Unravelling the mystery of migraine and cluster headache: insights into the genetics and biochemistry of these neurological disorders

Harder, A.V.E.

7

Cluster headache genome-wide association study and meta-analysis identifies eight loci and implicates smoking as causal risk factor

Bendik S. Winsvold*, Aster V.E. Harder*, Caroline Ran*, Mona A. Chalmer*, Maria Carolina Dalmasso*, Egil Ferkingstad*, Kumar Parijat Tripathi, Elena Bacchelli, Sigrid Børte, Carmen Fourier, Anja S. Petersen, Lisanne S. Vijfhuizen, Emer O'Connor, Gyda Bjornsdottir, Paavo Häppölä, Yen-Feng Wang, Ida Callesen, Tim Kelderman, Victor J. Gallardo, Irene de Boer, Felicia Jennysdotter Olofsgård, Katja Heinze, Nunu Lund, Laurent F. Thomas, Chia-Lin Hsu, Matti Pirinen, Heidi Hautakangas, Marta Ribasés, Simona Guerzoni, Prasanth Sivakumar, Janice Yip, Axel Heinze, Fahri Küçükali, Sisse R. Ostrowski, Ole B. Pedersen, Espen S. Kristoffersen, Amy E. Martinsen, María S. Artigas, Susie Lagrata, Maria Michela Cainazzo, Joycee Adebimpe, Olivia Quinn, Carl Göbel, Anna Cirkel, Alexander E. Volk, Stefanie Heilmann-Heimbach, Anne Heidi Skogholt, Maiken E. Gabrielsen, Leopoldine A. Wilbrink, Daisuke Danno, Dwij Mehta, Daníel F. Guðbjartsson, *HUNT All-In Headache*, *The International Headache Genetics Consortium*, *DBDS Genomic Consortium*, Frits R. Rosendaal, Ko Willems van Dijk, Rolf Fronczek, Michael Wagner, Martin Scherer, Hartmut Göbel, Kristel Sleegers, Olafur A. Sveinsson, Luca Pani, Michele Zoli, Josep A. Ramos-Quiroga, Efthimios Dardiotis, Anna Steinberg, Steffi Riedel-Heller, Christina Sjöstrand, Thorgeir E. Thorgeirsson, Hreinn Stefansson, Laura Southgate, Richard C. Trembath, Jana Vandrovcova, Raymond Noordam, Koen Paemeleire, Kari Stefansson, Cathy Shen-Jang Fann, Elisabet Waldenlind, Erling Tronvik, Rigmor H. Jensen, Shih-Pin Chen, Henry Houlden, Gisela M. Terwindt, Christian Kubisch, Elena Maestrini, Michail Vikelis, Patricia Pozo-Rosich, Andrea C. Belin*, Manjit Matharu*, Arn M.J.M. van den Maagdenberg*, Thomas F. Hansen*, Alfredo Ramirez*, John-Anker Zwart* for the International Consortium for Cluster Headache Genetics.

* shared authorship

Ann Neurol 2023; 10.1002/ana.26743

Abstract

Objective: Aggregating data for the first genome-wide association study meta-analysis of cluster headache, to identify genetic risk variants and gain biological insights.

Methods: A total of 4,777 cases (3,348 men and 1,429 women) with clinically diagnosed cluster headache were recruited from ten European and one East Asian cohorts. We first performed an inverse-variance genome-wide association meta-analysis of 4,043 cases and 21,729 controls of European ancestry. In a secondary trans-ancestry meta-analysis we included 734 cases and 9,846 controls of East Asian ancestry. Candidate causal genes were prioritized by five complementary methods: expression quantitative trait loci, transcriptome-wide association, fine-mapping of causal gene sets, genetically driven DNA methylation, and effects on protein structure. Gene set and tissue enrichment analyses, genetic correlation, genetic risk score analysis and Mendelian randomization were part of the downstream analyses.

Results: The estimated SNP-based heritability of cluster headache was 14.5%. We identified nine independent signals in seven genome-wide significant loci in the primary meta-analysis, and one additional locus in the trans-ethnic meta-analysis. Five of the loci were previously known. The 20 genes prioritized as potentially causal for cluster headache showed enrichment to artery and brain tissue. Cluster headache was genetically correlated with cigarette smoking, risk-taking behavior, ADHD, depression and musculoskeletal pain. Mendelian randomization analysis indicated a causal effect of cigarette smoking intensity on cluster headache. Three of the identified loci were shared with migraine.

Interpretation: This first genome-wide association study meta-analysis gives clues to the biological basis of cluster headache and indicates that smoking is a causal risk factor.

Introduction

Cluster headache (CH) is a primary headache disorder that affects 0.1% of the population and is four times more common in men than in women.[1] It is characterized by episodes of excruciating unilateral pain centered around the eye or the temple.[2] The large majority of patients are either current or previous smokers and there is a higher prevalence of illicit drug use, depression and sleep disorders among patients with CH than in the general population.[1, 3]

Much is unknown about the pathophysiology of CH, but hypothalamic, trigeminovascular, and autonomic nervous system dysfunction are likely involved.[1, 4] Previous twin- and family-based studies have suggested the involvement of genetic factors,[5] and two recent genome-wide association studies (GWAS) in individuals of European ancestry[6, 7] demonstrated robust genetic associations for CH, independently identifying four genetic risk loci on chromosome 1 (near the gene *DUSP10*), chromosome 2 (within *MERTK* and near *SATB2*), and chromosome 6 (within *FHL5*), with odds ratios (ORs) ranging from 1.30 to 1.61. A third GWAS in Han Chinese individuals replicated two of these loci (*MERTK* and *SATB2*) and reported an additional locus in the gene *CAPN2*.[8]

To identify additional genetic factors and increase power for functional interpretation of the genetic signals, we established the International Consortium for Cluster Headache Genetics (CCG) and analyzed data from ten European and one East Asian CH cohorts; those used in the four previous GWASs of CH[6, 7, 9] and five additional cohorts, increasing the sample size for analysis 3.2-fold compared to the largest previous CH GWAS.[7]

Methods

Cohorts and phenotyping

For reference, acronyms are listed in **Table S1**. Data were obtained from ten European and one East Asian cohorts (**Table 1**), with a combined sample size of 4,777 patients with CH (3,348 men and 1,429 women) and 31,575 controls, of which 4,043 patients (85%) were of European and 734 (15%) of East Asian ancestry. Cases were recruited between 2005 and 2022 through specialized headache clinics and diagnosed according to standardized ICHD criteria.[2, 10] Details on the recruitment and phenotyping in each cohort is provided in **Table S1**. All studies were approved by local research ethics committees, and written informed consent was obtained from each study participant.

GWAS and meta-analysis

A standardized quality control (QC) and analysis protocol was applied to each individual GWAS, while allowing for adaptations to comply with local data sharing regulations and analysis pipelines. Details are given in **Table S3**. Samples in each cohort were genotyped on genome-wide arrays,

and QC was performed on each dataset prior to imputation. Only variants with an imputation quality of ≥ 0.3[11] and a minor allele count of ≥ 12 were kept for further analysis. For X chromosome analyses males were coded as diploid. Prior to the meta-analysis, the per-study allele labels and allele frequencies were compared with those of the imputation reference panels using EasyQC,[11] and removed or reconciled mismatches. The analysis of the Taiwanese cohort was performed separately.[8]

We first conducted, an inverse variance weighted fixed-effects meta-analysis of European-ancestry cohorts using METAL,[12] without genomic control. A total of 14,860,930 variants were present in at least one cohort and included in the meta-analysis, and 5,199,189 (35%) variants were present in all ten cohorts. To identify additional loci we next conducted a secondary trans-ancestry GWAS meta-analysis that also included the East Asian ancestry cohort, using MR-MEGA with default settings,[13] which accounts for allelic heterogeneity between ancestries. Of 15,425,163 variants analyzed, 3,792,160 were present in the East Asian cohort. Of these, 3,225,258 (85%) were also present in at least one European cohort. Genome-wide significance was set to $p < 5 \times 10^{-8}$.

