.5E-01	1.01	0.485	0.463	2.8E-03	1.09
31	rs7723716	5	67,246,186	ANZ	A	G	Gen	0.289	0.326	1.4E-04	0.84	0.304	0.311	2.3E-01	0.97	0.298	0.320	1.5E-03	0.90
32	rs6595125	5	117,776,599	ANZ	G	A	Gen	0.528	0.485	4.8E-05	1.19	0.481	0.489	2.2E-01	0.97	0.477	0.505	5.9E-04	0.90
33	rs331085	5	127,788,959	cand	G	A	Gen	0.093	0.115	9.6E-04	0.79	0.110	0.115	2.2E-01	0.95	0.102	0.115	4.4E-03	0.87
34	rs3024389	6	6,194,619	ANZ	A	G	Gen	0.414	0.377	3.8E-04	1.17	0.399	0.389	1.7E-01	1.04	0.405	0.382	1.6E-03	1.10
35	rs2064187	6	17,987,197	cand	A	T	Imp	0.432	0.392	8.6E-05	1.18	0.405	0.403	4.3E-01	1.01	0.416	0.397	7.1E-03	1.08
36	rs9271366	6	32,694,832	prev	G	A	Gen	0.329	0.155	6.9E-89	2.68	0.339	0.151	8.7E-100	2.88	0.335	0.153	7.0E-184	2.78
37	rs321329	6	45,830,966	ANZ	A	G	Gen	0.406	0.448	8.0E-05	0.84	0.413	0.426	9.3E-02	0.95	0.410	0.439	2.0E-04	0.89
38	rs1478804	6	129,471,387	ANZ, cand	G	A	Gen	0.241	0.205	3.9E-05	1.23	0.211	0.211	5.2E-01	1.00	0.224	0.208	4.3E-03	1.11
39	rs265387	6	129,606,637	ANZ, cand	G	A	Gen	0.269	0.233	1.1E-04	1.21	0.235	0.248	9.3E-01	0.93	0.249	0.239	8.9E-02	1.06
40	rs12189801	6	142,683,299	ANZ	G	A	Gen	0.133	0.166	1.4E-05	0.77	0.155	0.148	8.3E-01	1.06	0.146	0.159	1.7E-02	0.90
41	rs11155242	6	142,733,242	ANZ	C	A	Gen	0.166	0.203	9.6E-06	0.78	0.189	0.179	8.7E-01	1.06	0.179	0.194	1.8E-02	0.91
42	rs9644677	8	9,449,218	ANZ	G	A	Gen	0.302	0.264	7.0E-05	1.21	0.290	0.273	3.7E-02	1.09	0.295	0.268	5.1E-05	1.14
43	rs17121574	8	12,799,052	ANZ	C	G	Imp	0.021	0.007	1.1E-08	2.86	0.019	0.023	9.3E-01	0.80	0.020	0.014	1.8E-03	1.42
44	rs13261147	8	13,948,127	ANZ	A	G	Gen	0.219	0.252	3.2E-04	0.83	0.246	0.236	8.6E-01	1.06	0.234	0.245	7.7E-02	0.94
45	rs16920169	8	55,473,854	ANZ	T	C	Imp	0.021	0.040	4.4E-08	0.51	0.040	0.035	9.1E-01	1.16	0.032	0.038	5.0E-02	0.84
46	rs7834421	8	55,481,112	ANZ	A	C	Gen	0.021	0.041	4.4E-07	0.51	0.039	0.035	8.5E-01	1.12	0.032	0.038	6.6E-03	0.79
47	rs2305046	8	80,875,655	cand	A	G	Gen	0.109	0.091	4.3E-	1.22	0.093	0.110	1.0E+00	0.84	0.100	0.098	9.2E-01	1.01

