



A Meta-Analysis of the Genome-Wide Association Studies on Two Genetically Correlated Phenotypes Suggests Four New Risk Loci for Headaches

Weihua Meng^{1,2} · Parminder S. Reel² · Charvi Nangia² · Aravind Lathika Rajendrakumar² · Harry L. Hebert² · Qian Guo¹ · Mark J. Adams³ · Hua Zheng⁴ · Zen Haut Lu⁵ · 23andMe Research Team⁶ · Debashree Ray^{7,8} · Lesley A. Colvin² · Colin N. A. Palmer² · Andrew M. McIntosh³ · Blair H. Smith²

Received: 14 May 2022 / Revised: 16 September 2022 / Accepted: 21 September 2022 / Published online: 18 November 2022
© The Author(s) 2022

Abstract

Headache is one of the commonest complaints that doctors need to address in clinical settings. The genetic mechanisms of different types of headache are not well understood while it has been suggested that self-reported headache and self-reported migraine were genetically correlated. In this study, we performed a meta-analysis of genome-wide association studies (GWAS) on the self-reported headache phenotype from the UK Biobank and the self-reported migraine phenotype from the 23andMe using the Unified Score-based Association Test (metaUSAT) software for genetically correlated phenotypes ($N=397,385$). We identified 38 loci for headaches, of which 34 loci have been reported before and four loci were newly suggested. The *LDL receptor related protein 1 (LRP1)*—*Signal Transducer and Activator of Transcription 6 (STAT6)*—*Short chain Dehydrogenase/Reductase family 9C member 7 (SDR9C7)* region in chromosome 12 was the most significantly associated locus with a leading p value of 1.24×10^{-62} of rs11172113. The *One Cut homeobox 2 (ONECUT2)* gene locus in chromosome 18 was the strongest signal among the four new loci with a p value of 1.29×10^{-9} of rs673939. Our study demonstrated that the genetically correlated phenotypes of self-reported headache and self-reported migraine can be meta-analysed together in theory and in practice to boost study power to identify more variants for headaches. This study has paved way for a large GWAS meta-analysis involving cohorts of different while genetically correlated headache phenotypes.

Keywords Headache · Migraine · Unified Score-based Association Test · Correlated phenotypes · Meta-analysis · Genome-wide association study

✉ Weihua Meng
w.meng@dundee.ac.uk; weihua.meng@nottingham.edu.cn

¹ Nottingham Ningbo China Beacons of Excellence Research and Innovation Institute, University of Nottingham Ningbo China, Ningbo 315100, China

² Division of Population Health and Genomics, Ninewells Hospital and Medical School, University of Dundee, Dundee DD2 4BF, UK

³ Division of Psychiatry, Edinburgh Medical School, University of Edinburgh, Edinburgh EH10 5HF, UK

⁴ Department of Anaesthesiology, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430030, China

⁵ PAPRSB Institute of Health Sciences, Universiti Brunei Darussalam, Bandar Seri Begawan BE1410, Brunei Darussalam

⁶ 23andMe, Inc., Sunnyvale, CA 94086, USA

⁷ Department of Epidemiology, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD 21205, USA

⁸ Department of Biostatistics, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD 21205, USA

Introduction

Headache is one of the commonest symptoms that present to clinicians in general practice or in specialist neurology clinics (Fuller and Kaye 2007). Its lifetime prevalence in individuals is as high as 93% (Boardman et al. 2003). Globally, around 46% of the adult population suffers with an active headache disorder (Stovner et al. 2007). According to the current definitions of the International Headache Society, headaches can be classified into three categories: (1) primary headaches (including migraine, tension-type headache, and trigeminal autonomic cephalalgias); (2) secondary headaches (including headaches attributed to other disorders such as trauma, infection); (3) painful cranial neuropathies, other facial pain and other headaches (Headache Classification Committee of the International Headache Society 2018).

Among all types of headache, tension-type headache is the commonest form, causing over 40% of all headaches in the general population, while migraine is the most disabling type at population level, with a prevalence of around 10% of all headaches (Riesco et al. 2017). It is important to note that an individual can experience more than one type of headache at the same time (Fuller and Kaye 2007).

According to the Global Burden of Diseases 2019 study, headache disorders represent the 14th leading cause of disability-adjusted life years (DALYs) when considering all ages and both genders in 174 causes (GBD 2019 Diseases and Injuries Collaborators 2020). It was estimated in 2003 that migraine alone costs the UK over two billion pounds a year (Steiner et al. 2003). Migraine is still one of the leading causes of disability among over 300 diseases (Steiner et al. 2020).