Due to heterogeneity in allele frequencies and differences in LD structure between European and East Asian populations, which complicates LD modeling, we focused subsequent fine-mapping and functional analyses on data from the European-ancestry GWAS.

SNP-based heritability was calculated using LDSC[14] after excluding variants that (1) were not present in the HapMap 3 reference panel, (2) explained > 1% of phenotype variation, or variants in LD ($r^2 > 0.1$) with these, and (3) were in the major histocompatibility complex region. Heritability estimates were converted to the liability scale assuming a population prevalence of CH of 0.1%.[1]

Fine-mapping for significant loci was performed using PICS2[15] with 1000 Genomes EUR LD reference. Next, a stepwise conditional analysis was performed using FINEMAP[16] [17] Only biallelic, non-indel variants were included, and a $p < 5 \times 10^{-8}$ was used to define SNPs that were conditionally independent from the lead variant.

Candidate gene mapping

To prioritize candidate genes for a causal association to CH, five methods were applied: (1) expression quantitative trait locus (eQTL) analysis, (2) transcriptome-wide association (FUSION), (3) fine-mapping of causal gene sets (FOCUS), (4) association to genetically driven DNAm (MetaMeth), and (5) genes affected by protein-altering variants in high LD with the lead CH variants.

eQTL analysis

Association between variants and gene expression (*cis*-eQTL) was estimated based on RNA sequencing and genotype data from 59,327 individuals (**Table S4**).[18] For each CH variant it was tested whether the variant itself, or variants in high LD ($r^2 \geq 0.8$), associated with one or more top *cis*-eQTLs, defined as the variant with the lowest *p* value within a distance of 1 Mb from the gene

for each gene and tissue. The significance threshold was determined at $p < 1 \times 10^{-9}$. Details on data sources and methods are described previously.[17,18]

Transcriptome-wide association study analysis (TWAS-FUSION)

To identify genes whose expression is significantly associated with CH, the CH meta-analysis results were integrated with gene expression data from single tissues (**Table S5**) using TWAS-FUSION.[19] TWAS expression weights were computed using five linear models (**Table S5**), followed by cross-validation to determine the best performing model for a given gene. The imputed gene expression was then used to test for association with CH, taking into account the LD structure and Bonferroni correcting for the number of genes tested for the given tissue. A joint/conditional analysis was performed to test for the significance of GWAS signals after removing TWAS-significant signals (expression weight from TWAS). Each variant association from the CH GWAS meta-analysis was conditioned on the joint model and a *p* value for conditional analysis results was obtained by permutation testing.

Fine-mapping of causal gene sets (FOCUS)

FOCUS[20] took as input the CH meta-analysis results, the previously calculated TWAS expression prediction weights and LD-information for all SNPs in the risk regions, and estimated the probability for any given set of genes to explain the respective TWAS signal. FOCUS was run for chromosomes 1,2,6,7 and 17, in which TWAS-Fusion showed suggestive association of genes with tissues.

Genetically driven DNA methylation scan (MetaMeth)

Association between CH and genetically driven DNA methylation (DNAm) was assessed using the MetaMeth function in EstiMeth (v1.1).[21] EstiMeth includes 86,710 models reflecting a robust genetically driven signal at methylation of 5'-C-phosphate- G-3' (CpG) sites in whole blood.[21] The approach was applied to the CH meta-analysis results, and significance was set at p value < 0.05 after false discovery rate (FDR) correction. Each CpG was paired with its annotated gene(s) and represented in a Miami plot using the R-project (https://www.R-project.org/) ggplot package.[22]

Protein-altering variants (VEP-Ensembl)

At deCODE Genetics (Iceland), for each of the lead CH variants it was determined if it was in high LD ($r^2 > 0.80$, based on the Icelandic genotype data) with protein-altering (coding or splice) variants with moderate or high impact, as annotated using release 100 of the Ensembl Variant Effect Predictor (VEP-Ensembl) tool.[23]

Gene set and tissue enrichment analyses

Genes prioritized by at least one of the five methods were used as input to the GENE2FUNC tool implemented in FUMA[24] to examine enrichment in differentially expressed gene (DEG)

sets for 54 tissues from GTEX v8,[25] and in biological pathways and functional categories from MsigDB, WikiPathways and the NHGRI GWAS catalog.[24] *P* values < 9.26×10^{-4} (0.05/54 tests) were considered statistically significant.[24] We also applied two approaches based on variant-level summary statistics: (1) DEPICT v1.194 analysis[26] applied to independent variants with a nominal association to CH ($p < 1 \times 10^{-6}$), and (2) LD-Score Regression applied to specifically expressed genes (LDSC-SEG) v1.0.1.[27] applied to the full set of summary statistics from the meta-analysis. Both methods were run with default settings. FDR < 0.05 was considered statistically significant.

Drug target identification

For genes prioritized by at least one of the five methods, we examined their druggability status using the dataset from Finan et al.[28] (**Table S6**). For detailed structured information about drugs and drug targets we integrated information from the DrugBank online database (https://www.drugbank.com)[29] (version 5.1.9, released 2022-01-04).

Genetic risk score analysis

Genetic risk scores (GRS) were based on summary statistics from the meta-analysis of all European ancestry cohorts except the given cohort to create independent test samples. In three cohorts (Dutch, Swedish cohort 1 and Danish) GRS were calculated with LDpred2,[30] which uses the whole discovery dataset without applying a *p* value threshold. In the German cohort GRS were calculated using PRSice2,[31] (**Tables S7**). Sample-specific GRSs were normalized using the target sample mean and standard deviation. Using linear regression, adjusting for sex and the first 4-6 principal components, we examined the association of GRS in each cohort to case-control status, and among cases to episodic vs. chronic CH, male vs. female patients, age at onset, currently smoking yes vs. no and ever vs. never smoked was examined for each cohort. *P* values < 0.0024 (0.05/21 tests) were considered statistically significant.

Genetic correlation

In a hypothesis-free fashion, LDSC (v1.0.1.)[14] was used to calculate pairwise genetic correlations between CH and 1,150 phenotypes from published GWAS (**Table S8)** based on GWAS summary statistics. Applying a stringent Bonferroni correction (0.05/1,150), the significance threshold was set at ($p < 4.35 \times 10^{-5}$). To evaluate differences in the correlation profiles for CH and migraine, the genetic correlation was calculated between migraine (48,975 migraine cases and 540,381 controls from Hautakangas et al.,[17] not including 23andMe) and each of the traits that were significantly correlated with CH, while Bonferroni correcting for the number of tests (0.05/84, $p < 5.95 \times 10^{-5}$).

Colocalization analysis

To test whether CH loci that were in close proximity to previously reported migraine loci share causal variants for both CH and migraine, the Bayesian colocalization procedure implemented in

the R package 'coloc' (v5.1.0) was used with default settings[32] and the migraine dataset described above. Colocalization was tested for the region between the two nearest recombination hotspots (https://bitbucket.org/nygcresearch/ldetect-data/src/master/EUR/).