48	rs7821689	8	80,882,930	cand	A	C	Gen	0.129	0.108	1.6E-03	1.23	0.112	0.132	1.0E+00	0.83	0.119	0.118	1.0E+00	1.00
49	rs39502	8	90,845,949	ANZ, cand	C	A	Gen	0.463	0.422	1.0E-04	1.18	0.442	0.437	3.0E-01	1.02	0.451	0.428	2.0E-03	1.10
50	rs447618	8	90,854,822	ANZ, cand	G	A	Gen	0.393	0.431	2.0E-04	0.85	0.405	0.418	9.9E-02	0.95	0.400	0.426	4.8E-04	0.90
51	rs39767	8	90,889,143	ANZ, cand	A	G	Gen	0.414	0.460	1.1E-05	0.83	0.428	0.445	5.0E-02	0.93	0.422	0.454	2.5E-05	0.88
52	rs6984045	8	131,161,595	ANZ, MSGC	C	T	Imp	0.035	0.018	5.0E-06	1.98	0.030	0.024	3.6E-02	1.26	0.032	0.021	1.6E-06	1.59
53	rs12685827	9	21,570,763	cand	A	G	Gen	0.124	0.105	5.0E-03	1.20	0.107	0.112	7.5E-01	0.96	0.114	0.108	1.3E-01	1.07
54	rs9410321	9	91,220,006	cand	G	A	Gen	0.362	0.327	6.4E-04	1.17	0.352	0.341	1.4E-01	1.05	0.356	0.333	1.5E-03	1.11
55	rs2567719	9	116,007,084	ANZ	T	G	Imp	0.324	0.366	6.6E-06	0.83	0.326	0.336	1.7E-01	0.96	0.325	0.353	2.6E-05	0.88
56	rs10983396	9	118,693,400	ANZ	T	A	Imp	0.035	0.051	1.6E-05	0.68	0.054	0.065	1.5E-02	0.83	0.046	0.056	9.2E-04	0.81
57	rs505922	9	135,139,050	cand	G	A	Gen	0.300	0.333	1.0E-03	0.86	0.311	0.304	7.7E-01	1.03	0.306	0.321	6.9E-02	0.94
58	rs2104286	10	6,139,051	prev	G	A	Imp	0.234	0.263	2.1E-03	0.86	0.241	0.272	2.6E-04	0.85	0.238	0.267	7.4E-06	0.86
59	rs7908011	10	61,673,395	ANZ	G	A	Gen	0.221	0.251	1.1E-03	0.85	0.245	0.245	4.8E-01	1.00	0.235	0.249	1.9E-02	0.92
60	rs2224665	10	69,831,639	ANZ, MSGC	C	T	Imp	0.158	0.195	6.5E-06	0.78	0.190	0.188	6.3E-01	1.02	0.177	0.192	1.2E-02	0.91
61	rs1539316	10	69,964,780	ANZ, MSGC	A	C	Gen	0.156	0.180	3.4E-03	0.84	0.183	0.180	6.4E-01	1.02	0.172	0.180	7.0E-02	0.93
62	rs11000204	10	73,608,572	cand	A	G	Gen	0.499	0.468	3.8E-03	1.13	0.500	0.482	4.5E-02	1.07	0.500	0.474	1.1E-03	1.10
63	rs1152686	10	126,778,181	ANZ	G	A	Gen	0.121	0.142	4.7E-03	0.83	0.128	0.123	7.6E-01	1.05	0.125	0.135	1.2E-01	0.93
64	rs1152690	10	126,786,372	ANZ	G	A	Imp	0.079	0.104	4.9E-06	0.74	0.087	0.092	2.4E-01	0.95	0.084	0.099	2.0E-04	0.83
65	rs7937045	11	1,686,169	ANZ	A	G	Gen	0.318	0.360	3.1E-05	0.83	0.342	0.338	6.5E-01	1.02	0.332	0.351	8.1E-03	0.92
66	rs13447665	11	93,831,882	ANZ	C	T	Imp	0.005	0.008	6.6E-06	0.61	0.007	0.010	7.8E-02	0.72	0.006	0.009	1.8E-02	0.66
67	rs10876994	12	56,351,004	ANZ, MSGC	C	A	Imp	0.219	0.256	4.9E-05	0.82	0.228	0.276	6.1E-08	0.78	0.224	0.264	2.7E-10	0.80
68	rs12368653	12	56,419,523	ANZ, MSGC	A	G	Gen	0.521	0.477	4.6E-05	1.19	0.510	0.474	2.6E-04	1.15	0.515	0.476	1.0E-07	1.17
69	rs703842	12	56,449,006	ANZ, MSGC	G	A	Gen	0.281	0.326	4.1E-06	0.81	0.293	0.339	1.4E-06	0.81	0.288	0.331	5.4E-11	0.81
70	rs2365736	12	93,088,049	ANZ, MSGC	G	A	Gen	0.401	0.365	5.6E-04	1.17	0.374	0.372	4.3E-01	1.01	0.386	0.368	9.8E-03	1.08
71	rs8022640	14	76,594,340	ANZ	A	G	Gen	0.341	0.302	9.4E-05	1.20	0.329	0.338	8.2E-01	0.96	0.334	0.316	3.9E-02	1.07

