

## MANBA, CXCR5, SOX8, RPS6KB1 and ZBTB46 are genetic risk loci for multiple sclerosis

International Multiple Sclerosis Genetics Consortium<sup>#</sup>

<sup>#</sup>A list of all authors and their affiliations can be found in Appendix 1;

Correspondence to: Prof. Dr. Frauke Zipp,  
Rhine Main Neuroscience Network (rmn2),  
Johannes Gutenberg University Centre Mainz,  
Neurology Department,  
Langenbeckstr. 1,  
55131 Mainz, Germany,  
E-mail: frauke.zipp@unimedizin-mainz.de

Correspondence may also be addressed to: Dr. Christina M. Lill, Max Planck Institute for Molecular Genetics, Department of Vertebrate Genomics, Neuropsychiatric Genetics Group, Ihnestr. 63-73, 14195 Berlin, Germany, E-mail: lill@molgen.mpg.de

A recent genome-wide association study reported five loci for which there was strong, but sub-genome-wide significant evidence for association with multiple sclerosis risk. The aim of this study was to evaluate the role of these potential risk loci in a large and independent data set of ~20 000 subjects. We tested five single nucleotide polymorphisms rs228614 (MANBA), rs630923 (CXCR5), rs2744148 (SOX8), rs180515 (RPS6KB1), and rs6062314 (ZBTB46) for association with multiple sclerosis risk in a total of 8499 cases with multiple sclerosis, 8765 unrelated control subjects and 958 trios of European descent. In addition, we assessed the overall evidence for association by combining these newly generated data with the results from the original genome-wide association study by meta-analysis. All five tested single nucleotide polymorphisms showed consistent and statistically significant evidence for association with multiple sclerosis in our validation data sets (rs228614: odds ratio = 0.91,  $P = 2.4 \times 10^{-6}$ ; rs630923: odds ratio = 0.89,  $P = 1.2 \times 10^{-4}$ ; rs2744148: odds ratio = 1.14,  $P = 1.8 \times 10^{-6}$ ; rs180515: odds ratio = 1.12,  $P = 5.2 \times 10^{-7}$ ; rs6062314: odds ratio = 0.90,  $P = 4.3 \times 10^{-3}$ ). Combining our data with results from the previous genome-wide association study by meta-analysis, the evidence for association was strengthened further, surpassing the threshold for genome-wide significance ( $P < 5 \times 10^{-8}$ ) in each case. Our study provides compelling evidence that these five loci are genuine multiple sclerosis susceptibility loci. These results may eventually lead to a better understanding of the underlying disease pathophysiology.

**Keywords:** multiple sclerosis; complex genetics; genetic risk; immunogenetics; genetic association

**Abbreviations:** GWAS = genome-wide association study; SNP = single nucleotide polymorphism

### Introduction

Multiple sclerosis is the most common inflammatory demyelinating disease of the CNS that is likely caused by an interplay of genetic and environmental risk factors. Apart from several independent association signals in the *HLA* (human leukocyte antigen) region

on chromosome 6p21, a recent genome-wide association study (GWAS) in multiple sclerosis has reported 52 loci exerting small to moderate risk effects (IMSGC and WTCCC2, 2011). In addition, five additional loci provided strong support for association ( $P < 5 \times 10^{-7}$ ) in that GWAS, but failed to meet current criteria for genome-wide significance ( $P < 5 \times 10^{-8}$ ). The most

significantly associated single-nucleotide polymorphisms (SNPs) in these regions were rs228614 in *MANBA* (mannosidase, beta A, lysosomal), rs630923 upstream of *CXCR5* (chemokine C-X-C motif receptor 5), rs2744148 downstream of *SOX8* (sex determining region Y-box 8), rs180515 downstream of *RPS6KB1* (ribosomal protein S6 kinase, 70 kDa, polypeptide 1), and rs6062314 in *ZBTB46* (zinc finger and BTB domain containing 46) (IMSGC and WTCCC2, 2011). Given the lack of genome-wide significance, independent validation efforts are needed to further discern the putative role of these loci in multiple sclerosis risk. To this end, we have tested these five SNPs for association with multiple sclerosis risk in a multicentric study comprising 20 138 subjects of European descent who were independent of the original GWAS sample (IMSGC and WTCCC2, 2011).

## Materials and methods

### Power analysis

Power was estimated using the Genetic Power Calculator (Purcell *et al.*, 2003) assuming a one-sided  $\alpha$  of 0.01 and a disease prevalence of 0.1%.