Mendelian randomization analysis

To test for a causal effect of smoking on CH, we performed a summary statistics-based two-sample inverse-variance weighted (IVW) Mendelian randomization analysis,[33] using as instrumental variables 40 independent variants significantly ($p < 5 \times 10^{-8}$) associated with "Cigarettes smoked per day" in a previous GWAS,[34] as an indication for smoking intensity (**Table S9**). Since the IVW method assumes the absence of horizontal pleiotropy, several sensitivity analyses were employed to exclude pleiotropy. Cochran's Q tests were used to detect heterogeneity.[35] In addition, the MR-Egger intercept was used to detect directional pleiotropy.[35, 36] Both models were fit using robust regression and assuming a *t*-distribution of the fitted parameters. Analyses were performed using the MendelianRandomization package (version 0.5.1) in R (version 3.6.3). To verify the causality between smoking and CH, we applied a latent causal variable (LCV) model to estimate the genetic causality proportion (GCP).[37] Here, a latent variable mediates the genetic correlation, avoiding false positives due to genetic correlations when determining causality. A GCP of 0 is interpreted as no, and GCP of 1 as complete, genetic causality.

Results

European-ancestry GWAS meta-analysis

Seven independent genome-wide significant CH associated ($p < 5 \times 10^{-8}$) risk loci (**Table 2**, **Figure 1** and **2**) were identified. Associations were consistent across the ten cohorts (heterogeneity $p > 0.10$, **Tables 2** and **S10**). Named by their nearest protein-coding gene, four of risk loci were previously reported[6, 7] (*DUSP10*, *MERTK*, *FTCDNL1* and *FHL5*), while three are novel (*WNT2*, *PLCE1*, *LRP1*). A stepwise conditional analysis using FINEMAP[16] revealed that two of the identified loci (*MERTK* and *WNT2*) contained additional independent signals, increasing the number of independent association signals to nine (**Table S11**). Fine-mapping with PICS2[15] suggested that the lead signal in the *LRP1* locus (rs11172113) is most likely the causal variant (posterior probability 65.8%). Five other variants in three other loci had PICS2 posterior probability > 10% for being causal (**Table S12**).

The genomic inflation factor (λ) was 1.086, while the LD score regression intercept was 1.004 (SE 0.007), with a ratio of 0.033 (SE 0.062), indicating that 96.7% of the observed signal is caused by true polygenic heritability rather than confounding factors, such as population stratification. The estimated SNP-based heritability (h^2) of CH was 14.5% (SE 1.74%) on the liability scale.

One additional genome-wide significant CH locus, in *CAPN2*, was identified when adding the East Asian cohort in an ancestry-adjusted GWAS meta-analysis (**Table 3**, **Table S13**, **Figure**

3). This locus, previously reported and internally replicated within the East Asian cohort.[8] was exclusively driven by the same cohort in our analysis (see **Table S13**). However, a nearby locus reached nominal significance in the European-ancestry meta-analysis, with lead variant rs68046706 (OR 1.76, 95% CI 1.10 - 1.26, p = 3.86 x10^{-6}) 86 kb away from rs10916600. The *WNT2* locus identified in the European-ancestry meta-analysis, for which the lead variant was not present in the East Asian cohort, fell below significance (p = 5.91 x 10^{-7}). At the *PLCE1* locus, the new lead variant was a missense variant (rs2274224) in *PLCE1*. Cohort-wise associations for all the identified loci are given in **Tables S10** and **S13**.

All the five previously reported GWAS-significant loci were re-identified in our study, while none of the associations reported from candidate gene studies were replicated (**Table S14**).

Table 1 Cluster headache GWAS studies included in the meta-analysis.

Study	Cases (n)	Controls (n)
Dutch Cluster Headache Cohort[a]	943	1,424
UK Cluster Headache Cohort[b]	852	5,614
Swedish Cluster Headache Cohort 1[b]	591	1,134
German Cluster Headache Cohort	477	938
Danish Cluster Headache Cohort	492	9,658
Swedish Cluster Headache Cohort 2	255	241
Trondheim Cluster Headache Cohort[a]	144	1,800
Greek Cluster Headache Cohort	99	91
Barcelona Cluster Headache Cohort	97	482
Italian Cluster Headache Cohort	93	347
Total	4,043	21,729

[a] Previously published in whole or in part by Harder et al.[6]; [b] Previously published by O'Conner et al.[7]

The subsequent downstream analyses were based on the European-ancestry meta-analysis. To prioritize candidate genes for a causal association with CH, we applied five methods. (1) eQTL analysis found that at the *MERTK* locus, three variants in high LD (r^2 > 0.92) with the lead variant rs13399108 modulate the expression of *TMEM87B* (in fibroblasts and aortic artery) and *SLC20A1* (in whole blood). At the *FHL5* locus, two variants (r^2 > 0.84 with the lead variant rs9486725) associate with the expression of *UFL1* (in whole blood, white blood cells and tibial artery). At the *LRP1* locus, the T allele of lead variant rs11172113 associates with an increased *LRP1* mRNA expression in aortic artery, adipose tissue and tibial artery (**Table S15**). (2) The transcriptome-wide association study (TWAS-FUSION) identified eight candidate genes at five loci with a significant TWAS p value ≤ 1.0 x 10^{-6} (**Table S5**). (3) Fine mapping by FOCUS identified eight candidate genes based on posterior inclusion probability (PIP) > 0.5 (**Table S16**). Four genes (*MERTK*, *TMEM87B*, *SATB2* and *CFTR*) were prioritized by both TWAS-FUSION and FOCUS with high confidence (PIP > 0.99 in the same tissue in both analyses). (4) Using MetaMeth, 13 CpG sites at nine genes were predicted to be hypo- or hypermethylated in

CH (**Table S17**, **Figure 4**). (5) At two loci, the lead variant was in high LD with protein-altering missense variants. That is, at the *FHL5* locus, the intronic lead variant rs9486725 is in strong LD ($r^2 \geq 0.98$) with p.Arg204Gly (rs2273621) and p.Ser243Arg (rs9373985 in *FHL5*; and at the *PLCE1* locus the intronic lead variant rs57866767 is in strong LD ($r^2 = 1$) with a p.Arg1267Pro (rs2274224) in *PLCE1* (**Table S18**).

Table 2 Summary of the genomic loci associated with cluster headache.

