Genome-wide meta-analysis identifies new susceptibility loci for migraine

Migraine is the most common brain disorder, affecting approximately 14% of the adult population, but its molecular mechanisms are poorly understood. We report the results of a meta-analysis across 29 genome-wide association studies, including a total of 23,285 individuals with migraine (cases) and 95,425 population-matched controls. We identified 12 loci associated with migraine susceptibility ($P < 5 \times 10^{-8}$).

Five loci are new: near AJAP1 at 1p36, near TSPAN2 at 1p13, within FH5 at 6q16, within C7orf10 at 7p14 and near MMP16 at 8q21. Three of these loci were identified in disease subgroup analyses. Brain tissue expression quantitative trait locus analysis suggests potential functional candidate genes at four loci: APOA1BP, TBC1D7, FUT9, STAT6 and ATP5B.

Recently, significant progress has been made in the identification of common genetic variants associated with migraine susceptibility through genome-wide association studies (GWAS) of clinic-based migraine with aura, migraine in the general population and memory formation. The associated locus also overlaps UFL1 (also known as KIAA0776), which encodes a hypothetical protein. At 7p14, mutations in C7orf10 have been found in phenotypically mild or even clinically asymptomatic forms of glutaric aciduria type III (ref. 9), a rare metabolic abnormality leading to persistent excretion of glutaric acid.

The newly associated loci at 1p36, 1p13 and 8q21 are located outside of known transcripts. At 1p36, rs10915437 is located approximately 500 kb telomeric of AJAP1 and approximately 300 kb centromeric of a gene cluster encoding the apoptosis-related proteins DFFB and TP73 as well as centrosomal protein CEP104. AJAP1 is expressed in the brain (Supplementary Fig. 6) and has been associated with tumor invasion and the regulation of metalloproteinase activity. At 1p13, rs12134493 is located 87 kb upstream of TSPAN2, a member of the tetraspanin family, encoding a cell surface protein that mediates signal transduction events involved in the regulation of cell development, activation, growth and motility. TSPAN2 has further been shown to act as a regulator of metalloproteinase activity. At 8q21, rs10504861 is located 200 kb telomeric to the matrix metalloproteinase MMP16. Members of the metalloproteinase family are widely expressed in human tissues and are involved in the breakdown of extracellular matrix in normal physiological processes. Notably, the protein encoded by MMP16 (MT-MMP2) cleaves LRP1 (ref. 12), encoded by previously reported migraine-associated gene2. In addition, MMP16 has recently been shown to be involved in basal NgR1 (Nogo-66 receptor) shedding in cortical neurons, thereby increasing axonal and synaptic plasticity.

A full list of authors and affiliations appears at the end of the paper.

Received 7 November 2012; accepted 30 May 2013; published online 23 June 2013; doi:10.1038/ng.2676
Four of the 12 migraine-associated loci (near AJAP1, near TGFBR2, in PHACTR1 and near MMP16), including 2 of the new associations, were identified exclusively in the subgroup analyses (Table 1 and Supplementary Table 4). Two of the loci (rs9349379 in PHACTR1 at 6p21 and rs10504861 near MMP16 at 8q21) had associations that reached genome-wide significance only in migraine without aura, whereas no SNPs had signals that reached genome-wide significance in migraine with aura (Fig. 2). The lowest P value in migraine with aura was observed for rs7015657 (P = 7.88 × 10^{-5}), which is located approximately 582 kb downstream of GFRA2, a member of the glial cell line–derived neurotrophic factor family.

A similar subgroup analysis was performed in only those samples that originated from specialized migraine clinics (Fig. 2). Two loci with suggestive associations in the primary analysis, rs6790925 (near TGFBR2) and rs6478241 (in ASTN2), reached genome-wide significance in the clinic-based subgroup (Table 1 and Supplementary Tables 6 and 7). To assess whether there are consistent differences in the effect sizes estimated from clinic-based samples versus the primary analysis of all samples (Supplementary Table 6 and 7), we performed a 2-tailed binomial test assessing which set had estimated larger effect sizes for the 12 significant loci. For all 12 of the genome-wide significant loci, the clinic-based sample effect sizes were estimated to be larger, which is significantly different from chance (P = 4.88 × 10^{-4}), suggesting higher genetic loading for clinic-based samples. When all reported loci (P < 1 × 10^{-5}) were considered, only the clinic-based group showed a number of associated SNPs with higher effect sizes (odds ratio (OR) > 1.2) at low frequency (minor allele frequency (MAF) < 0.05) (Supplementary Fig. 7). Thus, clinic-based migraine samples may represent a promising subgroup to help prioritize loci in the search for low-frequency variants with moderate effects. Overall, of the 146 loci identified, there were twice as many that had causative minor alleles than had protective ones (with a higher ratio of causative to protective alleles at lower MAFs).

