

Rare variants with large effects provide functional insights into the pathology of migraine subtypes, with and without aura

Received: 2 December 2022

A list of authors and their affiliations appears at the end of the paper

Accepted: 18 September 2023

Published online: 26 October 2023

 Check for updates

Migraine is a complex neurovascular disease with a range of severity and symptoms, yet mostly studied as one phenotype in genome-wide association studies (GWAS). Here we combine large GWAS datasets from six European populations to study the main migraine subtypes, migraine with aura (MA) and migraine without aura (MO). We identified four new MA-associated variants (in *PRRT2*, *PALMD*, *ABO* and *LRRK2*) and classified 13 MO-associated variants. Rare variants with large effects highlight three genes. A rare frameshift variant in brain-expressed *PRRT2* confers large risk of MA and epilepsy, but not MO. A burden test of rare loss-of-function variants in *SCN11A*, encoding a neuron-expressed sodium channel with a key role in pain sensation, shows strong protection against migraine. Finally, a rare variant with *cis*-regulatory effects on *KCNK5* confers large protection against migraine and brain aneurysms. Our findings offer new insights with therapeutic potential into the complex biology of migraine and its subtypes.

Migraine is a complex neurovascular disease characterized by recurrent, disabling headache attacks that are difficult to treat. It is among the most common pain disorders worldwide, with prevalence of up to 20% in adult populations and affecting three times more females than males¹. Two main subtypes are clinically distinguished, migraine with aura (MA) and migraine without aura (MO)². MO is characterized by severe headache attacks accompanied by nausea and hypersensitivity to light and sound, whereas MA is characterized by gradually spreading, fully reversible focal neurological symptoms, collectively called aura, that are usually followed by headache¹. An estimated 30% of migraineurs have MA, and the most frequently experienced aura involves visual disturbances (for example, flashes of bright light and blurred vision)³. During MA attacks, characteristic regional brain blood flow changes indicate that MA is caused by cortical spreading depression, a transient wave of neuronal depolarization of the cortex^{4,5}. Such findings are not observed in MO^{6,7}, suggesting divergent pathogenesis of these migraine subtypes. A rare and clinically distinct subtype of MA is familial hemiplegic migraine (FHM)². Three genes have been linked to FHM—one encoding a membrane protein involved in maintaining gradients of sodium and potassium ions across plasma membranes

(*ATPIA2*), and two genes encoding sodium and calcium channels expressed in brain (*SCN1A* and *CACNA1A*, respectively)⁸.

More is known about the genetics and biology of migraine than any other pain disorder, leading to recent treatment advances such as those targeting the calcitonin gene-related peptide (CGRP) activation of the trigeminovascular system^{9,10}. The largest genome-wide association studies (GWAS) meta-analysis of migraine to date identified 123 migraine risk loci, among them a locus including genes encoding CGRP (*CALCA* and *CALCB*)¹¹. However, the pathophysiology of migraine is not fully understood, and a substantial subset of patients has treatment-resistant migraine¹². In the study reporting 123 common (minor allele frequency (MAF) > 2%) migraine variants, subtype analysis showed that 5 associate specifically with migraine subtypes—3 with MA (in or near *CACNA1A*, *HMOX2* and *MPPED2*) and 2 with MO (near *SPINK2* and *FECH*)^{11,13}. These findings suggest that the genetics of MA and MO should be studied separately and with more emphasis on detecting rare variants.

To identify both distinct and common biological underpinnings of these migraine subtypes, we performed GWAS meta-analyses of clinically defined MA, MO and overall migraine, using six datasets and analyzing variants down to 0.001% in frequency. We used samples from

✉ e-mail: gyda.bjornsdottir@decode.is; kari.stefansson@decode.is

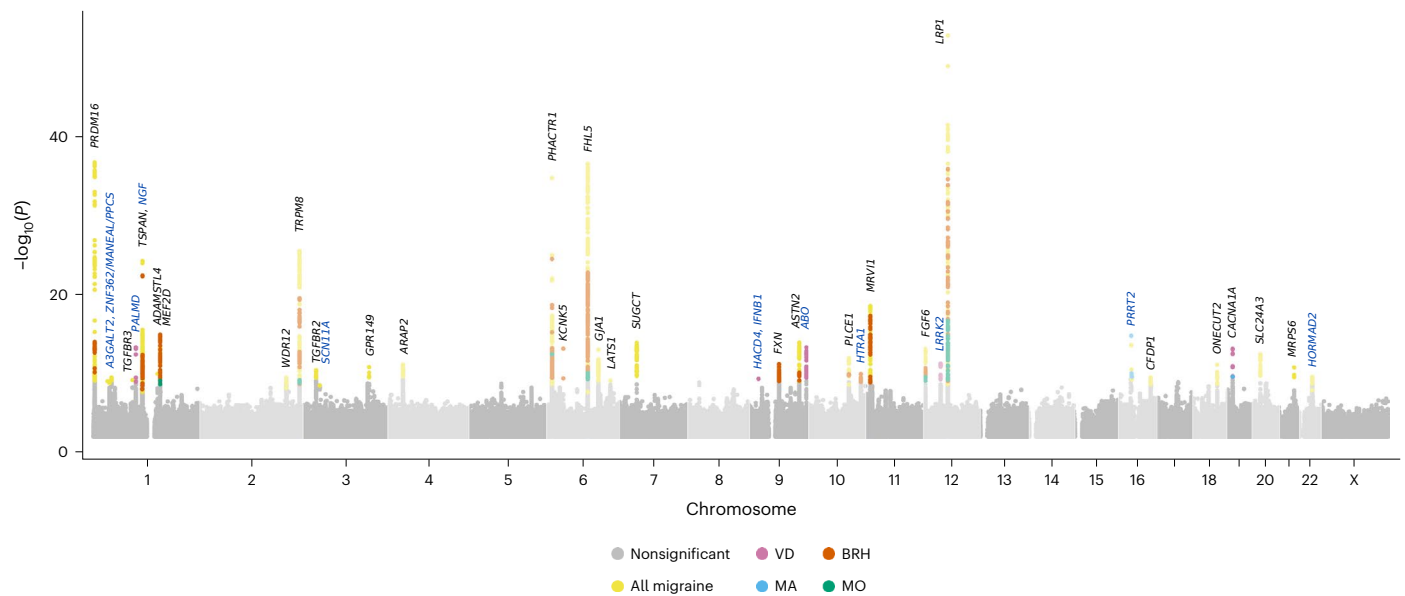


Fig. 1 | Manhattan plot of GWAS meta-analysis results for all studied phenotypes. The graph shows data for migraine ($n_{\text{case/control}} = 74,495/1,259,808$), MA ($n_{\text{case/control}} = 16,603/1,336,517$), MO ($n_{\text{case/control}} = 11,718/1,330,747$), VD ($n_{\text{case/control}} = 30,297/86,134$) and BRH ($n_{\text{case/control}} = 51,803/123,732$). See Supplementary Table 1 for $n_{\text{case/control}}$ for each cohort. On the x axis, variants are plotted along the 22 autosomes and the X chromosome. On the y axis is the statistical significance of their association with the respective phenotypes from meta-analyses using a fixed-effects inverse-variance method based on effect

estimates and s.e. under the additive model, in which each dataset was assumed to have a common OR but allowed to have different population frequencies for alleles and genotypes. Gray dots are not significant variants. Variant associations that reach the P threshold weighted by variant annotation²¹ are represented by color-coded dots. Adjacent chromosomes are presented in different shades of gray. Known migraine loci are represented by gene names in black text, and new loci are represented by gene names in blue text.

about 1.3 million individuals, of which 12,000 have MO, 17,000 have MA and 80,000 have migraine. Because migraine and especially its subtypes are considerably underdiagnosed¹⁴, and to obtain measures of specific symptoms and severity, we also assessed self-reported proxy phenotypes representing severe and recurrent migraine headaches (52,000 cases) as well as migraine's most distinctive subtype, headaches preceded by visual aura (30,000 cases). Here we report 4 new MA-associated variants and show that 13 known migraine variants associate with MO over MA. In all, we observed associations with 44 lead variants, 12 of which are new for migraine, and we found functional evidence implicating 22 genes—3 in MA, 3 in MO and the remainder in overall migraine. Among the findings are rare variants with large effects providing new insights into biological underpinnings of distinct characteristics of migraine, with and without aura.

Results

We conducted GWAS meta-analyses of clinically defined migraine, MA and MO, using datasets from Iceland (deCODE Genetics), Denmark (Copenhagen Hospital Biobank (CHB)¹⁵ and Danish Blood Donor Study (DBDS)¹⁶), the United Kingdom (UK; UK Biobank¹⁷), the United States (US; Intermountain Health¹⁸), Norway (the Hordaland Health Study (HUSK)¹⁹) and Finland (FinnGen²⁰). We also performed GWAS meta-analyses of two self-reported proxy phenotypes available in three datasets (Iceland, UK and Denmark)—an MA proxy represented by experiencing visual disturbances (VD) preceding headaches, and a severe migraine proxy represented by bad and recurrent headaches (BRH). In total, we analyzed data on 1.3 million individuals, including 16,603 with MA, 11,718 with MO, 79,495 with any migraine, 30,297 with VD and 51,803 with BRH (Methods; Supplementary Table 1). We analyzed up to 85 million variants, and using a significance threshold weighted by variant impact²¹, we found associations with 44 lead variants at 39 loci (Fig. 1, Tables 1 and 2 and Supplementary Tables 2–7). Two variants associate with MA (one new), five with the MA-proxy VD (four new) and six with MO. The remaining variants associate with overall migraine or

BRH. In all, we report 12 new migraine variants (regional plots shown in Supplementary Figs. 1 and 2).

Using cross-trait linkage disequilibrium (LD) score regression²², we calculated genetic correlations in nonoverlapping samples (Methods) showing that VD correlates genetically with clinically defined MA ($r_g = 0.65$, $P = 4.0 \times 10^{-23}$) but not MO ($r_g = -0.09$, $P = 0.21$), and BRH correlate strongly with clinically defined migraine ($r_g = 0.85$, $P = 7.4 \times 10^{-91}$; Supplementary Table 8 and Supplementary Fig. 3). Further supporting VD as an MA proxy, the GWAS meta-analysis of VD reveals an association with a variant (**rs11085837-A**) in high LD ($r^2 = 0.96$) with the reported MA variant in *CACNA1A*, **rs10405121-A¹¹** (Fig. 1 and Table 1). Its VD effect (odds ratio (OR) = 0.926, $P = 8.8 \times 10^{-14}$) is consistent with its MA effect (OR = 0.930, $P = 1.8 \times 10^{-9}$), and no association is detected with MO (OR = 0.983, $P = 0.22$). In Supplementary Table 9, we list associations with all migraine phenotypes of the current study with the recently published 123 migraine variants¹¹, finding support ($P < 0.05$) in our data for all but 9 variants (Supplementary Note 1).

A rare loss-of-function *PRRT2* variant associates with MA

The top MA association is with a rare insertion in *PRRT2* leading to frameshift (**rs587778771-GCC**, p.Arg217ProfsTer8; OR = 5.446, $P = 5.6 \times 10^{-16}$). This variant also associates with VD (OR = 3.634, $P = 0.0037$) but not MO ($P = 0.97$; Table 3). It is detected in only three cohorts, with a founder effect observed in Iceland (frequency = 0.117%), compared to UK and US (frequency = 0.013% and 0.0051%, respectively). It is detected at even lower frequencies in samples from Denmark, with no carriers detected in Norway or Finland. This variant has been reported in case studies of rare neurological disorders, including benign infantile seizures and paroxysmal kinesigenic dyskinesia (PKD)²³. In a few carriers, FHM has also been detected⁸. Among six Danish heterozygous carriers identified, five are in the same family, of which three have FHM.

The p.Arg217ProfsTer8 insertion is located in an unstable DNA site^{24,25} where we find another rarer (0.024%) deletion

Table 1 | Lead variants associated with migraine subtypes and headache-related visual disturbances (MA proxy)

Phenotypes	Locus	Position hg38	Variants	OA	EA	EAF (%)	Nearest genes	Variant annotation	OR (95% CI)	P	P_{bonf}	P_{het}	SNP previously reported at locus (r^2 if correlated SNP)
MA	16p11.2	29813694	rs587778771	GC	GCC	0.05	<i>PRRT2</i>	Frameshift	5.446 (3.626, 8.148)	5.6×10^{-16}	6.7×10^{-10}	–	–
MA	19p13.13	13228314	rs10405121	G	A	27.3	<i>CACNA1A</i>	Intron	0.927 (0.905, 0.949)	2.5×10^{-10}	0.03	0.24	rs10405121 ^a
VD	1p21.2	99579683	rs11166276	T	C	49.0	<i>PALMD</i>	TF-binding site	0.926 (0.907, 0.945)	5.1×10^{-14}	2.2×10^{-6}	0.87	–
VD	9p21.3	21047562	rs77778288	A	C	12.9	<i>HACD4/IFNB1</i>	Regulatory region	1.097 (1.065, 1.129)	4.9×10^{-10}	0.02	0.77	–
VD	9q34.2	133257521	rs8176719	T	TC	32.9	<i>ABO</i>	Frameshift	1.081 (1.059, 1.104)	3.0×10^{-13}	1.2×10^{-7}	0.07	–
VD	12q12	40221267	rs10748014	C	T	46.8	<i>LRRK2</i>	Upstream	1.073 (1.052, 1.094)	5.6×10^{-12}	0.00012	0.11	–
VD	19p13.13	13234712	rs11085837	G	A	45.6	<i>CACNA1A</i>	Intron	0.926 (0.907, 0.945)	8.8×10^{-14}	3.7×10^{-6}	0.18	rs10405121 (0.96) ^a
MO	1q22	156460957	rs750439 ^c	T	C	33.9	<i>MEF2D</i>	Downstream	1.092 (1.062, 1.123)	8.7×10^{-10}	0.017	0.06	rs2274319 (0.64) ^a , rs1925950 (0.86) ^b
MO	2q37.1	233937757	rs12470426 ^c	G	A	9.1	<i>TRPM8</i>	Intron	0.853 (0.812, 0.897)	6.0×10^{-10}	7.6×10^{-2}	0.56	rs10166942 (0.39) ^a , rs10166942 (0.96) ^b
MO	6p24.1	12903725	rs9349379 ^c	A	G	41.7	<i>PHACTR1</i>	3' UTR	0.904 (0.879, 0.929)	3.8×10^{-13}	8.0×10^{-6}	0.35	rs9349379 ^{ab}
MO	6q16.1	96610677	rs2273621 ^c	A	G	32.3	<i>FHL5</i>	Missense	1.096 (1.065, 1.128)	3.0×10^{-10}	1.9×10^{-3}	0.61	rs11153082 (1.0) ^a , rs67338227 (0.54) ^b
MO	12p13.32	4418156	rs2160875 ^c	T	C	49.1	<i>FGF6</i>	Regulatory region	1.088 (1.059, 1.118)	8.2×10^{-10}	3.4×10^{-2}	0.85	rs2160875 ^a , rs1024905 (0.83) ^b
MO	12q13.3	57132863	rs4759276 ^c	G	A	39.9	<i>LRP1</i>	Intron	0.889 (0.865, 0.914)	9.8×10^{-17}	1.3×10^{-8}	0.32	rs11172113 (0.84) ^{ab}

Supplementary Table 1 shows $n_{\text{case/control}}$ per cohort and Supplementary Table 7 shows associations of these variants with all migraine. Discovery phenotype is in the first column: MA, headache-related VD proxy for MA, MO. Effect allele frequency (EAF) is the average frequency of EA in the cohorts studied (Supplementary Table 1; Methods). OR and P value for inverse-variance weighted meta-analysis of association results for all cohorts. P_{bonf} is the P value after a variant class-specific Bonferroni adjustment²¹. P_{het} is the heterogeneity P value from a likelihood ratio test. Results per cohort and for all phenotypes are in Supplementary Tables 2–7. Associations of these and correlated variants ($r^2 > 0.8$) with various traits listed in the GWAS Catalog (<https://www.ebi.ac.uk/gwas/>) are in Supplementary Table 12. Bold are variants that associate primarily with MO, over MA or VD (Fig. 3). CI, confidence interval; EA, effect allele; OA, other allele. ^aSNPs previously reported in ref. 11. ^bSNPs previously reported in ref. 69. ^cMO-associated variants that also correlate with migraine variants in Table 2; [rs750439](#) ($r^2 = 0.64$) with [rs1925950](#), [rs12470426](#) ($r^2 = 0.39$) with [rs1003540](#), [rs2160875](#) ($r^2 = 1.0$) with [rs7957385](#), [rs4759276](#) ($r^2 = 0.84$) with [rs11172113](#) and [rs9349379](#) and [rs2273621](#) also associate with migraine.

