

2011-08

Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis

Sawcer, S

<http://hdl.handle.net/10026.1/12039>

10.1038/nature10251

Nature

Springer Science and Business Media LLC

All content in PEARL is protected by copyright law. Author manuscripts are made available in accordance with publisher policies. Please cite only the published version using the details provided on the item record or document. In the absence of an open licence (e.g. Creative Commons), permissions for further reuse of content should be sought from the publisher or author.

Published in final edited form as:

Nature. ; 476(7359): 214–219. doi:10.1038/nature10251.

Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis

The International Multiple Sclerosis Genetics Consortium (IMSGC), Wellcome Trust Case Control Consortium 2 (WTCCC2)*, Stephen Sawcer^{1,*}, Garrett Hellenthal^{2,*}, Matti Pirinen^{2,*}, Chris C.A. Spencer^{2,*}, Nikolaos A. Patsopoulos^{3,5}, Loukas Moutsianas⁶, Alexander Dilthey⁶, Zhan Su², Colin Freeman², Sarah E. Hunt⁷, Sarah Edkins⁷, Emma Gray⁷, David R. Booth⁸, Simon C. Potter⁷, An Goris⁹, Gavin Band², Annette Bang Oturai¹⁰, Amy Strange², Janna Saarela¹¹, Céline Bellenguez², Bertrand Fontaine¹², Matthew Gillman⁷, Bernhard Hemmer¹³, Rhian Gwilliam⁷, Frauke Zipp^{14,15}, Alagurevathi Jayakumar⁷, Roland Martin¹⁶, Stephen Leslie¹⁷, Stanley Hawkins¹⁸, Eleni Giannoulidou², Sandra D'Alfonso¹⁹, Hannah Blackburn⁷, Filippo Martinelli Boneschi²⁰, Jennifer Liddle⁷, Hanne F. Harbo^{21,22}, Marc L. Perez⁷, Anne Spurkland²³, Matthew J Waller⁷, Marcin P. Mycko²⁴, Michelle Ricketts⁷, Manuel Comabella²⁵, Naomi Hammond⁷, Ingrid Kockum²⁶, Owen T. McCann⁷, Maria Ban¹, Pamela Whittaker⁷, Anu Kempainen¹, Paul Weston⁷, Clive Hawkins²⁷, Sara Widaa⁷, John Zajicek²⁸, Serge Dronov⁷, Neil Robertson²⁹, Suzannah J. Bumpstead⁷, Lisa F. Barcellos^{30,31}, Rathi Ravindrarajah⁷, Roby Abraham²⁷, Lars Alfredsson³², Kristin Ardlie⁴, Cristin Aubin⁴, Amie Baker¹, Katharine Baker²⁹, Sergio E. Baranzini³³, Laura Bergamaschi¹⁹, Roberto Bergamaschi³⁴, Allan Bernstein³¹, Achim Berthele¹³, Mike Boggild³⁵, Jonathan P. Bradfield³⁶, David Brassat³⁷, Simon A. Broadley³⁸, Dorothea Buck¹³, Helmut Butzkueven^{39,42}, Ruggero Capra⁴³, William M. Carroll⁴⁴, Paola Cavalla⁴⁵, Elisabeth G. Celius²¹, Sabine Cepok¹³, Rosetta Chiavacci³⁶, Françoise Clerget-Darpoux⁴⁶, Kathleen Clysters⁹, Giancarlo Comi²⁰, Mark Cossburn²⁹, Isabelle Cournu-Rebeix¹², Mathew B. Cox⁴⁷, Wendy Cozen⁴⁸, Bruce A.C. Cree³³, Anne H. Cross⁴⁹, Daniele Cusi⁵⁰, Mark J. Daly^{4,51,52}, Emma Davis⁵³, Paul I.W. de Bakker^{3,4,54,55}, Marc Debouverie⁵⁶, Marie Beatrice D'hooghe⁵⁷, Katherine Dixon⁵³, Rita Dobosi⁹, Bénédicte Dubois⁹, David Ellinghaus⁵⁸, Irina Elovaara^{59,60}, Federica Esposito²⁰, Claire Fontenille¹², Simon Foote⁶¹, Andre Franke⁵⁸, Daniela Galimberti⁶², Angelo Ghezzi⁶³, Joseph Glessner³⁶, Refujia Gomez³³, Olivier Gout⁶⁴, Colin Graham⁶⁵, Struan F.A. Grant^{36,66,67}, Franca Rosa Guerini⁶⁸, Hakon Hakonarson^{36,66,67}, Per Hall⁶⁹, Anders Hamsten⁷⁰, Hans-Peter Hartung⁷¹, Rob N. Heard⁸, Simon Heath⁷², Jeremy Hobart²⁸, Muna Hoshi¹³, Carmen Infante-Duarte⁷³, Gillian Ingram²⁹, Wendy Ingram²⁸, Talat Islam⁴⁸, Maja Jagodic²⁶, Michael Kabesch⁷⁴, Allan G. Kermode⁴⁴, Trevor J. Kilpatrick^{39,40,75}, Cecilia Kim³⁶, Norman Klopp⁷⁶, Keijo Koivisto⁷⁷, Malin Larsson⁷⁰, Mark Lathrop⁷², Jeannette S. Lechner-Scott^{47,78}, Maurizio A. Leone⁷⁹, Virpi Leppä^{11,80}, Ulrika Liljedahl⁸¹, Izaura Lima Bomfim²⁶, Robin R. Lincoln³³, Jenny Link²⁶, Jianjun Liu⁸², Åslaug R. Lorentzen^{22,83}, Sara Lupoli^{50,84}, Fabio Macchiardi^{50,85}, Thomas Mack⁴⁸, Mark Marriott^{39,40}, Vittorio Martinelli²⁰, Deborah Mason⁸⁶, Jacob L. McCauley⁸⁷,

Correspondence and requests for materials should be addressed to D.A.S.C. (alastair.compston@medschl.cam.ac.uk) and P.D. (donnelly@well.ox.ac.uk) on behalf of the IMSGC and WTCCC2 respectively.

*A full list of authors appears at the end of this article. Membership of both consortia and details of individual contributions are listed in the Supplementary Material.

All authors reviewed and approved the manuscript.

*These authors contributed equally.

†These authors jointly directed the study

Supplementary Information is linked to the online version of the paper at www.nature.com/nature

Reprints and permissions information is available at www.nature.com/reprints.

The authors declare no competing financial interests.

Frank Mentch³⁶, Inger-Lise Mero^{21,83}, Tania Mihalova²⁷, Xavier Montalban²⁵, John Mottershead^{88,89}, Kjell-Morten Myhr^{90,91}, Paola Naldi⁷⁹, William Ollier⁵³, Alison Page⁹², Aarno Palotie^{7,11,93,94}, Jean Pelletier⁹⁵, Laura Piccio⁴⁹, Trevor Pickersgill²⁹, Fredrik Piehl²⁶, Susan Pobywajlo⁵, Hong L. Quach³⁰, Patricia P. Ramsay³⁰, Mauri Reunanen⁹⁶, Richard Reynolds⁹⁷, John D. Rioux⁹⁸, Mariaemma Rodegher²⁰, Sabine Roesner¹⁶, Justin P. Rubio³⁹, Ina-Maria Rückert⁷⁶, Marco Salvetti⁹⁹, Erika Salvi^{50,100}, Adam Santaniello³³, Catherine A. Schaefer³¹, Stefan Schreiber^{58,101}, Christian Schulze¹⁰², Rodney J. Scott⁴⁷, Finn Sellebjerg¹⁰, Krzysztof W. Selmaj²⁴, David Sexton¹⁰³, Ling Shen³¹, Brigid Simms-Acuna³¹, Sheila Skidmore¹, Patrick M.A. Sleiman^{36,66}, Cathrine Smestad²¹, Per Soelberg Sørensen¹⁰, Helle Bach Søndergaard¹⁰, Jim Stankovich⁶¹, Richard C. Strange²⁷, Anna-Maija Sulonen^{11,80}, Emilie Sundqvist²⁶, Ann-Christine Syvänen⁸¹, Francesca Taddeo¹⁰⁰, Bruce Taylor⁶¹, Jenefer M. Blackwell^{104,105}, Pentti Tienari¹⁰⁶, Elvira Bramon¹⁰⁷, Ayman Tourbah¹⁰⁸, Matthew A. Brown¹⁰⁹, Ewa Tronczynska²⁴, Juan P. Casas¹¹⁰, Niall Tubridy^{40,111}, Aiden Corvin¹¹², Jane Vickery²⁸, Janusz Jankowski¹¹³, Pablo Villoslada¹¹⁴, Hugh S. Markus¹¹⁵, Kai Wang^{36,66}, Christopher G. Mathew¹¹⁶, James Wason¹¹⁷, Colin N.A. Palmer¹¹⁸, H-Erich Wichmann^{76,119,120}, Robert Plomin¹²¹, Ernest Willoughby¹²², Anna Rautanen², Juliane Winkelmann^{13,123,124}, Michael Wittig^{58,125}, Richard C. Trembath¹¹⁶, Jacqueline Yaouanq¹²⁶, Ananth C. Viswanathan¹²⁷, Haitao Zhang^{36,66}, Nicholas W. Wood¹²⁸, Rebecca Zuvich¹⁰³, Panos Deloukas⁷, Cordelia Langford⁷, Audrey Duncanson¹²⁹, Jorge R. Oksenberg³³, Margaret A. Pericak-Vance⁸⁷, Jonathan L. Haines¹⁰³, Tomas Olsson²⁶, Jan Hillert²⁶, Adrian J. Ivinson^{51,130}, Philip L. De Jager^{4,5,51}, Leena Peltonen^{7,11,80,93,94}, Graeme J. Stewart⁸, David A. Hafler^{4,131}, Stephen L. Hauser³³, Gil McVean², Peter Donnelly^{2,6,†}, and Alastair Compston^{1,†}