To explore the biological contexts of the identified loci, we examined the properties of the most proximal genes to the 12 top-associated SNPs reaching genome-wide significance (Table 1). In expression data from 55,269 samples profiled using the Affymetrix HG-U133 Plus 2.0 microarray (including 1,990 brain and 384 endothelial samples), 11 of the 12 genes nearest to the identified loci (all except FH5) were at least moderately expressed (with > 20% of the tissue samples showing a normalized log2 expression value greater than 6; Online Methods) in disease-relevant brain regions (Supplementary Fig. 6). In contrast, only TGFBR2 and MEF2D showed moderate or greater expression in the endothelial samples. Possibly reflecting known comorbidity between migraine and cardiovascular disease14, 2 of the 12 most proximal genes (TGFBR2 and PHACTR1) have also been associated with cardiovascular traits: TGFBR2 mutations have been reported to cause monogenic Marfan syndrome15 and to be involved in early-onset myocardial infarction17. TSPAN2 (ref. 18), MEF2D19,

---

**Table 1 Results of association analyses**

| SNP     | Chr. | Position (bp) | Location | Gene | Minor allele | MAF   | P value | OR (95% CI) | q value | Group or subgroup with the lowest P value | Additional subgroup with significant association |
|---------|------|---------------|----------|------|--------------|-------|---------|------------|---------|-----------------------------------------|--------------------------------------------------|
| rs2651899 | 1    | 3073572       | Genic    | PRDM16 | C             | 0.41  | 3.28 × 10^{-14} | 1.09 (1.07–1.12) | 0.214  | 20 | All samples |
| rs10915437 | 1    | 4082866      | Intergenic | Near AJAP1 | G  | 0.36 | 2.81 × 10^{-8} | 0.86 (0.82–0.91) | 0.108  | 47 | Clinics only |
| rs12134493 | 1    | 115479469    | Intergenic | Near TSPAN2 | A  | 0.46 | 6.71 × 10^{-14} | 1.14 (1.10–1.18) | 0.480  | 14 | All samples |
| rs2274316 | 1    | 154712866    | Genic    | MEF2Db | C  | 0.37 | 2.81 × 10^{-8} | 1.07 (1.04–1.09) | 0.021  | 45 | All samples |
| rs7577262 | 2    | 234483608    | Genic    | TRPM8b | A  | 0.10 | 3.27 × 10^{-13} | 0.87 (0.84–0.90) | 0.070  | 33 | All samples Clinics only, MO only |
| rs6790925 | 3    | 30455089     | Intergenic | Near TGFBR2 | T  | 0.38 | 2.16 × 10^{-11} | 1.15 (1.10–1.21) | 0.780  | 0 | Clinics only |
| rs9349379 | 6    | 13011943     | Intergenic | PHACTR1 | G  | 0.40 | 2.81 × 10^{-10} | 0.86 (0.82–0.90) | 0.443  | 0 | MO only All samples |
| rs13208321 | 6    | 96967075     | Genic    | FH5 | A  | 0.22 | 2.15 × 10^{-12} | 1.18 (1.13–1.24) | 0.168  | 0 | All samples MO only |
| rs4379368 | 7    | 40432725    | Genic    | c7orf10 | T  | 0.12 | 2.46 × 10^{-8} | 1.11 (1.08–1.15) | 0.441  | 2 | All samples MO only |
| rs10504861 | 8    | 89617048    | Intergenic | Near MMP16 | T  | 0.16 | 1.32 × 10^{-8} | 0.86 (0.81–0.90) | 0.755  | 0 | MO only |
| rs6478241 | 9    | 118292450    | Genic    | ASTN2b | A  | 0.38 | 1.04 × 10^{-9} | 1.16 (1.11–1.22) | 0.646  | 0 | Clinics only All samples |
| rs11172113 | 12   | 55831350    | Genic    | LRP1b | C  | 0.43 | 2.69 × 10^{-19} | 0.90 (0.88–0.92) | 0.188  | 22 | All samples MO only |

When multiple subgroups showed significant association, P values and ORs are shown for the analysis with the lowest P value. ORs are reported for the minor allele. Chromosomal positions are based on NCBI Build 36. Chr., chromosome; CI, confidence interval; q value, P value from Cochran's heterogeneity statistic; P, heterogeneity index; MO, migraine without aura.

bFor genic SNPs (located within the coding region of a gene), the relevant gene is listed; for intergenic SNPs, the gene closest to the locus is listed. bPreviously identified migraine-associated locus.
**TRPM8** (ref. 20), **TGFBR2** (ref. 21), **PHACTR1** (ref. 22), **MMP16** (ref. 23), **ASTN2** (ref. 24) and **LRP1** (ref. 25) have been suggested to have functions in synapse formation or regulation. **PRDM16** has been linked to the oxidative stress response, and **AJAP1** has been implicated in maintaining tissue borders (Supplementary Fig. 8).