(p.Arg217GlufsTer12) that also leads to premature PRRT2 truncation²⁵. This variant also shows a founder effect in Iceland, being tenfold more frequent than in the UK (frequency of 0.0025%), and not detected in other cohorts. It was previously reported in a single case study of a homozygous carrier with severe PKD that responded to carbamazepine, an epilepsy drug that reduces the generation of rapid action potentials in the brain²⁶ and is also used to treat migraine. We found p.Arg217GlufsTer12 in 38 heterozygous carriers in Iceland, mainly in two families where it segregates with migraine and epilepsy. Of 38 carriers, 11 (29%) are diagnosed with migraine (without subtype), six (16%) with epilepsy and one with MA and epilepsy.

For these rare variants, we looked for associations with other phenotypes. Apart from the MA and migraine associations, p.Arg217ProfsTer8 associates only with epilepsy (OR = 7.077, $P = 1.9 \times 10^{-35}$; Table 3 and Supplementary Table 10). We find epilepsy moderately genetically correlated with migraine ($r_g = 0.28$, $P = 9.4 \times 10^{-6}$) and VD ($r_g = 0.28$, $P = 2.8 \times 10^{-4}$), but not with MO ($r_g = 0.05$, $P = 0.90$). We tested 30 epilepsy variants²⁷ in our data and found that only two also impact migraine (at $P < 3.3 \times 10^{-4} = 0.05/30$ variants \times 5 phenotypes). The common (23.3%) intron variant [rs59237858](#)-T in *SCN1A* that confers protection against epilepsy²⁷ confers risk of migraine (OR = 1.031, $P = 8.6 \times 10^{-6}$) in our data, and [rs62151809](#)-T (44.7%) near *TMEM182* confers risk of epilepsy²⁷ and of VD in our data (OR = 1.047, $P = 8.5 \times 10^{-6}$). None of the 30 epilepsy variants associate with MO or BRH (Supplementary Table 11). Conversely, of the 44 variants reported here, only p.Arg217ProfsTer8 associates with epilepsy.

GWAS meta-analysis of MA-proxy phenotype yields new MA-associated loci

Besides the known MA-associated variant in *CACNA1A*, we found four other variants associating with the MA-proxy VD, all new to migraine (Table 1). The first, [rs11166276](#)-C, is in a TF-binding site near *PALMD* (OR = 0.926, $P = 5.1 \times 10^{-14}$). It is in complete LD with [rs7543130](#) that also associates protectively with aortic valve stenosis²⁸. Secondly, in *ABO*, the frameshift variant [rs8176719-TC associates with VD \(OR = 1.081, \$P = 3.0 \times 10^{-13}\$ \). This variant contributes to determining the non-O blood groups²⁹, and variants in high LD associates with various coagulation factors and risk of venous thromboembolism \(Supplementary Table 12\). This variant associates with MA \(OR = 1.030, \$P = 0.015\$ \) and overall migraine \(OR = 1.020, \$P = 1.5 \times 10^{-3}\$; Supplementary Table 7\). Thirdly, a variant upstream of *LRRK2*, \[rs10748014\]\(#\)-T, associates with VD \(OR = 1.073, \$P = 5.6 \times 10^{-12}\$ \). *LRRK2* encodes leucine-rich repeat kinase 2, a gene harboring common risk variants for inherited Parkinson's disease \(PD\)³⁰, none of which are in LD with \[rs10748014\]\(#\) \(Supplementary Table 12\). This variant also associates with MA \(OR = 1.065, \$P = 8.4 \times 10^{-8}\$ \) and weakly with overall migraine \(OR = 1.012, \$P = 0.048\$ \), and we detected no association with MO or PD. Finally, in a regulatory region near *HACD4/IFNB1* is an association with \[rs77778288\]\(#\)-C \(frequency = 12.9%, OR = 1.097, \$P = 4.9 \times 10^{-10}\$ \). *IFNB1* encodes interferon \$\beta\$ 1, which is used to treat multiple sclerosis and can induce headaches³¹.](#)

We compared the effects of these VD variants on MA and all migraine in effect–effect plots (Fig. 2). Based on the slope derived from a weighted regression through the origin, overall MA and migraine effect estimates are 73% and 29%, respectively, of VD effect estimates,

Table 2 | Variants identified in association with all migraine (M) or migraine proxy (BRH)

Pheno-types	Locus	Position hg38	Variants	OA	EA	EAF (%)	Nearest genes	Variant annotation	OR (95% CI)	P	P_{bonf}	P_{het}	SNP previously reported at locus (r^2 if correlated SNP)
M	1p36.32	3155918	rs10797381	T	A	22.7	<i>PRDM16</i>	Intron	1.094 (1.079, 1.109)	2.1×10^{-37}	8.7×10^{-30}	0.08	rs10218452 (1.0) ^a , rs2651899 (0.37) ^b
M	1p36.1	33302206	rs933718575	A	G	0.01	<i>A3GALT2</i>	Downstream	11.032 (5.11, 23.8)	9.7×10^{-10}	0.020	–	
M	1p34.3	37790755	rs71642605	T	C	25.3	MANEAL	Upstream	1.042 (1.028, 1.056)	1.1×10^{-9}	0.023	0.89	–
M	1p34.2	42465863	rs11799356	G	A	34.2	<i>PPCS</i>	Downstream	1.039 (1.026, 1.052)	6.0×10^{-10}	0.013	0.67	–
M	1p22.1	91731541	rs12070846	T	C	23.0	<i>TGFBR3</i>	Intron	1.044 (1.030, 1.058)	6.8×10^{-10}	0.029	0.16	rs11165300 (0.88) ^a
M	1p13.2	115135325	rs12134493 ^c	C	A	12.0	<i>TSPAN2</i>	TF-binding site	1.112 (1.092, 1.132)	1.7×10^{-30}	7.2×10^{-23}	0.03	rs2078371 (1.0) ^a , rs12134493 ^b
M	1p13.2	115286692	rs6330 ^c	G	A	46.7	<i>NGF</i>	Missense	1.035 (1.023–1.048)	2.1×10^{-8}	0.041	0.06	–
M	1q21.1	150538184	rs6693567	T	C	26.0	ADAMSTL4	Regulatory region	1.044 (1.031, 1.058)	1.1×10^{-10}	0.0046	0.64	rs6693567^{a, b}
M	1q22	156480948	rs1925950	A	G	36.0	MEF2D	Synonymous	1.047 (1.034, 1.059)	6.3×10^{-14}	1.3×10^{-6}	0.08	rs2274319 (1.0)^a, rs3790455 (1.0)^b
M	2q33.2	202901033	rs35212307	T	C	12.6	<i>WDR12</i>	Missense	0.949 (0.933, 0.966)	6.7×10^{-9}	0.013	0.54	rs149163995 (0.99) ^b
M	2q37.1	233917239	rs1003540	A	G	19.4	<i>TRPM8</i>	Upstream	0.923 (0.910, 0.937)	3.3×10^{-26}	6.9×10^{-19}	0.50	rs10166942 (1.0) ^a
M	3p24.1	30424073	rs4955309	C	A	31.9	<i>TGFBR2</i>	Intergenic	1.042 (1.030, 1.055)	4.0×10^{-11}	0.005	0.10	rs7371912 (0.91) ^a , rs7640543 (0.97) ^b
M	3p22.2	38894643	rs33985936	C	T	25.0	<i>SCN11A</i>	Missense	1.041 (1.027, 1.054)	3.4×10^{-9}	0.0065	0.32	–
M	3q25.2	154572157	rs13078967	A	C	3.5	<i>GPR149</i>	Regulatory region	0.892 (0.862, 0.922)	1.6×10^{-11}	0.00066	0.27	rs13078967 ^{a, b}
M	4p15.1	35563301	rs74992952	G	A	17.9	ARAP2	Intergenic	0.949 (0.935, 0.963)	8.8×10^{-12}	0.00037	0.45	rs73805934 (0.92)^a
M	6p24.1	12903725	rs9349379	A	G	41.9	PHACTR1	3' UTR	0.928 (0.917, 0.939)	1.9×10^{-35}	3.9×10^{-28}	0.43	rs9349379^{a, b}
BRH	6p21.2	39280316	rs72854118	A	G	0.67	<i>KCNK5</i>	TF-binding site	0.697 (0.634, 0.766)	7.6×10^{-14}	3.2×10^{-5}	0.10	–
M	6q16.1	96610677	rs2273621	A	G	32.3	<i>FHL5</i>	Missense	1.082 (1.069, 1.096)	1.1×10^{-36}	2.1×10^{-30}	0.14	rs11153082 (0.99) ^a , rs11759769 (0.55) ^b
M	6q22.31	121487928	rs7743275	G	A	19.9	<i>GJA1</i>	Regulatory region	1.060 (1.044, 1.077)	9.7×10^{-14}	1.2×10^{-5}	0.17	rs28455731 (0.73) ^{a, b}
M	6q25.1	149721026	rs1359155039	TAAA AAAA	TAAA AAAAA	32.8	<i>LATS1</i>	Upstream	0.958 (0.945, 0.971)	8.1×10^{-10}	0.017	0.38	rs9383843 (0.87) ^a
M	7p14.1	40367277	rs186166891	A	T	10.4	SUGCT	Intron	1.084 (1.062, 1.106)	1.1×10^{-14}	1.4×10^{-6}	0.40	rs10234636 (0.91)^a, rs4379368 (0.91)^b
BRH	9q21.11	69099647	rs34965002	G	A	43.3	<i>FXN/TJP2</i>	Regulatory region	1.056 (1.039, 1.072)	6.6×10^{-12}	0.000277	0.19	rs7034179 (0.87) ^a
M	9q33.1	116479356	rs12684144 ^d	T	C	22.3	ASTN2	Intron	1.055 (1.041, 1.070)	1.3×10^{-14}	5.4×10^{-7}	0.01	rs3891689 (0.91)^a, rs6478241 (0.57)^b
M	10q23.33	94279840	rs2274224	G	C	41.5	PLCE1	Missense	0.959 (0.948, 0.970)	2.7×10^{-12}	5.1×10^{-6}	0.04	rs2274224^a, rs11187838 (1.0)^b
BRH	10q26.13	122470997	rs12252027	G	T	11.4	<i>HTRA1</i>	Intron	0.926 (0.904, 0.948)	1.2×10^{-10}	0.0149	0.73	–
M	11p15.4	10652192	rs4909945	C	T	33.0	MRVI1	Missense	0.945 (0.934, 0.957)	3.1×10^{-19}	5.9×10^{-13}	0.61	rs4910165 (1.0)^{a, b}
M	12p13.32	4416380	rs7957385	G	A	48.6	FGF6	Intergenic	1.045 (1.033, 1.057)	8.2×10^{-14}	1.0×10^{-5}	0.03	rs2160875 (1.0)^a, rs140668749 (1.0)^b

Table 2 (continued) | Variants identified in association with all migraine (M) or migraine proxy (BRH)

Pheno-types	Locus	Position hg38	Variants	OA	EA	EAF (%)	Nearest genes	Variant annotation	OR (95% CI)	P	P_{bonf}	P_{het}	SNP previously reported at locus (r^2 if correlated SNP)
M	12q13.3	57133500	rs11172113	T	C	42.3	LRP1	Intron	0.912 (0.901, 0.923)	1.8×10^{-53}	7.4×10^{-46}	0.42	rs11172113^{a,b}
M	16p11.2	29813694	rs587778771	GC	GCC	0.05	PRRT2	Frameshift	3.038 (2.320, 3.977)	6.6×10^{-16}	1.0×10^{-9}	0.83	–
M	16q23.1	75289942	rs17685540	C	T	41.0	CFDP1	Downstream	1.037 (1.025, 1.049)	1.4×10^{-9}	0.029	0.31	rs8046696 (0.98) ^a , rs77505915 (0.91) ^b
M	18q21.31	57494932	rs7233335	C	G	20.4	ONECUT2	Downstream	0.954 (0.941, 0.968)	1.8×10^{-10}	0.0038	0.01	rs8087942 (0.45) ^a
M	20p11.23	19494370	rs3827986	G	A	24.5	SLC24A3	Intron	1.050 (1.036, 1.064)	4.6×10^{-13}	1.9×10^{-5}	0.52	rs4814864 (1.0)^a
M	21q22.11	34221526	rs28451064	G	A	13.6	MRPS6	Regulatory region	0.943 (0.927, 0.959)	1.8×10^{-11}	0.00075	0.2	rs28451064^a
M	22q12.2	30076759	rs5753008	T	C	35.6	HORMAD2	Upstream	1.039 (1.027, 1.051)	4.0×10^{-10}	0.0084	0.5	–

Effect allele frequency (EAF) is the average frequency of EA in the cohorts studied. OR and P value for inverse-variance-weighted meta-analysis of association results for all cohorts (Supplementary Table 1; Methods). P_{bonf} is the P value after a variant class-specific Bonferroni adjustment²¹. P_{het} is the heterogeneity P value from a likelihood ratio test. Bold are variants that associate primarily with MO, not MA or VD, or with larger effects on MO than on MA or VD (Fig. 3). ^aSNPs previously reported in ref. 11. ^bSNPs previously reported in ref. 69. ^cResults presented are after adjusting for the respective effects of these uncorrelated ($r^2=0.02$) variants at this locus. Results per cohort for all studied phenotypes are in Supplementary Tables 2–7. ^drs12684144-C confers protection against VD and risk against MO.

and no associations were detected for MO, which is in line with our estimates of genetic correlation between these traits.