¹University of Cambridge, Department of Clinical Neurosciences, Addenbrooke's Hospital, BOX 165, Hills Road, Cambridge, CB2 0QQ, UK ²Wellcome Trust Centre for Human Genetics, Roosevelt Drive, Oxford OX3 7BN, UK ³Division of Genetics, Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02115, USA ⁴Broad Institute of Harvard University and Massachusetts Institute of Technology, Cambridge, MA, USA ⁵Center for Neurologic Diseases, Department of Neurology, Brigham & Women's Hospital, Boston, MA 02115, USA ⁶Dept Statistics, University of Oxford, Oxford OX1 3TG, UK ⁷Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK ⁸Westmead Millennium Institute, University of Sydney, Australia ⁹Laboratory for Neuroimmunology, Section of Experimental Neurology, Katholieke Universiteit Leuven, 3000 Leuven, Belgium ¹⁰Danish Multiple Sclerosis Center, Department of Neurology, Copenhagen University Hospital, Rigshospitalet, 2100 Copenhagen, Denmark ¹¹Institute for Molecular Medicine Finland (FIMM), University of Helsinki, Helsinki, 00290, Finland ¹²INSERM UMR S 975 CRICM, UPMC, Département de neurologie Pitié-Salpêtrière, AP-HP, Paris, France ¹³Department of Neurology, Klinikum Rechts der Isar der Technischen Universität, Ismaninger Strasse 22, 81675 Munich, Germany ¹⁴Department of Neurology, University Medicine Mainz, Johannes Gutenberg University Mainz, Langenbeckstr. 1, 55131 Mainz, Germany ¹⁵Max Delbrueck Center for Molecular Medicine, Robert-Rössle-Str. 10, 13092 Berlin, Germany ¹⁶Institute for Neuroimmunology and Clinical MS Research (inims), Centre for Molecular Neurobiology, Falkenried 94, D-20251 Hamburg, Germany ¹⁷Department of Clinical Pharmacology, University of Oxford, Old Road Campus Research Building, Old Road Campus, Oxford, OX3 7DQ, UK ¹⁸Queen's University Belfast, University Road, Belfast, BT7 1NN, Northern Ireland, UK ¹⁹Department of Medical Sciences and Interdisciplinary Research Center of Autoimmune Diseases (IRCAD), University of Eastern Piedmont, Novara, Italy ²⁰Department of Neurology, Institute of Experimental Neurology (INSPE), Division of Neuroscience, San Raffaele Scientific Institute, Via Olgettina 58, 20132, Milan, Italy ²¹Department of Neurology, Oslo University Hospital, N-0407 Oslo, Norway ²²Department of Neurology, University of Oslo, N-0318 Oslo, Norway ²³Institute of Basal Medical Sciences, University of Oslo, N-0317 Oslo, Norway ²⁴Department of Neurology, Laboratory of

Neuroimmunology, Medical University of Lodz, Kopcynskiego 22, 90-153 Lodz, Poland ²⁵Clinical Neuroimmunology Unit, Multiple Sclerosis Center of Catalonia (CEM-Cat), Vall d'Hebron University Hospital, Barcelona, Spain ²⁶Department of Clinical Neurosciences, Centre for Molecular Medicine CMM, L8:04, Karolinska Institutet, Karolinska Hospital, 171 76 Stockholm, Sweden ²⁷Keele University Medical School, Stoke-on-Trent, UK ²⁸Peninsula College of Medicine and Dentistry, Universities of Exeter and Plymouth, Clinical Neurology Research Group, Tamar Science Park, Plymouth, PL6 8BX, UK ²⁹Department of Neurology, University Hospital of Wales, Heath Park, Cardiff, CF14 4XW, UK ³⁰Genetic Epidemiology and Genomics Laboratory, Division of Epidemiology, School of Public Health, University of California, Berkeley, CA 94720-7356, USA ³¹Kaiser Permanente Northern California Division of Research, 2000 Broadway, Oakland, CA, 94612, USA ³²Institute of Environmental Medicine, Karolinska Institutet, Box 210, 171 77 Stockholm, Sweden ³³Department of Neurology, University of California San Francisco, 505 Parnassus Avenue, S-256, San Francisco, CA 94143-0435, USA ³⁴Neurological Institute C. Mondino, IRCCS, Pavia, Italy ³⁵The Walton Centre for Neurology and Neurosurgery, Liverpool, UK ³⁶Center for Applied Genomics, The Children's Hospital of Philadelphia, 3615 Civic Center Blvd., Philadelphia, PA, 19104, USA ³⁷INSERM U 563 et Pôle Neurosciences, Hopital Purpan, Toulouse, France ³⁸School of Medicine, Griffith University, Australia ³⁹Florey Neuroscience Institutes, University of Melbourne, Victoria, Australia 3010 ⁴⁰Royal Melbourne Hospital, Parkville, Victoria, Australia, 3050 ⁴¹Box Hill Hospital, Box Hill 3128, Australia ⁴²Department of Medicine, RMH Cluster, University of Melbourne, Victoria, Australia 3010 ⁴³Multiple Sclerosis Centre, Department of Neurology, Ospedali Civili di Brescia, Brescia, Italy ⁴⁴Centre for Neuromuscular and Neurological Disorders, University of Western Australia, Perth WA 6009, Australia ⁴⁵Department of Neurosciences, University of Turin, A.O.U. San Giovanni Battista, Turin, Italy ⁴⁶INSERM U535, Univ Paris-Sud, Villejuif, France ⁴⁷University of Newcastle, University Drive, Callaghan NSW 2308, Australia ⁴⁸Department of Preventive Medicine, Keck School of Medicine, University of Southern California, 1540 Alcazar St. NOR 4453, Los Angeles, CA 90033 ⁴⁹Department of Neurology, Washington University, St Louis MO, USA ⁵⁰University of Milan, Department of Medicine, Surgery and Dentistry, AO San Paolo, University of Milan, c/o Filarete Foundation - Viale Ortles 22/4 - 20139 Milano, Italy ⁵¹Harvard Medical School, Boston, MA, USA ⁵²Center for Human Genetic Research, Massachusetts General Hospital, USA ⁵³The UK DNA Banking Network, Centre for Integrated Genomic Medical Research, University of Manchester, UK ⁵⁴Department of Medical Genetics, Division of Biomedical Genetics, University Medical Center Utrecht, Utrecht, The Netherlands ⁵⁵Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, The Netherlands ⁵⁶Service de Neurologie, Hôpital Central, Nancy, France ⁵⁷National Multiple Sclerosis Center, 1820 Melsbroek, Belgium ⁵⁸Institute for Clinical Molecular Biology, Christian-Albrechts-University, Kiel, Germany ⁵⁹Department of Neurology, Tampere University Hospital, Tampere, Finland ⁶⁰University of Tampere, Medical School, Tampere, Finland ⁶¹Menzies Research Institute, Locked Bag 23, Hobart, Tasmania, Australia 7000 ⁶²Department of Neurological Sciences, Centro Dino Ferrari, University of Milan, Fondazione Cà Granda, Ospedale Maggiore Policlinico, Milan, Italy ⁶³Centro Studi Sclerosi Multipla, Ospedale di Gallarate, Gallarate (VA), Italy ⁶⁴Service de Neurologie, Fondation Ophthalmologique Adolphe de Rothschild, Paris, France ⁶⁵Belfast Health and Social Care Trust, City Hospital, Belfast BT9 7AB, Northern Ireland, UK ⁶⁶Division of Genetics, The Children's Hospital of Philadelphia, 3615 Civic Center Blvd., Philadelphia, PA, 19104, USA ⁶⁷Department of Pediatrics, University of Pennsylvania School of Medicine, 3615 Civic Center Blvd., Philadelphia, PA, 19104, USA ⁶⁸Laboratory of Molecular Medicine and Biotechnology, Don C. Gnocchi Foundation IRCCS, S. Maria Nascente, Milan, Italy ⁶⁹Department of Medical Epidemiology and Biostatistics, Karolinska Institute, 17177 Stockholm, Sweden ⁷⁰Atherosclerosis Research Unit, Department of Medicine Solna, Karolinska Institutet, Center for Molecular Medicine, L8:03, Karolinska University Hospital Solna, S-171 76 Stockholm, Sweden ⁷¹Department of Neurology, Heinrich-Heine-University, Düsseldorf, Germany ⁷²Centre National de Genotypage, 2 rue Gaston