To identify possible genes underlying these associations that were not included in proximity-based analyses, we examined expression quantitative trait locus (eQTL) data for 394 samples of brain tissue from the North American Brain Expression Consortium and the UK Brain Expression Consortium. Of the 12 regions with significant associations with migraine, 4 were found to contain significant eQTLs (Online Methods) identified in the SNP-probe pairs in the tested brain samples (4 in the frontal cortex and 1 in the cerebellum; Table 2). On chromosome 1, rs12136718 is an eQTL for **APOA1BP**, rather than the gene closest to it, **MEF2D**. **APOA1BP** is widely expressed and is potentially linked to cholesterol efflux from cells26. On chromosome 6, rs35128104 is an eQTL for **FUT9**, encoding α1,3-fucosyl transferase IX. The FUT9 enzyme synthesizes the Lewis X carbohydrate structure, which has been implicated in neurite outgrowth in several types of neuronal cells in the brain27,28. At the chromosome 12 locus, two different eQTLs in brain tissue were found within the association peak, including for **STAT6** (rs4559) and **ATP5B** (rs113953523), the former at a very robust P value (2.16 × 10−22), **STAT6** family members are known for transducing activation signals to transcription factors in macrophages30, and **STAT6** phosphorylation has recently been shown to be an intermediate step in the astrocyte response to oxidative stress, resulting in prostaglandin release31. The second eQTL-regulated gene at the same locus, **ATP5B**, is the β subunit of the mitochondrial ATP synthase, but a specific role in neuronal or vascular cells is not known. Finally, on chromosome 6, rs9349379 is an eQTL in cerebellar tissue for **TBC1D7**, which potentially downregulates the tuberous sclerosis gene **TSC1** through positive regulation of the mTOR signaling pathway32. In the central nervous system, **TSC1**-**TSC2** signaling contributes to neural connectivity through multifaceted roles33. On the basis of the available data, it is not possible to decide whether the gene identified by eQTL analysis or the gene closest to the strongest association is the most relevant gene contributing to migraine pathogenesis.

In an analysis of hypersensitivity sites, associated SNPs in the loci for migraine were found to occur substantially more often in DNa1 hypersensitivity sites in a number of tissues (Supplementary Fig. 9).

### Table 2 Significant eQTLs observed in 394 samples of 2 different brain tissues at loci significantly associated with migraine

| Chr. | Position (bp) | Locus | Probe | Gene name | SNPs with significant eQTL with probe | SNP with lowest eQTL P value | Lowest eQTL P value | Highest LD with reported SNP | SNPs with significant eQTL with probe | SNP with lowest eQTL P value | Lowest eQTL P value | Highest LD with reported SNP |
|------|---------------|-------|-------|-----------|--------------------------------------|----------------------------|------------------------|-------------------------------|--------------------------------------|----------------------------|------------------------|-------------------------------|
| 1    | 154830352     | MEF2D | ILMN,1729533 | APR1BP   | rs12136718                           | 2.18 × 10−5                | 0.38                   | rs3995224 2.06 × 10−4        | 1                                   | rs3995224 2.06 × 10−4        | 0.451                   |
| 6    | 13413333      | PHACTR1 | ILMN,1661622 | TBC1D7   | 0                                    | 0                          | 0                      | rs3995224 2.06 × 10−4        | 1                                   | rs3995224 2.06 × 10−4        | 0.451                   |
| 6    | 96769593      | FHL5  | ILMN,1878007 | FUT9     | rs35128104                           | 5.96 × 10−5                | 0.74                   | 0                            | rs3995224 2.06 × 10−4        | 0.451                   |
| 12   | 55318337      | LRP1  | ILMN,1772132 | ATP5B    | rs113953523                          | 1.62 × 10−4                | 0.39                   | 0                            | rs3995224 2.06 × 10−4        | 0.451                   |
| 12   | 55775568      | LRP1  | ILMN,1763198 | STAT6    | rs4559                               | 2.16 × 10−22               | 1                      | rs3995224 2.06 × 10−4        | 0.451                   |

Chr., chromosome. All base-pair positions are given for NCBI Build 36 (hg18). Owing to multiple overlapping signals, the human leukocyte antigen (HLA) region was excluded from the analysis. False discovery rate (FDR)-corrected significance thresholds for eQTL SNP–probe pairs were P < 0.0001668 for frontal cortex and P < 0.0002187 for cerebellum.