Migraine subtype classification of lead variants

We used a similar approach discussed in ref. 11 to study the effects of 43 lead variants on the migraine subtypes adjusting for sample overlap (PRRT2 excluded as it has larger effects than other variants and is shown to be an MA-associated variant; Methods). We find that the new variants in *ABO*, *LRRK2* and *PALMD*, and the previously reported¹¹ MA-associated variant in *CACNA1A* are classified as MA-associated variants, and 13 variants are classified as MO-associated variants (bold in Tables 1 and 2; Fig. 3 and Supplementary Fig. 4). All MO-associated variants are in known migraine loci except the new MO-associated variant rs71642605-C in *MANEAL*. We find that one of the MO-associated variants, rs12684144-C in *ASTN2*, confers protection against VD (OR = 0.956, $P = 0.00017$) but risk of MO (OR = 1.073, $P = 1.5 \times 10^{-5}$). In line with only 30% of migraineurs experiencing aura³, its association with overall migraine confers risk (OR = 1.055, $P = 1.3 \times 10^{-14}$).

Protein-altering variants in *NGF* and *SCN11A*

Among new variants associated with overall migraine is the common missense variant rs6330-A (p.Ala35Val) in *NGF* (OR = 1.035, $P = 2.1 \times 10^{-8}$). *NGF* encodes nerve growth factor that is involved in regulating growth and differentiation of sympathetic and certain sensory neurons (<https://www.ncbi.nlm.nih.gov/gene>). *NGF* is at 1p13.2 and nearby is *TSPAN2*, harboring a previously reported¹¹ migraine-associated variant (rs2078371) that is, however, uncorrelated ($r^2 = 0.02$) with rs6330. Conditional analysis shows that the effects of rs6330-A on migraine are significant when adjusting for rs2078371 (Table 2).

In *SCN11A*, another common (25%) missense variant, rs33985936-T (p.Val909Ile), associates with overall migraine (OR = 1.041, $P = 3.4 \times 10^{-9}$). *SCN11A* encodes $\text{Na}_v1.9$, which is highly expressed in nociceptive neurons of dorsal root and trigeminal ganglia^{32,33}. Rare loss-of-function (LOF) variants in *SCN11A* can lead to both extremely painful and completely pain-insensitive disorders^{32,33}. We looked for LOF variants in *SCN11A* and found them at very low frequency in all datasets studied, with the highest in the UK at a combined frequency of 0.13%, which is two orders of magnitude higher than in other cohorts. We used a genome-wide burden test combining the effects of these

rare variants on migraine in the UK, and at a threshold of $P = 2.5 \times 10^{-6}$ ($P = 0.05/20,000$ genes³⁴ tested), they associate with strong protection against overall migraine (OR = 0.650, $P = 3.9 \times 10^{-7}$) and other severe headaches and are not driven by a single variant (Table 4 and Supplementary Note 2).

A rare variant targeting *KCNK5* with protective effects

In the GWAS meta-analysis of BRH, there is an association with a large protective effect (OR = 0.697, $P = 7.6 \times 10^{-14}$) with the rare (0.67%) intergenic variant rs72854118-G located in a regulatory region between two potassium channel genes, *KCNK5* and *KCNK17*. The variant also protects against clinically defined migraine (OR = 0.836, $P = 9.7 \times 10^{-7}$), but does not associate with migraine subtypes, MA, MO or VD ($P > 0.05$). Two additional variants in high LD are at this locus, rs72854120 and rs72851880 (Supplementary Fig. 2). A common (28.1%) intronic variant in *KCNK5* was previously reported¹¹ to be associated with migraine (rs10456100, OR = 1.051, $P = 9.2 \times 10^{-19}$), but is uncorrelated with rs72854118 ($r^2 = 0.002$). rs72854118-G is reported in weak association with decreased diastolic blood pressure ($\beta = -0.07$, $P = 2.7 \times 10^{-7}$)³⁵, and in a GWAS meta-analysis of self-reported migraine and headaches combined, one of two correlated SNPs, rs72854120-C, shows borderline association, more so with headaches than migraine ($Z_{\text{migraine}} = -2.68$, $Z_{\text{headache}} = -5.49$, $P = 2.8 \times 10^{-8}$)³⁶. Inspection of effect-effect plots of BRH versus clinically defined migraine for all 44 lead variants shows that rs72854118-G effects on BRH far exceed its migraine effects (Fig. 4 and Supplementary Fig. 5). We performed a phenocan in 1,000 GWAS meta-analyses at deCODE Genetics (P threshold = $0.05/1,000 = 5.0 \times 10^{-5}$) and observed that rs72854118-G also confers substantial protection against brain aneurysms (OR = 0.470, $P = 1.8 \times 10^{-8}$) and coronary artery disease (CAD) requiring bypass surgery (OR = 0.725, $P = 9.3 \times 10^{-8}$), but associates more weakly with CAD in general (OR = 0.900, $P = 1.9 \times 10^{-5}$) and systolic blood pressure (effect = -0.054 s.d., $P = 2.0 \times 10^{-5}$; Supplementary Table 15). Of 17 known brain aneurysm variants³⁷, 3 are in migraine loci (*FHL5*, *SLC24A3* and *PLCE1*). Plotting effects of the brain aneurysm variants (including rs72854118) on brain aneurysms versus effects on migraine and BRH, we find this variant is an outlier in both and confers larger protective effects against brain aneurysms than other brain aneurysm variants (Supplementary Fig. 5).

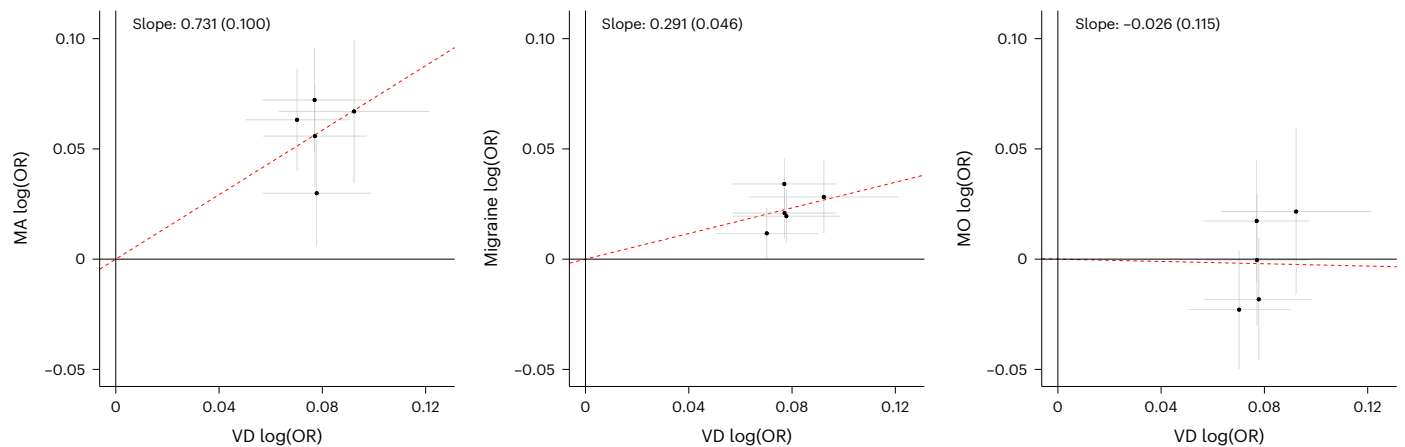


Fig. 2 | Effects of SNPs associated with self-reported headache-related VD in clinically defined MA, overall migraine and MO. The x axis (VD, $n_{\text{case/control}} = 30,297/86,134$) and the y axis (MA, $n_{\text{case/control}} = 16,603/1,336,517$; migraine, $n_{\text{case/control}} = 74,495/1,259,808$ and MO, $n_{\text{case/control}} = 11,718/1,330,747$) show the logarithmic estimated odds ratios, log(OR), for the associations with the respective phenotypes from meta-analyses using a fixed-effects inverse-variance method based on effect estimates and s.e. under the additive model,

in which each dataset was assumed to have a common OR but allowed to have different population frequencies for alleles and genotypes. All effects are shown for the VD risk allele, and black crosses indicate 95% CIs. The dashed red lines represent slope (s.d.) based on a simple linear regression through the origin using $1/s.e.$ as weights. Effect estimates are 73%, 29% and 0% of VD effect estimates for MA, migraine and MO, respectively.

Colocalization highlights new migraine and aura genes

We performed systemic functional annotation of the 44 lead variants and variants in high LD ($r^2 \geq 0.8$) and studied their association with mRNA sequence data (expression quantitative trait loci (eQTL)) and with protein levels in plasma³⁸ (protein quantitative trait loci (pQTL); Methods; Supplementary Tables 16–19). Results are summarized in Supplementary Fig. 6. For the lead variants, we find 144 eQTLs, of which 16 implicate a specific gene (Supplementary Table 17). Variant **rs4768221-G**, in complete LD with **rs10748014-T** (VD association OR = 1.073, $P = 1.2 \times 10^{-12}$) upstream of *LRRK2*, consistently associates with VD and is the top ranking eQTL for this gene in blood. The allele associated with increased risk of VD associates with reduced *LRRK2* expression in blood ($\beta = -0.74$ s.d., $P = 1.3 \times 10^{-1.260}$).

The lead BRH variant near *KCNK5* **rs72854118**, but not the other correlated variants at this locus, is found within a distal enhancer-like sequence (dELS) as defined by ENCODE's catalog of candidate *cis*-regulatory elements³⁹, and the gene target for this regulatory element is *KCNK5* (Supplementary Tables 20 and 21 and Supplementary Note 3). The variant is too rare to be studied in Genotype-Tissue Expression (GTEx, which includes only three carriers; Supplementary Fig. 7), and its expression coverage in tissues available to us is too low for conclusive results.

Three variants (or variants with $r^2 \geq 0.8$) represent top *cis* pQTLs at their respective loci in Icelandic SomaScan plasma protein association data and two variants in the UK Olink data (Supplementary Table 19). These proteomic methods differ in protein profiles, but in both datasets are pQTL variants correlating with the migraine variant **rs1359155039-TAAAAAAAAA** upstream of *LATS1* that associates with reduced migraine risk and increased LRP11 plasma levels ($\beta = 0.58$ s.d., $P = 10^{-1.140}$ and $\beta = 0.59$ s.d., $P = 10^{-2.140}$ in Iceland and UK, respectively). LRP11 is predicted to be located in plasma membrane and involved in several processes, including response to heat and cold (<https://www.ncbi.nlm.nih.gov/gene>).

We do not have RNA expression or protein data for enough carriers of the rare *PRRT2* variants to detect transcription or protein associations. However, on the basis of previous functional studies⁴⁰, the gene's known function as a key component of the Ca^{2+} -dependent neurotransmitter release machinery⁴¹, and its reported links to rare paroxysmal brain disorders including infantile convulsions, the movement disorder PKD and FHM⁴², in addition to the findings in this current study, we

conclude that *PRRT2* is also a risk gene for the common forms of MA and epilepsy. Finally, we scanned the GWAS catalog (<https://www.ebi.ac.uk/gwas/>) for associations with lead variants identified in this study (or $r^2 \geq 0.8$). Results are presented in Supplementary Table 12.

Pathway analysis highlights NGF-related processes

For the 22 genes with evidence supporting their role in migraine or subtypes, we performed a protein network analysis (<https://reactome.org>). Among the top 67 relevant pathways identified, 13 involve NGF processing, including TrkA activation by NGF, previously studied in the context of pain and pain therapeutics⁴³. Interestingly, pathways involved in phase-4 resting potential and cardiac conduction involve the products of both *KCNK5* and *SCN11A*, with the products of both *LRRK2* and *LRP1* interacting in the cardiac conduction pathway (Supplementary Data and Supplementary Table 22).

Genetic drug target analysis

We performed a genetic drug target analysis for the 22 genes for which we have evidence of function pointing to the gene in addition to the established MA gene *CACNA1A*. Drugs at various levels of development target four genes that associate with MA (*PRRT2*, *ABO*, *LRRK2* and *CACNA1A*), none associated with MO, and four genes that associate with overall migraine or severe headaches (*KCNK5*, *NGF*, *SCN11A* and *TRPM8*; Supplementary Table 23 and Supplementary Note 5). Targeting *PRRT2* is bryostatin, a powerful protein kinase C agonist that was originally developed to prevent tumor growth, but in preclinical studies has also shown promising effects as a restorative synapse drug that is currently in trials to treat Alzheimer's disease⁴⁴. Several voltage-gated Ca^{+2} channel blockers have been developed against *CACNA1A*, but have not been tested in migraine. Targeting *TRPM8*, cutaneous menthol treatment has been found to alleviate migraine headaches⁴⁵. Targeting *SCN11A* (and other voltage-gated sodium transporter genes), intranasal lidocaine can be effective in treating acute migraine⁴⁶, and intravenous lidocaine infusion is suggested for treating refractory chronic migraine⁴⁷. Drugs targeting other genes have not been tested for migraine, but β -nerve growth factor inhibitors (antibodies) that target *NGF* (fasinumab, tanezumab and fulranumab) are widely studied in the context of various other chronic pain conditions (for example, sciatica, low back pain and abdominal pain; www.ClinicalTrials.gov).

Table 3 | GWAS meta-analysis results for *PRRT2* frameshift variant (p.Arg217ProfsTer8)

Phenotypes	Iceland (MAF=0.117%)		UK (MAF=0.013%)		US (MAF=0.0051%)		Combined		
	OR (95% CI)	P	OR (95% CI)	P	OR (95% CI)	P	OR (95% CI)	P	P_{het}
Epilepsy	7.482 (5.398, 10.370)	1.3×10^{-33}	4.284 (1.548, 11.859)	0.0051	5.455 (0.407, 73.054)	0.20	7.077 (5.197, 9.635)	1.9×10^{-35}	0.58
MA	5.534 (3.631, 8.434)	1.7×10^{-15}	3.019 (0.283, 32.163)	0.36	5.869 (0.712, 48.348)	0.10	5.446 (3.626, 8.148)	5.6×10^{-16}	0.88
Migraine	3.129 (2.333, 4.196)	2.6×10^{-14}	2.482 (1.202, 5.125)	0.014	3.553 (0.489, 25.791)	0.21	3.038 (2.320, 3.977)	6.6×10^{-16}	0.83
BRH	5.276 (2.104, 13.227)	3.9×10^{-4}	1.981 (0.857, 4.581)	0.11	–	–	3.091 (1.664, 5.742)	3.6×10^{-4}	0.12
VD	8.344 (1.952, 35.662)	4.2×10^{-3}	2.274 (0.764, 6.771)	0.14	–	–	3.634 (1.519, 8.696)	3.7×10^{-3}	0.16
MO	1.025 (0.283, 3.712)	0.97	0.017 (0.000, 4357619.32)	0.68	0.017 (0.000, 15176.02)	0.56	0.972 (0.271, 3.489)	0.97	0.78

The table shows OR with 95% CI and two-sided P values from GWAS results derived from a logistic regression of selected phenotypes in the three cohorts where p.Arg217ProfsTer8 was detected at sufficient frequency. Combined OR and two-sided P are results from inverse-variance-weighted meta-analyses of GWAS results. P values after a variant class-specific Bonferroni adjustment²¹. P_{het} is the heterogeneity P value from a likelihood ratio test. See Supplementary Table 1 for cohort descriptions and Supplementary Table 10b for other neurological associations with both rare *PRRT2* frameshift variants (p.Arg217ProfsTer8 and p.Arg217GlufsTer12).