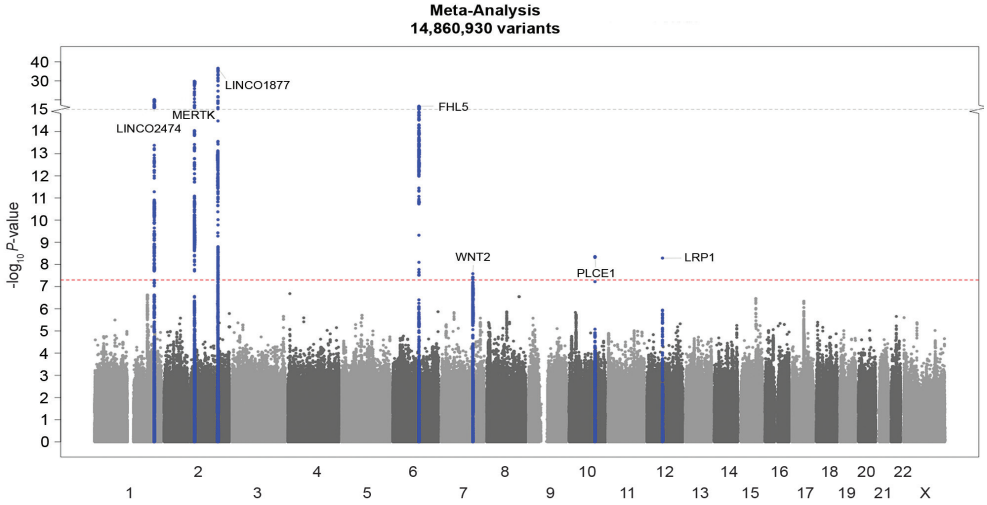
Locus name	Lead variant (Chr:Pos)	EA/ NEA (EAF)	OR (95% CI)	*p* value (Het *p*)	Variant type [Prioritized genes]
DUSP10	rs17011182 (1:222164327)	A/G (0.793)	1.38 (1.29-1.48)	7.76×10^{-21} (0.58)	regulatory region [*DUSP10*]
MERTK	rs13399108 (2:112747123)	A/G (0.373)	1.41 (1.33-1.50)	1.74×10^{-30} (0.16)	intron [***MERTK***, ***TMEM87B***, ***FBLN7***, *SLC20A1*]
FTCDNL1	rs6714578 (2:200485487)	A/G (0.655)	1.53 (1.43-1.63)	2.83×10^{-37} (0.65)	intergenic [***SATB2***]
FHL5	rs9486725 (6:97061159)	T/C (0.346)	1.29 (1.21-1.36)	2.50×10^{-17} (0.29)	intron [***UFL1***, *FHL5*, *KLHL32*, *NDUFAF4*]
WNT2	rs2402176 (7:116908448)	C/G (0.291)	1.20 (1.12-1.27)	2.61×10^{-8} (0.51)	intergenic [***CFTR***, *CAPZA2*, *ST7*]
PLCE1	rs57866767 (10:96023077)	T/C (0.588)	1.18 (1.12-1.25)	4.45×10^{-9} (0.51)	intron [*PLCE1*]
LRP1	rs11172113 (12:57527283)	T/C (0.600)	1.18 (1.12-1.25)	5.15×10^{-9} (0.52)	intron [***LRP1***]

Locus name = the closest protein-coding gene within a 250-Kb window. Chr = chromosome. Pos = position (hg19). EA = effect allele, which here is set to correspond with the risk allele. NEA = non-effect allele. EAF = effect allele frequency. OR = odds ratio. CI = confidence interval. Het p = p value from Cochran's Q-test for heterogeneity. Prioritized genes = genes prioritized by at least one of five complementary methods: (1) expression quantitative trait (eQTL) analysis, (2) transcriptome-wide association analysis using FUSION, (3) fine mapping of causal gene sets (FOCUS), (4) association to genetically driven DNAm (MetaMeth), and (5) protein-altering variants in high LD (r2 > 0.8) with index variant. Genes identified by ≥ 2 of the methods are marked in bold.Candidate gene mapping and functional characterization

Twenty genes were prioritized by at least one of the five methods. A summary of the gene prioritization results is given in **Table S19**.

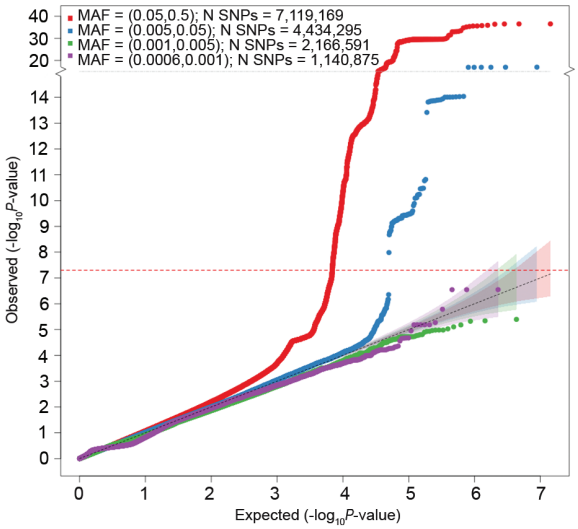
When considering the 20 prioritized genes, FUMA[24] found a significant enrichment for genes differentially expressed in artery (tibial artery) and brain (substantia nigra) (**Figure 5** and **Table S20**), and a significant overlap with genes reported in the GWAS catalog for 10 traits, most significantly for headache and migraine (**Table S21**). The summary statistics-based enrichment analyses DEPICT and LDSC-SEG did not yield significant enrichment for gene sets or tissues after correcting for multiple testing (**Tables S22-26**). Of the 20 prioritized genes (**Table S19**), ten are highlighted as druggable in the druggable genome database.[28] Of these, five encode targets of 33 existing drugs registered in DrugBank[29] (**Table S6**), including three genes that were implicated in CH by at least two gene prioritization methods (i.e. *MERTK*, *CFTR* and *LRP1*). Calpain 2, encoded by *CAPN2* in the trans-ancestry locus, was not registered in DrugBank.

Figure 1 Manhattan plot showing genome-wide significant loci associated with cluster headache (4,043 cases, 21,729 controls).



The horizontal axis shows the chromosomal position and the vertical axis shows the significance (-log10 p value) of tested markers. Each dot represents a genetic variant. The threshold for genome-wide significance ($p < 5 \times 10^{-8}$) is indicated by a red dotted line, and genome-wide significance loci are shown in blue.

Figure 2 Quantile-quantile (Q-Q) plot for association with cluster headache.

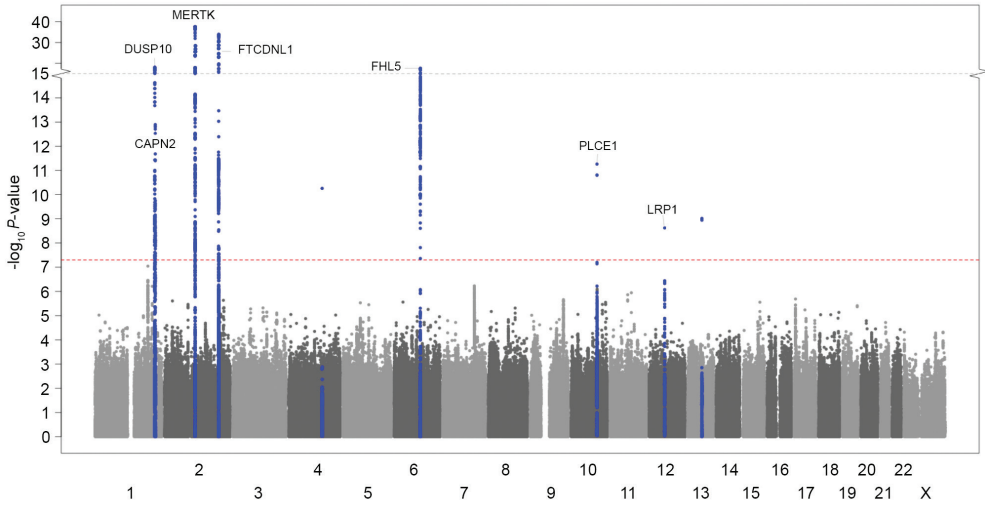


The horizontal axis shows -log10 p values expected under the null distribution. The vertical axis shows observed -log10 p values. Genomic inflation factor (λ) = 1.086. Red = common variants (MAF ≥ 5%), blue = low frequency variants (MAF = 0.5 - 5%), green = rare variants (MAF = 0.1 - 0.5%), purple = very rare variants (MAF < 0.1%). MAF = minor allele frequency; SNPs = single nucleotide polymorphisms.

Genetic risk score analysis

GRS for CH were associated with case-control status in leave-one-out analyses in each of the four tested independent cohorts. Among cases with CH, no association was seen between GRS and episodic vs. chronic CH, age-at-onset, sex, current smoking or ever smoking (**Table S7**).

Figure 3 Manhattan plot showing genome-wide significant loci associated with cluster headache in trans-ancestry meta-analysis (4,777 cases, 31,575 controls).



