COMMUNICATIONS

Genome-wide significant association of *ANKRD55* rs6859219 and multiple sclerosis risk

Christina M Lill,[1,2] Brit-Maren M Schjeide,[1] Christiane Graetz,[1,2] Tian Liu,[3] Vincent Damotte,[4] Denis A Akkad,[5] Paul Blaschke,[6] Lisa-Ann Gerdes,[7] Antje Kroner,[8,9] Felix Luessi,[2] Isabelle Cournu-Rebeix,[4,10] Sabine Hoffjan,[5] Alexander Winkelmann,[6] Emmanuel Touze,[11] Fernando Pico,[12] Philippe Corcia,[13] David Otaegui,[14] Alfredo Antigüedad,[15] Antonio Alcina,[16] Manuel Comabella,[17] Xavier Montalban,[17] Javier Olascoaga,[18] Fuencisla Matesanz,[16] Thomas Dörner,[19] Shu-Chen Li,[3,20] Elisabeth Steinhagen-Thiessen,[21] Ulman Lindenberger,[3] Andrew Chan,[22] Peter Rieckmann,[8,23] Hans-Peter Hartung,[24] Orhan Aktas,[24] Peter Lohse,[25] Mathias Buttmann,[8] Tania Kümpfel,[7] Christian Kubisch,[26] Uwe K Zettl,[6] Joerg T Epplen,[5] Bertrand Fontaine,[4,10] Frauke Zipp,[2] Koen Vandenbroeck,[27,28] Lars Bertram[1]

For numbered affiliations see end of article.

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

Received 9 November 2012
Revised 10 December 2012
Accepted 12 December 2012
Published Online First 12 January 2013

To cite: Lill CM, Schjeide B-MM, Graetz C, *et al*. *J Med Genet* 2013;**50**:140–143.

Multiple sclerosis (MS) is a genetically complex disease that shares a substantial proportion of risk loci with other autoimmune diseases.[1] Along these lines, *ANKRD55*, originally implicated in rheumatoid arthritis, was recently reported as a potential novel MS risk gene (rs6859219, $p=1.9\times10^{-7}$).[2] Here, we comprehensively validated this effect in independent datasets comprising 8846 newly genotyped subjects from Germany and France as well as 5003 subjects from two genome-wide association studies (GWAS). Upon meta-analysis of all available data (19 686 subjects), *ANKRD55* rs6859219 now shows compelling evidence for association with MS at genome-wide significance (OR=1.19, $p=3.1\times10^{-11}$). Our study adds *ANKRD55* to the list of established MS risk loci and extends previous evidence suggesting an overlapping genetic foundation across autoimmune diseases.

Ankyrin repeats are abundant in a large number of different proteins in humans and mediate protein–protein interactions. DNA-sequence variants in ankyrin repeat domain-containing proteins have been linked to a wide range of diseases; for example, *KRIT1* mutations causative for cerebral cavernous malformations,[3] *NOTCH3* mutations in cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy, and *RFXANK* mutations in the bare lymphocyte syndrome.[4] *ANKRD55* (located on chromosome 5q11.2) encodes the 'ankyrin repeat domain-containing protein 55' the function of which is currently unknown. Single nucleotide polymorphism rs6859219 in *ANKRD55* was implicated in a recent GWAS meta-analysis on rheumatoid arthritis.[5] Furthermore, a joint analysis of datasets on rheumatoid arthritis and coeliac disease also indicated a role of *ANKRD55* in the latter.[6] Given the augmenting evidence suggesting an overlap in the genetic architecture of autoimmune diseases including MS, we have previously investigated 10 'autoimmune loci' in 2895 Spanish MS cases and 2942 controls.[2] In that study, rs6859219 emerged as a putative new MS locus albeit at subgenome-wide significance ($p=1.9\times10^{-7}$).[2] Our failure to establish genome-wide significance was likely owing to the comparatively small sample size; thus, we set out to corroborate our initial association finding in additional independent datasets and to assess the overall evidence for association by meta-analysis.