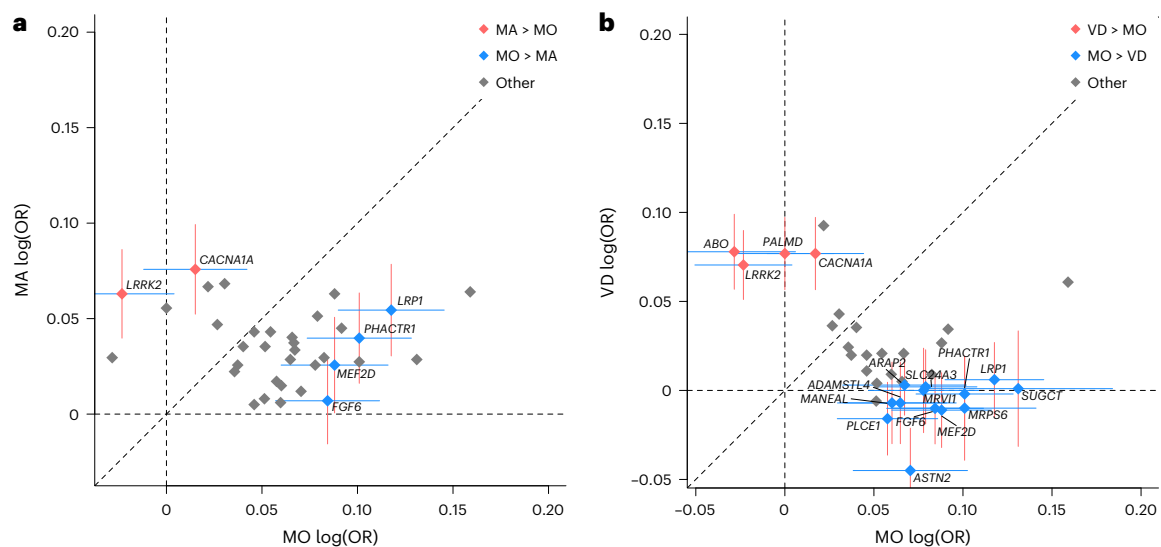


Fig. 3 | Subtype classification of lead variants. Effect plots for all lead variants except the MA variant in *PRRT2*. Effects are from meta-analyses using a fixed-effects inverse-variance method based on effect estimates and s.e. under the additive model, in which each dataset was assumed to have a common OR but allowed to have different population frequencies for alleles and genotypes. Data are presented as additive effect estimates (center) with 95% CI (crosses) for the annotated variants. **a**, Axes show logarithm of odds ratios ($\log(\text{OR})$) for MO (x axis; $n_{\text{case/control}} = 11,718/1,330,747$) and MA (y axis; $n_{\text{case/control}} = 16,603/1,336,517$). **b**, Axes show MO (x axis; $n_{\text{case/control}} = 11,718/1,330,747$) and VD (y axis; $n_{\text{case/control}} = 30,297/86,134$). $\log(\text{OR})$ is calculated for the effect allele. The effects

of variants that have been colored and annotated with gene names differ between the migraine subtypes at a significance threshold of $0.0012 = 0.05/43$. The 95% CIs for the $\log(\text{ORs})$ are shown for annotated variants. Effects are adjusted with sample overlap (r_{ij}) estimated from counts of cases, controls and the counts of overlaps in these groups between phenotypes⁷⁰ from all cohorts except FinnGen (for which we only have summary statistics). The parameter representing sample overlap between MO and MA is $r_{ij} = 0.23$ and MO and VD is $r_{ij} = 0.012$. Dashed lines show the coordinate axes, the diagonal and a line through the origin with slope = 1 (Methods; see Supplementary Tables 13 and 14 and Supplementary Fig. 4 for VD versus MA plot).

Discussion

Whether MA and MO are different diseases or part of a migraine continuum has long been debated^{48,49}. Little is known about the genetics underlying migraine subtypes as most prior studies have focused on migraine in general. Here we have identified several new associations supporting the distinct pathogenesis of MA and MO. In terms of MA, variants in *PRRT2*, *PALMD*, *CACNA1A*, *ABO* and *LRRK2* associate with MA (VD) over MO. Of these, two genes have the highest expression in the cerebellum (*PRRT2* and *CACNA1A*), and in both are rare autosomal dominant variants reported to cause rare forms of movement disorders and hemiplegic migraine (<https://www.omim.org/>). This is of interest in light of the characteristic cortical spreading depression observed in MA but not MO^{4,5}. Both *ABO* and *PALMD* are widely expressed in tissues, and both harbor variants associated with cardiovascular disorders. Indeed, the link between migraine and cardiovascular disease is well established⁵⁰. Drugs targeting these genes are in various phases of

development, but for indications other than migraine. Five drugs target *CACNA1A* for seven indications, including anxiety, insomnia and cardiovascular disease, and targeting *LRRK2* is a trial drug DNL201 (ClinicalTrials.gov identifier: NCT0371070, <https://clinicaltrials.gov/study/NCT03710707>) that shows promising therapeutic potential against PD⁵¹. *LRRK2* is especially abundant in dopamine-innervated areas and dopaminergic neurons of the substantia nigra³⁰. Increased *LRRK2* kinase activity is thought to impair lysosomal function and thus contribute to the pathogenesis of PD⁵². However, consistent with our results showing that the variant in *LRRK2* associates with increased risk of VD (MA) and with reduced *LRRK2* mRNA expression, the main adverse effects of this *LRRK2* inhibitor in healthy individuals were headache (40% of participants) and nausea (13%), the main symptoms of migraine, and dizziness (in 13%)⁵¹. While *LRRK2*'s expression is highest in brain areas associated with PD pathology, it is also expressed in other neurons and glial cells of the human brain³³. Considerable pleiomorphism can occur

Table 4 | Results of *SCN11A* LOF variant burden tests in the respective cohorts for association with migraine

Cohorts	Unique LOF variants combined	Combined frequency (%)	OR	P value	n_{cases}	n_{controls}
UK Biobank	127	0.129	0.650	3.90×10^{-7}	22,082	408,965
US	26	0.0254	0.751	0.63	7,427	50,785
Denmark	26	0.0183	0.629	0.43	14,371	266,473
Iceland	5	8.79×10^{-3}	0.882	0.83	24,604	319,066
–	Combined	0.0454	0.660	2.90×10^{-7}	68,484	1,045,289

The table shows a number of unique LOF variants tested in each cohort. We classified as high-impact variants those predicted as start-lost, stop-gain, stop-lost, splice donor, splice acceptor or frameshift. We used logistic regression under an additive model to test for association between LOF gene burdens and phenotypes using likelihood ratio test to compute two-sided *P* values (Methods; see Supplementary Note 2 for other headache associations in UK Biobank data).

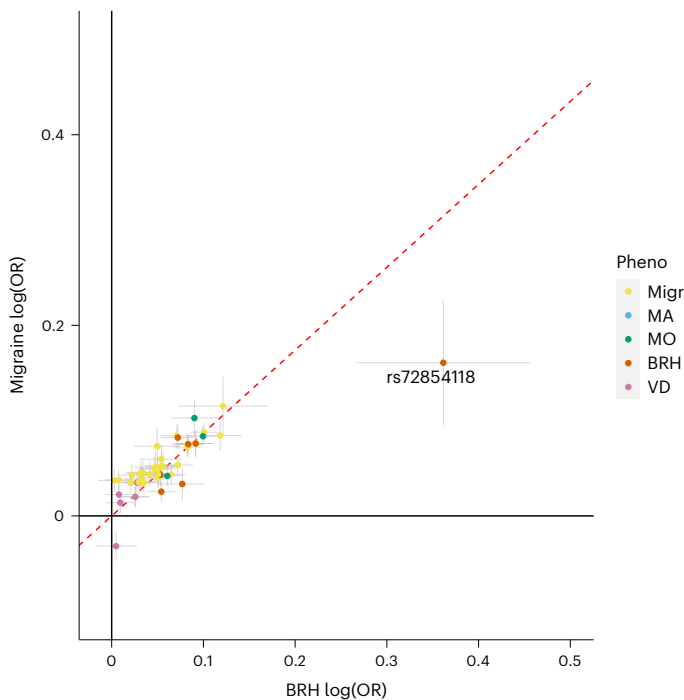


Fig. 4 | Rare variant rs72854118 in regulatory region targeting *KCNK5* associates with BRH. Effect–effect plot of clinically defined migraine ($n_{\text{case/control}} = 74,495/1,259,808$) vs. self-reported BRH ($n_{\text{case/control}} = 51,803/123,732$) effects for 42 lead variants identified in this study (excluding high-impact variants in *PRRT2* and *A3GALT2*; see Supplementary Table 7 for their associations with the respective phenotypes). Effects are from meta-analyses using a fixed-effects inverse-variance method based on effect estimates and s.e. under the additive model, in which each dataset was assumed to have a common OR but allowed to have different population frequencies for alleles and genotypes. The x axis and the y axis show the logarithmic estimated ORs for the associations with the respective phenotypes. Error bars represent 95% CI. The dashed red lines represent slope (s.d.) based on a simple linear regression through the origin using $1/\text{s.e.}$ as weights. Cohort descriptions are in Supplementary Table 1. Variants are colored according to their primary associations in this study. The red dot outlier depicts the variant rs72854118-G near *KCNK5*, its effects on BRH exceeding its effects on all migraine. Pheno, phenotype; Migr, migraine.

among *LRRK2* carriers sharing the same pathogenic variant, even within the same family⁵⁴. Indeed, *LRRK2* has been dubbed the ‘Rosetta stone’ of Parkinsonism, perhaps providing a common link between various neurological diseases⁵⁵.

Our GWAS meta-analysis identified six variants associated with MO, all in previously reported migraine loci. However, by the subtype stratification of all lead variants, we detect 13 variants that impact MO over MA. These MO-associated variants are in or near genes with various functions, such as muscle cell development and differentiation

(*MEF2D*, *FGF6* and *LRPI*) and intracellular calcium homeostasis (*MRVII* and *SLC24A3*). Several are in genes highly expressed in arteries (*MEF2D*, *LRPI*, *ADAMTSL4*, *SUGCT*, *MRVII* and *MRPS6*) and in brain (*MEF2D*, *ARAP2*, *PHACTRI* and *SLC24A3*). Of these, only *LRPI* is currently a drug target (<https://platform.opentargets.org>). *LRPI* encodes low-density lipoprotein receptor-related protein 1, and an *LRPI* binding agent is in trials to treat various brain tumors.

Our results highlight three genes in or near which rare variants show large and informative effects. Firstly, the rare insertion (p.Arg217ProfsTer8) in *PRRT2* that associates with large effects on epilepsy and MA provides new insights into these comorbid⁵⁶ and genetically correlated diseases. *PRRT2* is a four-exon gene that encodes a 340 amino acid protein with two predicted transmembrane domains²⁵. Both the insertion and rarer deletion lead to premature termination of around one-third of *PRRT2*, resulting in nonsense-mediated decay⁴⁰. Due to the founder effect in Iceland, we have power to show the pleiotropic effect of these LOF variants. Not only can they lead to rare neurological disorders, but they also confer substantial risk of common forms of MA and epilepsy, both of which are paroxysmal brain diseases frequently experienced with aura^{57,58}. *PRRT2* is widely expressed in the brain, particularly in the cerebellum^{25,59}. It is enriched in presynaptic terminals, is regulated by Ca^{+2} release and interacts with SNAP-25 and synaptogamin⁴¹. The mutant *PRRT2* of the truncating variants leads to increased glutamate release and subsequent neuronal hyperexcitability⁶⁰. A study of three Na_v 1 subunits (Na_v 1.1 encoded by *SCN1A*, Na_v 1.2 encoded by *SCN2A* and Na_v 1.6 encoded by *SCN8A*) expressed in human embryonic kidney cell lines (HEK-293) demonstrated that *PRRT2* directly interacts with and negatively modulates Na_v 1.2 and Na_v 1.6, which generate action potentials in excitatory neurons, but does not affect Na_v 1.1 channels, which generate action potentials in inhibitory neurons⁶¹. Lack of *PRRT2* leads to hyperactivity of Na_v 1.2 and Na_v 1.6 in homozygous *PRRT2* knockout (human and mouse) neurons⁶¹. The authors of that study suggest that the lack of *PRRT2* effects on Na_v 1.1 may enhance excitation/inhibition imbalance and trigger hyper-synchronized activity in neuronal networks⁶¹. Interestingly, we find that the only epilepsy variant in our data that also associates with migraine is rs59237858 in *SCN1A*, the gene that encodes Na_v 1.1.

Secondly, in the context of Na_v 1 channels, it is of interest that we find both common and rare variants in *SCN11A* that impact migraine risk. *SCN11A* encodes Na_v 1.9 that is expressed in primary sensory neurons in peripheral and trigeminal ganglia⁶² and is known to have a substantial role in pain perception⁶². Compared to other sodium channels, Na_v 1.9 generates a persistent current regulated by G-protein pathways⁶³. Whether Na_v 1.9 is also affected by *PRRT2*, like Na_v 1.2 and Na_v 1.6 (ref. 61), is not known. Currently in various stages of development are 63 drugs targeting *SCN11A* (most unspecific blockers of all Na_v subtypes), with 341 indications, including headache, epilepsy and pain in general (<https://genetics.opentargets.org/gene/ENSG00000168356>). Increasing specificity of Na_v subtype channel blockers and studying their protein interactions seems key to harnessing their therapeutic potential^{64,65}.

Thirdly, the rare intergenic **rs72854118-G** near *KCNK5* and *KCNK17* is another variant providing insight into the pathogenesis of migraine. Previous studies have assigned this variant to *KCNK17* and reported weak associations with reduced blood pressure³⁵ and protection against self-reported headaches and migraine³⁶. However, we find that **rs72854118**, but not its correlated variants at this locus, is in a *cis*-regulatory region targeting *KCNK5*. *KCNK5* encodes TWIK-related acid-sensitive potassium channel 2, primarily expressed in kidney (GTEx, <https://gtexportal.org>) but also in T cells, suggesting a role in the immune system⁶⁶. We find that the variant also confers protection against brain aneurysms and severe occlusive CAD, but associates weakly with blood pressure. Although hypertension is a risk factor for both aneurysms and CAD, it is not a conclusive risk factor for migraine⁶⁷. The observed association with brain aneurysms begs the question whether in some cases undetected brain aneurysms could be misclassified as migraine⁶⁸. According to the Open Targets Platform, no drugs are in development that target *KCNK5*.

In all, our findings are consistent with the results of previous GWAS analyses that have established migraine as a complex neurovascular brain disorder^{13,69}. However, our results also highlight several distinct biological pathways involved in MA and MO that warrant further study. In summary, we contribute new insights into both general and specific mechanisms underlying migraine and its subtypes, especially to the visual aura associated with migraine attacks. Our results also emphasize the importance of assessing disease subtypes and proxies to improve understanding of complex genetic signals.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41588-023-01538-0>.