### Data sets

The current study included a total of 8805 multiple sclerosis cases and 8981 unrelated control subjects of self-reported European descent from Germany, Spain, France, The Netherlands, and Australia, as well as 963 trios from the UK. Subjects were selected specifically to be non-overlapping with the original study (IMSGC and WTCCC2, 2011). Diagnosis of multiple sclerosis was established according to standard diagnostic criteria (Poser *et al.*, 1983; McDonald *et al.*, 2001). All samples were collected with informed written consent and appropriate ethical approval at the respective sites. The effective sample size after quality control comprised 8499 multiple sclerosis cases, 8765 unrelated control subjects, and 958 trios (see below and Supplementary Table 1).

### Genotyping and quality control

Genotyping for the German, Spanish and British samples was performed at the individual sites using single-assay allelic discrimination assays based on TaqMan<sup>®</sup> chemistry following the manufacturer's instructions (Applied Biosystems, Inc.). The French subjects were TaqMan<sup>®</sup> genotyped using the multiplex 'OpenArray' platform

(Applied Biosystems, Inc.), the Australian subjects were genotyped using the MassARRAY iPLEX system (Sequenom, Inc.), and the Dutch genotypes were generated on the Human610-Quad Bead GWAS array (Illumina, Inc.). Samples with missing genotypes for more than two SNPs were excluded before analysis [applicable to a total of 115 samples (0.5%) across all data sets]. Information on sex and/or age at examination was available for >90% of subjects in all case-control data sets except in the sample from Central Spain. Samples with missing information in these categories ( $n = 407$ ) were excluded. The threshold for genotyping efficiency per SNP and data set was set to >95%. Hardy–Weinberg equilibrium was assessed in control subjects and in unaffected founders of the nuclear families. Deviations from Hardy–Weinberg equilibrium were defined as  $P < 0.05$  based on Pearson's  $\chi^2$  as implemented in PLINK v1.07 (Purcell *et al.*, 2007).

### Association analyses

All association analyses were performed using PLINK. For the case-unrelated control data sets, logistic regression with adjustment for age at examination and/or sex was performed where available (Supplementary Table 1) using an additive transmission model. Transmission equilibrium testing was applied to the UK trio data set. Odds ratios (OR) are displayed for the allele dosage of the minor allele as defined by the frequency in the overall data set. Meta-analyses across all validation data sets were based on fixed-effect models. The threshold for nominal significance was set to  $P < 0.01$  (i.e. applying a conservative Bonferroni correction for five tests). All  $P$ -values are one-sided with regard to the expected direction of effect based on the original study (IMSGC and WTCCC2, 2011). Between-study heterogeneity was quantified using the  $I^2$  metric, and statistical significance was assessed by the Q-test statistic. Forest plots were generated using a customized version of the 'rmeta' package in R language (Lill *et al.*, 2012). Two-sided unweighted  $P$ -values of the original GWAS (IMSGC and WTCCC2, 2011) and of this study were combined using METAL (Willer *et al.*, 2010).