*Maximum extent of LD observed between a significant eQTL SNP-probe pair (Online Methods) and a SNP passing the reporting threshold (P < 1 × 10−5) for association with migraine.
This finding suggests that the loci associated with migraine are enriched for actively transcribed regions, supporting the notion that the variants have regulatory roles. Both neuronal and vascular tissue types carry an enriched set of sites within the detected loci. In addition, in querying the RegulomeDB (see URLs), several of the associated SNPs were found to overlap directly with known transcription factor binding motifs (Supplementary Table 8).

The number of significantly associated loci is still modest for forming a broad understanding of disease susceptibility, and any proposed functional hypothesis from the identified loci must thus be approached with caution. Some functional hypotheses could be inferred from the results of this study, as the majority of the identified loci harbor genes that can be linked to neuronal function.

The eQTL analysis further supported the notion that regulatory effects in brain tissue may underlie several of the association signals. However, the lack of replication across brain regions suggests that our ability to use eQTL data to pinpoint the most functionally significant gene within each locus is limited.

The larger number of significant loci in the migraine without aura group compared to the migraine with aura group (6 versus 0, respectively, Fig. 2), despite reasonably similar sample sizes, was somewhat unexpected. Migraine with aura has been shown to have a considerably higher heritability estimate and sibling recurrence risk than migraine without aura (3.8 versus 1.9) and has thus been considered to be the more heritable of the two common migraine types\(^3\). One possible explanation for the observed difference is that genetic susceptibility to migraine with aura is mediated more by rare variants with larger effect sizes, although this notion remains speculative. Another explanation for the difference may be a higher degree of heterogeneity among the cases of migraine with aura (due to genetically distinct subgroups, for example). No common variants specifically predisposing to migraine with aura were identified by this study (the lowest observed association \(P\) value was \(7.88 \times 10^{-8}\)).

In summary, we conducted a large migraine meta-analysis and identified 12 loci significantly associated with migraine susceptibility, including 5 new loci, as well as 134 additional suggestive loci. An eQTL analysis of brain tissue highlighted a further five genes potentially implicated in migraine susceptibility. Two of the 12 loci showed association only in the clinic-based sample group, suggesting higher specificity for severe migraine headache, and one locus was identified solely in the migraine without aura group. Seven previously reported loci for migraine susceptibility were replicated in independent samples in this study. Observed differences in the number of identified loci and the strength of association suggest that the genetic background of migraine with aura is considerably less influenced by common variants than that of migraine without aura, contrary to previous expectations. Finally, although pathway analysis of the 146 loci showed no concentration of candidate causal genes in any particular pathway or tissue, 8 of the 12 identified loci are located in or immediately adjacent to genes with known functions in synaptic or neuronal regulation, and several exert regulation, including over one another.

**URLs**: DAPPLE, http://www.broadinstitute.org/mpg/dapple/dapple.php; GRAIL, http://www.broadinstitute.org/mpg/grail/; GWAMA, http://www.well.ox.ac.uk/gwama/; IMPUTE2, http://mathgen.stats.ox.ac.uk/impute/; International Headache Genetics Consortium, http://www.headache.genetics.org/; MAGENTA, http://www.broadinstitute.org/mpg/magenta/; ProbABEL, http://www.genabel.org/packages/ProbABEL; RegulomeDB, http://regulome.stanford.edu/; SNAP, http://www.broadinstitute.org/mpg/snap/; SNPTEST, https://mathgen.stats.ox.ac.uk/genetics_software/snptest/snptest.html.

**METHODS**

Methods and any associated references are available in the online version of the paper.

**Note:** Supplementary information is available in the online version of the paper.

**ACKNOWLEDGMENTS**

Study-specific acknowledgments appear in the Supplementary Note. We wish to thank A. Coffey, S. Hunt, R. Gwillian, P. Whittaker, S. Potter and A. Tashakkori-Ghanbarian for their invaluable help with this study, and we collectively thank everyone who has contributed to the collection, genotyping and analysis of the individual cohorts, as well as all the study participants.