Cremieux, CO 5721, 91057 Evry Cedex, France ⁷³Experimental and Clinical Research Center, Charité – Universitätsmedizin Berlin and Max Delbrueck Center for Molecular Medicine, Berlin, Germany ⁷⁴Clinic for Paediatric Pneumology, Allergology and Neonatology, Hannover Medical School, Germany ⁷⁵Centre for Neuroscience, University of Melbourne, Victoria, Australia 3010 ⁷⁶Institute of Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, Ingolstädter Landstrasse 1, 85764 Neuherberg, Munich, Germany ⁷⁷Seinäjäki Central Hospital, Seinäjoki, Finland ⁷⁸Hunter Medical Research Institute, John Hunter Hospital, Lookout Road, New Lambton NSW 2305, Australia ⁷⁹SCDU Neurology, Maggiore della Carità Hospital, Novara, Italy ⁸⁰Unit of Public Health Genomics, National Institute for Health and Welfare, Helsinki, 00290, Finland ⁸¹Molecular Medicine, Department of Medical Sciences, Uppsala University, Entrance 70, 3rd Floor, Res Dept 2, Univeristy Hospital, S-75185, Uppsala, Sweden ⁸²Human Genetics and Cancer Biology, Genome Institute of Singapore, Singapore 138672 ⁸³Institute of Immunology, Oslo University Hospital, N-0027 Oslo, Norway ⁸⁴Institute of Experimental Neurology (INSPE), San Raffaele Scientific Institute, Via Olgettina 58, 20132, Milan, Italy ⁸⁵Dept of Psychiatry and Human Behavior, University of California, Irvine (UCI), 5251 California Av, S.te 240, Irvine CA, 92617 - USA ⁸⁶Christchurch School of Medicine, University of Otago, Christchurch, New Zealand ⁸⁷John P. Hussman Institute for Human Genomics and The Dr. John T Macdonald Foundation Department of Human Genetics, University of Miami, Miller School of Medicine, 1501 NW 10th Avenue, Miami, FL 33136, USA ⁸⁸Greater Manchester Centre for Clinical Neurosciences, Hope Hospital, Salford, UK ⁸⁹The Department of Neurology, Dunedin Public Hospital, Otago, NZ ⁹⁰The Multiple Sclerosis National Competence Centre, Department of Neurology, Haukeland University Hospital, N-5021 Bergen, Norway ⁹¹Department of Clinical Medicine, University of Bergen, N-5021 Bergen, Norway ⁹²Plymouth Hospitals NHS Trust, Department of Neurology, Derriford Hospital, Plymouth, PL6 8DH, UK ⁹³Department of Medical Genetics, University of Helsinki and University Central Hospital, Helsinki, Finland ⁹⁴Program in Medical and Population Genetics and Genetic Analysis Platform, The Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA ⁹⁵Pôle Neurosciences Cliniques, Service de Neurologie, Hôpital de la Timone, Marseille, France ⁹⁶Department Neurology, Oulu University Hospital, Oulu, Finland ⁹⁷UK MS Tissue Bank, Wolfson Neuroscience Laboratories, Imperial College London, Hammersmith Hospital, London, W12 0NN ⁹⁸Université de Montréal & Montreal Heart Institute, Research Center, 5000 rue Belanger, Montreal, Quebec H1T 1C8, Canada ⁹⁹Neurology and Center for Experimental Neurological Therapy (CENTERS), Sapienza University of Rome, Italy ¹⁰⁰KOS Genetic Srl, Via Podgora, 7 - 20123 Milan - Italy ¹⁰¹Department of General Internal Medicine, University Hospital, Schleswig-Holstein, Christian-Albrechts-University, Kiel, Germany ¹⁰²Systems Biology and Protein-Protein Interaction, Center for Molecular Neurobiology, Falkenried 94, D-20251 Hamburg, Germany ¹⁰³Center for Human Genetics Research, Vanderbilt University Medical Center, 519 Light Hall, Nashville, TN 37232, USA ¹⁰⁴Telethon Institute for Child Health Research, Centre for Child Health Research, University of Western Australia, 100 Roberts Road, Subiaco, Western Australia 6008 ¹⁰⁵Cambridge Institute for Medical Research, University of Cambridge School of Clinical Medicine, Cambridge CB2 0XY, UK ¹⁰⁶Department of Neurology, Helsinki University Central Hospital and Molecular Neurology Programme, Biomedicum, University of Helsinki, Helsinki, Finland ¹⁰⁷Division of Psychological Medicine and Psychiatry, Biomedical Research Centre for Mental Health at the Institute of Psychiatry, King's College London and The South London and Maudsley NHS Foundation Trust, Denmark Hill, London SE5 8AF, UK ¹⁰⁸Service de Neurologie et Faculté de Médecine de Reims, Université de Reims Champagne-Ardenne, Reims, France ¹⁰⁹University of Queensland Diamantina Institute, Princess Alexandra Hospital, Brisbane, Australia ¹¹⁰Dept Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, London WC1E 7HT, UK ¹¹¹St. Vincent's University Hospital, Dublin, Ireland ¹¹²Neuropsychiatric Genetics Research Group, Institute of Molecular Medicine, Trinity College Dublin, Dublin 2, Eire ¹¹³Centre for Gastroenterology, Bart's and the London School of Medicine and Dentistry, London E1 2AT, UK ¹¹⁴Department of

Neurosciences, Institute of Biomedical Research August Pi Sunyer (IDIBAPS), Hospital Clinic of Barcelona, Spain ¹¹⁵Clinical Neurosciences, St George's University of London, London SW17 0RE ¹¹⁶Dept Medical and Molecular Genetics, King's College London School of Medicine, Guy's Hospital, London SE1 9RT, UK ¹¹⁷Medical Research Council Biostatistics Unit, Robinson Way, Cambridge, CB2 0SR, UK ¹¹⁸Biomedical Research Institute, University of Dundee, Ninewells Hospital and Medical School, Dundee, DD1 9SY ¹¹⁹Institute of Medical Informatics, Biometry and Epidemiology, Ludwig-Maximilians-Universität, 81377 Munich, Germany ¹²⁰Klinikum Grosshadern, Munich, Germany ¹²¹King's College London, Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, Denmark Hill, London SE5 8AF, UK ¹²²Department of Neurology, Auckland City Hospital, Grafton Road, Auckland, New Zealand ¹²³Institut für Humangenetik, Technische Universität München, Germany ¹²⁴Institut für Humangenetik, Helmholtz Zentrum München, Germany ¹²⁵Popgen Biobank, Christian-Albrechts University Kiel, Kiel, Germany ¹²⁶Pôle Recherche et Santé Publique, CHU Pontchaillou, Rennes, France ¹²⁷NIHR Biomedical Research Centre for Ophthalmology, Moorfields Eye Hospital NHS Foundation Trust and UCL Institute of Ophthalmology, London EC1V 2PD, UK ¹²⁸Dept Molecular Neuroscience, Institute of Neurology, Queen Square, London WC1N 3BG, UK ¹²⁹Molecular and Physiological Sciences, The Wellcome Trust, London NW1 2BE ¹³⁰Harvard NeuroDiscovery Center, Harvard Medical School, Boston, MA, USA ¹³¹Department of Neurology & Immunology, Yale University Medical School, New Haven, CT, USA