The horizontal axis shows the chromosomal position and the vertical axis shows the significance (-log10 *p* value) of tested markers. Each dot represents a genetic variant. The threshold for genome-wide significance ($p < 5 \times 10^{-8}$) is indicated by a red dotted line, and genome-wide significance loci are shown in blue. Three genome-wide significant variants (rs9307511 on chr4 and rs338106 and rs747974 on chr 13) were considered spurious associations as they lacked a supporting LD structure, were driven by the East Asian cohort alone, and were previously interpreted as being spurious associations in this cohort.[8]

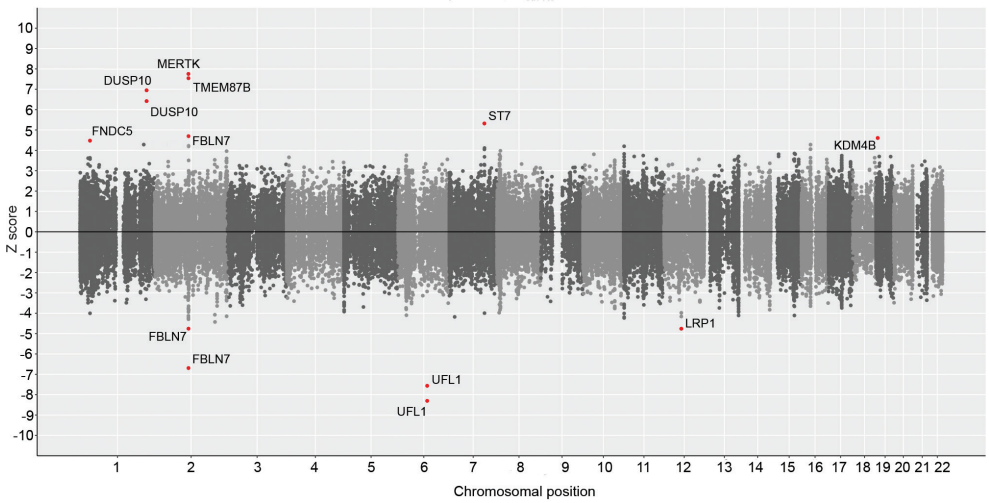
Genetic correlation

After correcting for multiple testing, CH was genetically correlated with 84 traits (**Table S8**). The strongest correlation was with 'cigarettes per day'[34] (rg = 0.36, $p = 6.32 \times 10^{-18}$). Notably, ten (12%) of the correlated traits were related to smoking behavior. CH was also positively correlated with measures of risk-taking behavior, ADHD, mood disorders, musculoskeletal pain, migraine, and with unfavorable lifestyle factors including low physical activity, low nutritional diet and lower educational attainment (**Table S8**). When examining the correlation of the same 84 traits to migraine, the genetic correlations to pain, depression and ADHD were similar to those seen for CH, while no correlation was observed between migraine and smoking traits or measures of risk-taking behavior.

Three of the CH loci are near previously identified risk loci for migraine (i.e. *FHL5*, *PLCE1*, *LRP1*).[17] (**Table S27**). Colocalization analysis indicated that CH and migraine are caused by the

same causal variant at each of the three loci (posterior probability 98.6% for *FHL5* locus, 99.6% for *PLCE1* locus and 100% for *LRP1* locus). Effect sizes were, however, consistently higher for CH (ORs 1.29, 1.18, 1.18) than for migraine (1.09, 1.06, 1.11) with non-overlapping confidence intervals for the ORs (**Table S28**). Among 122 loci associated with migraine in the most recent GWAS,[17] no other migraine variant was associated with CH after Bonferroni correction (**Table S29**). The effect sizes (beta) for association to migraine and CH were not significantly correlated (Pearson r = 0.16, p = 0.074) for the remaining 119 variants, after excluding the three overlapping loci.

Figure 4 Miami plot of genetically driven DNA methylation genes in cluster headache.

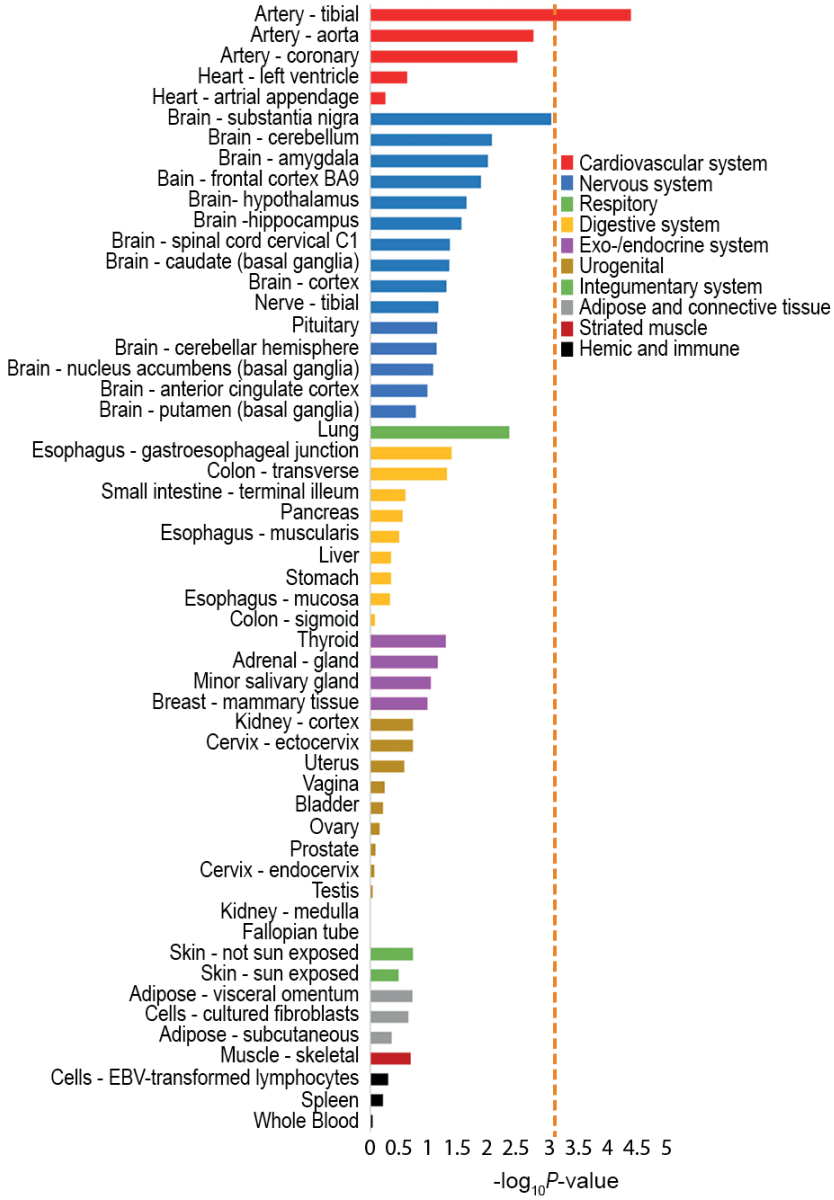


Computational prediction of genetically driven CpG methylation associated with cluster headache, using MetaMeth. Genes annotated to significant CpGs are shown (FDR-corrected p value < 0.05). Horizontal axis shows the chromosomal position and the vertical axis shows significance (-log10 pvalue). The top panel shows predicted hypermethylation, while the bottom panel shows predicted hypomethylation.