**AUTHOR CONTRIBUTIONS**

A. Palotie, D.I.C., D.R.N., M.J.D., K.S., G.D.S., M.W., M.D., A.J.M., M.J.D., C.K., T.K., D.P.S., L.C., J.-A.Z., M.-R.J., C.V.D., D.I.B., J.K., L.Q., and G.T. jointly supervised research. A. Palotie, D.I.C., D.R.N., V. Anttila, B.S.W., M.J.D., M.W., M.D., A.J.M., M.J.D., C.K., L.C., J.-A.Z., M.-R.J., C.V.D., D.I.B., J.K., T.K., M. Kallela, R.M., B.D.V., G.T., L.Q., M.A.I., L.L., E.H., M.S., H.S., K.S., E.J., T.F., and B.M. conceived and designed the study; V. Anttila, B.S.W., A. Palotie, D.R.N., G.D.S., M.W., M.D., A.J.M., M.J.D., C.K., J.-A.Z., M.-R.J., O.R., C.V.D., D.I.B., J.K., E.B., M. Kallela, B.D.V., G.T., E.H., T.F., R.R.F., N.G.M., A.G.U., TM, and J.G.E. performed the experiments. V. Anttila, B.S.W., P.G., D.I.C., D.R.N., M.J.D., D.P.S., E.B., J.R.G., S.B.R.J., T.K., F.B., G.M., R.M., M.D.V., S.E.M., L.Q., M.A.I., L.L., J.W., P.P., M.S., S. Steinberg, T.F. and B.M.-M. performed statistical analysis. V. Anttila, B.S.W., P.G., A. Palotie, D.I.C., D.R.N., L.C., J.R.G., S.B.R.J., K.L., T.K., F.B., G.M., R.M., B.D.V., S.E.M., L.Q., M.A.I., L.L., J.W., P.P., M.S., S. Steinberg, H.S., T.F., N.A., B.M.-M. and D.T. analyzed the data. D.I.C., M.J.D., M.D., A. Palotie, G.D.S., M.W., M.D., A.J.M., M.J.D., M.D., C.K., D.P.S., L.C., J.-A.Z., M.-R.J., O.R., C.V.D., D.I.B., T.W.P., J.K., E.B., J.R.G., K.L., C.V., V. Anttila, B.S.W., P.G., T.K., F.B., G.M., M. Kallela, R.M., B.D.V., G.T., U.T., W.L.M., L.Q., M. Koiranta, M.A.I., L.L., A.H.S., L.L., I.D., B.M.W., M.S., L.M.R., J.E.B., P.M.R., S. Steinberg, H.S., F.J., D.A.L., D.M.E., S.M.R., M.F., V. Arttia, M.A.K., T.F., J.S., R.R.F., N.P., C.M.W., R.Z., A.C.H., PA.F.M., G.W.M., N.G.M., G.B., H.G., A. Heinze, K.H.-K., F.M.K.W., A.-L.H., A. Pouta, J.v.D.E., A.G.U., A. Hofman, J.-H., J.M.V., K.H., M.A., B.M.-M., S. Schreiber, T.M., H.E.W., A.A., J.G.E., B.J.T. and D.T. contributed reagents and/or analysis of the individual cohorts, as well as all the study participants.

**COMPETING FINANCIAL INTERESTS**

The authors declare no competing financial interests.

Reprints and permissions information is available online at http://www.nature.com/reprints/index.html.