Abstract

Multiple sclerosis (OMIM 126200) is a common disease of the central nervous system in which the interplay between inflammatory and neurodegenerative processes typically results in intermittent neurological disturbance followed by progressive accumulation of disability.¹ Epidemiological studies have shown that genetic factors are primarily responsible for the substantially increased frequency of the disease seen in the relatives of affected individuals;^{2,3} and systematic attempts to identify linkage in multiplex families have confirmed that variation within the Major Histocompatibility Complex (MHC) exerts the greatest individual effect on risk.⁴ Modestly powered Genome-Wide Association Studies (GWAS)⁵⁻¹⁰ have enabled more than 20 additional risk loci to be identified and have shown that multiple variants exerting modest individual effects play a key role in disease susceptibility.¹¹ Most of the genetic architecture underlying susceptibility to the disease remains to be defined and is anticipated to require the analysis of sample sizes that are beyond the numbers currently available to individual research groups. In a collaborative GWAS involving 9772 cases of European descent collected by 23 research groups working in 15 different countries, we have replicated almost all of the previously suggested associations and identified at least a further 29 novel susceptibility loci. Within the MHC we have refined the identity of the *DRB1* risk alleles and confirmed that variation in the *HLA-A* gene underlies the independent protective effect attributable to the Class I region. Immunologically relevant genes are significantly over-represented amongst those mapping close to the identified loci and particularly implicate T helper cell differentiation in the pathogenesis of multiple sclerosis.

Keywords

multiple sclerosis; GWAS; genetics

We performed a large GWAS as part of the Wellcome Trust Case Control Consortium 2 (WTCCC2) project. Cases were recruited through the International Multiple Sclerosis Genetics Consortium (IMSGC) and compared with the WTCCC2 common control set^{12,13} supplemented by data from the control arms of existing GWAS. We introduced a number of

novel quality control (QC) methods for processing these datasets (see Supplementary Information), which ultimately provided reliable information from 9772 cases and 17376 controls (see Figure 1A). Following single nucleotide polymorphism (SNP) based QC, data from 441547 autosomal SNPs, common to all internally and externally generated datasets, were available for analysis.

The multi-population nature of our study (Figure 1 A and B) afforded an opportunity to assess various published approaches for controlling the potential confounding effects of population structure, several of which (in the event) proved unhelpful (see Supplementary Information). Whilst not common in primary GWAS undertaken to date, the challenge of combining data across populations, in contexts where not all case samples have controls available from the same population (thus precluding standard meta-analytical techniques), may become more routine as study sizes increase.

We attempted analyses of the non-United Kingdom (UK) data with the now widespread technique of using principal components as covariates to correct for structure. However, even use of all seven top principal components which captured genome-wide effects in our data resulted in an unacceptably high genomic inflation: for example, the genomic control factor (λ)¹⁴ was $\lambda = 1.2$. We tried to reduce the genomic inflation by discarding the case samples that seemed least well matched to control sets. Removal of half the available cases in this fashion only reduced λ to 1.1. In another approach to handling structure, statistical clustering algorithms were successful in identifying subgroups of the data within which cases and controls appeared well-matched for ancestry (See Supplementary figure S17). However tests within these subgroups combined via fixed-effects meta-analysis also yielded unacceptably high genomic inflation ($\lambda > 1.4$) in an analysis with seven matched sub-groups of cases and controls. Finally, we applied a novel variance components method (similar to Kang et al.¹⁵), separately to the UK and non-UK datasets, that explicitly accounts for correlations among the phenotypes of individuals resulting from relatedness, allowing us to deal successfully with all sources of structure in our samples (see Supplementary Information for details of the linear mixed model we used). For example, the genomic inflation was reduced to $\lambda = 0.995$ in the UK and 1.016 in the non-UK data (see also Supplementary Information). After fixed effects meta-analysis of the results from the UK and non-UK datasets, the inflation factor was $\lambda = 1.045$. We adopted this approach for all subsequent non-MHC association analyses.

Outside the MHC we identified 95 distinct regions having at least one SNP associated with multiple sclerosis at $p_{\text{GWAS}} < 1 \times 10^{-4.5}$; in six of these 95 regions conditional analysis revealed an additional SNP showing association to the same locus (one locus containing two such SNPs). In total we took all 102 SNPs forward to replication, which we performed using data from previously reported multiple sclerosis GWAS^{8,9} and the iControl database (excluding any WTCCC controls previously used in these studies). In total, the replication analysis included data from 4218 cases and 7296 controls. These were considered in six independent strata after which results were combined through a fixed effects meta-analysis. For 98 of the 102 SNPs, the same allele was over-represented in cases compared to controls. Twenty three of the 26 previously known or strongly suggested multiple sclerosis associated loci were replicated in our primary GWAS with $p_{\text{GWAS}} < 1 \times 10^{-3}$. Our GWAS and replication also revealed another 29 novel associated regions (defined as having $p_{\text{GWAS}} < 1 \times 10^{-4.5}$, one-sided $p_{\text{Replication}} < 0.05$, and $p_{\text{Combined}} < 5 \times 10^{-8}$), and a further 5 regions with strong evidence for association (with $p_{\text{GWAS}} < 1 \times 10^{-4.5}$, one sided $p_{\text{Replication}} < 0.05$, and $p_{\text{Combined}} < 5 \times 10^{-7}$). In one previously reported locus and two novel loci, additional SNPs were identified as being conditionally important in explaining risk. Just over one third of the identified loci overlap with regions already confirmed as associated with at least one other autoimmune disease (according to the GWAS catalog, <http://www.genome.gov/>

[gwastudies/](#)). Results both for the previously established and novel loci are shown in Figure 2 and Supplementary Tables 1-3; and details of all 102 SNPs taken to replication are available in the Supplementary Data file.

In order to assess objectively the collective evidence across the associated regions for particular classes of genes, we performed statistical analyses to look for enrichment of genes with similar function. We first identified the nearest gene to the lead SNP in each of the (52) regions of association and used the Gene Ontology (GO) database¹⁶ to define sets of functionally related genes (GO terms). We then tested whether the set of nearest-genes was enriched for particular GO terms using Fisher's exact test. The GO terms having the most significant enrichment include genes linked to lymphocyte function ($p = 3.2 \times 10^{-11}$, OR = 35.96) and in particular those with a role in T cell activation and proliferation ($p = 1.85 \times 10^{-9}$, OR = 40.85). These are representative of a larger group associated with various components of the GO 'immune system process' ($p = 8.6 \times 10^{-11}$, OR = 9.12). A similar analysis based on all genes in or near association regions showed similar enrichment, as did independent analyses based on nearest-gene or all genes in our next tier of signals, the 42 regions taken to replication but not meeting the thresholds above for association (see Supplementary file.) Although GO immune system genes only account for 7% of human genes, in 30% of our association regions the nearest gene to the lead SNP is an immune system gene. As an illustration, Figure 3 shows a schematic of genes involved in the T helper cell differentiation pathway; a striking number show strong evidence for association with multiple sclerosis particularly those acting as cell surface receptors. We infer from this pathway analysis of our GWAS signals that specific classes of immune system genes are especially important in the pathogenesis of multiple sclerosis.