Mendelian randomization analysis

Using the random-effect inverse variant weighted (IVW) method, we observed a strong association between the instrumental variables for smoking intensity and CH (β = 1.11, SE = 0.43, p = 6.3 x 10^{-6}). The direction and magnitude were similar in the MR-Egger analysis (β = 1.04, SE = 0.55, p = 4.6 x 10^{-4}). The Cochran´s Q test statistic was significant (p = 0.03), indicative of some heterogeneity, but the MR-Egger intercept showed no evidence for bias caused by directional pleiotropy (p = 0.79). Mendelian randomization may, however, yield false positive results in the presence of genetic correlation between the two traits examined.[37] To test for this, we performed a latent causal variable model, finding that smoking intensity had a nearly full (> 0.6) genetic causality with CH (p_{LCV} = 8.57 x 10^{-10}, GCP = 0.74 ± 0.18). Combined, the results strongly support a causal effect of smoking intensity on CH. Full results are presented in **Tables S30-32**.

Figure 5 Tissue enrichment for the putative causal genes.



Enrichment of the 20 genes with supportive evidence for implication in cluster headache in differentially expressed gene (DEG) sets for 54 tissues from GTEXv8. The analysis was performed using FUMA and based on pre-calculated DEG sets defined by a two-sided *t*-test per tissue versus all other tissues. The red line shows the significance threshold after adjustment for multiple testing by Bonferroni correction (p = 0.05/54 tests = 9.26 x 10^{-4}).

Discussion

In a GWAS meta-analysis for CH in European-ancestry cohorts we identified nine independent associations in seven risk loci and confirm the strong associations at four loci (ORs 1.29 - 1.53) reported in recent smaller GWAS.[6-8] One additional locus, previously reported and internally replicated in the East Asian cohort,[8] was identified in a subsequent trans-ancestry GWAS meta-analysis that included this cohort.

We estimate that common genetic variants explain 14.5% of CH's phenotypic variance. Twenty genes were prioritized as candidates for being involved in CH. These showed enrichment for arterial tissue, in addition to brain, fueling the idea that CH may have a vascular involvement.[1] Still, since no significant tissues were identified by summary statistics-based enrichment analyses (using DEPICT and LDSC-SEG), more evidence is needed to draw definite conclusions. Several of the 20 prioritized genes encode targets for existing drugs, and may represent candidates for repurposing studies. The clinical utility of GRS remains to be explored. We found no association between GRS and specific clinical phenotypes, suggesting that the signal is not driven by any of the subgroups.

Differences in CH clinical presentation between Asian and European populations, such as reduced restlessness and circadian rhythmicity, may indicate distinct genetic predispositions.[38] The CAPN2 locus was selectively driven by the East Asian cohort, and may exemplify how the contribution of individual risk loci varies between populations. Future well-powered trans-ancestral studies should further explore ancestry-related risk loci, and whether these are related to differences in clinical presentation.

In our hypothesis-free genetic correlation analysis CH was correlated with several traits, including smoking, risk-taking behavior, ADHD, mood disorders, musculoskeletal pain and migraine. The strongest genetic correlation was with smoking, which is consistent with the observation that as many as 70 - 90% of patients with CH smoke,[1,3,39] seen also in our cohorts (**Table S1**). The high proportion of smokers among patients with CH may theoretically be explained by smoking causing CH or *vice versa*, or because they have shared causal factors. Whether smoking is causing CH is heavily debated. On the one hand, smoking initiation typically predates the onset of CH[3] and among those with CH who have never smoked the majority were exposed to parental smoking in childhood.[40] Furthermore, it seems that smoking is associated with more severe manifestations of CH[1] and some data suggest that the prevalence of CH has followed trends in smoking prevalence.[39] On the other hand, arguments against a causal effect of smoking include the typically long latency between smoking onset and CH debut (> 15 years).[3] Also, in retrospective studies patients with CH who stopped smoking several years earlier did not experience an improvement in their CH.[1,39]

To investigate the potential causality of smoking on CH, we performed a Mendelian randomization and LCV analysis.[41] The analyses indicated a causal effect of smoking intensity on CH, with

high statistical confidence. Of note, the high observed proportion of smokers among cases with CH is expected if smoking is a causal risk factor. Since cases were recruited independently of smoking status, and the proportion of smokers is similar to previous reports, we find it unlikely that recruitment bias explains the results.

While our study cannot give definite answers regarding mechanisms linking smoking to CH, we note that several of the prioritized genes are influenced by smoking. Cigarette smoking leads to overexpression of *MERTK*[42] and reduced expression and function of *CFTR* in airway tissues.[43] Notably, our TWAS also revealed an increased expression for *MERTK* and reduced expression for *CFTR* in CH. It has been shown that smoking can induce epigenetic changes that persist even 30 years after smoking cessation,[44] therefore, the observation that patients who stop smoking do not experience an improvement of their CH might be explained by stable epigenetic modifications. In a large study, DNA methylation at 2,568 CpG sites related to 1,450 genes were found to be associated with *former* smoking at FDR < 0.05.[44] Four of our prioritized genes are among these (i.e. *FBLN7*, *SLC20A1*, *KDM4B*, *ST7*), that is 4 of 20 vs. 1,450 of 23,300 genes (*post hoc* one-tailed binomial $p = 0.033$). More detailed molecular studies in relevant tissues are needed to identify mechanisms linking smoking to CH.

The suggestion that smoking is a causal risk factor for CH has potential clinical implications. Smoking is a modifiable risk factor, and it gives a further impetus to promoting smoking cessation in this group of patients. The long-term effect of smoking cessation on CH should be carefully revisited by well-designed prospective studies.

Notably, CH was to some extent genetically correlated with measures of risk-taking behavior apart from smoking. While our results support a causal effect of smoking on the development of CH, it is possible that patients with CH are also more likely to start smoking because of a tendency toward risk-taking, as has been suggested.[39, 45] The genetic correlations to smoking and risk-taking behavior were not seen for migraine.

While primary headache disorders are among the top causes of disability worldwide,[46] it is unknown to what extent they represent biologically distinct disorders or rather variations in clinical presentation with a shared biological basis.[47] Migraine is the only other primary headache disorder that has been explored in well-powered GWAS.[17] We found that three of the eight risk loci for CH are shared with migraine, and colocalization analyses give a high probability that the same causal variants in these loci give rise to both disorders. Notably, the remaining five CH loci show no association to migraine (p values > 0.10). Likewise, apart from the three overlapping loci, none of the other 119 known migraine loci[17] show association with CH. Our results suggest, therefore, that CH and migraine have a partly shared and partly distinct genetic basis, likely reflecting partly shared and partly distinct biological mechanisms. This corresponds well with the clinical impression of the two disorders as being distinct entities, but with certain shared clinical characteristics, including unilateral headache cranial autonomic symptoms, and response to some

of the same medications.[47, 48] Future studies with deep phenotyping should explore if the shared genetic risk factors are directly related to shared clinical features, such as prominent autonomic symptoms in some migraine patients.[49]

We note that for all three shared loci, the effect sizes were higher for CH (ORs 1.18 - 1.29) than for migraine (1.06 - 1.11) with non-overlapping confidence intervals. Even for the most consistently identified migraine risk locus, *LRP1* (*p* value 1.38 x10^{-90} in the latest migraine GWAS),[17] the effect size was higher for CH (1.18 vs. 1.11). This holds true also when comparing to GWAS of clinic-based migraine cohorts (OR = 1.11).[50] The larger effect sizes suggest that the three shared loci are stronger drivers of disease susceptibility in CH than in migraine, and also makes it unlikely that the observed associations are a result of misclassification of migraine patients as having CH.