---

1. Anttila, V. et al. Genome-wide association study of migraine implicates a common susceptibility variant on Bq22.1. *Nat. Genet.* 42, 869–873 (2010).
2. Freilinger, T. et al. Genome-wide association study reveals three susceptibility loci for common migraine in the general population. *Nat. Genet.* 43, 695–698 (2011).
3. Ligthart, L. et al. Meta-analysis of genome-wide association for migraine in six population-based European cohorts. *Eur. J. Hum. Genet.* 19, 901–907 (2011).
4. Dash, P.K., Hoehner, B. & Kandel, E.R. Injection of the cAMP-responsive element coactivators: tissue-specific expression and selective activation of CREB and CREM. *Nat. Rev. Neurosci.* 10, 126–140 (2009).
5. Dash, P.K., Hoehner, B. & Kandel, E.R. Injection of the cAMP-responsive element coactivators: tissue-specific expression and selective activation of CREB and CREM. *Nat. Rev. Neurosci.* 10, 126–140 (2009).
6. Firia, G.M., De Cesare, D. & Sassone-Corsi, P. A family of LIM-only transcriptional coactivators: tissue-specific expression and selective activation of CREB and CREM. *Mol. Cell. Biol.* 20, 8613–8622 (2000).
7. Freeman, E.A. & Greenberg, S. Phosphorylation of CREB by CaMKII. *Science* 280, 785–787 (1998).
8. Freeman, E.A. & Greenberg, S. Phosphorylation of CREB by CaMKII. *Science* 280, 785–787 (1998).
9. Freeman, E.A. & Greenberg, S. Phosphorylation of CREB by CaMKII. *Science* 280, 785–787 (1998).
10. Freeman, E.A. & Greenberg, S. Phosphorylation of CREB by CaMKII. *Science* 280, 785–787 (1998).
Medical Centre, Rotterdam, The Netherlands. 23Department of Clinical Chemistry, Finnlab Laboratories and University of Tampere School of Medicine, Tampere, Finland. 24Department of Biological Psychology, VU University, Amsterdam, The Netherlands. 25EMGO+ Institute for Health and Care Research, VU University Medical Centre, Amsterdam, The Netherlands. 26Department of Public Health, Hjielt Institute, University of Helsinki, Helsinki, Finland. 27European Bioinformatics Institute, European Molecular Biology Laboratory, Wellcome Trust Genome Campus, Cambridge, UK. 28Department of Neurology, University Hospital Essen, Essen, Germany. 29Department of Epidemiology, Harvard Medical School, Boston, Massachusetts, USA. 30Department of Neurology, Landspitali University Hospital, Reykjavik, Iceland. 31Folkhälsoan Research Center, Helsinki, Finland. 32Department of Neurology, Klinikum der Universität München, Munich, Germany. 33Headache Research Unit, Department of Neurology and Groupe Interdisciplinaire de Génoprotéomique Appliquée (GIQA)-Neurosciences, Libra University, Liège, Belgium. 34Department of Psychiatry, Washington University School of Medicine, St. Louis, Missouri, USA. 35Kiel Pain and Headache Center, Kiel, Germany. 36Department of Clinical Sciences/Oncology and Gynecology, University Hospital of Oulu, Oulu, Finland. 37Medical Research Center, University of Oulu, Oulu, Finland. 38Department of Children, Young People and Families, National Institute for Health and Welfare, Helsinki, Finland. 39Department of Internal Medicine, Erasmus Medical Centre, Rotterdam, The Netherlands. 40Department of Genomics, Life & Brain Center, University of Bonn, Bonn, Germany. 41Institute of Human Genetics, University of Bonn, Bonn, Germany. 42Max Planck Institute of Psychiatry, Munich, Germany. 43Munich Cluster for Systems Neurology (SyNergy), Munich, Germany. 44Institute of Clinical Molecular Biology, Christian Albrechts University, Kiel, Germany. 45Department of Internal Medicine I, Christian Albrechts University, Kiel, Germany. 46Institute of Human Genetics, Helmholtz Center Munich, Neuherberg, Germany. 47Institute of Human Genetics, Klinikum Rechts der Isar, Technische Universität München, Munich, Germany. 48Institut für Medizinische Informationsverarbeitung, Biometrie und Epidemiologie, Ludwig-Maximilians-Universität München, Munich, Germany. 49Institute of Epidemiology I, Helmholtz Center Munich, Neuherberg, Germany. 50Klinikum Großhadern, Ludwig-Maximilians-Universität München, Munich, Germany. 51National Institute for Health and Welfare, Helsinki, Finland. 52Department of General Practice, Helsinki University Central Hospital, Helsinki, Finland. 53Vaasa Central Hospital, Vaasa, Finland. 54Department of General Practice and Primary Health Care, University of Helsinki, Helsinki, Finland. 55Laboratory of Neurogenetics, National Institute on Aging, US National Institutes of Health, Bethesda, Maryland, USA. 56Department of Molecular Neuroscience, Institute of Neurology, University College London, London, UK. 57Department of Genetics, King Faisal Specialist Hospital and Research Centre, Riyadh, Saudi Arabia. 58Details appear in the Supplementary Note. 59Center for Human Genetic Research, Massachusetts General Hospital, Boston, Massachusetts, USA. 60Pediatric Surgical Research Laboratories, Massachusetts General Hospital for Children, Massachusetts General Hospital, Boston, Massachusetts, USA. 61Center for Protein Research, University of Copenhagen, Copenhagen, Denmark. 62Center for Protein Research, University of Copenhagen, Copenhagen, Denmark. 63Center for Protein Research, University of Copenhagen, Copenhagen, Denmark. 64Department of Mental Health and Alcohol Research, National Institute for Health and Welfare, Helsinki, Finland. 65Department of Psychiatry, Leiden University Medical Centre, Leiden, The Netherlands. 66Department of Psychiatry, University Medical Centre Groningen, Groningen, The Netherlands. 67Research Centre of Applied and Preventive Cardiovascular Medicine, University of Turku, Turku University Hospital, Turku, Finland. 68Department of Clinical Physiology and Nuclear Medicine, Turku University Hospital, Turku, Finland. 69Department of Epidemiology and Biostatistics, School of Public Health, MRC–Health Protection Agency (HPA) Centre for Environment and Health, Faculty of Medicine, Imperial College, London, UK. 70Biocenter Oulu, University of Oulu, Oulu, Finland. 71Division of Population Health Sciences and Education, St George’s, University of London, London, UK. 72School of Medicine, University of Iceland, Reykjavik, Iceland. 73These authors contributed equally to this work. Correspondence should be addressed to A. Palotie (ap8@sanger.ac.uk) or V. Anttila (anttila@atgu.mgh.harvard.edu).
ONLINE METHODS