Our screen not only implicates a multitude of genes coding for cytokine pathway (*CXCR5*, *IL2RA*, *IL7R*, *IL7*, *IL12RB1*, *IL22RA2*, *IL12A*, *IL12B*, *IRF8*, *TNFRSF1A*, *TNFRSF14*, *TNFSF14*), co-stimulatory (*CD37*, *CD40*, *CD58*, *CD80*, *CD86*, *CLECL1*) and signal transduction (*CBLB*, *GPR65*, *MALTI*, *RGS1*, *STAT3*, *TAGAP*, *TYK2*) molecules of immunological relevance, but also relates to previously reported environmental risk factors such as Vitamin D^{9,17} (*CYP27B1*, *CYP24A1*) and therapies for multiple sclerosis including Natalizumab¹⁸ (*VCAM1*) and Daclizumab¹⁹ (*IL2RA*). There is a relative absence of genes relevant to potential pathways for neurodegeneration independent of inflammation (*GALC*, *KIF21B*).

To refine our understanding of the MHC associations in multiple sclerosis we imputed classical Human Leukocyte Antigen (HLA) types at six loci (*A*, *B*, *C*, *DQA1*, *DQB1* and *DRB1*)²⁰ and analysed these alongside the SNPs (see Supplementary Information for validation; at alleles responsible for the major signals described below, estimated specificity was at least 0.99 and sensitivity was at least 0.98, except for DRB1*13:03, where it was 0.88). Primary discovery was focused on the UK cohort with candidate signals being validated through support from additional case-control cohorts. Because of the extensive linkage disequilibrium within the MHC, we identified associated alleles in a stepwise manner, selecting the most strongly associated to include in a general model, in turn, if $p_{UK} < 10^{-4}$ and $p_{combined} < 10^{-9}$ (Supplementary Information). At each stage we explored possible interactions and departures from the simple model in which risk increases multiplicatively with each additional copy of the relevant allele (additive increase on the log-odds scale) within the logistic risk framework.

By this approach we found that DRB1*15:01 has the strongest association with multiple sclerosis amongst all classical and SNP alleles, with a consistent effect between cohorts ($p < 1 \times 10^{-320}$; Figure 4A). The data are consistent with an additive effect on the log-odds scale for each additional allele. Conditioning on DRB1*15:01, we confirmed the presence of a

protective Class I allele and identified the signal as being driven by HLA-A*02:01 (as previously suggested),²¹ with a consistent effect size across cohorts ($p = 9.1 \times 10^{-23}$; Figure 4A). Again, we found no strong evidence for departure from additivity on the log-odds scale or statistical interaction with DRB1*15:01. Conditioning on both DRB1*15:01 and A*02:01 revealed additional risk associated with the strongly linked alleles DRB1*03:01 and DQB1*02:01 ($p = 3.6 \times 10^{-10}$; Figure 4A; note that we cannot separate these alleles but for simplicity refer only to DRB1*03:01 below). Further conditioning identified an additional *DRB1* risk allele DRB1*13:03 ($p = 1.3 \times 10^{-11}$; Figure 4A). Although no other classical alleles meet the above criteria, we did observe several SNPs providing independent signals, the strongest coming from rs9277535_G (combined OR 1.28, $p = 2.2 \times 10^{-22}$), an allele known to be in linkage disequilibrium with DPB1*03:01 ($r^2 = 0.37$).²²

Analysis of the MHC SNP data using a genealogical method (GENECLUSTER)²³ offers an alternate means of relating our results to classical HLA alleles that provides additional insight into the underlying genetic architecture (see Supplementary Information). Figure 4B shows genealogical trees relating the classical alleles at *DRB1* and *HLA-A*, together with the estimated evolutionary position of the mutations predicted by GENECLUSTER as most completely modelling the association. At *HLA-DRB1*, three mutations are predicted, each of which implicates a clade of haplotypes carrying particular *DRB1* alleles. All of the *DRB1* alleles we have shown to be independently associated are included in these clades, each corresponding to a particular mutation. In addition, the analysis also explains why those haplotypes carrying the *08:01 allele have previously been shown to increase risk^{24,25} since they carry the same mutation as those bearing *13:03. At *HLA-A*, the predicted protective mutation is also concordant with our regression analysis of classical alleles in implicating *02:01 but, in addition, predicts that *68:01, *02:05, and *02:06 carry the same protective allele. All of these secondary predictions (increased risk from DRB1*08:01 and protection from HLA-A*68:01, *02:05, and *02:06) are supported in our regression analysis of classical alleles but the power to detect them in the primary analyses is limited because each allele occurs at a very low frequency.

We found no evidence for genetic associations with clinical course, severity of disease or month of birth, and no evidence of interaction with gender or DRB1*15:01 in any part of the genome (see Supplementary Information). However, analysis with respect to age at onset replicated the previously suggested association with the DRB1*15:01 allele.²⁶ Although no other part of the genome contained individual SNPs showing strong evidence for association, risk alleles determining susceptibility are collectively more closely associated with age at onset than expected by chance, suggesting that individual genetic susceptibility is inversely correlated with age at onset.

Our GWAS - large for any complex trait having a prevalence of 1:1000 and involving diverse populations of European descent - has identified 29 novel susceptibility loci. Four mutations, one from Class I and three from Class II, with effects modelled in a simple multiplicative manner within and across loci are sufficient to account for most of the risk attributable to the MHC (see Supplementary file). Although our data do not address the issue of which components within the nervous system are initially damaged by the inflammatory response the over-representation of genes that influence T cell maturation provides independent and compelling evidence that the critical disease mechanisms primarily involve immune dysregulation.

More generally, our study reinforces the view that the GWAS design, combined with very large experimental sample sizes and careful statistical analysis, provides valuable insights into the genetic architecture of common complex diseases. Here, this approach has identified many associated genetic variants close to genes, which are both individually interesting and

collectively illuminate the roles of key biological pathways. It also provides indirect evidence that many more common variants of small effect contribute to genetic susceptibility for multiple sclerosis. Simple models, in which the previously-known and newly-identified variants affect risk multiplicatively, both within and across loci, explain a meaningful proportion (~20%, see Supplementary Information) of genetic risk for the disease. Important challenges lie ahead, in understanding overlap between the genetic basis for susceptibility in the context of different autoimmune diseases, and in uncovering the functional mechanisms underlying these associations.

Materials and Methods

Details of case ascertainment, processing and genotyping, together with sample and genotyping quality control are provided in the Supplementary Information. Statistical methods developed for testing the reliability of externally generated data sets, detecting samples with non-European ancestry, correcting for structure, classical HLA imputation and meta-analysis are also outlined in the Supplementary Information. Results for all scans and all reported loci are described in detail in the Supplementary Information.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The principal funding for this study was provided by the Wellcome Trust (085475/B/08/Z, 085475/Z/08/Z, 075491/Z/04/Z and 068545/Z/02). The work was also supported by National Institutes of Health (AI076544, NS032830, NS049477, NS19142, NS049510, NS26799, NS43559, NS067305, CA104021, RR020092, RR024992 and K23N/S048869), US National Multiple Sclerosis Society (RG 4201-A-1), Nancy Davis Foundation, Cambridge NIHR Biomedical Research Centre, UK Medical Research Council (G0700061, G0000934), Multiple Sclerosis Society of Great Britain and Northern Ireland (898/08), Wolfson Royal Society Merit Award, Peter Doherty fellowship, Lagrange Fellowship, Harry Weaver Neuroscience Scholarships, Australian National Health and Medical Research Council (NHMRC), Australian Research Council Linkage Program Grant, JHH Charitable Trust Fund, Multiple Sclerosis Research Australia, Health Research Council New Zealand, National MS Society of New Zealand, Wetenschappelijk Onderzoek Multiple Sclerose, Bayer Chair on Fundamental Genetic Research regarding the Neuroimmunological aspects of Multiple Sclerosis, Biogen Idec Chair Translational Research in Multiple Sclerosis, FWO-Vlaanderen, Belgian Neurological Society, Danish Multiple Sclerosis Society, Neuropromise EU grant (LSHM-CT-2005-018637), Center of Excellence for Disease Genetics of the Academy of Finland, Sigrid Juselius Foundation, Helsinki University Central Hospital Research Foundation, Bundesministerium für Bildung und Technologie (KKNMS consortium Control MS), Deutsche Forschungsgemeinschaft, Institut National de la Santé et de la Recherche Médicale (INSERM), Association pour la Recherche sur la Sclérose En Plaques (ARSEP), Association Française contre les Myopathies (AFM), Italian Foundation for Multiple Sclerosis (2002/R/40, 2005/R/10, 2008/R/11 and 2008/R/15), Italian Ministry of Health (grant Giovani Ricercatori 2007 - D.lgs 502/92), Regione Piemonte (grants 2003, 2004, 2008, 2009), CRT Foundation, Turin, Moorfields / UCL Institute of Ophthalmology NIHR Biomedical Research Centre, Norwegian MS Register and Biobank, Research Council of Norway, South-Eastern and Western Norway regional Health Authorities, Ullevål University Hospital Scientific Advisory Council, Haukeland University Hospital, Amici Centro Sclerosi Multipla del San Raffaele (ACESM), Association of British Neurologists, Spanish Ministry of Health (FISPI060117), Bibbi and Niels Jensens Foundation, Montel Williams foundation, Hjärnfonden and Swedish medical research council (8691), Stockholm County Council (562183), Swedish Council for Working life and Social Research, Gemeinnützige Hertie Stiftung, Northern California Kaiser Permanente members and Polpharma Foundation, and Washington University Institute of Clinical and Translational Sciences - Brain, Behavioral and Performance Unit.