We genotyped rs6859219 in 5106 MS cases and 3740 healthy control subjects of self-reported European descent from Germany and France[7 8] (table 1) using a commercially available assay ('TaqMan', Applied Biosystems, Inc.). Furthermore, we obtained, reanalysed and included data on 1868 cases and 3135 controls for rs6859219 from two publicly available GWAS ('IMSGC'[9] and 'GeneMSA';[10] in the latter, rs6859219 was analysed following imputation). GWAS quality control, imputation and analysis protocols were followed as described previously.[8] Combined, these replication datasets comprised 6974 cases and 6875 controls and had ~94% power to detect an OR of 1.20 at $\alpha=1\times10^{-4}$. Power to detect association at genome-wide significance ($\alpha=5\times10^{-8}$) using all available data (9869 cases and 9817 controls, ie, after including the Spanish datasets of the original study) was ~96%.

Genotyping efficiency and accuracy (based on 5% duplicate samples) in the newly genotyped datasets were 99.0% and 100%, respectively. Genotypes in controls were distributed according to Hardy–Weinberg equilibrium ($p=0.209$ using Pearson's χ^2). Logistic regression analyses based on an additive model were adjusted for age and sex in the German and French datasets, and for principal components (PC 1–3) in IMSGC and GeneMSA to account for population substructure as previously described.[8] Fixed-effect meta-analysis revealed significant association of the *ANKRD55* rs6859219 C-allele with increased risk for MS across all

Table 1 Demographic details of the German and French case-control datasets genotyped for ANKRD55 rs6859219

Sites	N cases (% females)	N controls (% females)	Mean AAE (±SD) cases	Mean AAE (±SD) controls	Mean AAO (±SD) cases
Germany	3762 (71)	2972 (60)	41 (11)	42 (17)	30 (10)
Bochum/Essen	1070 (71)	404 (43)	42 (11)	43 (12)	32 (10)
Duesseldorf/Koeln	257 (72)	829 (62)	39 (10)	44 (16)	–
Mainz/Berlin	787 (69)	869 (65)	38 (10)	34 (15)	29 (10)
Munich	595 (71)	400 (50)	42 (12)	42 (16)	30 (10)
Rostock	526 (74)	470 (70)	42 (12)	52 (19)	32 (11)
Wuerzburg	527 (69)	0 (0)	39 (11)	–	28 (10)
France (Paris)	1344 (72)	768 (60)	44 (12)	40 (13)	32 (10)
All	5106 (71)	3740 (60)	41 (12)	39 (16)	31 (10)

AAE, age at examination; AAO, age at onset; N, number.

replication datasets (OR=1.15, 95% CI 1.08 to 1.23, $p=1.0\times10^{-5}$, figure 1). Inclusion of the Spanish case-control datasets now exceeded the threshold for genome-wide significance by more than three orders of magnitude (OR=1.19, 95% CI 1.13 to 1.25, $p=3.1\times10^{-11}$, figure 1). While we found some evidence for heterogeneity of effect size estimates across datasets ($I^2=62$, 95% CI 20 to 83, p Q statistic=0.0093; figure 1), all dataset-specific ORs suggested a risk effect for the C-allele, except for Central Spain. This indicates that heterogeneity was nearly entirely due to variance of effect size estimates at the same side of the null (OR<1).

Despite the compelling evidence now adding *ANKRD55* to the list of established MS risk loci, the following limitations should be considered when interpreting our results. First, since determination of ethnic origin was based on self-report in the German, French and Spanish datasets, the possibility exists that results in these samples are affected by more subtle population substructure. However, appropriately adjusting for potential substructure effects in the GWAS datasets did not show any substantial change in results as compared with the unadjusted datasets. Hence, it is unlikely that population substructure has had a notable influence on our association results with rs6859219. Second, not all MS GWAS datasets published to date are publicly available and could therefore not all be included in the current study. This applies to two datasets in particular ('Australia and New Zealand Multiple Sclerosis Genetics Consortium' and 'Brigham and Women's Hospital') included in a recent GWAS meta-analysis that reported an association between rs6859219 and MS at nominal significance.[11] However, combining the summary results reported in that study[11] with our data (while excluding the GeneMSA and IMSGC results calculated here) does not appreciably change our overall meta-analysis results (OR=1.16, 95% CI 1.11 to 1.22, $p=2.9\times10^{-10}$). Finally, the pathophysiological mechanisms underlying the association between rs6859219 in *ANKRD55* and MS remain elusive. Rs6859219 is located in intron 7 of the gene, which is highly expressed in CD4 effector memory cells[12] but whose function remains largely unknown. As described previously, the linkage disequilibrium pattern (LD) around *ANKRD55* is rather narrow,[2 5] and rs6859219 shows noteworthy LD ($r^2\geq0.3$) only with other intronic *ANKRD55* variants (based on 1000G CEU data; assessed with the SNAP software, http://www.broadinstitute.org/mpg/snap/). The LD block does not appear to extend to the two neighbouring genes (*IL6ST* encoding interleukin 6 signal transducer and *IL31RA* encoding interleukin 31 receptor A), which may be other immediate candidates interfering with the underlying autoimmune process in MS. Interestingly, recent in vitro ChIP-seq data generated by the