References

- Lipton, R. B. & Bigal, M. E. The epidemiology of migraine. *Am. J. Med.* **118**, 3S–10S (2005).
- Headache Classification Committee of the International Headache Society (IHS) The International Classification of Headache Disorders, 3rd edition. *Cephalalgia* **38**, 1–211 (2018).
- Rasmussen, B. K. & Olesen, J. Migraine with aura and migraine without aura: an epidemiological study. *Cephalalgia* **12**, 221–228 (1992).
- Lauritzen, M. Pathophysiology of the migraine aura. The spreading depression theory. *Brain* **117**, 199–210 (1994).
- Lai, J. & Dilli, E. Migraine aura: updates in pathophysiology and management. *Curr. Neurol. Neurosci. Rep.* **20**, 17 (2020).
- Olesen, J., Tfelt-Hansen, P., Henriksen, L. & Larsen, B. The common migraine attack may not be initiated by cerebral ischaemia. *Lancet* **2**, 438–440 (1981).
- Sanchez del Rio, M. et al. Perfusion weighted imaging during migraine: spontaneous visual aura and headache. *Cephalalgia* **19**, 701–707 (1999).
- Riant, F. et al. Hemiplegic migraine associated with *PRRT2* variants: a clinical and genetic study. *Neurology* **98**, e51–e61 (2022).
- De Vries, T., Villalon, C. M. & MaassenVanDenBrink, A. Pharmacological treatment of migraine: CGRP and 5-HT beyond the triptans. *Pharmacol. Ther.* **211**, 107528 (2020).
- Dodick, D. W. et al. ARISE: a phase 3 randomized trial of erenumab for episodic migraine. *Cephalalgia* **38**, 1026–1037 (2018).
- Hautakangas, H. et al. Genome-wide analysis of 102,084 migraine cases identifies 123 risk loci and subtype-specific risk alleles. *Nat. Genet.* **54**, 152–160 (2022).
- Sacco, S. et al. Burden and attitude to resistant and refractory migraine: a survey from the European Headache Federation with the endorsement of the European Migraine & Headache Alliance. *J. Headache Pain* **22**, 39 (2021).
- Hautakangas, H. et al. A genome-wide meta-analysis of migraine with over 102,000 cases identifies 124 risk loci and provides first genetic insights to new migraine therapeutics targeting CGRP pathway. In *Proceedings of 2019 Annual Meeting of American Society of Human Genetics* (ASHG, 2019).
- Katsarava, Z., Mania, M., Lampl, C., Herberhold, J. & Steiner, T. J. Poor medical care for people with migraine in Europe—evidence from the EuroLight study. *J. Headache Pain* **19**, 10 (2018).
- Sorensen, E. et al. Data resource profile: the Copenhagen Hospital Biobank (CHB). *Int. J. Epidemiol.* **50**, 719–720 (2021).
- Hansen, T. F. et al. DBDS Genomic Cohort, a prospective and comprehensive resource for integrative and temporal analysis of genetic, environmental and lifestyle factors affecting health of blood donors. *BMJ Open* **9**, e028401 (2019).
- Bycroft, C. et al. The UK Biobank resource with deep phenotyping and genomic data. *Nature* **562**, 203–209 (2018).
- Azriel, E. et al. Utilizing public health frameworks and partnerships to ensure equity in DNA-based population screening. *Front. Genet.* **13**, 886755 (2022).
- Refsum, H. et al. The Hordaland Homocysteine Study: a community-based study of homocysteine, its determinants, and associations with disease. *J. Nutr.* **136**, 1731S–1740S (2006).
- Kurki, M. I. et al. FinnGen provides genetic insights from a well-phenotyped isolated population. *Nature* **613**, 508–518 (2023).
- Sveinbjornsson, G. et al. Weighting sequence variants based on their annotation increases power of whole-genome association studies. *Nat. Genet.* **48**, 314–317 (2016).
- Bulik-Sullivan, B. K. et al. LD score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat. Genet.* **47**, 291–295 (2015).
- Ebrahimi-Fakhari, D., Saffari, A., Westenberger, A. & Klein, C. The evolving spectrum of *PRRT2*-associated paroxysmal diseases. *Brain* **138**, 3476–3495 (2015).
- Park, B. M., Kim, Y. O., Kim, M. K. & Woo, Y. J. A novel frameshift mutation of *PRRT2* in a family with infantile convulsions and choreathetosis syndrome: c.640delinsCC (p.Ala214ProfsTer11). *J. Genet. Med.* **16**, 19–22 (2019).
- Chen, W. J. et al. Exome sequencing identifies truncating mutations in *PRRT2* that cause paroxysmal kinesigenic dyskinesia. *Nat. Genet.* **43**, 1252–1255 (2011).
- Kaushik, J. S., Bala, K. & Dubey, R. Paroxysmal kinesigenic dyskinesia. *Indian Pediatr.* **55**, 74 (2018).
- International League Against Epilepsy Consortium on Complex Epilepsies. GWAS meta-analysis of over 29,000 people with epilepsy reveals 26 risk loci and subtype-specific genetic architecture. *Nat. Genet.* **55**, 1471–1482 (2023).
- Helgadottir, A. et al. Genome-wide analysis yields new loci associating with aortic valve stenosis. *Nat. Commun.* **9**, 987 (2018).
- Wang, M., Gao, J., Liu, J., Zhao, X. & Lei, Y. Genomic association vs. serological determination of ABO blood types in a Chinese cohort, with application in Mendelian randomization. *Genes (Basel)* **12**, 959 (2021).
- Gandhi, P. N., Wang, X., Zhu, X., Chen, S. G. & Wilson-Delfosse, A. L. The Roc domain of leucine-rich repeat kinase 2 is sufficient for interaction with microtubules. *J. Neurosci. Res.* **86**, 1711–1720 (2008).
- Elmazny, A. et al. Interferon- β -induced headache in patients with multiple sclerosis: frequency and characterization. *J. Pain. Res.* **13**, 537–545 (2020).

32. Ginanneschi, F. et al. SCN11A variant as possible pain generator in sensory axonal neuropathy. *Neurol. Sci.* **40**, 1295–1297 (2019).
33. Leipold, E. et al. A de novo gain-of-function mutation in SCN11A causes loss of pain perception. *Nat. Genet.* **45**, 1399–1404 (2013).
34. Lonsdale, J. et al. The Genotype-Tissue Expression (GTEx) project. *Nat. Genet.* **45**, 580–585 (2013).
35. Hoffmann, T. J. et al. Genome-wide association analyses using electronic health records identify new loci influencing blood pressure variation. *Nat. Genet.* **49**, 54–64 (2017).
36. Meng, W. et al. A meta-analysis of the genome-wide association studies on two genetically correlated phenotypes suggests four new risk loci for headaches. *Phenomics* **3**, 64–76 (2022).
37. Bakker, M. K. & Ruigrok, Y. M. Genetics of intracranial aneurysms. *Stroke* **52**, 3004–3012 (2021).
38. Ferkingstad, E. et al. Large-scale integration of the plasma proteome with genetics and disease. *Nat. Genet.* **53**, 1712–1721 (2021).
39. ENCODE Project Consortium et al. Expanded encyclopaedias of DNA elements in the human and mouse genomes. *Nature* **583**, 699–710 (2020).
40. Wu, L. et al. PRRT2 truncated mutations lead to nonsense-mediated mRNA decay in paroxysmal kinesigenic dyskinesia. *Parkinsonism Relat. Disord.* **20**, 1399–1404 (2014).
41. Valente, P. et al. PRRT2 is a key component of the Ca²⁺-dependent neurotransmitter release machinery. *Cell Rep.* **15**, 117–131 (2016).
42. Zhao, S. Y. et al. Functional study and pathogenicity classification of PRRT2 missense variants in PRRT2-related disorders. *CNS Neurosci. Ther.* **26**, 39–46 (2020).
43. Watson, J. J., Allen, S. J. & Dawbarn, D. Targeting nerve growth factor in pain: what is the therapeutic potential? *BioDrugs* **22**, 349–359 (2008).
44. Sun, M.-K. & Alkon, D. L. Bryostatin-1: pharmacology and therapeutic potential as a CNS drug. *CNS Drug Rev.* **12**, 1–8 (2006).
45. Borhani Haghighi, A. et al. Cutaneous application of menthol 10% solution as an abortive treatment of migraine without aura: a randomised, double-blind, placebo-controlled, crossed-over study. *Int J. Clin. Pract.* **64**, 451–456 (2010).
46. Chi, P. W. et al. Intranasal lidocaine for acute migraine: a meta-analysis of randomized controlled trials. *PLoS ONE* **14**, e0224285 (2019).
47. Schwenk, E. S. et al. Lidocaine infusions for refractory chronic migraine: a retrospective analysis. *Reg. Anesth. Pain Med.* **47**, 408–413 (2022).
48. Olesen, J. The international classification of headache disorders. *Headache* **48**, 691–693 (2008).
49. Olesen, J. ICHD-3 β is published. Use it immediately. *Cephalalgia* **33**, 627–628 (2013).
50. Winsvold, B. S. et al. Shared genetic risk between migraine and coronary artery disease: a genome-wide analysis of common variants. *PLoS ONE* **12**, e0185663 (2017).
51. Jennings, D. et al. Preclinical and clinical evaluation of the LRRK2 inhibitor DNL201 for Parkinson's disease. *Sci. Transl. Med.* **14**, eabj2658 (2022).
52. Schapansky, J. et al. Familial knockin mutation of LRRK2 causes lysosomal dysfunction and accumulation of endogenous insoluble α -synuclein in neurons. *Neurobiol. Dis.* **111**, 26–35 (2018).
53. Miklossy, J. et al. LRRK2 expression in normal and pathologic human brain and in human cell lines. *J. Neuropathol. Exp. Neurol.* **65**, 953–963 (2006).
54. Zimprich, A. et al. Mutations in LRRK2 cause autosomal-dominant parkinsonism with pleomorphic pathology. *Neuron* **44**, 601–607 (2004).
55. Dachselt, J. C. & Farrer, M. J. LRRK2 and Parkinson disease. *Arch. Neurol.* **67**, 542–547 (2010).
56. Kanner, A. M. Management of psychiatric and neurological comorbidities in epilepsy. *Nat. Rev. Neurol.* **12**, 106–116 (2016).
57. Noebels, J. L., Avoli, M., Rogawski, M. A., Olsen, R. W. & Delgado-Escueta A. V. (eds.) *Jasper's Basic Mechanisms of the Epilepsies* 4th edn (National Center for Biotechnology Information, 2012).
58. Baldin, E., Ludvigsson, P., Mixa, O. & Hesdorffer, D. C. Prevalence of recurrent symptoms and their association with epilepsy and febrile seizure in school-aged children: a community-based survey in Iceland. *Epilepsy Behav.* **23**, 315–319 (2012).
59. Lee, H. Y. et al. Mutations in the gene PRRT2 cause paroxysmal kinesigenic dyskinesia with infantile convulsions. *Cell Rep.* **1**, 2–12 (2012).
60. Li, M. et al. PRRT2 mutant leads to dysfunction of glutamate signaling. *Int. J. Mol. Sci.* **16**, 9134–9151 (2015).
61. Fruscione, F. et al. PRRT2 controls neuronal excitability by negatively modulating Na⁺ channel 1.2/1.6 activity. *Brain* **141**, 1000–1016 (2018).
62. Baker, M. D. & Nassar, M. A. Painful and painless mutations of SCN9A and SCN11A voltage-gated sodium channels. *Pflugers Arch.* **472**, 865–880 (2020).
63. Cummins, T. R. et al. A novel persistent tetrodotoxin-resistant sodium current in SNS-null and wild-type small primary sensory neurons. *J. Neurosci.* **19**, RC43 (1999).
64. Braden, K., Stratton, H. J., Salvemini, D. & Khanna, R. Small molecule targeting NaV1.7 via inhibition of the CRMP2-Ubc9 interaction reduces and prevents pain chronification in a mouse model of oxaliplatin-induced neuropathic pain. *Neurobiol. Pain* **11**, 100082 (2022).
65. Cai, S. et al. Selective targeting of NaV1.7 via inhibition of the CRMP2-Ubc9 interaction reduces pain in rodents. *Sci. Transl. Med.* **13**, eabh1314 (2021).
66. Bittner, S. et al. Upregulation of K2P5.1 potassium channels in multiple sclerosis. *Ann. Neurol.* **68**, 58–69 (2010).
67. Hagen, K. et al. Blood pressure and risk of headache: a prospective study of 22 685 adults in Norway. *J. Neurol. Neurosurg. Psychiatry* **72**, 463–466 (2002).
68. Lebedeva, E. R., Gurary, N. M., Sakovich, V. P. & Olesen, J. Migraine before rupture of intracranial aneurysms. *J. Headache Pain* **14**, 15 (2013).
69. Gormley, P. et al. Meta-analysis of 375,000 individuals identifies 38 susceptibility loci for migraine. *Nat. Genet.* **48**, 856–866 (2016).
70. Bhattacharjee, S. et al. A subset-based approach improves power and interpretation for the combined analysis of genetic association studies of heterogeneous traits. *Am. J. Hum. Genet.* **90**, 821–835 (2012).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023

Gyda Bjornsdottir^{1,33}✉, Mona A. Chalmer^{2,33}, Lilja Stefansdottir¹, Astros Th. Skuladottir¹, Gudmundur Einarsson¹, Margret Andresdottir¹, Doruk Beyter¹, Egil Ferkingstad¹, Solveig Gretarsdottir¹, Bjarni V. Halldorsson^{1,3}, Gisli H. Halldorsson^{1,4}, Anna Helgadottir¹, Hannes Helgason^{1,4}, Grimur Hjorleifsson Eldjarn¹, Adalbjorg Jonasdottir¹, Aslaug Jonasdottir¹, Ingileif Jonsdottir^{1,5}, Kirk U. Knowlton⁶, Lincoln D. Nadauld⁷, Sigrun H. Lund^{1,8}, Olafur Th. Magnusson¹, Pall Melsted^{1,4}, Kristjan H. S. Moore¹, Asmundur Oddsson¹, Pall I. Olason¹, Asgeir Sigurdsson¹, Olafur A. Stefansson¹, Jona Saemundsdottir¹, Gardar Sveinbjornsson¹, Vinicius Tragante¹, Unnur Unnsteinsdottir¹, G. Bragi Walters¹, Florian Zink¹, Linn Rødevand⁹, Ole A. Andreassen⁹, Jannicke Igland^{10,11}, Rolv T. Lie^{10,12}, Jan Haavik^{13,14}, Karina Banasik¹⁵, Søren Brunak¹⁵, Maria Didriksen¹⁶, Mie T. Bruun¹⁷, Christian Erikstrup^{18,19}, Lisette J. A. Kogelman², Kaspar R. Nielsen^{20,21}, Erik Sørensen¹⁶, Ole B. Pedersen^{22,23}, Henrik Ullum²⁴, DBDS Genetic Consortium*, Gisli Masson¹, Unnur Thorsteinsdottir^{1,5}, Jes Olesen², Petur Ludvigsson²⁵, Olafur Thorarensen²⁵, Anna Bjornsdottir²⁶, Gudrun R. Sigurdardottir²⁷, Olafur A. Sveinsson^{27,28}, Sisse R. Ostrowski^{16,23}, Hilma Holm¹, Daniel F. Gudbjartsson^{1,4}, Gudmar Thorleifsson¹, Patrick Sulem¹, Hreinn Stefansson¹, Thorgeir E. Thorgeirsson¹, Thomas F. Hansen^{2,15,34} & Kari Stefansson^{1,5,34}✉