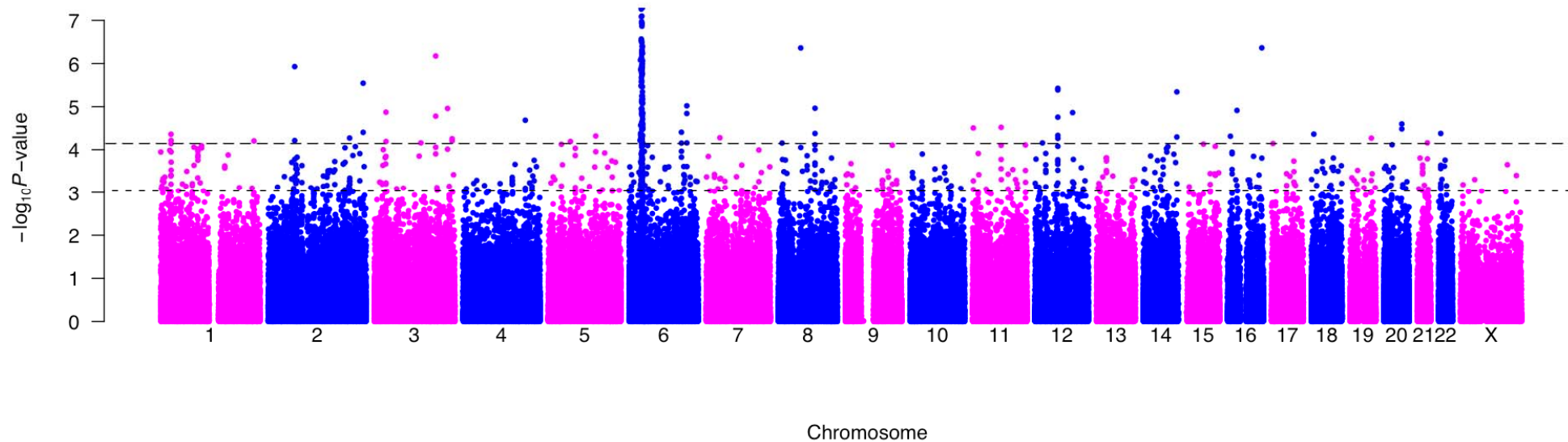
72	rs732787	14	99,286,789	ANZ, cand	A	G	Gen	0.390	0.351	2.0E-04	1.18	0.371	0.376	6.8E-01	0.98	0.379	0.361	2.1E-02	1.07
73	rs941726	14	102,319,429	ANZ	A	G	Gen	0.190	0.154	4.5E-06	1.30	0.172	0.163	1.2E-01	1.07	0.180	0.157	5.0E-05	1.17
74	rs6500842	16	6,808,287	ANZ	G	A	Gen	0.464	0.507	4.9E-05	0.84	0.489	0.498	2.0E-01	0.96	0.479	0.503	6.0E-04	0.90
75	rs8049603	16	22,974,761	ANZ	T	G	Imp	0.239	0.194	4.0E-07	1.31	0.226	0.213	6.9E-02	1.08	0.231	0.202	1.4E-06	1.19
76	rs886038	16	22,999,942	ANZ	A	G	Gen	0.311	0.278	6.8E-04	1.17	0.292	0.280	1.0E-01	1.06	0.300	0.279	1.0E-03	1.11
77	rs12952314	17	37,398,449	cand	A	G	Gen	0.102	0.082	6.8E-04	1.28	0.081	0.083	6.2E-01	0.98	0.090	0.082	2.7E-02	1.12
78	rs2051322	18	65,701,566	prev	A	G	Gen	0.102	0.116	3.8E-02	0.87	0.121	0.112	9.1E-01	1.09	0.113	0.114	6.5E-01	0.98
79	rs8112449	19	10,381,064	prev	A	G	Gen	0.293	0.328	5.2E-04	0.85	0.321	0.327	2.6E-01	0.97	0.309	0.327	3.5E-03	0.91
80	rs754292	19	16,560,813	IMSGC	G	A	Gen	0.471	0.441	6.1E-03	1.13	0.470	0.447	1.3E-02	1.10	0.470	0.444	4.1E-04	1.11
81	rs2422975	20	4,710,869	ANZ, IMSGC, cand	G	A	Imp	0.308	0.346	1.0E-05	0.84	0.322	0.330	2.2E-01	0.97	0.316	0.340	3.7E-04	0.90
82	rs2422979	20	4,712,682	ANZ, IMSGC, cand	A	G	Gen	0.414	0.453	2.4E-04	0.85	0.421	0.434	9.8E-02	0.95	0.418	0.445	4.2E-04	0.90
83	rs6078167	20	11,331,327	ANZ	A	G	Imp	0.419	0.466	8.2E-06	0.83	0.449	0.463	8.8E-02	0.94	0.437	0.464	1.3E-04	0.90