## Results

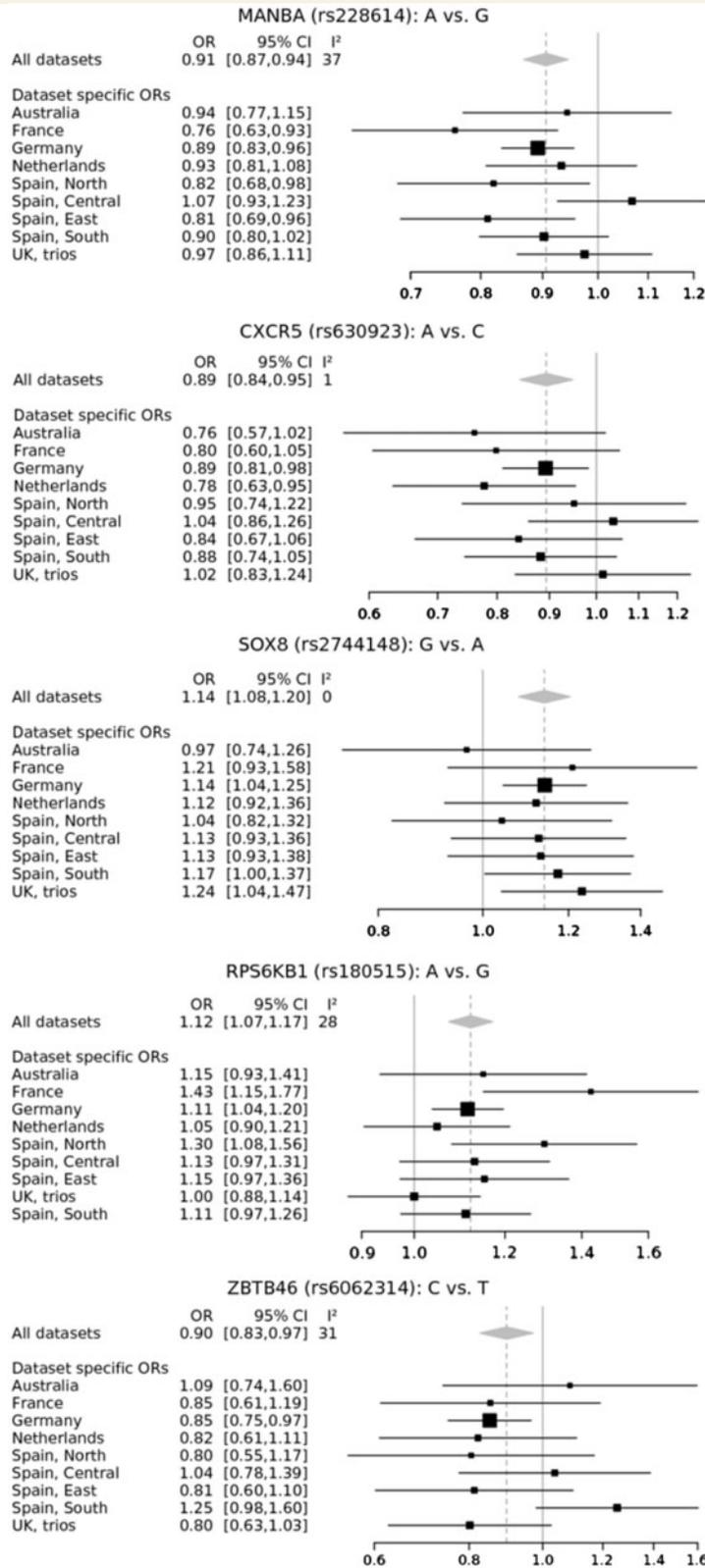
The combined effective replication data sets of 8499 cases, 8765 unrelated controls, and 958 trios had ~80% power to detect an odds ratio of 1.10 down to allele frequencies of 0.13. Control genotypes in all data sets were distributed according to Hardy–Weinberg equilibrium ( $P > 0.05$ ) for all SNPs. Total genotyping efficiency was >98% for each SNP (Table 1).

**Table 1 Association results for the five loci and multiple sclerosis assessed in 20 138 subjects of European descent**

SNP	Location (hg19)	Nearest gene	Validation data sets							Original study		Combined P**
			Eff.	MAF	OR (95% CI)	P*	$I^2$ (95% CI)	$P_Q$	OR	P**		
rs228614 (A/G)	chr4:103,578,637	<i>MANBA</i> (intronic)	99.0	47.6	0.91 (0.87–0.95)	$2.4 \times 10^{-6}$	37 (0–71)	0.120	0.92	$1.4 \times 10^{-7}$	$3.4 \times 10^{-12}$	
rs630923 (A/C)	chr11:118,754,353	<i>CXCR5</i> (122 bp 5')	98.3	15.7	0.89 (0.84–0.95)	$1.2 \times 10^{-4}$	1 (0–65)	0.425	0.89	$2.8 \times 10^{-7}$	$4.7 \times 10^{-10}$	
rs2744148 (G/A)	chr16:1,073,552	<i>SOX8</i> (36,573 bp 3')	99.5	16.8	1.14 (1.08–1.20)	$1.8 \times 10^{-6}$	0 (0–14)	0.915	1.12	$8.4 \times 10^{-8}$	$1.6 \times 10^{-12}$	
rs180515 (G/A)	chr17:58,024,275	<i>RPS6KB1</i> (3' UTR)	99.1	35.5	1.12 (1.07–1.17)	$5.2 \times 10^{-7}$	28 (0–66)	0.198	1.09	$8.8 \times 10^{-8}$	$2.3 \times 10^{-13}$	
rs6062314 (C/T)	chr20:62,409,713	<i>ZBTB46</i> (intronic)	99.1	7.9	0.90 (0.83–0.97)	$4.3 \times 10^{-3}$	31 (0–68)	0.169	0.86	$1.3 \times 10^{-7}$	$2.3 \times 10^{-8}$	