We acknowledge use of data from the British 1958 Birth Cohort, the UK National Blood Service, the popgen biobank, the KORA and MONICA Augsburg studies, the Accelerated Cure Project, the Brigham & Women's Hospital PhenoGenetic Project, the Swedish CAD project, the Norwegian Bone Marrow Donor Registry, the Children's Hospital of Philadelphia (CHOP), the Swedish Breast Cancer study, BRC-REFGENSEP (Pitié-Salpêtrière Centre d'Investigation Clinique (CIC) and Généthon) and HYPERGENES (HEALTH-F4-2007-201550). Projects which received support from the German Ministry of Education and Research, the Helmholtz Zentrum München–National Research Center, the German National Genome Research Network (NGFN), the LMUinnovativ, the Knut and Alice Wallenberg Foundation, the Center for Applied Genomics from

the Children's Hospital of Philadelphia Development Award, the Agency for Science & Technology and Research of Singapore and the Susan G Komen Breast Cancer Foundation.

We thank S. Bertrand, J. Bryant, S.L. Clark, L. Collimedaglia, G. Coniglio, J.S. Conquer, B. Colombo, T. Dibling, G. Eckstein, J.C. Eldred, G. Fischer, S. Gamble, P. Gregersen, R. Guerrero, C. Hind, P. Lichtner, L. Moiola, H. Mousavi, R. Naismith, R.J. Parks, R. Pearson, V. Pilato, M. Radaelli, E. Scarpini, C.R. Stribling, T. Strom, S. Taylor, D. Vukcevic, A. Wilk, for their help and support.

Detailed acknowledgements are available in the supplementary information file.

This manuscript is dedicated to the memory of Leena Peltonen, a member of both the IMSGC and WTCCC2, in recognition of her many contributions to, and her leadership in, human genetics.

References

1. Compston A, Coles A. Multiple sclerosis. *Lancet*. 2008; 372:1502–1517. [PubMed: 18970977]
2. Dyment DA, Yee IM, Ebers GC, Sadovnick AD. Multiple sclerosis in stepsiblings: recurrence risk and ascertainment. *J Neurol Neurosurg Psychiatry*. 2006; 77:258–259. [PubMed: 16421134]
3. Hemminki K, Li X, Sundquist J, Hillert J, Sundquist K. Risk for multiple sclerosis in relatives and spouses of patients diagnosed with autoimmune and related conditions. *Neurogenetics*. 2009; 10:5–11. [PubMed: 18843511]
4. The International Multiple Sclerosis Genetics Consortium (IMSGC). A high-density screen for linkage in multiple sclerosis. *Am J Hum Genet*. 2005; 77:454–467. [PubMed: 16080120]
5. The International Multiple Sclerosis Genetics Consortium (IMSGC). Risk Alleles for Multiple Sclerosis Identified by a Genomewide Study. *N Engl J Med*. 2007; 357:851–862. [PubMed: 17660530]
6. The Wellcome Trust Case Control Consortium (WTCCC). Association scan of 14,500 nonsynonymous SNPs in four diseases identifies autoimmunity variants. *Nat Genet*. 2007; 39:1329–1337. [PubMed: 17952073]
7. Baranzini SE, et al. Genome-wide association analysis of susceptibility and clinical phenotype in multiple sclerosis. *Hum Mol Genet*. 2009; 18:767–778. [PubMed: 19010793]
8. De Jager PL, et al. Meta-analysis of genome scans and replication identify CD6, IRF8 and TNFRSF1A as new multiple sclerosis susceptibility loci. *Nat Genet*. 2009; 41:776–782. [PubMed: 19525953]
9. The ANZgene Consortium. Genome-wide association study identifies new multiple sclerosis susceptibility loci on chromosomes 12 and 20. *Nat Genet*. 2009; 41:824–828. [PubMed: 19525955]
10. Sanna S, et al. Variants within the immunoregulatory CBLB gene are associated with multiple sclerosis. *Nat Genet*. 2010; 42:495–497. [PubMed: 20453840]
11. The International Multiple Sclerosis Genetics Consortium (IMSGC). Evidence for polygenic susceptibility to multiple sclerosis--the shape of things to come. *Am J Hum Genet*. 2010; 86:621–625. [PubMed: 20362272]
12. The UK Parkinson's Disease Consortium and the Wellcome Trust Case Control Consortium 2. Dissection of the genetics of Parkinson's disease identifies an additional association 5' of SNCA and multiple associated haplotypes at 17q21. *Hum Mol Genet*. 2011; 20:345–353. [PubMed: 21044948]
13. The Genetic Analysis of Psoriasis (GAP) Consortium and the Wellcome Trust Case Control Consortium 2. A genome-wide association study identifies new psoriasis susceptibility loci and an interaction between HLA-C and ERAP1. *Nat Genet*. 2010; 42:985–990. [PubMed: 20953190]
14. Devlin B, Roeder K. Genomic control for association studies. *Biometrics*. 1999; 55:997–1004. [PubMed: 11315092]
15. Kang HM, et al. Variance component model to account for sample structure in genome-wide association studies. *Nat Genet*. 2010; 42:348–354. [PubMed: 20208533]
16. Ashburner M, et al. The Gene Ontology Consortium. Gene ontology: tool for the unification of biology. *Nat Genet*. 2000; 25:25–29. [PubMed: 10802651]
17. Pierrot-Deseilligny C, Souberbielle JC. Is hypovitaminosis D one of the environmental risk factors for multiple sclerosis? *Brain*. 2010; 133:1869–1888. [PubMed: 20584945]

18. Steinman L. A molecular trio in relapse and remission in multiple sclerosis. *Nat Rev Immunol.* 2009; 9:440–447. [PubMed: 19444308]
19. Bielekova B, et al. Effect of anti-CD25 antibody daclizumab in the inhibition of inflammation and stabilization of disease progression in multiple sclerosis. *Arch Neurol.* 2009; 66:483–489. [PubMed: 19364933]
20. Leslie S, Donnelly P, McVean G. A statistical method for predicting classical HLA alleles from SNP data. *Am J Hum Genet.* 2008; 82:48–56. [PubMed: 18179884]
21. Brynedal B, et al. HLA-A confers an HLA-DRB1 independent influence on the risk of multiple sclerosis. *PLoS One.* 2007; 2:e664. [PubMed: 17653284]
22. Field J, et al. A polymorphism in the HLA-DPB1 gene is associated with susceptibility to multiple sclerosis. *PLoS One.* 2010; 5:e13454. [PubMed: 21049023]
23. Su Z, Cardin N, Donnelly P, Marchini J. A Bayesian method for detecting and characterizing allelic heterogeneity and boosting signals in genome-wide association studies. *Statistical Science.* 2009; 24:430–450.
24. Barcellos LF, et al. Heterogeneity at the HLA-DRB1 locus and risk for multiple sclerosis. *Hum Mol Genet.* 2006; 15:2813–2824. [PubMed: 16905561]
25. Dymant DA, et al. Complex interactions among MHC haplotypes in multiple sclerosis: susceptibility and resistance. *Hum Mol Genet.* 2005; 14:2019–2026. [PubMed: 15930013]
26. Masterman T, et al. HLA-DR15 is associated with lower age at onset in multiple sclerosis. *Ann Neurol.* 2000; 48:211–219. [PubMed: 10939572]