<b>84</b>	<b>rs6131010</b>	<b>20</b>	<b>44,157,712</b>	<b>ANZ, cand</b>	<b>A</b>	<b>G</b>	<b>Imp</b>	<b>0.291</b>	<b>0.253</b>	<b>6.7E-05</b>	<b>1.21</b>	<b>0.287</b>	<b>0.262</b>	<b>3.0E-03</b>	<b>1.14</b>	<b>0.289</b>	<b>0.257</b>	<b>8.5E-07</b>	<b>1.18</b>
<b>85</b>	<b>rs6074022</b>	<b>20</b>	<b>44,173,603</b>	<b>ANZ, cand</b>	<b>G</b>	<b>A</b>	<b>Gen</b>	<b>0.287</b>	<b>0.247</b>	<b>2.5E-05</b>	<b>1.22</b>	<b>0.275</b>	<b>0.245</b>	<b>4.6E-04</b>	<b>1.17</b>	<b>0.280</b>	<b>0.247</b>	<b>1.3E-07</b>	<b>1.20</b>
<b>86</b>	<b>rs1569723</b>	<b>20</b>	<b>44,175,471</b>	<b>ANZ, cand</b>	<b>C</b>	<b>A</b>	<b>Gen</b>	<b>0.287</b>	<b>0.248</b>	<b>3.3E-05</b>	<b>1.22</b>	<b>0.274</b>	<b>0.245</b>	<b>7.4E-04</b>	<b>1.16</b>	<b>0.279</b>	<b>0.247</b>	<b>2.9E-07</b>	<b>1.19</b>
87	rs10482977	21	27,014,402	ANZ	G	A	Gen	0.069	0.090	2.3E-04	0.74	0.082	0.082	5.0E-01	1.00	0.077	0.087	9.3E-03	0.87
88	rs8135758	22	20,406,405	ANZ, cand	A	G	Gen	0.209	0.180	6.1E-04	1.20	0.190	0.177	6.0E-02	1.09	0.198	0.179	4.4E-04	1.14
89	rs5999264	22	20,425,920	ANZ, cand	A	G	Gen	0.310	0.270	4.2E-05	1.21	0.302	0.297	3.0E-01	1.03	0.305	0.281	1.1E-03	1.11
90	rs3788333	22	20,616,791	ANZ, cand	G	A	Gen	0.099	0.124	2.5E-04	0.78	0.121	0.111	9.3E-01	1.10	0.112	0.119	1.3E-01	0.93
91	rs9609163	22	29,664,518	cand	A	G	Gen	0.168	0.140	2.4E-04	1.24	0.164	0.152	6.3E-02	1.09	0.166	0.145	2.5E-04	1.16