Fixed effect meta-analysis results for the SNPs tested across all validation data sets were performed using PLINK. The association results from the original study (IMSGC and WTCCC2, 2011) and this study were combined using METAL. Allele names are displayed as minor/major allele based on frequencies in the entire validation data set. Brackets following the gene name list the location of the SNP relative to the gene, base pairs (bp) indicate upstream (5') or downstream (3') distance to the primary transcript (as annotated on the UCSC Genome Browser).

CI = confidence interval; Eff. = genotyping efficiency (in %); hg19 = human genome build 19; MAF = minor allele frequency in controls (in %); OR = odds ratio; UTR = untranslated region; \* = one-sided; \*\* = two-sided.



**Figure 1** Meta-analysis of validation data sets assessing the association between the *MANBA*, *CXCR5*, *SOX8*, *RPS6KB1* and *ZBTB46* loci and multiple sclerosis risk in populations of European descent. The x-axis depicts the odds ratio (OR). Study-specific odds ratios (black squares) and 95% confidence intervals (CIs, lines) were calculated using an additive model. The summary odds ratios and 95% confidence intervals (grey diamonds) were calculated based on fixed-effect meta-analysis.

Fixed-effect meta-analysis across all validation data sets revealed highly significant associations of all five tested SNPs and multiple sclerosis risk in the validation data sets, i.e. rs228614 (*MANBA*, OR = 0.91,  $P = 2.4 \times 10^{-6}$ ), rs630923 (*CXCR5*, OR = 0.89,  $P = 1.2 \times 10^{-4}$ ), rs2744148 (*SOX8*, OR = 1.14,  $P = 1.8 \times 10^{-6}$ ), rs180515 (*RPS6KB1*, OR = 1.12,  $P = 5.2 \times 10^{-7}$ ), and rs6062314 (*ZBTB46*, OR = 0.90,  $P = 4.3 \times 10^{-3}$ ). Effect estimates were similar to those originally reported (IMSGC and WTCCC2, 2011). There was no evidence for substantial between-study heterogeneity for any of the five SNPs (Fig. 1 and Table 1). Combining our results with *P*-values from the original GWAS (IMSGC and WTCCC2, 2011) increased the statistical support of our findings to genome-wide significance for each of the five tested SNPs: rs228614 (*MANBA*),  $P = 3.4 \times 10^{-12}$ , rs630923 (*CXCR5*),  $P = 4.7 \times 10^{-10}$ , rs2744148 (*SOX8*),  $P = 1.6 \times 10^{-12}$ , rs180515 (*RPS6KB1*),  $P = 2.3 \times 10^{-13}$ , and rs6062314 (*ZBTB46*),  $P = 2.3 \times 10^{-8}$  (Table 1).

## Discussion

Our study shows that common genetic variants in or near *MANBA*, *CXCR5*, *SOX8*, *RPS6KB1*, and *ZBTB46*, are associated with multiple sclerosis risk at genome-wide significance. Our results, thus, provide compelling evidence that these loci represent genuine genetic risk factors for multiple sclerosis.

As is the case for the majority of genetic associations, the precise molecular genetic mechanisms underlying these results still remain to be assessed. That is, future studies need to clarify whether the SNPs tested here are directly involved in altering gene expression/protein function or whether such effects are exerted by other correlated variants, possibly located in neighbouring genes. For instance, SNP rs180515 in the 3' UTR of *RPS6KB1* is located in the seed region of a predicted micro-RNA binding site for hsa-miR-3616-5p and hsa-miR-573 and may thus directly alter *RPS6KB1* translation (Supplementary Fig. 1) (Schilling, 2012). The intronic SNP rs228614 in *MANBA* is in substantial linkage disequilibrium with two non-synonymous SNPs in the same gene [rs2866413 (p.Thr701Met),  $r^2 = 0.87$ , and rs227368 (p.Val253Leu),  $r^2 = 0.74$ , based on 1000 Genomes Pilot 1 CEU data (1000 Genomes Project Consortium, 2010)], which may affect protein function. However, and possibly more importantly, of all five loci tested here *MANBA* is the only one to contain SNPs (including rs228614) showing strong *cis* effects on messenger RNA expression based on recently published data (Yang *et al.*, 2010) (Supplementary Fig. 2). The intronic SNP rs6062314 in *ZBTB46* displays only moderate linkage disequilibrium to potentially functional variants, i.e. a non-synonymous SNP in *LIME1* [Lck interacting transmembrane adaptor 1, SNP rs1151625 (p.Pro211Leu),  $r^2 = 0.39$ ], and in *ZGPAT* [zinc finger, CCCH-type with G patch domain, SNP rs1291212 (p.Ser61Arg),  $r^2 = 0.31$ ]. In addition, *TNFRSF6B* (tumour necrosis factor receptor superfamily, member 6b, decoy) is also located in this chromosomal region and would represent a compelling candidate based on its implications on T cell function (e.g. Zhang *et al.*, 2001). However, the only coding sequence SNP displaying noteworthy linkage disequilibrium with rs6062314 in this gene does not invoke an amino acid change (rs2738787,  $r^2 = 0.36$ ). Finally,