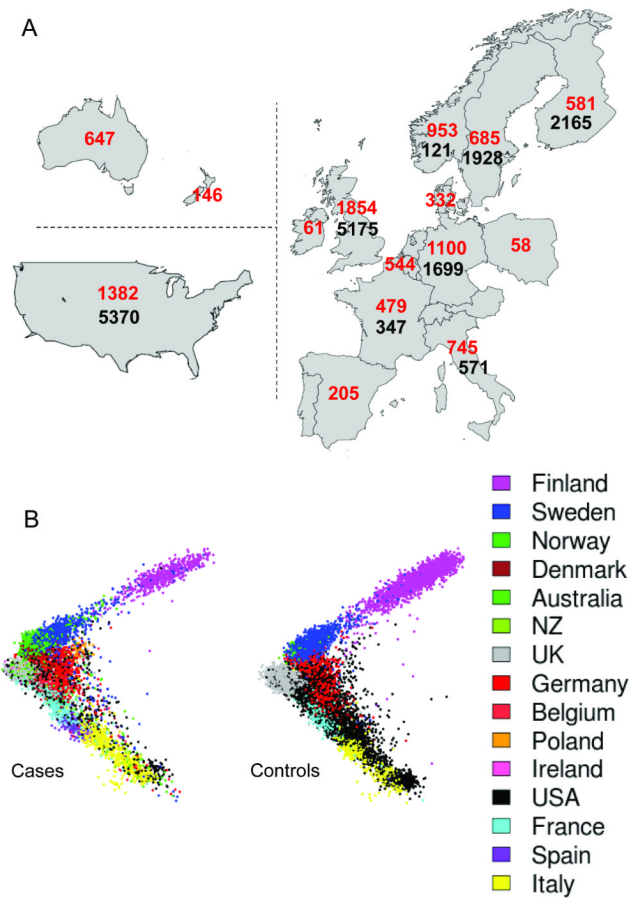


Figure 1.

Distribution of cases and controls. All cases and controls were drawn from populations with European ancestry; cases from 15 countries and controls from 8. **A:** numbers of case (red) and control (black) samples from each country. **B:** The projection of samples onto the first two principal components of genetic variation, with cases shown on the left and controls on the right. The axes are orientated to approximate to the geography, and samples are colour coded as indicated in the legend. We genotyped the cases (9772) and some Swedish controls (527) using the Illumina Human 660-Quad platform, and the UK controls (5175, the WTCCC2 common control set^{12,13}) using the Illumina 1.2M platform. All other controls were genotyped externally using various Illumina genotyping systems (see Supplementary Information).

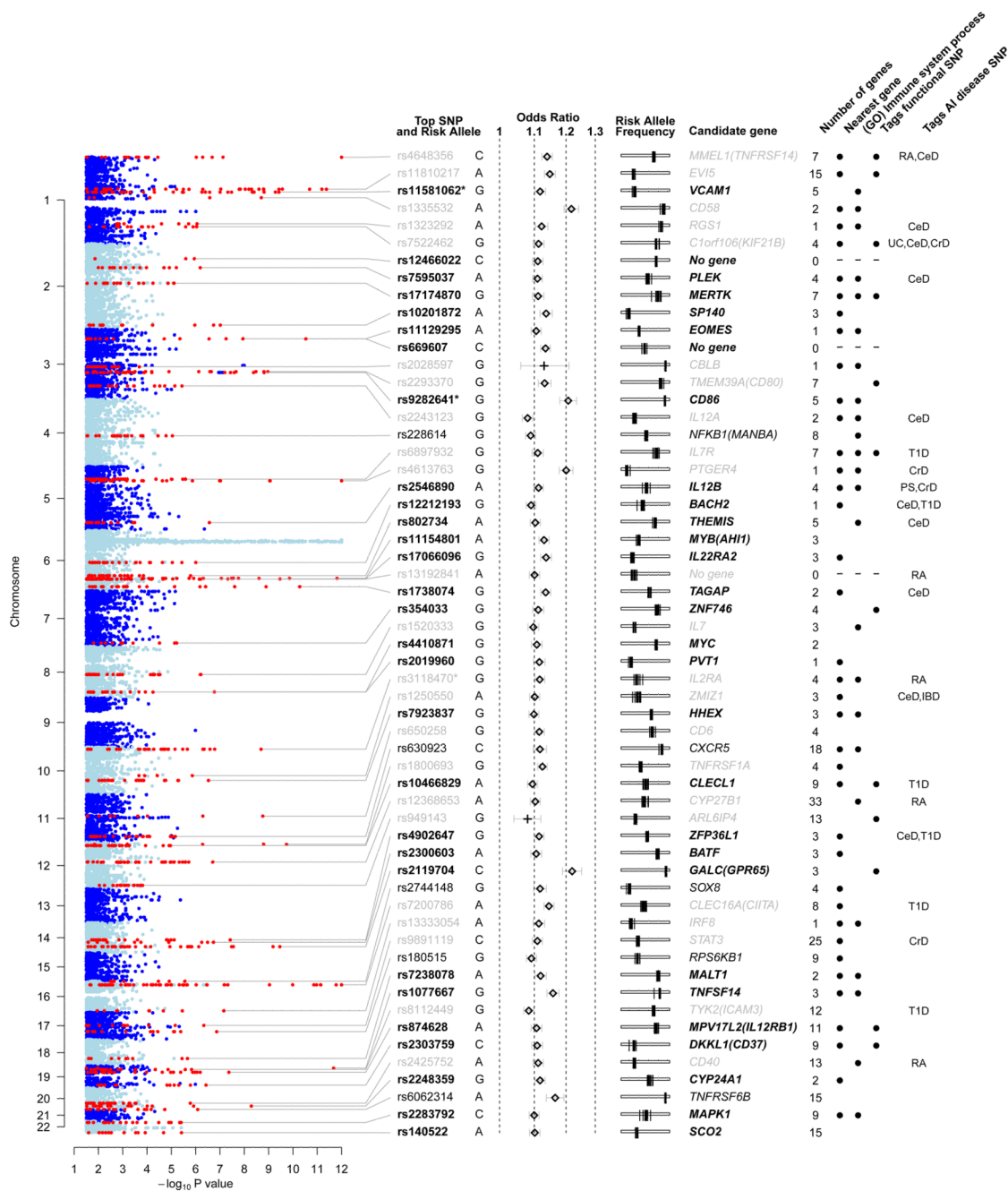


Figure 2. Regions of the genome showing association to multiple sclerosis. Columns from left to right: evidence for association from the linear mixed model analysis of the discovery data (thresholded at $-\log_{10}(p) = 12$). Non-MHC regions containing associated SNPs are shown in red and are labelled with the rsID (bold for newly identified loci, black for strong evidence, grey for previously reported) and risk allele of the most significant SNP. * indicates that the locus contains a secondary SNP signal. Odds ratio and 95% confidence intervals estimated from the meta-analysis of the discovery and replication data (+ indicates estimates for previously-known loci from discovery data only). Risk allele frequency estimates in each of the control populations used in the study (each is shown as a vertical bar on a scale from 0 to

1 going left to right). For each region of association the number of genes is reported, and where non-zero a candidate gene is given. Black dots indicate that the candidate gene is physically the nearest gene included in the “immune system process” GO term. When the most-significant SNP tags a SNP predicted to have an impact on the function of the candidate gene this is indicated. Where such a SNP exists, the gene involved is selected as the candidate gene; otherwise the nearest gene is selected unless there are strong biological reasons for a different choice. The final column indicates SNPs which are correlated ($r^2 > 0.1$) with SNPs reported to be associated with other autoimmune (AI) diseases (abbreviations: RA = Rheumatoid arthritis; CeD = Celiac disease; UC = Ulcerative colitis; CrD = Crohns’s disease; T1D = Type 1 diabetes; PS = Psoriasis). An interactive version of the figure is available at www.well.ox.ac.uk/wtccc2/ms.