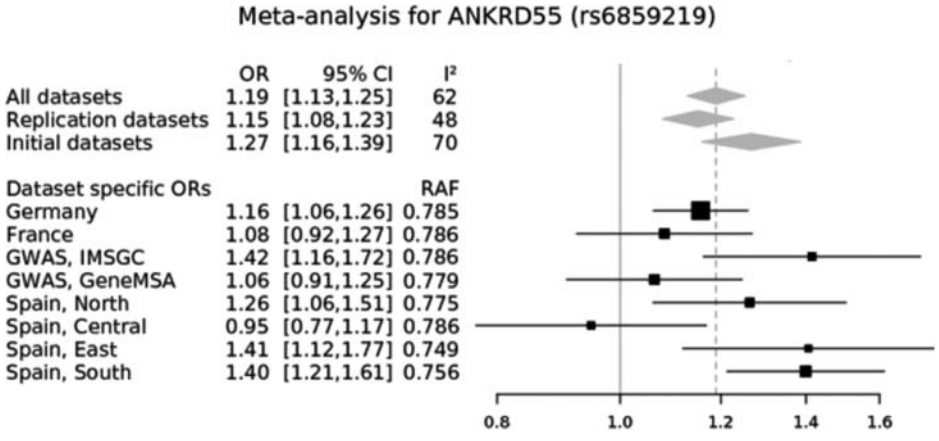

Figure 1 Meta-analysis of datasets assessing the association between *ANKRD55* rs6859219 and multiple sclerosis susceptibility in populations of European descent. Study-specific ORs (black squares) and 95% CIs (lines) were calculated using an additive model. The x-axis depicts the OR with regard to the risk allele dosage, that is, the C-allele. The summary ORs and 95% CIs (grey diamonds) were calculated based on fixed-effect meta-analysis combining all datasets as well as after stratification for the initial datasets and the validation datasets as indicated. GeneMSA, Genetic Multiple Sclerosis Associations[10]; GWAS, genome-wide association studies; IMSGC, International Multiple Sclerosis Genetics Consortium[9]; RAF, risk allele frequency in controls in the individual datasets.

international ENCODE project indicate that rs6859219, located ~182 kb 5′ of the *IL6ST* transcription start site, lies within a target site for several transcription factors including activating enhancer binding protein 2 and early B cell factor 1[13] (accessible via http://genome.ucsc.edu/cgi-bin/hgGateway). Thus, functional genetic studies have to assess whether the association between rs6859219 and risk for MS and other autoimmune diseases is due to a dysfunction of *ANKRD55*, for example, via affecting mRNA splicing, or whether the effects may be caused by an altered transcriptional regulation of*IL6ST* and/or *IL31RA*.