rs630923 maps into a potential *CXCR5* transcription factor binding site for the nuclear factor of kappa light polypeptide gene enhancer in B-cells (NFkB) in a region of DNase I hypersensitivity (ENCODE Project Consortium, 2012). Rs630923 is predicted to alter NFkB binding (Boyle *et al.*, 2012) and could thus, potentially affect *CXCR5* transcription.

It should be emphasized that the abovementioned potential functional consequences are based on *in silico* assessments and require experimental testing and validation. It is also possible that hitherto unknown, rare DNA sequence variants underlie or contribute to the observed association signals.

In summary, our study provides compelling evidence that the list of established multiple sclerosis risk genes can now be extended by five additional loci, all of which show genome-wide significant association with disease risk. Further fine-mapping and functional studies are required to elucidate the biochemical and pathophysiological mechanisms underlying these associations.

## Acknowledgements

We are grateful to the individuals participating in this study. We acknowledge use of the cohort of the CRB-REFGENSEP and thank ICM, CIC Pitié-Salpêtrière, Généthron and REFGENSEP's members for their help and support.

## Funding

This project was funded by grants from the German Ministry for Education and Research (BMBF) and German Research Foundation (DFG; to F.Z.), the BMBF (grant 16SV5538) and the Cure Alzheimer's Fund (to L.B.), the Walter- and Ilse-Rose-Stiftung and the BMBF (grant 01GM1203A; to H.-P.H. and O.A.), the BMBF (grant NBL3 to U.K.Z.; grant 01UW0808 to U.L., and E.S.-T.), the Innovation Fund of the Max Planck Society (M.FE.A.LD0002 to U.L.), the Fondo de Investigaciones Sanitarias FEDER-FIS PI10/1985 and Fundación Alicia Koplowitz (to E.U.), and the FEDER-FIS grant number PI12/00555 and MINECO-FEDER grant number SAF2009-11491 (to F.M. and A.A.). The research leading to these results has received funding from the program 'Investissements d'avenir' ANR-10-IAIHU-06. This project was supported by INSERM, ARSEP, AFM, GIS-IBISA, the National Institutes of Health (grant RO1NS049477) and the Cambridge NIHR Biomedical Research Centre. C.M.L. was supported by the Fidelity Biosciences Research Initiative. E.U. works for the Fundación de Investigación Biomédica del Hospital Clínico San Carlos-IdISSC. S.M. is an early stage researcher funded by the European Community's Seventh Framework Programme ([FP7/2007-2013] under grant agreement n° 212877 (UEPHA\*multiple sclerosis).

## Supplementary material

Supplementary material is available at *Brain* online.