Criteria for selection of SNPs for replication genotyping are: ‘ANZ’ denotes SNPs ranked in the top 50 most associated non-MHC SNPs that were genotyped, or the top 100 imputed SNPs in the current study. ‘IMSGC’ indicates SNPs located in a genomic region showing evidence of association ( $p < 0.05$ ) in a GWAS performed by the International MS Genetics Consortium (IMSGC)<sup>6</sup>. ‘Cand’ refers to SNPs that are located either within or close to a gene that was considered a biologically plausible candidate. ‘Prev’ denotes SNPs associated with previously identified MS susceptibility loci. SNPs in bold-type are located in regions on chromosome 12q13-14 and 20q13 showing the evidence of association with risk of MS  $p < 5 \times 10^{-7}$ .

Supplementary Figure 1a: Genomic control analysis

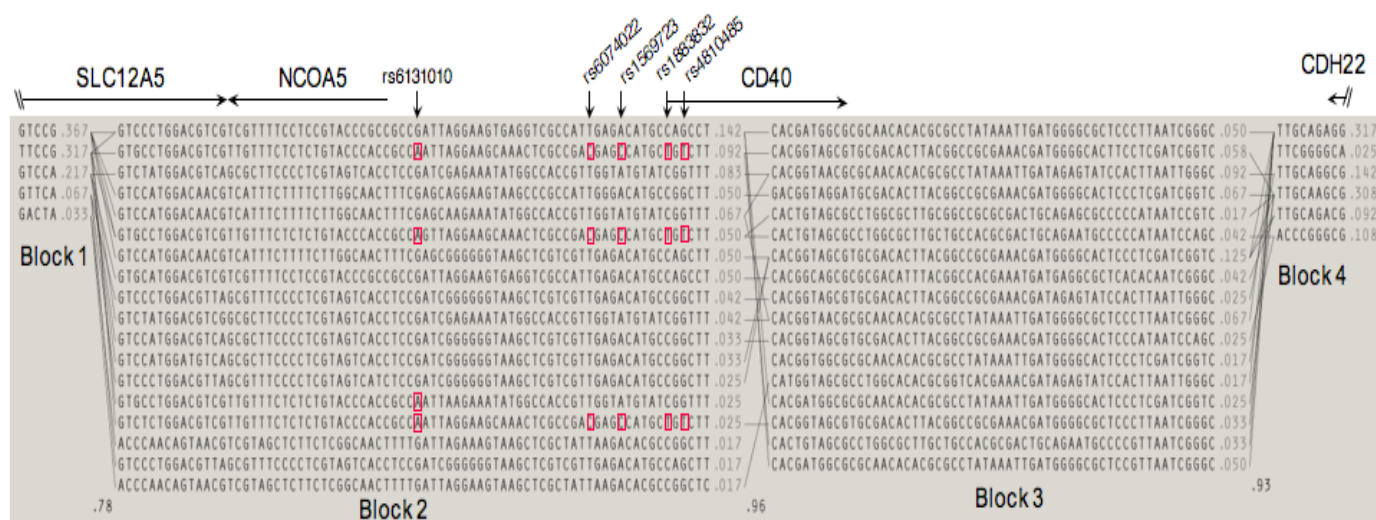


Supplementary Figure 1b: A Genome-wide association scan for multiple sclerosis susceptibility genes