Overall study design. For this meta-analysis, we used SNP marker data from 23,285 cases and 95,425 controls of European descent from 29 studies, including 5 clinic-based studies, compared to population-matched control samples with unknown migraine status, as well as 14 entirely population-based cohorts. Four of the population-based cohorts (BSBC, NFBBC1966, Young Finns and FinnTwin; see Supplementary Note for further details) were birth cohorts. The data sets for the meta-analysis included previously genotyped GWAS data from migraine-specific studies by the International Headache Genetics Consortium (see URLs) and the Women’s Genome Health Study, as well as a number of pre-existing population-based GWAS (for a complete list of references, see Supplementary Table 2). Local research ethics committees approved the individual studies, and informed consent was obtained from all participants when necessary (see the Supplementary Note for full details of ethics and consent procedures for each study). Additional details on sample recruitment and phenotypes and summary details for each collection are given in the Supplementary Note and in Supplementary Tables 1 and 2. Genome-wide SNP genotyping was performed independently in each cohort with the use of various standard genotyping technologies, and genotypes were imputed when necessary (see the Supplementary Note for full details of ethics and consent procedures for each study). Addo-on details on sample recruitment and phenotypes and summary details for each collection are given in the Supplementary Note and in Supplementary Tables 1 and 2. Genome-wide SNP genotyping was performed independently in each cohort with the use of various standard genotyping technologies, and genotypes were imputed for each study with reference to HapMap release 21 or 22 CEU (Utah residents of Northern and Western European ancestry) phased genotypes.

Study phenotypes. The primary phenotype analyzed was migraine of any type, regardless of source. This initial step was followed by a subgroup analysis consisting of (i) only the clinical samples, (ii) only samples satisfying criteria for migraine with aura and (iii) only samples satisfying criteria for migraine without aura. Population-based samples were not analyzed across the genome as a subgroup because they constituted 78% of cases and 85% of controls in the main analysis, but associations were calculated for significant SNPs for comparative purposes. In the clinical cohorts, a headache specialist assigned a migraine diagnosis on the basis of an interview in person or over the telephone or through the use of an extensive questionnaire. For the population studies, migraine status for individuals in a study sample was determined using a questionnaire (Supplementary Note).

Statistical analysis of GWAS data. Each study contributed summary statistical data from an association analysis performed using a frequentist additive model based on an expected allelic dosage model for SNP markers, adjusting for sex (using either SNPTTEST or ProbABEL (see URLs)). SNPs were filtered on the study level with inclusion criteria of MAF of >0.1% and imputation quality measure (I_q) of >0.6 (IMPUTE 2)\(^{36}\) or r\(^2\) of >0.3 (MACH)\(^{37}\). Four of the included studies contained new genotyping (HUNT) or imputation (the Finnish, German and Dutch MA studies and HUNT). In the meta-analysis, combined association data for ~2.3 million imputed and genotyped autosomal SNPs with a heterogeneity coefficient \(\lambda = 1.15; \) Supplementary Fig. 4. For estimating genome-wide significance, we used the commonly accepted thresholds of 5 × 10\(^{-8}\) for primary loci\(^{38}\) and 1 × 10\(^{-5}\) for secondary loci, in accordance with the reporting threshold for the GWAS catalog\(^{39}\). At secondary loci, to limit spurious associations, at least two SNPs within a 50-kb window were required to pass the significance threshold. We also estimated the robustness of this threshold using the FDR method of Benjamini and Hochberg\(^{40}\), showing that the \(P\)-value threshold corresponding to FDR of <0.05 was 2.33 × 10\(^{-5}\). The quantile-quantile plot of the meta-analysis \(P\) values (Supplementary Fig. 4) showed a marked excess of association signals, well beyond the number expected by chance, below the threshold for suggestive association used for reporting. The loci with significant associations were visualized using the LocusZoom interface\(^{41}\). For the heterogeneity analyses for migraine type (Supplementary Table 9), owing to shared controls in some of the sets, the available study samples were divided into groups that were as equally sized as possible in terms of effective study size, and the data were then analyzed using the sex heterogeneity analysis method\(^{42}\) (– sex option) of GWAMA\(^{43}\), with a dummy variable coding for migraine with aura and migraine without aura instead of sex. For the heterogeneity analyses for sex, the same method was used to compare \(P\) values from male-only and female-only analyses.