## References

- Boyle AP, Hong EL, Hariharan M, Cheng Y, Schaub MA, Kasowski M, et al. Annotation of functional variation in personal genomes using RegulomeDB. *Genome Res* 2012; 22: 1790–7.
- ENCODE Project Consortium, Dunham I, Kundaje A, Aldred SF, Collins PJ, Davis CA, et al. An integrated encyclopedia of DNA elements in the human genome. *Nature* 2012; 489: 57–74.
- IMSGC, WTCCC2. Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. *Nature* 2011; 476: 214–19.
- Lill CM, Roehr JT, McQueen MB, Kavvoura FK, Bagade S, Schjeide BM, et al. Comprehensive research synopsis and systematic meta-analyses in Parkinson's disease genetics: The PDGene database. *PLoS Genet* 2012; 8: e1002548.
- McDonald WI, Compston A, Edan G, Goodkin D, Hartung HP, Lublin FD, et al. Recommended diagnostic criteria for multiple sclerosis: guidelines from the International Panel on the diagnosis of multiple sclerosis. *Ann Neurol* 2001; 50: 121–7.
- Poser CM, Paty DW, Scheinberg L, McDonald WI, Davis FA, Ebers GC, et al. New diagnostic criteria for multiple sclerosis: guidelines for research protocols. *Ann Neurol* 1983; 13: 227–31.
- Purcell S, Cherny SS, Sham PC. Genetic power calculator: design of linkage and association genetic mapping studies of complex traits. *Bioinformatics* 2003; 19: 149–50.
- Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007; 81: 559–75.
- Schilling M. In silico assessment of the effects of single nucleotide polymorphisms on miRNA-mRNA interactions. Bachelor of Science in Bioinformatics, 2012. Max Planck Institute for Molecular Genetics, Free University Berlin.
- Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* 2010; 26: 2190–1.
- Yang T-P, Beazley C, Montgomery SB, Dimas AS, Gutierrez-Arcelus M, Stranger BE, et al. Genevar: a database and Java application for the analysis and visualization of SNP-gene associations in eQTL studies. *Bioinformatics* 2010; 26: 2474–6.
- Zhang J, Salcedo TW, Wan X, Ullrich S, Hu B, Gregorio T, et al. Modulation of T-cell responses to alloantigens by TR6/DcR3. *J Clin Invest* 2001; 107: 1459–68.
- 1000 Genomes Project Consortium. A map of human genome variation from population-scale sequencing. *Nature* 2010; 467: 1061–73.
- Catherine Lubetzki,<sup>7,18</sup> Felix Luessi,<sup>2</sup> Sunny Malhotra,<sup>6</sup> Julia Mescheriakova,<sup>10</sup> Xavier Montalban,<sup>6</sup> Caroline Papeix,<sup>7,18</sup> Lidia F. Paredes,<sup>5</sup> Peter Rieckmann,<sup>16,†</sup> Elisabeth Steinhagen-Thiessen,<sup>27</sup> Alexander Winkelmann,<sup>15</sup> Uwe K. Zettl,<sup>15</sup> Rogier Hintzen,<sup>10</sup> Koen Vandenbroeck,<sup>9,28</sup> Graeme Stewart,<sup>8</sup> Bertrand Fontaine,<sup>7,18</sup> Manuel Comabella,<sup>6</sup> Elena Urcelay,<sup>5</sup> Fuencisla Matesanz,<sup>4</sup> Stephen Sawcer,<sup>3</sup> Lars Bertram<sup>1,‡</sup> and Frauke Zipp<sup>2,‡</sup>, on behalf of the International Multiple Sclerosis Genetics Consortium.
- <sup>1</sup>Neuropsychiatric Genetics Group, Department of Vertebrate Genomics, Max Planck Institute for Molecular Genetics, Berlin, Germany; <sup>2</sup>Department of Neurology, Focus Program Translational Neuroscience, University Medical Center of the Johannes Gutenberg-University Mainz, Mainz, Germany; <sup>3</sup>University of Cambridge, Department of Clinical Neurosciences, Addenbrookes Hospital, Cambridge, UK; <sup>4</sup>Instituto de Parasitología y Biomedicina 'López Neyra', Consejo Superior de Investigaciones Científicas (IPBLN-CSIC), Granada, Spain; <sup>5</sup>Immunology Department, Hospital Clínico San Carlos, Instituto de Investigación Sanitaria del Hospital Clínico San Carlos- IdISSC, Madrid, Spain; <sup>6</sup>Department of Neurology-Neuroimmunology, Centre d'Esclerosi Múltiple de Catalunya, Cemcat, Hospital Universitari Vall d'Hebron (HUVH), Barcelona, Spain; <sup>7</sup>UPMC-INSERM-CNRS-ICM, UMR 975-7225, Institut Cerveau Moelle Epinière (ICM), Hôpital Pitié-Salpêtrière, Paris, France; <sup>8</sup>Institute for Immunology and Allergy Research, Westmead Millennium Institute, University of Sydney, Sydney, Australia; <sup>9</sup>Neurogenomics Laboratory, Department of Neuroscience, University of the Basque Country UPV/EHU, Leioa, Spain; <sup>10</sup>MS Centre ErasMS, Erasmus MC, University Medical Center, Rotterdam, The Netherlands; <sup>11</sup>Department of Human Genetics, Ruhr University, Bochum, Germany; <sup>12</sup>Department of Neurology, Medical Faculty, Heinrich Heine University, Düsseldorf, Germany; <sup>13</sup>Servicio de Neurología, Hospital de Basurto, Bilbao, Spain; <sup>14</sup>Multiple Sclerosis Unit, Neurology Department, Hospital Clínico San Carlos, Instituto de Investigación Sanitaria del Hospital Clínico San Carlos- IdISSC, Madrid, Spain; <sup>15</sup>Department of Neurology, University of Rostock, Rostock, Germany; <sup>16</sup>Department of Neurology, University of Würzburg, Würzburg, Germany; <sup>17</sup>Department of Neurology, St. Josef-Hospital, Ruhr-University, Bochum, Germany; <sup>18</sup>Assistance Publique-Hôpitaux de Paris (AP-HP), Département de Neurologie, Hôpital Pitié-Salpêtrière, Paris, France; <sup>19</sup>Department of Medicine, Rheumatology, and Clinical Immunology & DRFZ, Charité University Medicine, Berlin, Germany; <sup>20</sup>Servicio de Neurología, Instituto de Neurociencias Clínicas, Hospital Regional Universitario Carlos Haya, Málaga, Spain; <sup>21</sup>Institute for Clinical Neuroimmunology, Ludwig Maximilian University, Munich, Germany; <sup>22</sup>Unidad de Esclerosi Múltiple, Hospital Virgen Macarena, Sevilla, Spain; <sup>23</sup>Institute of Human Genetics, University of Ulm, Ulm, Germany; <sup>24</sup>Max Planck Institute for Human Development, Berlin, Germany; <sup>25</sup>Department of Psychology, Technische Universität Dresden, Dresden, Germany; <sup>26</sup>Department of Clinical Chemistry, Ludwig Maximilian University, Munich, Germany; <sup>27</sup>Interdisciplinary Metabolic Center, Lipids Clinic, Charité University Medicine, Berlin, Germany; <sup>28</sup>IKERBASQUE, Basque Foundation for Science, Bilbao, Spain
- <sup>†</sup>Current address: Centre for Research in Neuroscience at McGill University, Montreal, Canada (Antje Kroner), Institute of Laboratory Medicine and Human Genetics, Singen, Germany (Peter Lohse), Department of Neurology, Sozialstiftung Bamberg Hospital, Bamberg, Germany (Peter Rieckmann)
- <sup>‡</sup>These authors contributed equally to this work