Author affiliations
[1]Neuropsychiatric Genetics Group, Department of Vertebrate Genomics, Max Planck Institute for Molecular Genetics, Berlin, Germany
[2]Department of Neurology, Focus Program Translational Neuroscience, University Medical Center of the Johannes Gutenberg-University Mainz, Mainz, Germany
[3]Max Planck Institute for Human Development, Berlin, Germany
[4]UPMC-INSERM-CNRS-UPMC-ICM, UMR 975–7225, Institut Cerveau Moelle Epinière (ICM), Hôpital Pitié-Salpêtrière, Paris, France
[5]Department of Human Genetics, Ruhr University, Bochum, Germany
[6]Department of Neurology, University of Rostock, Rostock, Germany
[7]Institute for Clinical Neuroimmunology, Ludwig Maximilian University, Munich, Germany
[8]Department of Neurology, University of Würzburg, Würzburg, Germany
[9]Centre for Research in Neuroscience at McGill University, Montreal, Canada
[10]Assistance Publique-Hôpitaux de Paris (AP-HP), Département de Neurologie, Hôpital Pitié-Salpêtrière, Paris, France
[11]Department of Neurology, Hôpital Sainte-Anne, Paris, France
[12]Department of Neurology, Centre Hospitalier de Versailles, Le Chesnay, France
[13]Department of Neurology, Centre Hospitalier Regional Universitaire, Tours, France
[14]Área de Neurociencias, Inst. Investigación Sanitaria Biodonostia, San Sebastián, Spain
[15]Servicio de Neurología, Hospital de Basurto, Bilbao, Spain
[16]Department of Biología Celular e Inmunología, Instituto de Parasitología y Biomedicina 'LópezNeyra' (IPBLN), CSIC, Granada, Spain
[17]Centre d'Esclerosi Múltiple de Catalunya, CEMCat, Unitat de Neuroimmunologia Clínica, Hospital Universitari Vall d'Hebron, Barcelona, Spain
[18]Servicio de Neurología, Unidad de Esclerosis Múltiple, Hospital Donostia, San Sebastián, Spain
[19]Department of Medicine, Rheumatology, and Clinical Immunology & DRFZ, Charité University Medicine, Berlin, Germany
[20]Department of Psychology, Technische Universität Dresden, Dresden, Germany
[21]Interdisciplinary Metabolic Center, Lipids Clinic, Charité University Medicine, Berlin, Germany
[22]Department of Neurology, St. Josef-Hospital, Ruhr-University, Bochum, Germany
[23]Department of Neurology, Sozialstiftung Bamberg Hospital, Bamberg, Germany
[24]Department of Neurology, Medical Faculty, Heinrich Heine University, Düsseldorf, Germany
[25]Department of Clinical Chemistry, Ludwig Maximilian University, Munich, Germany
[26]Institute of Human Genetics, University of Ulm, Ulm, Germany
[27]Neurogenomiks Laboratory, Department of Neuroscience, University of the Basque Country UPV/EHU, Leioa, Spain
[28]IKERBASQUE, Basque Foundation for Science, Bilbao, Spain

Acknowledgements We are grateful to the individuals participating in this study. We would like to thank investigators from the International Multiple Sclerosis Genetics Consortium and from the GeneMSA project for making their GWAS data available via the dbGaP platform. We acknowledge use of the cohort of the CRB-REFGENSEP and thank ICM, CIC Pitié-Salpêtrière, Généthon and REFGENSEP's members for their help and support.

Contributors Study design and supervision: CML, FZ, KV and LB. Data acquisition and performing of experiments: CML, B-MMS, CG, VD, DAA, PB, L-AG, AK, FL, IC-R, SH, AW, ET, FP, PC, DO, AAn, AAl, MC, XM, JO, FM, TD, S-CL, ES-T, UL, AC, PR, H-PH, OA, PL, MB, TK, CK, UKZ, JTE and BF. Data analysis and interpretation of results: CML, TL, KV and LB. Writing of manuscript: CML and LB with the help of all coauthors.

Funding This project was funded by grants from the German Ministry for Education and Research (BMBF) and German Research Foundation (DFG; to FZ), the BMBF and the Cure Alzheimer's Fund (to LB), the Walter- and Ilse-Rose-Stiftung (to H-PH and OA), the BMBF (grant NBL3 to UKZ; grant 01UW0808 to UL and ES-T), and the Innovation Fund of the Max Planck Society (M.FE.A.BILD0002 to UL). This project was supported by INSERM, ARSEP, AFM and GIS-IBISA. CML was supported by the Fidelity Biosciences Research Initiative.