Results of the genome-wide association analysis, comparing 302,098 SNPs between 1,618 MS cases and 3,413 controls. P-values are plotted on the  $-\log_{10}$  scale **(a)** in a quantile-quantile (QQ) plot, and **(b)** by chromosome according to genomic position. Blue dots in **(a)** show the distribution of  $-\log_{10}$  p-values before excluding SNPs located within and in close proximity to the MHC region (24 Mb–36 Mb) on chromosome 6 and pink dots show the distribution after excluding MHC SNPs. Green dots show the distribution of  $-\log_{10}$  p-values without MHC SNPs prior to excluding 178 outlier samples discarded in the principal components analysis. Prior to excluding these ancestry outliers the genomic inflation factor  $\lambda$  was 1.13 compared to 1.10 afterwards. The long- and short-dashed lines in **(b)** show the p-values of the 50<sup>th</sup> ( $-\log_{10} p=4.14$  or  $p=7.2 \times 10^{-5}$ ) and 500<sup>th</sup> ( $-\log_{10} p=3.04$  or  $p=9.2 \times 10^{-4}$ ) ranked non-MHC SNP, respectively. These thresholds were used during the selection of SNPs for the replication phase. The y-axis of this plot has been curtailed at  $p=1 \times 10^{-7}$  for clarity.

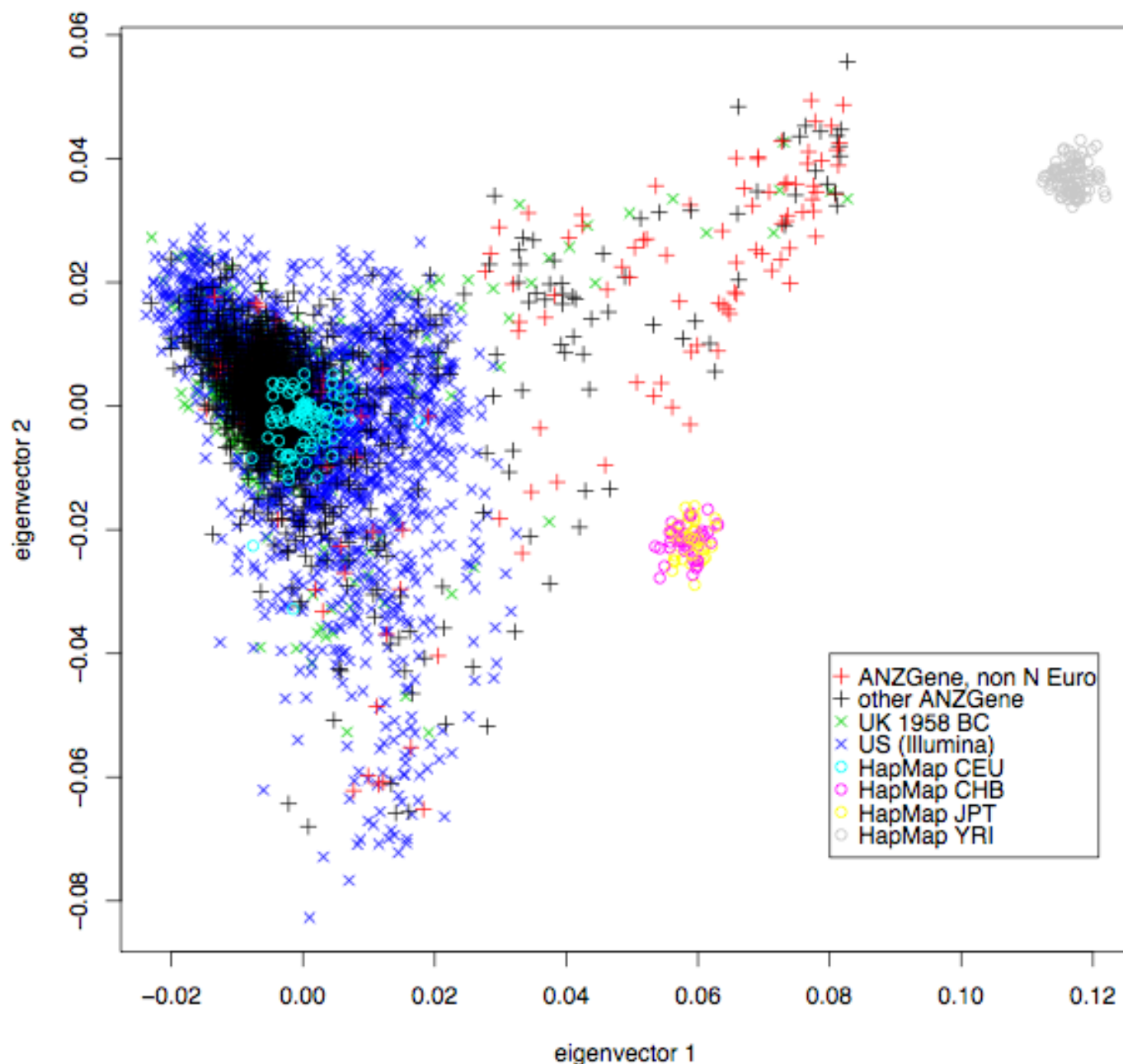
Supplementary Figure 2: Haplotype structure at the chromosome 20 locus