## Appendix 1

### List of authors

Christina M. Lill,<sup>1,2</sup> Brit-Maren M. Schjeide,<sup>1</sup> Christiane Graetz,<sup>1,2</sup> Maria Ban,<sup>3</sup> Antonio Alcina,<sup>4</sup> Miguel A. Ortiz,<sup>5</sup> Jennifer Pérez,<sup>6</sup> Vincent Damotte,<sup>7</sup> David Booth,<sup>8</sup> Aitzkoa Lopez de Lapuente,<sup>9</sup> Linda Broer,<sup>10</sup> Marcel Schilling,<sup>1</sup> Denis A. Akkad,<sup>11</sup> Orhan Aktas,<sup>12</sup> Iraide Alloza,<sup>9</sup> Alfredo Antigüedad,<sup>13</sup> Rafa Arroyo,<sup>14</sup> Paul Blaschke,<sup>15</sup> Mathias Buttman,<sup>16</sup> Andrew Chan,<sup>17</sup> Alastair Compston,<sup>3</sup> Isabelle Cournu-Rebeix,<sup>7,18</sup> Thomas Dörner,<sup>19</sup> Joerg T. Epplen,<sup>11</sup> Óscar Fernández,<sup>20</sup> Lisa-Ann Gerdes,<sup>21</sup> Léna Guillot-Noël,<sup>7</sup> Hans-Peter Hartung,<sup>12</sup> Sabine Hoffman,<sup>11</sup> Guillermo Izquierdo,<sup>22</sup> Anu Kempainen,<sup>3</sup> Antje Kroner,<sup>16,†</sup> Christian Kubisch,<sup>23</sup> Tania Kümpfel,<sup>21</sup> Shu-Chen Li,<sup>24,25</sup> Ulman Lindenberger,<sup>24</sup> Peter Lohse,<sup>26,†</sup>