Pathway analyses. MAGENTA. MAGENTA software\(^ {44}\) was used to conduct an analysis to evaluate whether markers located in genes involved in particular biological pathways (such enrichment of low \(P\) values for association with migraine, using pathway lists from GO, Panther, Ingenuity, KEGG, Reactome and Biocarta (Supplementary Table 10). In gene set enrichment analysis, \(P\) values were estimated via 10,000 permutations of genes evaluated at the 75th percentile of FDR (owing to the assumption of high polygenicity), manually corrected to account for FDR across all pathway sets.

DAPPLE. Using a refined database of high-confidence protein-protein interactions (InWeb\(^ {45,46}\), we employed DAPPLE\(^ {47}\) to assess the amount of physical interactions connecting the genes within 50 kb of the 146 reported migraine loci, as well as an analysis of only the 16 proteins encoded by genes directly underlying the 12 genome-wide significant loci (Supplementary Fig. 10 and Supplementary Table 11). Both direct and indirect (through first-order common interaction partners) were measured and compared to a random expectation over 10,000 permutations, and the resulting network was plotted.

GRAIL. The GRAIL web interface\(^ {48}\) was used to explore similarities in published PubMed articles (August 2012 freeze), using data from HapMap release 22 CEU samples and with gene size correction set to on. From the GRAIL results, only genes with significant \(P\) values (\(P < 0.05\)) for co-occurrence in published abstracts are shown, and the list of similar genes was capped to include only genes within the top 200 highest rankings (Supplementary Table 12).

Overlap with DNase I hypersensitivity sites. The positions of SNPs in migraine-associated loci were overlapped with DNase I hotspot regions from the Encyclopedia of DNA Elements (ENCODE) Project that mark generalized chromatin accessibility, mapped for each of 125 diverse cell lines and tissues\(^ {49}\). To assess the significance of overlap for the set of SNPs as a whole, 100 background sets of SNPs were chosen from the genome so that each migraine-associated SNP was matched in each set by a SNP within the same decile for MAP, distance to the nearest transcriptional start site and GC content of the 100-base region surrounding the SNP. Background SNP sets were overlapped with the DNase I hotspots, and the enrichment for overlap with the migraine-associated SNP set was expressed as the \(z\) score relative to the distribution of the background SNP set overlaps, separately for each cell line. In addition, migraine-associated SNPs were analyzed for other overlap with ENCODE data, including with transcription factor motifs, using RegulomeDB (see URLs)\(^ {50}\).

Tissue-based gene expression analysis. For the tissue analysis, a microarray-based analysis of gene expression was performed on a data set of 55,269 samples in the Gene Expression Omnibus (GEO) that were measured on the Affymetrix U133 Plus 2.0 array. Each sample in the raw expression data was first linearly transformed using a modified invariant set normalization method\(^ {51}\) based on a set of 80 control genes with stable expression on the U133 Plus 2.0 array. Expression data were log\(_2\) transformed to stabilize the variance and distribution of expression levels. Finally, data were quantile normalized to match the expression distribution of each sample. Expression values for genes with multiple probe sets were calculated by taking the median value of all probe sets for that gene. After normalization, a log\(_2\) expression value of four was considered baseline, and a log\(_2\) expression value of greater than six was considered to indicate expression. Sample annotations were curated on the basis of GEO descriptions provided by depositors. To account for variation in the number of samples representing each tissue in the data set, expression of a gene was plotted to show the fraction of samples of a tissue that exceeded a log\(_2\) expression value of six, with higher fractions indicating more ubiquitous expression in the tissue in question.

eQTL analysis. On the basis of the meta-analysis results for association with migraine, 146 regions of interest were queried against eQTL results from the North American Brain Expression Consortium and UK Brain Expression Consortium studies (GEO accession GSE36192 and database of Genotypes and Phenotypes (dbGaP) accession phs000249). These eQTL results are based on mRNA expression levels in cerebellum and frontal cortex tissue from 394
human subjects (Supplementary Note). Of the 146 associated regions, 134 were represented in the eQTL analysis. Within these regions, 831 mRNA expression traits and 222,668 cis SNP–expression trait pairs were considered in cerebellum, and 864 mRNA expression traits and 230,660 cis SNP–expression trait pairs were considered in the frontal cortex. Forty-five SNPs within the migraine-associated loci were found to have significant correlation (evaluated at FDR-corrected thresholds: 0.0001668 for frontal cortex and 0.0002187 for cerebellum) with the expression of 12 mRNA transcripts in both the cerebellum and the frontal cortex. A more detailed methods description can be found in the Supplementary Note. The extent of linkage disequilibrium between the SNPs associated with migraine and the SNPs in the tested eQTL SNP-probe pairs was evaluated using SNAP (see URLs).