Competing interests LA Gerdes reports to have received travel expenses and personal compensation from Merck Serono, Teva Pharmaceutical Industries, Bayer Schering Pharma, Novartis and Biogen Idec. T Kl reports to have received travel expenses and personal compensations from Bayer Schering Pharmacy, Teva, Merck-Serono, Novartis, Sanofi-Aventis and Biogen-Idec as well as grant support from Bayer-Schering AG. None of the other authors reports any disclosures.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1 Cotsapas C, Voight BF, Rossin E, Lage K, Neale BM, Wallace C, Abecasis GR, Barrett JC, Behrens T, Cho J, De Jager PL, Elder JT, Graham RR, Gregersen P, Klareskog L, Siminovitch KA, van Heel DA, Wijmenga C, Worthington J, Todd JA, Hafler DA, Rich SS, Daly MJ, FOCiS Network of Consortia. Pervasive sharing of genetic effects in autoimmune disease. *PLoS Genet* 2011;7:e1002254.
2 Alloza I, Otaegui D, de Lapuente AL, Antigüedad A, Varadé J, Núñez C, Arroyo R, Urcelay E, Fernandez O, Leyva L, Fedetz M, Izquierdo G, Lucas M, Oliver-Martos B, Alcina A, Saiz A, Blanco Y, Comabella M, Montalban X, Olascoaga J, Matesanz F, Vandenbroeck K. ANKRD55 and DHCR7 are novel multiple sclerosis risk loci. *Genes Immun* 2012;13:253–7.
3 Labauge P, Denier C, Bergametti F, Tournier-Lasserve E. Genetics of cavernous angiomas. *Lancet Neurol* 2007;6:237–44.
4 Mosavi LK, Cammett TJ, Desrosiers DC, Peng Z-Y. The ankyrin repeat as molecular architecture for protein recognition. *Protein Sci* 2004;13:1435–48.
5 Stahl EA, Raychaudhuri S, Remmers EF, Xie G, Eyre S, Thomson BP, Li Y, Kurreeman FA, Zhernakova A, Hinks A, Guiducci C, Chen R, Alfredsson L, Amos CI, Ardlie KG, Barton A, Bowes J, Brouwer E, Burtt NP, Catanese JJ, Coblyn J, Coenen MJ, Costenbader KH, Criswell LA, Crusius JB, Cui J, de Bakker PI, De Jager PL, Ding B, Emery P, Flynn E, Harrison P, Hocking LJ, Huizinga TW, Kastner DL, Ke X, Lee AT, Liu X, Martin P, Morgan AW, Padyukov L, Posthumus MD, Radstake TR, Reid DM, Seielstad M, Seldin MF, Shadick NA, Steer S, Tak PP, Thomson W, van der Helm-van Mil AH, van der Horst-Bruinsma IE, van der Schoot CE, van Riel PL, Weinblatt ME, Wilson AG, Wolbink GJ, Wordsworth BP, Wijmenga C, Karlson EW, Toes RE, de Vries N, Begovich AB, Worthington J, Siminovitch KA, Gregersen PK, Klareskog L, Plenge RM, BIRAC Consortium, YEAR Consortium. Genome-wide association study meta-analysis identifies seven new rheumatoid arthritis risk loci. *Nat Genet* 2010;42:508–14.
6 Zhernakova A, Stahl EA, Trynka G, Raychaudhuri S, Festen EA, Franke L, Westra HJ, Fehrmann RS, Kurreeman FA, Thomson B, Gupta N, Romanos J, McManus R, Ryan AW, Turner G, Brouwer E, Posthumus MD, Remmers EF, Tucci F, Toes R, Grandone E, Mazzilli MC, Rybak A, Cukrowska B, Coenen MJ, Radstake TR, van Riel PL, Li Y, de Bakker PI, Gregersen PK, Worthington J, Siminovitch KA, Klareskog L, Huizinga TW, Wijmenga C, Plenge RM. Meta-analysis of genome-wide association studies in celiac disease and rheumatoid arthritis identifies fourteen non-HLA shared loci. *PLoS Genet* 2011;7:e1002004.