A major strength of our study is the substantially larger sample size compared to previous studies, which allows for downstream functional analyses, and clinical diagnoses made according to ICHD criteria.[2, 10] This was made possible through the establishment of the International Consortium for Cluster Headache Genetics (CCG), which has brought together 16 headache research groups from 13 countries (www.clusterheadachegenetics.org). A limitation of the current study is that it included only a single non-European cohort, from east Asia, limiting, the possibility for conducting ancestry-specific meta-analyses and downstream analyses, for non-European ancestries. This highlights the need for future, well-powered trans-ancestry genetic studies in CH.

In conclusion, in this GWAS meta-analysis we identify nine independent associations in seven risk loci for CH in European-ancestry samples and one additional locus in East Asian samples. The prioritized genes show enrichment in arterial and brain tissues. CH shares certain risk loci with migraine, and is most strongly genetically correlated with smoking. Of clinical interest, Mendelian randomization analysis indicates a causal effect of cigarette smoking on the development of CH.

Supplementary Information

Supplementary Tables

https://onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1002%2Fana.26743&file=ana26743-sup-0001-Supinfo.xlsx

Acknowlegdements: We want to thank both all participating patients and their general practitioners for their good collaboration. Additional acknowledgement are found in Table S13. This work was funded by the South-Eastern Norway Regional Health Authority (#2020034); the Swedish National Infrastructure for Computing (SNIC) at Uppmax, Uppsala University partially funded by the Swedish Research Council through grant agreement no. 2018-05973; the Wellcome Trust under award 076113, 085475 and 090355; the Instituto de Salud Carlos III (PI18/01788, PI19/01224 and PI20/00041), and co-financed by the European Regional Development Fund (ERDF), the Biomedical Network Research Center on Mental Health (CIBERSAM, Madrid, Spain); the Agència de Gestió d'Ajuts Universitaris i de Recerca-AGAUR, Generalitat de Catalunya (2017SGR1461); the Netherlands Organization for Scientific Research, i.e. the Center of Medical System Biology established by the Netherlands Genomics Initiative/Netherlands Organization for Scientific Research (to Arn M.J.M. van den Maagdenberg); the EU-funded FP7 "EUROHEADPAIN" (grant no. 6026337 to Arn M.J.M. van den Maagdenberg); The NEO study, which comprised of the Dutch controls, was supported by the participating Departments, the Division and the Board of Directors of the Leiden University Medical Centre, and by the Leiden University, Research Profile Area 'Vascular and Regenerative Medicine'; Aster V.E. Harder was sponsored by the Leiden University Fund / Fonds Mr. J.J. van Enter 'Pro Universitate', www.luf.nl (grant W212163-2-64). Part of the genotyping of the German sample was funded by the JPND EADB grant to Alfredo Ramirez (German Federal Ministry of Education and Research (BMBF) garnt:01ED1619A; the BMBF (grants KND: 01GI0102, 01GI0420, 01GI0422, 01GI0423, 01GI0429, 01GI0431, 01GI0433, 01GI0434; grants KNDD: 01GI0710, 01GI0711, 01GI0712, 01GI0713, 01GI0714, 01GI0715, 01GI0716; grants Health Service Research Initiative: 01GY1322A, 01GY1322B, 01GY1322C, 01GY1322D, 01GY1322E, 01GY1322F, 01GY1322G); PainFACT (H2020-2020-848099) to Thorgeir E. Thorgeirsson; Italian Ministry of Health (RF2009-1549619); The Research Funding Pool at Rigshospitalet to Mona A.

Chalmer; The Swedish Brain Foundation and the Mellby Gård Foundation (FO2020-0006, FO2022-0001);), Karolinska Institutet Research Funds (2018-01738, 2020-01411, 2022-01781); The Swedish Research Council (2017-01096); the Region Stockholm (ALF project 20200095); the Brain Research Center, National Yang Ming Chiao Tung University from The Featured Areas Research Center Program within the framework of the Higher Education Sprout Project by the Ministry of Education (MOE) in Taiwan; the Ministry of Science and Technology, Taiwan [MOST-108-2314-B-010 -022 -MY3 & 110-2326-B-A49A-501-MY3]. The funders had no role in the design and conduct of the study; collection, management, analysis and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Author's contributions: B.S.W., A.V.E.H., C.R., M.A.C., M.C.D., G.M.T., P.P.R., A.C.B., M.M., A.M.J.M.v.d.M., T.F.H., A.R. and J.Z. contributed to the conception and design of the study; B.S.W., A.V.E.H., C.R., M.A.C., M.C.D., E.F., K.P.T., E.B., S.B., C.F., A.S.P., L.S.V., S.H.M., E.O., G.B., P.H., Y.F.W., I.C., T.K., V.J.G., I.d.B., F.J., K.H., N.L., L.F.T., C.L.H., M.P., H.Ha., M.R., S.G., P.S., J.Y., A.H., F.K., S.R.O., O.B.P., E.S.K., A.E.M., M.S.A., S.L., M.M.C., J.A., O.Q., C.G., A.C., A.E.V., S.H.H., A.H.S., M.E.G., L.A.W., D.D., D.M., D.F.G., F.R.R., K.W.v.D., R.F., M.W., M.S., H.G., K.Sl., O.A.S., L.P., M.Z., J.A.R.Q., E.D., A.S., S.R.H., C.S., T.E.T., H.S., L.S., R.C.T., J.V., R.N., K.P., K.St., C.S.J.F., E.W., E.T., R.H.J., S.C., H.Ho, G.M.T., C.K., E.M., M.V., P.P.R., A.C.B., M.M., A.M.J.M.v.d.M., T.F.H., A.R. and J.Z. contributed to the acquisition and analysis of data; B.S.W., A.V.E.H., C.R., M.A.C., M.C.D., E.F., K.P.T. and S.H.M. contributed to drafting the text or preparing the figures.

Members of ***"HUNT All-In Headache"*** are available in Table S34

Members of ***"The International Headache Genetics Consortium"*** are available in Table S35

Members of ***"DBDS Genomic Consortium"*** are available in Table S36

Potential conflicts of interest: Nothing to report.

Data availability: Summary statistics generated by the International Consortium for Cluster Headache Genetics are available for academic use from www.clusterheadachegenetics.org/access/