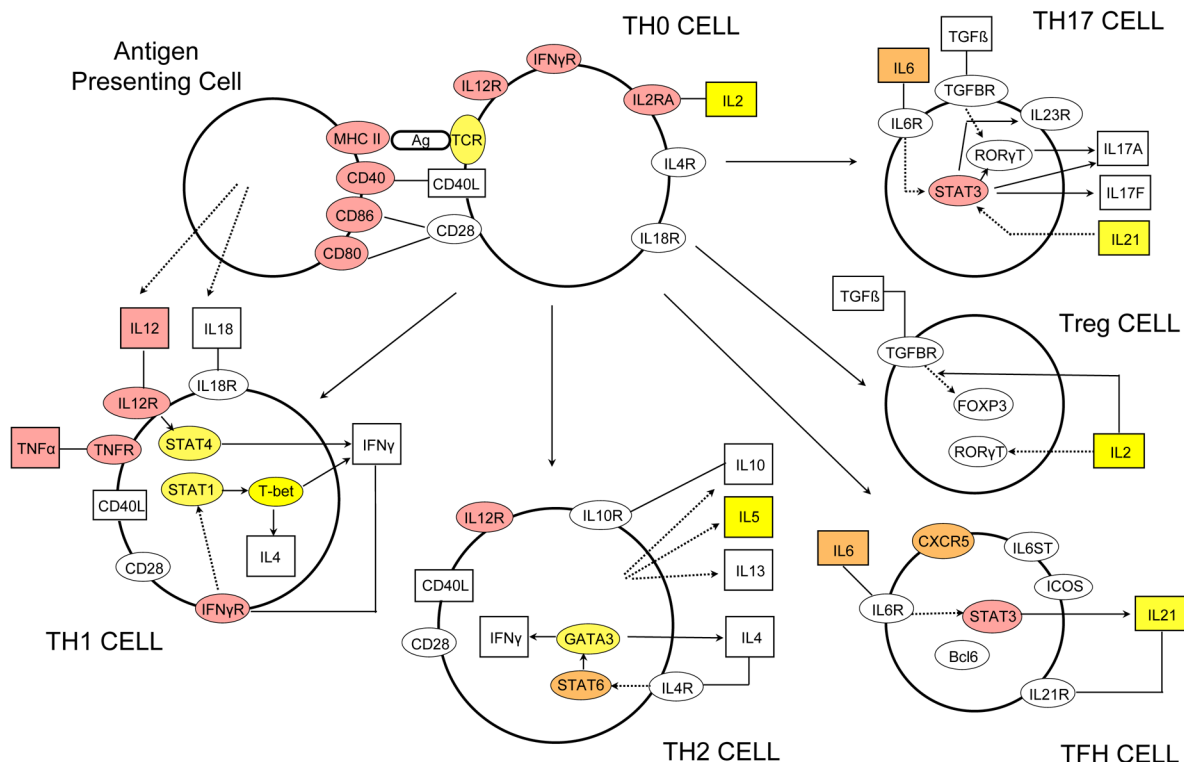


Figure 3. Graphical representation of the T helper cell differentiation pathway. The figure is derived from an image generated by Ingenuity Pathway Analysis (IPA) software version 8.8 (Ingenuity Systems, Inc., Redwood City, CA, USA). Alpha-numeric labels indicate the individual genes and gene complexes (nodes) included in the pathway (note some are included more than once). Coloured nodes are those containing a gene implicated by proximity to a SNP showing evidence of association. Red: in bold or grey in Figure 2 (plus MHC class II region and TNF α); Orange: other loci in Figure 2 or discovery P value < $1 \times 10^{-4.5}$ and consistent replication data. Yellow: Discovery P value < 1×10^{-3} . Other molecules (proteins, vitamins etc.) may also be of relevance in these processes but are not included here as they are not currently listed as being part of this particular pathway in the IPA database.

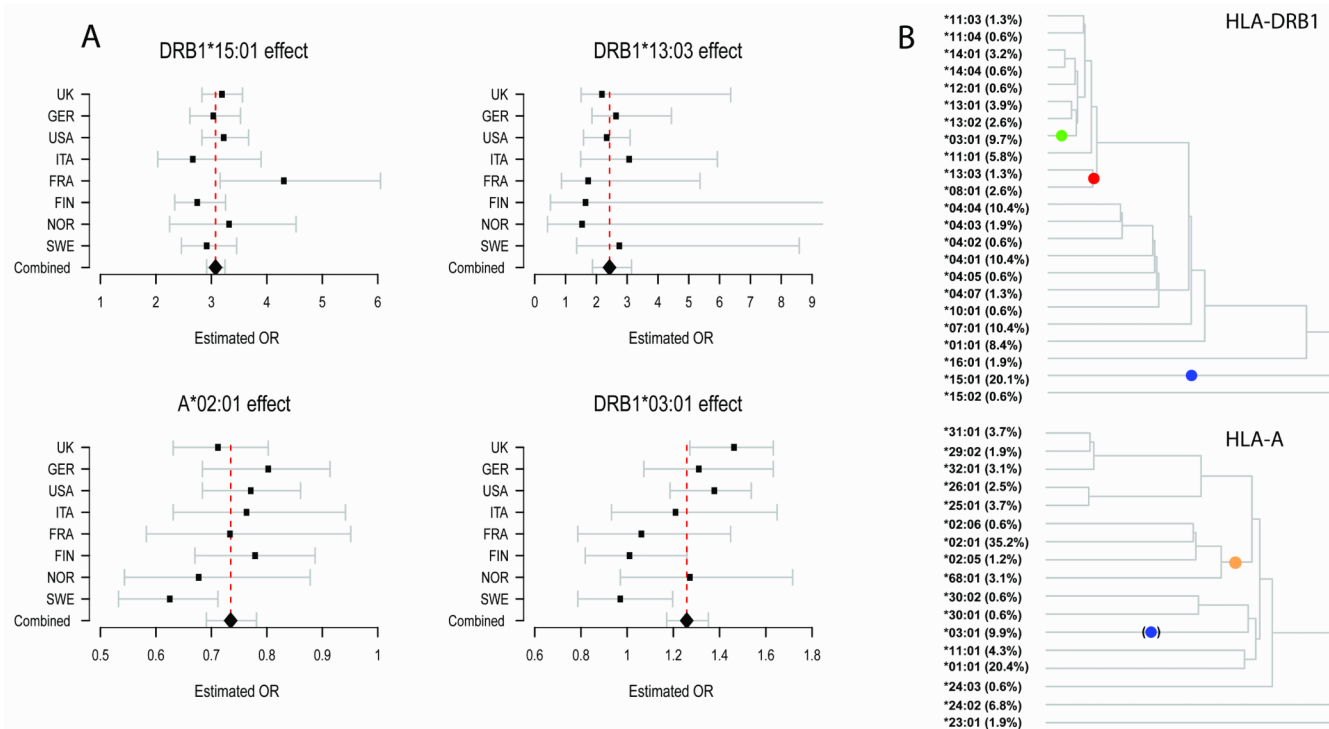


Figure 4. Results for the main MHC alleles. **A:** Forest plots for each of the primary HLA alleles (HLA-A*02:01, DRB1*15:01, DRB1*03:01 and DRB1*13:03) showing consistency of effect across the populations and combined OR of 0.73, 3.1, 1.26 and 2.4 respectively (whiskers indicate 95% confidence intervals). **B:** The genealogical trees estimated for *DRB1* (top) and *HLA-A* (bottom). These trees were constructed using classical HLA and SNP typing data available from the HapMap CEU haplotype data. Each left hand branch of the tree terminates on a set of haplotypes carrying a particular HLA allele. The coloured dots indicate the mostly likely locations for a disease-associated mutation as predicted by the GENECLUSTER program²³. In the *DRB1* tree, the blue dot captures a risk effect attributable to all haplotypes carrying the *15:01 allele. The green dot captures a risk effect carried by all haplotypes carrying the *03:01 allele and the red dot captures a risk effect on haplotypes carrying *13:03 or *08:01. In the *HLA-A* plot, the orange dot is a protective mutation lying at the root of all *02:01, *02:05, *02:06 and *68:01 alleles. The blue dot in brackets denotes a branch containing those *03:01 haplotypes that also carry DRB1*15:01; the GENECLUSTER prediction here is thus a reflection, due to linkage disequilibrium of the risk attributable to DRB1*15:01. The terminal branches are labelled with the allele carried by the haplotype and its frequency.