7 Lill CM, Schjeide BM, Akkad DA, Blaschke P, Winkelmann A, Gerdes LA, Hoffjan S, Luessi F, Dörner T, Li SC, Steinhagen-Thiessen E, Lindenberger U, Chan A, Hartung HP, Aktas O, Lohse P, Kümpfel T, Kubisch C, Epplen JT, Zettl UK, Bertram L, Zipp F. Independent replication of STAT3 association with multiple sclerosis risk in a large German case-control sample. *Neurogenetics* 2012;13:83–6.
8 Lill CM, Liu T, Schjeide BM, Roehr JT, Akkad DA, Damotte V, Alcina A, Ortiz MA, Arroyo R, Lopez de Lapuente A, Blaschke P, Winkelmann A, Gerdes LA, Luessi F, Fernadez O, Izquierdo G, Antigüedad A, Hoffjan S, Cournu-Rebeix I, Gromöller S, Faber H, Liebsch M, Meissner E, Chanvillard C, Touze E, Pico F, Corcia P, Dörner T, Steinhagen-Thiessen E, Baeckman L, Heekeren HR, Li SC, Lindenberger U, Chan A, Hartung HP, Aktas O, Lohse P, Kümpfel T, Kubisch C, Epplen JT, Zettl UK, Fontaine B, Vandenbroeck K, Matesanz F, Urcelay E, Bertram L, Zipp F, ANZgene Consortium. Closing the case of APOE in multiple sclerosis: no association with disease risk in over 29 000 subjects. *J Med Genet* 2012;49:558–62.
9 Hafler DA, Compston A, Sawcer S, Lander ES, Daly MJ, De Jager PL, de Bakker PI, Gabriel SB, Mirel DB, Ivinson AJ, Pericak-Vance MA, Gregory SG, Rioux JD, McCauley JL, Haines JL, Barcellos LF, Cree B, Oksenberg JR, Hauser SL, International Multiple Sclerosis Genetics Consortium. Risk alleles for multiple sclerosis identified by a genomewide study. *N Engl J Med* 2007;357:851–62.
10 Baranzini SE, Wang J, Gibson RA, Galwey N, Naegelin Y, Barkhof F, Radue EW, Lindberg RL, Uitdehaag BM, Johnson MR, Angelakopoulou A, Hall L, Richardson JC, Prinjha RK, Gass A, Geurts JJ, Kragt J, Sombekke M, Vrenken H, Qualley P, Lincoln RR, Gomez R, Caillier SJ, George MF, Mousavi H, Guerrero R, Okuda DT, Cree BA, Green AJ, Waubant E, Goodin DS, Pelletier D, Matthews PM, Hauser SL, Kappos L, Polman CH, Oksenberg JR. Genome-wide association analysis of susceptibility and clinical phenotype in multiple sclerosis. *Hum Mol Genet* 2009;18:767–78.
11 Patsopoulos NA, Esposito F, Reischl J, Lehr S, Bauer D, Heubach J, Sandbrink R, Pohl C, Edan G, Kappos L, Miller D, Montalbán J, Polman CH, Freedman MS, Hartung HP, Arnason BG, Comi G, Cook S, Filippi M, Goodin DS, Jeffery D, O'Connor P, Ebers GC, Langdon D, Reder AT, Traboulsee A, Zipp F, Schimrigk S, Hillert J, Bahlo M, Booth DR, Broadley S, Brown MA, Browning BL, Browning SR, Butzkueven H, Carroll WM, Chapman C, Foote SJ, Griffiths L, Kermode AG, Kilpatrick TJ, Lechner-Scott J, Marriott M, Mason D, Moscato P, Heard RN, Pender MP, Perreau VM, Perera D, Rubio JP, Scott RJ, Slee M, Stankovich J, Stewart GJ,