Reference

1. May A, Schwedt TJ, Magis D, et al. Cluster headache. *Nat Rev Dis Primers*. 2018;4:18006.
2. Headache Classification Committee of the International Headache Society (IHS). The International Classification of Headache Disorders, 3rd edition. *Cephalalgia*. 2018;38(1):1-211.
3. Lund N, Petersen A, Snoer A, Jensen RH, Barloese M. Cluster headache is associated with unhealthy lifestyle and lifestyle-related comorbid diseases: Results from the Danish Cluster Headache Survey. *Cephalalgia*. 2019;39(2):254-263.
4. Wei DY, Goadsby PJ. Cluster headache pathophysiology - insights from current and emerging treatments. *Nat Rev Neurol*. 2021;17(5):308-324.
5. Sjaastad O, Shen JM, Stovner LJ, Elsas T. Cluster headache in identical twins. *Headache*. 1993;33(4):214-217.
6. Harder AVE, Winsvold BS, Noordam R, et al. Genetic Susceptibility Loci in Genomewide Association Study of Cluster Headache. *Ann Neurol*. 2021;90(2):203-216.
7. O'Connor E, Fourier C, Ran C, et al. Genome-Wide Association Study Identifies Risk Loci for Cluster Headache. *Ann Neurol*. 2021;90(2):193-202.
8. Chen SP, Hsu CL, Wang YF, et al. Genome-wide analyses identify novel risk loci for cluster headache in Han Chinese residing in Taiwan. *J Headache Pain*. 2022;23(1):147.
9. Bacchelli E, Cainazzo MM, Cameli C, et al. A genome-wide analysis in cluster headache points to neprilysin and PACAP receptor gene variants. *J Headache Pain*. 2016;17(1):114.
10. Headache Classification Subcommittee of the International Headache S. The International Classification of Headache Disorders: 2nd edition. *Cephalalgia : an international journal of headache*. 2004;24 Suppl 1:9-160.
11. Winkler TW, Day FR, Croteau-Chonka DC, et al. Quality control and conduct of genome-wide association meta-analyses. *Nat Protoc*. 2014;9(5):1192-1212.
12. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics*. 2010;26(17):2190-2191.
13. Mägi R, Horikoshi M, Sofer T, et al. Trans-ethnic meta-regression of genome-wide association studies accounting for ancestry increases power for discovery and improves fine-mapping resolution. *Hum Mol Genet*. 2017;26(18):3639-3650.
14. Bulik-Sullivan B, Finucane HK, Anttila V, et al. An atlas of genetic correlations across human diseases and traits. *Nat Genet*. 2015;47(11):1236-1241.
15. Taylor KE, Ansel KM, Marson A, Criswell LA, Farh KK. PICS2: next-generation fine mapping via probabilistic identification of causal SNPs. *Bioinformatics*. 2021;37(18):3004-3007.
16. Benner C, Spencer CC, Havulinna AS, et al. FINEMAP: efficient variable selection using summary data from genome-wide association studies. *Bioinformatics*. 2016;32(10):1493-1501.
17. Hautakangas H, Winsvold BS, Ruotsalainen SE, et al. Genome-wide analysis of 102,084 migraine cases identifies 123 risk loci and subtype-specific risk alleles. *Nat Genet*. 2022; 54(2):152-160
18. Ferkingstad E, Sulem P, Atlason BA, et al. Large-scale integration of the plasma proteome with genetics and disease. *Nat Genet*. 2021;53(12):1712-1721.
19. Gusev A, Ko A, Shi H, et al. Integrative approaches for large-scale transcriptome-wide association studies. *Nat Genet*. 2016;48(3):245-252.
20. Mancuso N, Freund MK, Johnson R, et al. Probabilistic fine-mapping of transcriptome-wide association studies. *Nat Genet*. 2019;51(4):675-682.
21. Freytag V, Vukojevic V, Wagner-Thelen H, et al. Genetic estimators of DNA methylation provide insights into the molecular basis of polygenic traits. *Transl Psychiatry*. 2018;8(1):31.

22. Wickham H. *ggplot2 Elegant Graphics for Data Analysis*. Use R! Springer International Publishing; 2016.
23. McLaren W, Gil L, Hunt SE, et al. The Ensembl Variant Effect Predictor. *Genome Biol*. 2016;17(1):122.
24. Watanabe K, Taskesen E, van Bochoven A, Posthuma D. Functional mapping and annotation of genetic associations with FUMA. *Nat Commun*. 2017;8(1):1826.
25. Consortium GT. The Genotype-Tissue Expression (GTEx) project. *Nat Genet*. 2013;45(6):580-585.
26. Pers TH, Karjalainen JM, Chan Y, et al. Biological interpretation of genome-wide association studies using predicted gene functions. *Nature Communications*. 2015;6:5890
27. Finucane HK, Reshef YA, Anttila V, et al. Heritability enrichment of specifically expressed genes identifies disease-relevant tissues and cell types. *Nat Genet*. 2018;50(4):621-629.
28. Finan C, Gaulton A, Kruger FA, et al. The druggable genome and support for target identification and validation in drug development. *Sci Transl Med*. 2017;9(383) :eaag1166.
29. Wishart DS, Feunang YD, Guo AC, et al. DrugBank 5.0: a major update to the DrugBank database for 2018. *Nucleic Acids Res*. 2018;46(D1):D1074-d1082.
30. Privé F, Arbel J, Vilhjálmsson BJ. LDpred2: better, faster, stronger. *Bioinformatics*. 2020;36(22-23):5424-5431.
31. Choi SW, O'Reilly PF. PRSice-2: Polygenic Risk Score software for biobank-scale data. *Gigascience*. 2019;8(7):giz082.
32. Giambartolomei C, Vukcevic D, Schadt EE, et al. Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. *PLoS Genet*. 2014;10(5):e1004383.
33. Hemani G, Zheng J, Elsworth B, et al. The MR-Base platform supports systematic causal inference across the human phenome. *Elife*. 2018;7:e34408.
34. Liu M, Jiang Y, Wedow R, et al. Association studies of up to 1.2 million individuals yield new insights into the genetic etiology of tobacco and alcohol use. *Nat Genet*. 2019;51(2):237-244.
35. Bowden J, Del Greco MF, Minelli C, et al. A framework for the investigation of pleiotropy in two-sample summary data Mendelian randomization. *Stat Med*. 2017;36(11):1783-1802.
36. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *Int J Epidemiol*. 2015;44(2):512-525.
37. O'Connor LJ, Price AL. Distinguishing genetic correlation from causation across 52 diseases and complex traits. *Nat Genet*. 2018;50(12):1728-1734.
38. Peng KP, Takizawa T, Lee MJ. Cluster headache in Asian populations: Similarities, disparities, and a narrative review of the mechanisms of the chronic subtype. *Cephalalgia*. 2020;40(10):1104-1112.
39. Ferrari A, Zappaterra M, Righi F, et al. Impact of continuing or quitting smoking on episodic cluster headache: a pilot survey. *J Headache Pain*. 2013;14:48.
40. Rozen TD. Cluster headache as the result of secondhand cigarette smoke exposure during childhood. *Headache*. 2010;50(1):130-132.
41. Sanderson E, Glymour MM, Holmes MV, et al. Mendelian randomization. *Nature Reviews Methods Primers*. 2022;2(1):6.
42. Kazeros A, Harvey BG, Carolan BJ, et al. Overexpression of apoptotic cell removal receptor MERTK in alveolar macrophages of cigarette smokers. *Am J Respir Cell Mol Biol*. 2008;39(6):747-757.
43. Bodas M, Min T, Vij N. Critical role of CFTR-dependent lipid rafts in cigarette smoke-induced lung epithelial injury. *Am J Physiol Lung Cell Mol Physiol*. 2011;300(6):L811-820.
44. Joehanes R, Just AC, Marioni RE, et al. Epigenetic Signatures of Cigarette Smoking. *Circ Cardiovasc Genet*. 2016;9(5):436-447.

45. Lambru G, Castellini P, Manzoni GC, Torelli P. Mode of occurrence of traumatic head injuries in male patients with cluster headache or migraine: Is there a connection with lifestyle? *Cephalalgia*. 2010;30(12):1502-1508.

46. Global Burden of Disease Study C. Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990-2013: a systematic analysis for the Global Burden of Disease Study 2013. *Lancet (London, England)*. 2015;386(9995):743-800.

47. Vollesen AL, Benemei S, Cortese F, et al. Migraine and cluster headache - the common link. *J Headache Pain*. 2018;19(1):89.

48. Chwolka M, Goadsby PJ, Gantenbein AR. Comorbidity or combination - more evidence for cluster-migraine? *Cephalalgia*. 2023;43(1):3331024221133383.

49. Lai TH, Fuh JL, Wang SJ. Cranial autonomic symptoms in migraine: characteristics and comparison with cluster headache. *J Neurol Neurosurg Psychiatry*. 2009;80(10):1116-1119.

50. Gormley P, Anttila V, Winsvold BS, et al. Meta-analysis of 375,000 individuals identifies 38 susceptibility loci for migraine. *Nat Genet*. 2016;48(8):856-866.