¹deCODE Genetics/Amgen, Inc., Reykjavik, Iceland. ²Danish Headache Center, Department of Neurology, Copenhagen University Hospital, Rigshospitalet-Glostrup, Copenhagen, Denmark. ³Reykjavik University, School of Technology, Reykjavik, Iceland. ⁴School of Engineering and Natural Sciences, University of Iceland, Reykjavik, Iceland. ⁵Faculty of Medicine, School of Health Sciences, University of Iceland, Reykjavik, Iceland. ⁶Intermountain Heart Institute, Salt Lake City, UT, USA. ⁷Intermountain Healthcare, Saint George, UT, USA. ⁸Faculty of Physical Sciences, School of Engineering and Natural Sciences, University of Iceland, Reykjavik, Iceland. ⁹NORMENT, Centre for Mental Disorders Research, Division of Mental Health and Addiction, Oslo University Hospital, and Institute of Clinical Medicine, University of Oslo, Oslo, Norway. ¹⁰Department of Global Public Health and Primary Care, University of Bergen, Bergen, Norway. ¹¹Department of Health and Social Science, Centre for Evidence-Based Practice, Western Norway University of Applied Science, Bergen, Norway. ¹²Centre for Fertility and Health, Norwegian Institute of Public Health, Oslo, Norway. ¹³Department of Biomedicine, University of Bergen, Bergen, Norway. ¹⁴Division of Psychiatry, Haukeland University Hospital, Bergen, Norway. ¹⁵Novo Nordisk Foundation Center for Protein Research, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark. ¹⁶Department of Clinical Immunology, Copenhagen University Hospital, Rigshospitalet, Copenhagen, Denmark. ¹⁷Department of Clinical Immunology, Odense University Hospital, Odense, Denmark. ¹⁸Department of Clinical Immunology, Aarhus University Hospital, Aarhus, Denmark. ¹⁹Department of Clinical Medicine Health, Aarhus University, Aarhus, Denmark. ²⁰Department of Clinical Immunology, Aalborg University Hospital, Aalborg, Denmark. ²¹Department of Clinical Medicine, Aalborg University, Aalborg, Denmark. ²²Department of Clinical Immunology, Zealand University Hospital, Køge, Denmark. ²³Department of Clinical Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark. ²⁴Statens Serum Institut, Copenhagen, Denmark. ²⁵Department of Pediatrics, Landspítali University Hospital, Reykjavik, Iceland. ²⁶Heilsuklasinn Clinic, Reykjavik, Iceland. ²⁷Laeknasetrid Clinic, Reykjavik, Iceland. ²⁸Department of Neurology, Landspítali University Hospital, Reykjavik, Iceland. ³³These authors contributed equally: Gyda Bjornsdottir, Mona A. Chalmer. ³⁴These authors jointly supervised this work: Thomas F. Hansen, Kari Stefansson. *A list of members and affiliations appears at the end of the paper. ✉e-mail: gyda.bjornsdottir@decode.is; kari.stefansson@decode.is

DBDS Genetic Consortium

Karina Banasik¹⁵, Jakob Bay²², Jens K. Boldsen¹⁸, Thorsten Brodersen²², Søren Brunak¹⁵, Kristoffer Burgdorf^{15,16}, Mona A. Chalmer^{2,33}, Maria Didriksen¹⁶, Khoa M. Dinh¹⁸, Joseph Dowsett¹⁶, Christian Erikstrup^{18,19}, Bjarke Feenstra^{16,24}, Frank Geller^{16,24}, Daniel F. Gudbjartsson^{1,4}, Thomas F. Hansen^{2,15,34}, Lotte Hindhede¹⁸, Henrik Hjalgrim²⁹, Rikke L. Jacobsen¹⁶, Gregor Jemec³⁰, Katrine Kaspersen¹⁸, Bertram D. Kjerulf¹⁸, Lisette J. A. Kogelman², Margit A. H. Larsen¹⁶, Ioannis Louloudis¹⁵, Agnete Lundgaard¹⁵, Susan Mikkelsen¹⁸, Christina Mikkelsen¹⁶, Kaspar R. Nielsen^{20,21}, Ioanna Nissen¹⁶, Mette Nyegaard³¹, Sisse R. Ostrowski^{16,23}, Ole B. Pedersen^{22,23}, Alexander P. Henriksen¹⁵, Palle D. Rohde³¹, Klaus Rostgaard²⁹, Michael Swinn¹⁶, Kari Stefansson^{1,5,34}, Hreinn Stefansson¹, Erik Sørensen¹⁶, Unnur Thorsteinsdottir^{1,5}, Lise W. Thørner¹⁶, Mie T. Bruun¹⁷, Thomas Werge^{23,32} & David Westergaard¹⁵

²⁹Danish Cancer Society Research Center, Copenhagen, Denmark. ³⁰Department of Dermatology, Zealand University Hospital, Roskilde, Denmark.

³¹Department of Health Science and Technology, Faculty of Medicine, Aalborg University, Aalborg, Denmark. ³²Institute of Biological Psychiatry, Mental Health Centre, Sct. Hans, Copenhagen University Hospital, Roskilde, Denmark.

Methods

Ethics statement

All human research was approved by the relevant ethics review boards and conducted according to the Declaration of Helsinki. All participants provided written and informed consent as described per the study population below.

Study populations

Cases and controls were defined from six study populations.

Iceland. About 155,000, or close to half of the Icelandic population of 340,000, have participated in an ongoing nationwide research program at deCODE Genetics^{71,72}. Participants donated blood or buccal samples after signing informed consents allowing the use of their samples and data in various studies approved by the National Bioethics Committee (NBC). The data used here were analyzed under a study on the genetics of migraine (NBC; 19-158-V3, VSNb2019090003/03.01) following review by the Icelandic Data Protection Authority.

Denmark. Danish samples and data were obtained in collaboration with the Copenhagen Hospital Biobank Study¹⁵ and the DBDS¹⁶. CHB is a research biobank, which contains samples obtained during diagnostic procedures on hospitalized and outpatients in the Danish Capital Region hospitals. Data analysis within this study was performed under the 'Genetics of pain and degenerative diseases' protocol, approved by the Danish Data Protection Agency (P-2019-51) and the National Committee on Health Research Ethics (NVK-18038012). The DBDS Genomic Cohort is a nationwide study of ~110,000 blood donors¹⁶. The Danish Data Protection Agency (P-2019-99) and the National Committee on Health Research Ethics (NVK-1700407) approved the studies under which data on DBDS participants were obtained for this study.

UK. Since 2006, the UK Biobank resource has collected extensive phenotype and genotype data from ~500,000 participants recruited in the age range of 40–69 from across the UK after signing an informed consent for the use of their data in genetic studies¹⁷. The North West Research Ethics Committee reviewed and approved the UK Biobank's scientific protocol and operational procedures (REC Reference: 06/MRE08/65). This study was conducted using the UK Biobank Resource (application 42256).

Finland. The FinnGen study²⁰ consists of samples collected from the Finnish biobanks and phenotype data collected at Finland's national health registers. The Coordinating Ethics Committee of the Helsinki and Uusimaa Hospital District evaluated and approved the FinnGen research project. The project complies with existing legislation (in particular the Biobank Law and the Personal Data Act). The official data controller of the study is the University of Helsinki. The summary statistics for FinnGen's migraine GWAS were imported from a source available to consortium partners (Release 6: <https://r6.finnngen.fi/>).

US. Participants from the US were recruited via ongoing studies conducted at Intermountain Healthcare (<https://intermountainhealthcare.org>). These studies include the Intermountain Inspire Registry and the HerediGene: Population study¹⁸. The latter is a large-scale collaboration between Intermountain Healthcare, deCODE Genetics and Amgen. The Intermountain Healthcare Institutional Review Board approved this study, and all participants provided written informed consent and samples for genotyping.

Norway. Data on Norwegian migraine cases and controls were obtained from the HUSK study, a population-based study carried out in Hordaland county in Western Norway¹⁹. In 1992–1993, all Hordaland County residents born between 1950 and 1952, all Bergen residents born

between 1925 and 1927 and three neighboring municipalities and a random sample of individuals born between 1926 and 1949 were invited to participate. In total, 18,044 individuals participated, of which 17,561 provided blood samples for genotyping, of which 10,000 were genotyped at deCODE Genetics. All participants signed informed consents, and the study was approved and carried out by the National Health Screening Service, Oslo (now the Norwegian Institute of Public Health) in cooperation with the University of Bergen¹⁹.

Phenotype definitions

Cases with migraine and the migraine subtypes with and without aura were in all cohorts but Norway (using self-reported migraine from questionnaires), mainly defined by International Classification of Diseases 10th Revision (ICD-10) codes (or comparable codes from earlier versions of ICD) representing MA (code G43.1, MO (G43.0) and overall migraine (G43). Diagnostic codes were assigned by physicians and captured through both inpatient and outpatient diagnostic registries. As triptan medications (Anatomical Therapeutic Chemical code N02CC) are used to prevent/treat migraine attacks, individuals who had received triptan subscriptions were identified in data from drug registries (Iceland, Denmark, Finland and the UK) and added to migraine cases (without subtype).

Both proxy phenotypes used in this study were based on validated questionnaire items selected for the headache section of UK Biobank's pain questionnaire (https://biobank.ctsu.ox.ac.uk/crystal/ukb/docs/pain_questionnaire.pdf), which was designed in consultation with a group of leaders in pain research. The headache section is based on questions used in the American Migraine Prevalence and Prevention study⁷³. For the MA-proxy phenotype used in this study (VD preceding headaches), we defined cases and controls from questionnaire data obtained in the studies conducted in Iceland, Denmark and the UK Biobank. Questions used in Icelandic and Danish cohorts were comparable to the question answered by participants in the UK Biobank (data field 120065: data description: visual changes before or near the onset of headaches, Question: 'I develop visual changes such as spots, lines and heat waves or graying out of my vision'). Responses 'Yes' were compared to responses 'No.' Such defined cases with, and controls without, headache-related VD had all previously responded 'Yes' to a question on headaches as asked in the UK Biobank survey (data field 120053: data description: bad and/or recurring headaches at any time in life, Question: 'Have you ever had bad and/or recurring headaches at any time in your life?'). We used this UK Biobank data field 120053 as a migraine proxy, defining comparable severity qualified headache questions in Icelandic and Danish questionnaire datasets for the GWAS meta-analysis.

Genotyping and whole-genome sequencing

Iceland. At deCODE Genetics, 63,118 Icelandic samples have been whole-genome sequenced (WGS) using GAIIX, HiSeq, HiSeqX and NovaSeq Illumina technology^{71,72} to a mean depth of 38×. Genotypes of single-nucleotide polymorphisms (SNPs) and insertions/deletions (indels) were identified and called jointly by GraphTyper⁷⁴. The effects of sequence variants on protein-coding genes were annotated using the variant effect predictor (VEP) using protein-coding transcripts from RefSeq. Including all sequenced samples, 155,250 samples from Icelandic participants have been genotyped using various Illumina SNP arrays^{71,72}. The chip-typed individuals were long-range phased⁷⁵, and the variants identified in the WGS Icelanders imputed into the chip-typed individuals. Additionally, genotype probabilities for 285,644 ungenotyped close relatives of chip-typed individuals were calculated based on extensive encrypted genealogy data compiled by deCODE Genetics (an unencrypted version is publicly available to all Icelandic citizens at <https://www.islendingabok.is/english>). All variants tested were required to have imputation information over 0.8.

Denmark. Danish samples from both CHB and DBDS were genotyped at deCODE Genetics using Illumina Infinium Global Screening Array. Individual genotype arrays were discarded if the total yield was below 98%. Variants were derived from sequencing 25,215 Scandinavian samples (8,360 Danish) using NovaSeq Illumina technology. Only samples with a genome-wide average coverage of over 20× were used. The genotypes of SNPs and indels were called jointly by GraphTyper⁷⁴. Variants with a missing rate >2% were discarded. The genotyped samples were phased using Eagle (version 2.4.1) and high-quality variants imputed into 270,627 genotyped Danes using haplotype sharing in a Hidden Markov Model based on a Li and Stephens model⁷⁶ similar to the one used in IMPUTE2 (ref. 77).

UK. In the UK Biobank dataset, the first 50,000 participants were genotyped using a custom-made Affymetrix chip, UK BiLEVE Axiom⁷⁸, and the remaining participants using the Affymetrix UK Biobank Axiom array¹⁷. We used existing long-range phasing of the SNP chip-genotyped samples¹⁷. We excluded SNP and indel sequence variants in which at least 50% of samples had no coverage (genotype quality (GQ) score = 0), if the Hardy–Weinberg *P* value was <10⁻³⁰ or if heterozygous excess <0.05 or >1.5. At deCODE Genetics, a collaborative effort was recently performed to whole-genome sequence 150,119 samples from the UK Biobank, allowing us to create a haplotype reference panel, which was then imputed into the UK Biobank chip-genotyped dataset, as previously described elsewhere⁷⁹.

US. Samples from the US (Intermountain dataset) were genotyped using Illumina Global Screening Array chips (*n* = 28,279) and WGS using NovaSeq Illumina technology (*n* = 16,621). Samples were filtered on 98% variant yield and any duplicates were removed. Over 245 million high-quality sequence variants and indels, sequenced to a mean depth of 20×, were identified using GraphTyper⁷⁴. Quality-controlled chip genotype data were phased using SHAPEIT4 (ref. 80). A phased haplotype reference panel was prepared from the sequence variants using the long-range phased chip-genotyped samples using in-house tools and methods described previously^{71,72}.

Norway. Norwegian samples were genotyped on Illumina SNP arrays (OmniExpress or Global Screening Array). The chip-genotyping QC and imputation of the Norwegian dataset were performed at deCODE Genetics in Iceland using the same methods as described above for the Icelandic samples. The imputation for Norwegian samples is based on a haplotype reference panel of 25,215 samples of European ancestry, of which 3,336 are Norwegian.

Finland. A custom-made FinnGen ThermoFisher Axiom array (>650,000 SNPs) was used to genotype FinnGen samples at the Thermo Fisher Scientific genotyping service facility in San Diego. Genotype calls were made with the AxiomGT1 algorithm (<https://finngen.gitbook.io/documentation/methods/genotype-imputation>). The FinnGen Release 6 used in this study contains 260,405 genotyped individuals after quality control (QC). Individuals with ambiguous sex, high genotype missingness (>5%), excess heterozygosity (± 4 s.d.) or non-Finnish ancestry were excluded, as were variants with high missingness (>2%), low Hardy–Weinberg equilibrium ($<1 \times 10^{-6}$) or minor allele count (<3). Imputation was performed using the Finnish population-specific and high coverage (25–30 times) WGS backbone and the population-specific SISu v3 imputation reference panel with Beagle 4.1. More than 16 million variants have been imputed in the Finnish dataset (https://www.finngen.fi/en/access_results).

Genetic ancestry filtering and principal components

For the UK Biobank, we used a British–Irish ancestry subset defined previously⁷⁹. Procedures to account for ancestry in FinnGen²⁰ and Iceland⁷² have also been previously described. Genetic ancestry analysis to

identify subsets of individuals with similar ancestry was performed for the Danish, Intermountain and Norwegian datasets separately. ADMIXTURE (v1.23)⁸¹ was run in supervised mode using the 1000 Genomes populations⁸² CEU (Utah residents with Northern and Western European ancestry), CHB (Han Chinese in Beijing, China), ITU (Indian Telugu in the UK), PEL (Peruvian in Lima, Peru) and YRI (Yoruba in Ibadan, Nigeria) as training samples. These training samples had themselves been filtered for ancestry outliers using principal component analysis (PCA) and unsupervised ADMIXTURE.