Taylor BV, Tubridy N, Willoughby E, Wiley J, Matthews P, Boneschi FM, Compston A, Haines J, Hauser SL, McCauley J, Ivinson A, Oksenberg JR, Pericak-Vance M, Sawcer SJ, De Jager PL, Hafler DA, de Bakker PI, Bayer Pharma MS Genetics Working Group, Steering Committees of Studies Evaluating IFNβ-1b and a CCR1-Antagonist, ANZgene Consortium, GeneMSA, International Multiple Sclerosis Genetics Consortium. Genome-wide meta-analysis identifies novel multiple sclerosis susceptibility loci. *Ann Neurol* 2011;70:897–912.

12 Hu X, Kim H, Stahl E, Plenge R, Daly M, Raychaudhuri S. Integrating autoimmune risk loci with gene-expression data identifies specific pathogenic immune cell subsets. *Am J Hum Genet* 2011;89:496–506.

13 Dunham I, Kundaje A, Aldred SF, Collins PJ, Davis CA, Doyle F, Epstein CB, Frietze S, Harrow J, Kaul R, Khatun J, Lajoie BR, Landt SG, Lee BK, Pauli F, Rosenbloom KR, Sabo P, Safi A, Sanyal A, Shoresh N, Simon JM, Song L, Trinklein ND, Altshuler RC, Birney E, Brown JB, Cheng C, Djebali S, Dong X, Dunham I, Ernst J, Furey TS, Gerstein M, Giardine B, Greven M, Hardison RC, Harris RS, Herrero J, Hoffman MM, Iyer S, Kelllis M, Khatun J, Kheradpour P, Kundaje A, Lassman T, Li Q, Lin X, Marinov GK, Merkel A, Mortazavi A, Parker SC, Reddy TE, Rozowsky J, Schlesinger F, Thurman RE, Wang J, Ward LD, Whitfield TW, Wilder SP, Wu W, Xi HS, Yip KY, Zhuang J, Bernstein BE, Birney E, Dunham I, Green ED, Gunter C, Snyder M, Pazin MJ, Lowdon RF, Dillon LA, Adams LB, Kelly CJ, Zhang J, Wexler JR, Green ED, Good PJ, Feingold EA, Bernstein BE, Birney E, Crawford GE, Dekker J, Elinitski L, Farnham PJ, Gerstein M, Giddings MC, Gingeras TR, Green ED, Guigó R, Hardison RC, Hubbard TJ, Kellis M, Kent WJ, Lieb JD, Margulies EH, Myers RM, Snyder M, Starnatoyannopoulos JA, Tennebaum SA, Weng Z, White KP, Wold B, Khatun J, Yu Y, Wrobel J, Risk BA, Gunawardena HP, Kuiper HC, Maier CW, Xie L, Chen X, Giddings MC, Bernstein BE, Epstein CB, Shoresh N, Ernst J, Kheradpour P, Mikkelsen TS, Gillespie S, Goren A, Ram O, Zhang X, Wang L, Issner R, Coyne MJ, Durham T, Ku M, Truong T, Ward LD, Altshuler RC, Eaton ML, Kellis M, Djebali S, Davis CA, Merkel A, Dobin A, Lassmann T, Mortazavi A, Tanzer A, Lagarde J, Lin W, Schlesinger F, Xue C, Marinov GK, Khatun J, Williams BA, Zaleski C, Rozowsky J, Röder M, Kokocinski F, Abdelhamid RF, Alioto T, Antoshechkin I, Baer MT, Batut P, Bell I, Bell K, Chakrabortty S, Chen X, Chrast J, Curado J, Derrien T, Drenkow J, Dumais E, Dumais J, Duttagupta R, Fastuca M, Fejes-Toth K, Ferreira P, Foissac S, Fullwood MJ, Gao H, Gonzalez D, Gordon A, Gunawardena HP, Howald C, Jha S, Johnson R, Kapranov P, King B, Kingswood C, Li G, Luo OJ, Park E, Preall JB, Presaud K, Ribeca P, Risk BA, Robyr D, Ruan X, Sammeth M, Sandu KS, Schaeffer L, See LH, Shahab A, Skancke J, Suzuki AM, Takahashi H, Tilgner H, Trout D, Walters N, Wang H, Wrobel J, Yu Y, Hayashizaki Y, Harrow J, Gerstein M, Hubbard TJ, Reymond A, Antonarakis SE, Hannon GJ, Giddings MC, Ruan Y, Wold B, Carninci P, Guigó R, Gingeras TR, Rosenbloom KR, Sloan CA, Learned K, Malladi VS, Wong MC, Barber GP, Cline MS, Dreszer TR, Heitner SG, Karolchik D, Kent WJ, Kirkup VM, Meyer LR, Long JC, Maddren M, Raney BJ, Furey TS, Song L, Grasfeder LL, Giresi PG, Lee BK, Battenhouse A, Sheffield NC, Simon JM, Showers KA, Safi A, London D, Bhinge AA, Shestak C, Schaner MR, Kim SK, Zhang ZZ, Mieczkowski PA, Mieczkowska JO, Liu Z, McDaniell RM, Ni Y, Rashid NU, Kim MJ, Adar S, Zhang Z, Wang T, Winter D, Keefe D, Birney E, Iyer VR, Lieb JD, Crawford GE, Li G, Sandhu KS, Zheng M, Wang P, Luo OJ, Shahab A, Fullwood MJ, Ruan X, Ruan