Four haplotype blocks were determined using Haploview<sup>35</sup> according to parameters defined by the solid spine of LD. Haplotypes are indicated by horizontal contiguous collections of SNP alleles with haplotype frequencies estimated at the end of each block. Each SNP that meets the requirements for inclusion in a haplotype block is represented by a vertical column of A, C, G or T alleles. The SNP rs1883832C>T occurs in the Kozak consensus sequence adjacent to the ATG translation initiation codon of the CD40 gene at the -1 position, and the common C allele has been shown to enhance production CD40 protein<sup>27</sup>. The red boxes identify SNP alleles associated with risk of MS. Note that rs1883832C is predisposing for Graves' disease<sup>1</sup> and rs4810485G for RA<sup>2</sup>, but these alleles are predicted to be protective for MS. The position of genes is indicated relative to the SNPs they encompass and not physical size. Connections between haplotypes in different blocks are indicated by thin lines.

Note that genotyping assays were designed for the alternate DNA strand for rs6074022 and thus the G allele in **Table 2** of the manuscript is equivalent to the C allele in this figure.

### Supplementary Figure 3: Principal components analysis



Using the program EIGENSTRAT<sup>32</sup>, principal components analysis (PCA) was used to investigate hidden population structure in the combined remaining case and control data. PCA was conducted on a subset of 77,704 SNPs not in LD, which were common to all sample sets. Through successive iterations 178 outliers (35 ANZ cases and 1 control; 8 UK controls; 134 US controls) were excluded because they were located more than six standard deviations from the mean along the first 10 principal components. Cases and controls differed significantly in their mean positions along eigenvectors 1, 2 and 3. These eigenvectors also separate the three HapMap Phase 2 populations CEU, YRI and CHB/JPT, indicating that the eigenvectors represent ethnic variation and not genotyping batch effects. Positions along these eigenvectors were recorded for subsequent analyses.

In the figure, Eigenvectors 1 and 2 were both correlated with SNPs near the lactase gene on chromosome 2q21.3. MS cases self-reporting southern or central European ancestry (ANZgene, non N Euro) were often positioned at one end of the distribution along these eigenvectors.

**Supplementary note:****Association testing of SNPs genotyped in the replication phase for association with the clinical course of MS**

Of MS cases genotyped in both the GWAS and replication phase, 487 had a primary progressive disease course (PPMS), and 3387 had either relapsing-remitting or secondary progressive disease (relapsing MS) (**Supplementary Table 1**). Association testing was performed between MS cases stratified into PPMS and relapsing MS subtypes to determine whether any of the 100 SNPs associated with risk of MS in the GWAS, and genotyped in the replication phase, also showed evidence of association with clinical course. Three SNPs associated with risk of MS in the GWAS showed weak evidence of association ( $p \leq 0.05$ ) with clinical course in the combined sample (rs5999264,  $p = 0.005$ ; rs886038,  $p = 0.01$ ; rs1609798,  $p = 0.02$ ). None of these associations with clinical course would withstand correction for multiple testing, and should therefore be treated with caution.