For the Danish and Intermountain datasets, samples assigned <0.93 CEU were excluded. We performed a different filtering procedure for the Norwegian dataset to include individuals with Finnish and Saami ancestry, who are common in Norway⁸³. To identify such individuals, we first selected candidates those assigned between 0.5 and 0.93 CEU ancestry. We then merged these individuals with the Human Origins dataset and calculated *F* statistics⁸⁴ of the form f_3 (Mbuti; candidate individual, X), where X was each of the Human Origins populations Nganasan, Pima, Han and Norwegian. In these F_3 statistics, we identified a clear cluster of individuals with excess affinity to Nganasan and Norwegian over Pima and Han. In available metadata, we observed that these individuals were highly enriched for locations of residence in Finnmark and officially designated Saami villages. These genetic and demographic features match expectations for individuals of Saami or Finnish ancestry. Except for this cluster, we excluded all other Norwegian individuals assigned <0.93 CEU ancestry. Genetic principal components for use as covariates in association analysis were obtained using bigsnpr⁸⁵.

Association testing and meta-analysis

Using software developed at deCODE Genetics⁷², we applied logistic regression assuming an additive model to test for genome-wide associations between sequence variants and migraine phenotypes. Association results from FinnGen were imported (Release 6: <http://r6.finngen.fi>). For the Icelandic data, the model included sex, county of birth, current age or age at death (first-order and second-order terms included), blood sample availability for the individual and an indicator function for the overlap of the lifetime of the individual with the time span of phenotype collection. To include imputed but ungenotyped individuals, we used county of birth as a proxy covariate for the first PCs in our analysis because county of birth has been shown to be in concordance with the first PC in Iceland⁸⁶. For the Danish, Norwegian, UK and US data, the covariates were sex, age, expected allele count and 20 PCs to adjust for population stratification. The association analysis of the imported Finnish data was adjusted for sex, age, the genotyping batch and the first ten PCs. We used LD score regression intercepts²² to adjust the χ^2 statistics and avoid inflation due to cryptic relatedness and stratification, using a set of 1.1 million variants. *P* values were calculated from the adjusted χ^2 results. All statistical tests were two-sided unless otherwise indicated.

For the meta-analyses, we combined GWASs from the respective cohorts with summary statistics from Finland using a fixed-effects inverse-variance method based on effect estimates and s.e. in which each dataset was assumed to have a common OR but allowed to have different population frequencies for alleles and genotypes. The total number of variants included in the meta-analyses was between 68 and 80 million variants. Sequence variants were mapped to the NCBI Build 38 and matched on position and alleles to harmonize the datasets. The threshold for genome-wide significance was corrected for multiple testing with a weighted Bonferroni adjustment that controls for the family-wise error rate, using as weights the enrichment of variant classes with predicted functional impact among association signals²¹. The significance threshold then becomes 2.5×10^{-7} for high-impact variants (including stop-gained, frameshift, splice acceptor or donor), 5.0×10^{-8} for moderate-impact variants (including missense, splice-region variants and in-frame indels), 4.5×10^{-9}

for low-impact variants, 2.3×10^{-9} for other DNase I hypersensitivity sites (DHS) variants and 7.5×10^{-10} for other non-DHS variants²¹. In a random-effects method, a likelihood ratio test was performed in all genome-wide associations to test the heterogeneity of the effect estimate in the four datasets; the null hypothesis is that the effects are the same in all datasets, and the alternative hypothesis is that the effects differ between datasets.

The primary signal at each genomic locus was defined as the sequence variant with the lowest Bonferroni-adjusted *P* value using the adjusted significance thresholds described above. Conditional analysis was used to identify possible secondary signals within 500 kb from the primary signal. This was done using genotype data for the Icelandic, Norwegian, Danish, UK and US datasets and an approximate conditional analysis implemented in GCTA software⁸⁷ for the Finnish summary data. Adjusted *P* values and ORs were combined using a fixed-effects inverse-variance method. Class-specific genome-wide significance thresholds were also used for the secondary signals. Manhattan plots were generated using *topr* package in R.

For burden testing, we used the UK Biobank whole-exome sequenced dataset, consisting of 400,912 whole-exome sequenced White British (individuals identified by PCA analyses)^{88,89} who enrolled in the study between 2006 and 2010 throughout the UK and were aged 38–65 years at recruitment. A wide range of phenotypic data has been provided by the UK Biobank primarily from hospital records and increasingly from general practitioners from the UK. For the Icelandic, US and Danish cohorts, we used the phenotypes and WGS and imputation data previously described.

We used VEP⁹⁰ to attribute predicted consequences to the variants sequenced in each dataset. We classified as high-impact variants those predicted as start-lost, stop-gain, stop-lost, splice donor, splice acceptor or frameshift, collectively called LOF variants. For case-control analyses, we used logistic regression under an additive model to test for association between LOF gene burdens and phenotypes, in which disease status was the dependent variable and genotype counts as the independent variable, using likelihood ratio test to compute two-sided *P* values. Individuals were coded 1 if they carried any of the LOF variants in the autosomal gene being tested and 0 otherwise. For the UK Biobank association testing, 20 PCs were used to adjust for population substructure, and age and sex were included as covariates in the logistic regression model. We further included variables indicating sequencing batches to remove batch effects. For these analyses, we used software developed at deCODE Genetics⁷².

Genetic correlations

Using cross-trait LD score regression²², we estimated the genetic correlation between each of the migraine and proxy (BRH) and migraine subtype phenotypes (MO, MA and VD) defined in this study, in addition to epilepsy. In this analysis, we used results for about 1.2 million well-imputed variants, and for LD information, we used precomputed LD scores for European populations (downloaded from https://data.broadinstitute.org/alkesgroup/LDSCORE/eur_w_ld_chr.tar.bz2). To avoid bias due to sample overlap, we used the Icelandic and Danish cohorts combined to test for correlation with the respective phenotypes in the other remaining datasets combined. Finally, we meta-analyzed the results of the two correlation analyses for each correlation for a combined correlation estimation. The significance level for the correlation estimates was determined using a simple Bonferroni correction for the number of meta-analyzed correlations, and hence significance was set at $P < 0.0033$ (0.05/15).

Identification and confirmation of rare *PRRT2* variants

The variants in the *PRRT2* gene are in a stretch of nine C's, with one extra C in carriers of the insertion (p.Arg217ProfsTer8) and one missing C in carriers of the deletion (p.Arg217GlufsTer12). This imposes a technical challenge for accurate whole-genome sequence calling. Therefore, all

potential carriers of both variants were analyzed with Sanger sequencing. Primers were designed using Primer 3 software. Following PCR, cycle sequencing reactions were performed in both directions on MJ Research PTC-225 thermal cyclers, using the BigDye Terminator Cycle Sequencing Kit v3.1 (Life Technologies) and Ampure XP and CleanSeq kits (Agencourt) for cleanup of the PCR products and cycle sequencing reactions. Sequencing products were loaded onto the 3730 XL DNA Analyzer (Applied Biosystems) and analyzed with Sequencher 5.0 software (Gene Codes Corporation). Based on the sequencing results, the variants were then re-imputed into the respective cohorts.

Migraine subtype analysis of lead variants

To classify our lead variants by migraine subtype, we plotted their effects on MA versus MO and VD versus MO using the method applied in ref. 11. This method requires a correlation parameter between MO and MA (MO and VD) to account for sample overlap, and previously this parameter was estimated from GWAS summary statistics¹¹, using empirical Pearson correlation of effect size estimates of common variants (MAF > 0.05), which do not show a strong association with either of the migraine subtypes studied ($P > 1 \times 10^{-4}$)⁹¹. In our data, this estimate of the correlation parameter was $r_{ij} = 0.59$ between MO and MA and $r_{ij} = 0.198$ between MO and VD (estimated using 7,858,264 markers), which is considerably larger than if we estimated the sample overlap directly using counts of cases, controls and the counts of overlaps in these groups between phenotypes⁷⁰ (from all cohorts except the summary statistics from FinnGen), where we get $r_{ij} = 0.023$ for MO and MA and $r_{ij} = 0.012$ for MO and VD. As the latter estimates are more conservative, we used those in the subtype analysis. Finally, we tested whether the effect sizes between MA and MO (and VD and MO) were equal at a Bonferroni corrected significance threshold of $P = 0.05/43$ (as we excluded from the 44 lead variants the MA variant in *PRRT2*) performed by using normal approximation and accounting for the correlation in effect size difference estimators. As pointed out in ref. 11, this subtype classification method takes into account the different statistical power of the migraine subtype GWASs, which is an advantage compared to simply comparing subtype effects. For the subtype analysis, we followed the R code available at <https://github.com/mjpirinen/migraine-meta>.

Functional data and colocalization analysis

To highlight genes whose products potentially mediate the observed associations with migraine and migraine subtypes, we annotated the associations detected in this study (Tables 1 and 2) as well as variants in high LD ($r^2 \geq 0.8$ and within ± 1 Mb) that are predicted to affect coding or splicing of a protein (VEP using RefSeq gene set), mRNA expression (top local eQTL, *cis*-eQTL) in multiple tissues from deCODE, GTEx (<https://www.gtexportal.org>) and other public datasets (see Supplementary Table 18 for eQTL data sources) and/or plasma protein levels (top pQTL) identified in large proteomic datasets from Iceland and the UK. The Icelandic proteomics data were analyzed using the SomaLogic SOMAscan proteomics assay that scans 4,907 aptamers, measuring 4,719 proteins in samples from 35,559 Icelanders with the genetic information available at deCODE Genetics³⁸. Plasma protein levels were standardized and adjusted for year of birth, sex and year of sample collection (2000–2019)³⁸. The UK proteomics dataset was analyzed using the Olink proteomics assay characterizing 1,463 proteins in 54,306 participants in the UK Biobank⁹².

RNA sequencing was performed on whole blood from 17,848 Icelanders and on subcutaneous adipose tissue from 769 Icelanders, respectively³⁸. Gene expression was computed based on personalized transcript abundances using kallisto⁹³. Association between sequence variants and gene expression (*cis*-eQTL) was tested using a generalized linear regression, assuming additive genetic effect and normal quantile gene expression estimates, adjusting for measurements of sequencing artifacts, demographic variables, blood composition

and PCs⁹⁴. The gene expression PCs were computed per chromosome using a leave-one-chromosome-out method. All variants within 1 Mb of each gene were tested.

We performed gene-based enrichment analysis using the GEN-E2FUNC tool in FUMA⁹⁵. The genes were tested for over-representation in different gene sets, including Gene Ontology cellular components (MsigDB c5) and GWAS Catalog-reported genes.

Genetic drug target analysis

Using sources from the Drug-Gene Interaction Database⁹⁶, Open Targets⁹⁷ and the National Institutes of Health's Illuminating the Druggable Genome⁹⁸, we performed a genetic drug target analysis for the 22 genes for which we have evidence of function pointing to the gene (Supplementary Fig. 6), in addition to the established MA gene *CACNA1A*.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Our previously described Icelandic population whole-genome sequence data have been deposited at the European Variant Archive under accession PRJEB15197. The GWAS summary statistics for the migraine GWAS meta-analyses are available at <https://www.decode.com/summarydata/>. FinnGen data are publicly available and were downloaded from https://www.finnngen.fi/en/access_results. The UKB data were downloaded under application 42256. Proteomics data and protein mapping to UniProt identifiers and gene names were provided by SomaLogic and Olink. Other data generated or analyzed in this study are included in the article and its Supplementary Information. URLs for other external data used are as follows: precomputed LD scores for European populations, https://data.broadinstitute.org/alkesgroup/LDSCORE/eur_w_ld_chr.tar.bz2; GWAS Catalog, <https://www.ebi.ac.uk/gwas/>; GTEx project, <https://gtexportal.org/home/>. URL sources for expression data can be found in Supplementary Table 18.

Code availability

We used publicly available software that is available on request under the following URLs: GraphTyper (v2.0-beta, GNU GPLv3 license), <https://github.com/DecodeGenetics/graph typer>; Eagle (version 2.4.1), <http://www.hsph.harvard.edu/alkes-price/software/>; SHAPEIT4, <https://odelaneau.github.io/shapeit4/>; ADMIXTURE (v1.23), <https://dalexander.github.io/admixture/>; BOLT-LMM (v.2.1), <http://www.hsph.harvard.edu/alkes-price/software/>; R (version 3.6.3), <https://www.r-project.org/>; R package ggplot for visualization (version 3.3.3), <https://ggplot2.tidyverse.org/>; Ensembl v.87, <https://www.ensembl.org/index.html>; IMPUTE2 v.2.3.1, https://mathgen.stats.ox.ac.uk/impute/impute_v2.html; dbSNP v.140, <http://www.ncbi.nlm.nih.gov/SNP/>; kallisto v.0.46, <https://github.com/pachterlab/kallisto>; for subtype stratification analysis, we used R code available at <https://github.com/mjpirinen/migraine-meta>; MAGMA (v1.08), <http://ctglab.nl/software/magma>; VEP (release 100), <https://github.com/Ensembl/ensembl-vep>; FUMA, <https://fuma.ctglab.nl/>; Sequencher 5.0, <https://sequencher.software.informer.com/5.0/>; NCBI Build 38, <https://www.ncbi.nlm.nih.gov/>. No custom code was written for this study.

References

- Jonsson, H. et al. Whole genome characterization of sequence diversity of 15,220 Icelanders. *Sci. Data* **4**, 170115 (2017).
- Gudbjartsson, D. F. et al. Large-scale whole-genome sequencing of the Icelandic population. *Nat. Genet.* **47**, 435–444 (2015).
- Lipton, R. B. et al. Migraine prevalence, disease burden, and the need for preventive therapy. *Neurology* **68**, 343–349 (2007).
- Eggertsson, H. P. et al. GraphTyper enables population-scale genotyping using pangenome graphs. *Nat. Genet.* **49**, 1654–1660 (2017).
- Kong, A. et al. Detection of sharing by descent, long-range phasing and haplotype imputation. *Nat. Genet.* **40**, 1068–1075 (2008).
- Li, N. & Stephens, M. Modeling linkage disequilibrium and identifying recombination hotspots using single-nucleotide polymorphism data. *Genetics* **165**, 2213–2233 (2003).
- Howie, B. N., Donnelly, P. & Marchini, J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet.* **5**, e1000529 (2009).
- Wain, L. V. et al. Novel insights into the genetics of smoking behaviour, lung function, and chronic obstructive pulmonary disease (UK BiLEVE): a genetic association study in UK Biobank. *Lancet Respir. Med.* **3**, 769–781 (2015).
- Halldórsson, B. V. et al. The sequences of 150,119 genomes in the UK Biobank. *Nature* **607**, 732–740 (2022).
- Delaneau, O., Zagury, J. F., Robinson, M. R., Marchini, J. L. & Dermitzakis, E. T. Accurate, scalable and integrative haplotype estimation. *Nat. Commun.* **10**, 5436 (2019).
- Alexander, D. H., Novembre, J. & Lange, K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* **19**, 1655–1664 (2009).
- Auton, A. et al. A global reference for human genetic variation. *Nature* **526**, 68–74 (2015).
- Mattingsdal, M. et al. The genetic structure of Norway. *Eur. J. Hum. Genet.* **29**, 1710–1718 (2021).
- Patterson, N. et al. Ancient admixture in human history. *Genetics* **192**, 1065–1093 (2012).
- Privé, F., Aschard, H., Ziyatdinov, A. & Blum, M. G. B. Efficient analysis of large-scale genome-wide data with two R packages: bigstatsr and bigsnpr. *Bioinformatics* **34**, 2781–2787 (2018).
- Price, A. L. et al. The impact of divergence time on the nature of population structure: an example from Iceland. *PLoS Genet.* **5**, e1000505 (2009).
- Yang, J. et al. Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. *Nat. Genet.* **44**, S1–S3 (2012).
- Backman, J. D. et al. Exome sequencing and analysis of 454,787 UK Biobank participants. *Nature* **599**, 628–634 (2021).
- Bycroft, C. et al. The UK Biobank resource with deep phenotyping and genomic data. *Nature* **562**, 203–209 (2018).
- McLaren, W. et al. The Ensembl variant effect predictor. *Genome Biol.* **17**, 122 (2016).
- Cichonska, A. et al. metaCCA: summary statistics-based multivariate meta-analysis of genome-wide association studies using canonical correlation analysis. *Bioinformatics* **32**, 1981–1989 (2016).
- Sun, B. B. et al. Plasma proteomic associations with genetics and health in the UK Biobank. *Nature* **622**, 329–338 (2023).
- Bray, N. L., Pimentel, H., Melsted, P. & Pachter, L. Near-optimal probabilistic RNA-seq quantification. *Nat. Biotechnol.* **34**, 525–527 (2016).
- Stegle, O., Parts, L., Piipari, M., Winn, J. & Durbin, R. Using probabilistic estimation of expression residuals (PEER) to obtain increased power and interpretability of gene expression analyses. *Nat. Protoc.* **7**, 500–507 (2012).
- Watanabe, K., Taskesen, E., van Bochoven, A. & Posthuma, D. Functional mapping and annotation of genetic associations with FUMA. *Nat. Commun.* **8**, 1826 (2017).
- Freshour, S. L. et al. Integration of the drug-gene interaction database (DGIdb 4.0) with open crowdsourcing efforts. *Nucleic Acids Res.* **49**, D1144–D1151 (2021).