Y, Myers RM, Pauli F, Williams BA, Gertz J, Marinov GK, Reddy TE, Vielmetter J, Partridge EC, Trout D, Varley KE, Gasper C, Bansal A, Pepke S, Jain P, Amrhein H, Bowling KM, Anaya M, Cross MK, King B, Muratet MA, Antoshechkin I, Newberry KM, McCue K, Nesmith AS, Fisher-Aylor KI, Pusey B, DeSalvo G, Parker SL, Balasubramanian S, Davis NS, Meadows SK, Eggleston T, Gunter C, Newberry JS, Levy SE, Absher DM, Mortazavi A, Wong WH, Wold B, Blow MJ, Visel A, Pennachio LA, Elnitski L, Margulies EH, Parker SC, Petrykowska HM, Abyzov A, Aken B, Barrell D, Barson G, Berry A, Bignell A, Boychenko V, Bussotti G, Chrast J, Davidson C, Derrien T, Despacio-Reyes G, Diekhans M, Ezkurdia I, Frankish A, Gilbert J, Gonzalez JM, Griffiths E, Harte R, Hendrix DA, Howald C, Hunt T, Jungreis I, Kay M, Khurana E, Kokocinski F, Leng J, Lin MF, Loveland J, Lu Z, Manthravadi D, Mariotti M, Mudge J, Mukherjee G, Notredame C, Pei B, Rodriguez JM, Saunders G, Sboner A, Searle S, Sisu C, Snow C, Steward C, Tanzer A, Tapanan E, Tress ML, van Baren MJ, Walters N, Washieti S, Wilming L, Zadissa A, Zhengdong Z, Brent M, Haussler D, Kellis M, Valencia A, Gerstein M, Raymond A, Guigó R, Harrow J, Hubbard TJ, Landt SG, Frietze S, Abyzov A, Addleman N, Alexander RP, Auerbach RK, Balasubramanian S, Bettinger K, Bhardwaj N, Boyle AP, Cao AR, Cayting P, Charos A, Cheng Y, Cheng C, Eastman C, Euskirchen G, Fleming JD, Grubert F, Habegger L, Hariharan M, Harmanci A, Iyenger S, Jin VX, Karczewski KJ, Kasowski M, Lacroute P, Lam H, Larnarre-Vincent N, Leng J, Lian J, Lindahl-Allen M, Min R, Miotto B, Monahan H, Moqtaderi Z, Mu XJ, O'Geen H, Ouyang Z, Patacsil D, Pei B, Raha D, Ramirez L, Reed B, Rozowsky J, Sboner A, Shi M, Sisu C, Slifer T, Witt H, Wu L, Xu X, Yan KK, Yang X, Yip KY, Zhang Z, Struhl K, Weissman SM, Gerstein M, Farnham PJ, Snyder M, Tenebaum SA, Penalva LO, Doyle F, Karmakar S, Landt SG, Bhanvadia RR, Choudhury A, Domanus M, Ma L, Moran J, Patacsil D, Slifer T, Victorsen A, Yang X, Snyder M, White KP, Auer T, Centarin L, Eichenlaub M, Gruhl F, Heerman S, Hoeckendorf B, Inoue D, Kellner T, Kirchmaier S, Mueller C, Reinhardt R, Schertel L, Schneider S, Sinn R, Wittbrodt B, Wittbrodt J, Weng Z, Whitfield TW, Wang J, Collins PJ, Aldred SF, Trinklein ND, Partridge EC, Myers RM, Dekker J, Jain G, Lajoie BR, Sanyal A, Balasundaram G, Bates DL, Byron R, Canfield TK, Diegel MJ, Dunn D, Ebersol AK, Ebersol AK, Frum T, Garg K, Gist E, Hansen RS, Boatman L, Haugen E, Humbert R, Jain G, Johnson AK, Johnson EM, Kutyavin TM, Lajoie BR, Lee K, Lotakis D, Maurano MT, Neph SJ, Neri FV, Nguyen ED, Qu H, Reynolds AP, Roach V, Rynes E, Sabo P, Sanchez ME, Sandstrom RS, Sanyal A, Shafer AO, Stergachis AB, Thomas S, Thurman RE, Vernot B, Vierstra J, Vong S, Wang H, Weaver MA, Yan Y, Zhang M, Akey JA, Bender M, Dorschner MO, Groudine M, MacCoss MJ, Navas P, Stamatoyannopoulos G, Kaul R, Dekker J, Stamatoyannopoulos JA, Dunham I, Beal K, Brazma A, Flicek P, Herrero J, Johnson N, Keefe D, Lukk M, Luscombe NM, Sobral D, Vaquerizas JM, Wilder SP, Batzoglou S, Sidow A, Hussami N, Kyriazopoulou-Panagiotopoulou S, Libbrecht MW, Schaub MA, Kundaje A, Hardison RC, Miller W, Giardine B, Harris RS, Wu W, Bickel PJ, Banfai B, Boley NP, Brown JB, Huang H, Li Q, Li JJ, Noble WS, Bilmes JA, Buske OJ, Hoffman MM, Sahu AO, Kharchenko PV, Park PJ, Baker D, Taylor J, Weng Z, Iyer S, Dong X, Greven M, Lin X, Wang J, Xi HS, Zhuang J, Gerstein M, Alexander RP, Balasubramanian S, Cheng C, Harmanci A, Lochovsky L, Min R, Mu XJ, Rozowsky J, Yan KK, Yip KY, Birney E, ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. *Nature* 2012;489:57–74.