97. Ochoa, D. et al. The next-generation Open Targets Platform: reimaged, redesigned, rebuilt. *Nucleic Acids Res.* **51**, D1353–D1359 (2022).
98. Nguyen, D.-T. et al. Pharos: collating protein information to shed light on the druggable genome. *Nucleic Acids Res.* **45**, D995–D1002 (2016).

Acknowledgements

We thank all participants who contributed data and samples used in this study. Their contributions are essential for research such as reported here. We thank all investigators and colleagues who collaborated on the many aspects of this study, including data collection, sample handling, phenotypic characterization of clinical samples, genotyping and analysis of the whole-genome association data. We acknowledge participants and investigators of the FinnGen study²⁰ and the UK Biobank study. This research has been conducted using the UK Biobank Resource, a major biomedical database (application 42256, <https://www.ukbiobank.ac.uk/>). The financial support from the European Commission to the painFACT project to T.E.T. (H2020-2020-848099) is acknowledged, as is support from the Novo Nordisk Foundation, DBDS Consortium (grants NNF17OC0027594 and NNF14CC0001). The Genotype-Tissue Expression (GTEx) Project was supported by the Common Fund of the Office of the Director of the National Institutes of Health (commonfund.nih.gov/GTEx). Additional funds were provided by the NCI, NHGRI, NHLBI, NIDA, NIMH and NINDS. Donors were enrolled at Biospecimen Source Sites funded by NCI\Leidos Biomedical Research, Inc. subcontracts to the National Disease Research Interchange (10XS170), GTEx Project March 5, 2014 version Page 5 of 8 Roswell Park Cancer Institute (10XS171), and Science Care, Inc. (X10S172). The Laboratory, Data Analysis, and Coordinating Center (LDACC) was funded through a contract (HHSN268201000029C) to the The Broad Institute, Inc. Biorepository operations were funded through a Leidos Biomedical Research, Inc. subcontract to Van Andel Research Institute (10ST1035). Additional data repository and project management were provided by Leidos Biomedical Research, Inc. (HHSN261200800001E). The Brain Bank was supported supplements to University of Miami grant DA006227. Statistical Methods development grants were made to the University of Geneva (MH090941 and MH101814), the University of Chicago (MH090951, MH090937, MH101825 and MH101820), the University of North Carolina - Chapel Hill (MH090936), North Carolina State University (MH101819), Harvard University (MH090948), Stanford University (MH101782), Washington University (MH101810) and to the University of Pennsylvania (MH101822). The datasets used for the analyses described in this manuscript were obtained from dbGaP at

<http://www.ncbi.nlm.nih.gov/gap> through dbGaP accession number phs000424.v9.p2.

Author contributions

O.B.P. (olbp@regionsjaelland.dk) is the representative for the DBDS Genetic Consortium. G.B., M.A.C., L.S., A.Th.S., G.E., E.F., S.G., B.V.H., A.H., Adalbjorg Jonasdottir, Aslaug Jonasdottir, I.J., G.M., K.H.S.M., O.Th.M., P.I.O., A.S., O.A. Stefansson, G.S., V.T., U.U., G.B.W., F.Z., U.T., S.R.O., H. Holm, D.F.G., G.T., P.S., H.S., T.E.T., T.F.H. and K.S. designed the study, analyzed data and interpreted results. G.B., M.A., A.H., I.J., A.O., J.S., U.U., G.B.W., U.T., H. Holm, D.F.G., P.S., H.S., T.E.T. and K.S. collected and analyzed Icelandic phenotypes and samples for the study. G.B., A.Th.S., D.B., E.F., G.H.H., H. Helgason, S.H.L., P.M., A.S., O.A. Stefansson, H. Holm, G.H.E., D.F.G., G.T., P.S., H.S., T.E.T., T.F.H. and K.S. performed and/or interpreted results from functional studies, transcriptomics, proteomics and gene set enrichment. O.A.A., J.H., J.I., R.T.L. and L.R. designed, collected, contributed and interpreted Norwegian study data. The DBDS Genetic Consortium, M.A.C., K.B., S.B., M.D., M.T.B., C.E., L.J.A.K., K.R.N., E.S., O.B.P., H.U., J.O., S.R.O. and T.F.H. designed, collected, contributed and interpreted Danish study data. L.D.N. and K.U.K. designed, collected, contributed and interpreted the US study data. G.B., M.A.C., L.S., A.Th.S., E.F., S.G., A.H., Adalbjorg Jonasdottir, Aslaug Jonasdottir, A.S., A.B., A.O., G.R.S., P.L., O.T., O.A. Sveinsson, H. Holm, G.T., P.S., H.S., T.E.T., T.F.H. and K.S. drafted the manuscript with input and comments from other authors who all reviewed and contributed to the final version of the manuscript.

Competing interests

Authors affiliated with deCODE Genetics/Amgen declare competing financial interests as employees. The remaining authors declare no competing financial interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41588-023-01538-0>.

Correspondence and requests for materials should be addressed to Gyda Bjornsdottir or Kari Stefansson.

Peer review information *Nature Genetics* thanks Guy Rouleau and Ynte Ruigrok for their contribution to the peer review of this work.

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

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a | Confirmed

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
Give P values as exact values whenever suitable.
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection

Data analysis

In conjunction with methods developed at deCODE Genetics as described in the methods section, we used publicly available software that is available on request under the following URLs:
 GraphTyper (v2.0-beta, GNU GPLv3 license): <https://github.com/DecodeGenetics/graphtyper>,
 Eagle (version 2.4.1): <http://www.hsph.harvard.edu/alkes-price/software/>,
 Shapeit 4: <https://odelaneau.github.io/shapeit4/>
 ADMIXTURE (v1.23): <https://dalexander.github.io/admixture/>
 BOLT-LMM (v.2.1): <http://www.hsph.harvard.edu/alkes-price/software/>,
 R (version 3.6.3): <https://www.r-project.org/>,
 R package ggplot for visualization (version 3.3.3): <https://ggplot2.tidyverse.org/>,
 Ensembl v.87: <https://www.ensembl.org/index.html>,
 IMPUTE2 v.2.3.1: https://mathgen.stats.ox.ac.uk/impute/impute_v2.html,
 dbSNP v.140: <http://www.ncbi.nlm.nih.gov/SNP/>,
 kallisto v.0.46: <https://github.com/pachterlab/kallisto>,
 For subtype stratification analysis we used R code available at: <https://github.com/mjpirinen/migraine-meta>,
 MAGMA (v1.08): <http://ctglab.nl/software/magma>,
 Variant Effect Predictor (release 100): <https://github.com/Ensembl/ensembl-vep> and FUMA:at <https://fuma.ctglab.nl/>.
 Sequencher 5.0: <https://sequencher.software.informer.com/5.0/>
 No custom code was written for this study.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

Our previously described Icelandic population WGS data have been deposited at the European Variant Archive under accession PRJEB15197 (<https://www.ebi.ac.uk/ena/browser/view/PRJEB15197>). The GWAS summary statistics for the migraine GWAS meta-analyses are available at <https://www.decode.com/summarydata/>. FinnGen data are publicly available and were downloaded from https://www.finnngen.fi/en/access_results. The UKB data were downloaded under application no. 42256. Proteomics data and protein mapping to UniProt identifiers and gene names were provided by SomaLogic and Olink. Other data generated or analyzed in this study are included in the article and its Supplementary tables and information.

Human research participants

Policy information about [studies involving human research participants and Sex and Gender in Research](#).

Reporting on sex and gender

No sex-specific results are reported.

Population characteristics

Our study is based on data from study participants of European descent from 6 populations (Iceland, Denmark, Norway, UK, USA, Finland). A description of all population characteristics is included in the methods section. Genetic ancestry filtering and principal components determining European ancestry in each population are also described in methods. Cases with migraine and the migraine subtypes with and without aura, were in all cohorts but Norway (using self-reported migraine from questionnaires), mainly defined by International Classification of Diseases (ICD-10) codes (or comparable codes from earlier versions of ICD) representing migraine with aura (MA, code G43.1, migraine without aura (MO, G43.0) and overall migraine (G43). Diagnostic codes were assigned by physicians and captured through both inpatient and outpatient diagnostic registries. As triptan medications (Anatomical Therapeutic Chemical (ATC) code N02CC) are used to prevent/treat migraine attacks, subjects who had received triptan subscriptions were identified in data from drug registries (Iceland, Denmark, Finland and UK) and added to migraine cases (without subtype). For the MA-proxy phenotype used in this study (Visual disturbances preceding headaches, VD), we defined cases and controls from questionnaire data obtained in the studies conducted in Iceland, Denmark and the UK Biobank. Questions used in Icelandic and Danish cohorts were comparable to the question answered by participants in the UK Biobank (Data-Field 120065: Data description: Visual changes before or near the onset of headaches, Question: "I develop visual changes such as spots, lines and heat waves or greying out of my vision". Responses "Yes" were compared to responses "No". Such defined cases with, and controls without, headache-related visual disturbances had all previously responded "Yes" to a question on headaches as asked in the UK Biobank survey (Data-Field 120053: Data description: Bad and/or recurring headaches at any time in life, Question: "Have you ever had bad and/or recurring headaches at any time in your life?"). We used this UK Biobank data field 120053 as a migraine-proxy, defining comparable severity qualified headache questions in Icelandic and Danish questionnaire datasets for the GWAS meta-analysis.

Recruitment

The GWAS meta-analyses reported in the study were performed using data from 6 populations (Iceland, Denmark, Norway, UK, USA, Finland) and results across populations were compared. Recruitment and phenotype assessment information for each population is detailed in the methods section. In short:

Icelandic participants were recruited into the ongoing deCODE Genetics migraine studies and various other deCODE research projects. About 155K, or close to half of the Icelandic population of 340K has participated in various ongoing nationwide research programs at deCODE Genetics.

Danish samples and data were obtained in collaboration with the ongoing recruitment performed by the Copenhagen Hospital Biobank Study and the Danish Blood Donor Study (DBDS).
 Data on Norwegian migraine cases and controls were obtained from the Hordaland Health study (HUSK), a population-based study recruiting participants in Hordaland county in Western Norway.
 UK participants were recruited into the UK Biobank study that since 2006 has collected extensive phenotype and genotype data from about 500,000 participants, recruited widely from across the UK.
 Participants from the US were recruited via ongoing studies at Intermountain Healthcare in Utah.
 Data from Finland was obtained from the ongoing FinnGen study, that consists of samples collected from the Finnish biobanks and phenotype data collected at Finland's national health registers.

Ethics oversight

All data and samples on which this study is based, were collected under licenses obtained from the respective studies' local ethics and data privacy protection committees and under informed consent of participants, as described in detail in the methods section of the manuscript. In short:
 Icelandic data were analyzed under a study on the genetics of migraine approved by the National Bioethics Committee (NBC#; 19-158-V3, VSNb2019090003/03.01) following review by the Icelandic Data Protection Authority (DPA).
 Danish data were analyzed under the "Genetics of pain and degenerative diseases" protocol, approved by the Danish Data Protection Agency (P-2019-51) and the National Committee on Health Research Ethics (NVK-18038012). The Danish Data Protection Agency (P-2019-99) and the National Committee on Health Research Ethics (NVK-1700407) approved the studies under which data on DBDS participants were obtained for this study.
 Data from the UK Biobank were analyzed under approval from The North West Research Ethics Committee, which reviewed and approved UK Biobank's scientific protocol and operational procedures (REC Reference Number: 06/MRE08/65). This study was conducted under the approved application number 42256.
 The Finnish data were analyzed under approval from The Coordinating Ethics Committee of the Helsinki and Uusimaa Hospital District, which evaluated and approved the FinnGen research project. The FinnGen project complies with existing legislation (in particular the Biobank Law and the Personal Data Act). The official data controller of the FinnGen study is the University of Helsinki.
 Analyses of data from the US (Intermountain) were approved by The Intermountain Healthcare Institutional Review Board and all participants provided written informed consent and samples for genotyping.
 Norwegian data were analyzed under approval from the National Health Screening Service, Oslo (now the Norwegian Institute of Public Health) in cooperation with the University of Bergen.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences Behavioural & social sciences Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Sample sizes are reported in the article and correspond to all available data.
Data exclusions	No available data was excluded from the study, other than data from participants of non-European ethnicity as described for all cohorts in methods.
Replication	The GWAS meta-analyses reported in the study were performed using data from 6 populations (Iceland, Denmark, Norway, UK, USA, Finland) and association results across populations are provided in our results. We did not conduct replication since we had all study data available to us included in the GWAS meta-analysis. For all known migraine variants published by Hautakangas et al., 2022, we provide association results with all studied phenotypes.
Randomization	Randomization was not performed. Within each GWAS included in the GWAS meta-analyses, covariates were adjusted for to account for potential confounding. Covariate adjustments are reported in detail in methods.
Blinding	Group allocation was not relevant to this study, hence blinding was not performed.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

- | n/a | Included in the study |
|-------------------------------------|--------------------------------------------------------|
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Antibodies |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Eukaryotic cell lines |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Palaeontology and archaeology |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Animals and other organisms |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Clinical data |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Dual use research of concern |

Methods

- | n/a | Included in the study |
|-------------------------------------|-------------------------------------------------|
| <input checked="" type="checkbox"/> | <input type="checkbox"/> ChIP-seq |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Flow cytometry |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> MRI-based neuroimaging |