It has been confirmed that headaches such as migraine are heritable. The single nucleotide polymorphism (SNP) based-heritabilities of migraine and self-reported headache were 0.15 and 0.21 in Caucasians, respectively (Gormley et al. 2016; Meng et al. 2018). Genome-wide association studies (GWAS) have revealed that there are significant genetic components contributing to migraine (Anttila et al. 2010, 2013; Chasman et al. 2011; Freilinger et al. 2012; Ligthart et al. 2011). A GWAS meta-analysis paper consisting of 22 cohorts by Gormley et al. identified 38 genetic loci for migraine (Gormley et al. 2016). Our study based on the UK Biobank resource also revealed 28 risk loci for self-reported headache, of which 14 loci had been previously identified by Gormley et al. (2016) and 14 loci were newly reported (Meng et al. 2018). A recent large GWAS meta-analysis on migraine has suggested 123 migraine-related loci (Hautakangas et al. 2022).

Recently, researchers have been encouraged to perform GWAS meta-analysis on genetically correlated

phenotypes, due to the increasing recognition of pleiotropy in GWAS, to boost study power to detect more genetic components (Masotti et al. 2019). Pleiotropy refers to the phenomenon where a genetic variant or a gene has non-zero effect on multiple phenotypic traits and can contribute to genetic correlations among these traits (Stearns 2010). Successful examples of recent GWAS meta-analysis studies on genetically correlated phenotypes have included education and intelligence as well as different hypertension phenotypes (Hill et al. 2019; Zhu et al. 2015).

In a previous study, we reported that the self-reported headache phenotype from the UK Biobank and the self-reported migraine phenotype from the 23andMe were genetically correlated, with a high correlation value of 0.72 ($p = 1.66 \times 10^{-68}$, standard error = 0.04) (Meng et al. 2020). Therefore, we aimed to perform a joint GWAS meta-analysis study of these two different but highly genetically correlated phenotypes with a view to replicating previously identified genetic associations and identifying new associations arising from the increased power of this approach.

Materials and Methods

Cohorts' Information

The two sets of GWAS summary statistics used in this study were from the GWAS on self-reported headache based on the UK Biobank cohort and the GWAS on self-reported migraine provided by the 23andMe (Gormley et al. 2016; Meng et al. 2018).

The definitions of self-reported headache (UK Biobank) were as follows: cases ($N = 74,461$), defined as those who self-reported headache symptoms affecting daily lives within last month using the UK Biobank online questionnaire; controls ($N = 149,312$), defined as those who did not have any pain affecting daily lives within last month (UK Biobank code 6159). The corresponding GWAS analysis was performed using a linear mixed model adjusting for age, sex, nine population principal components, genotyping arrays, and assessment centers (Meng et al. 2018). The dataset contains 9,304,965 SNPs (minor allele frequency > 0.005, imputation score > 0.1).

The definitions of self-reported migraine (23andMe) were: cases ($N = 30,465$), defined as those who self-reported a migraine history (diagnosed by doctors or self-diagnosing) using the 23andMe online questionnaire; controls ($N = 143,147$), those who self-reported having no migraine. The corresponding GWAS was performed using a linear mixed model adjusting for age, sex, and five population principal components (Gormley et al. 2016). The dataset contains 19,023,436 SNPs (containing minor allele frequency

and imputation score information for all SNPs). This dataset was also used by Gormley et al. (2016).

All of the participants in both GWAS were of European descent. In addition, as the UK Biobank cohort only recruited within the UK, while the 23andMe mainly recruited from the USA, there was little sample overlap between the two cohorts (linkage disequilibrium score regression intercept = 0.009) (Meng et al. 2020). The detailed cohorts' information and the statistical methods of the two GWAS can be found in the original papers (Gormley et al. 2016; Meng et al. 2018).

The Preprocessing of the GWAS Summary Statistics

SNPs in both datasets were coded in a forward direction and according to the GRCh37 genome build. In total, 8,500,802 SNPs with minor allele frequency > 0.005 in both datasets were extracted. To ensure the datasets could be jointly meta-analysed, these SNPs were checked for same effect alleles and flipped accordingly in R (<https://www.r-project.org/>).

The Meta-Analysis Method

The Unified Score-based Association Test (metaUSAT) is a software package for performing GWAS meta-analysis studies on genetically correlated phenotypes (Ray and Boehnke 2018) (<https://github.com/RayDebashree/metaUSAT>). The metaUSAT software applies a multivariate meta-analysis approach instead of a univariate approach of analysing each related trait separately. Unlike traditional GWAS meta-analysis on a single trait, where several sets of summary statistics on a single trait are combined into a single summary measure for that trait, the multivariate meta-analysis implemented by metaUSAT does not combine the summary statistics; instead, a joint analysis is performed using summary statistics from related traits. It is a statistical inference approach that leverages related traits to provide a p value for the test of no association of any trait with a SNP against the alternative that at least one trait is associated with the SNP. Being a complex data-adaptive approach, the metaUSAT software does not output an overall effect size (Beta) and standard error (SE) values for each SNP. The metaUSAT software is robust to the association structure of correlated traits and potential sample overlap (Ray and Boehnke 2018).

The Annotation Method

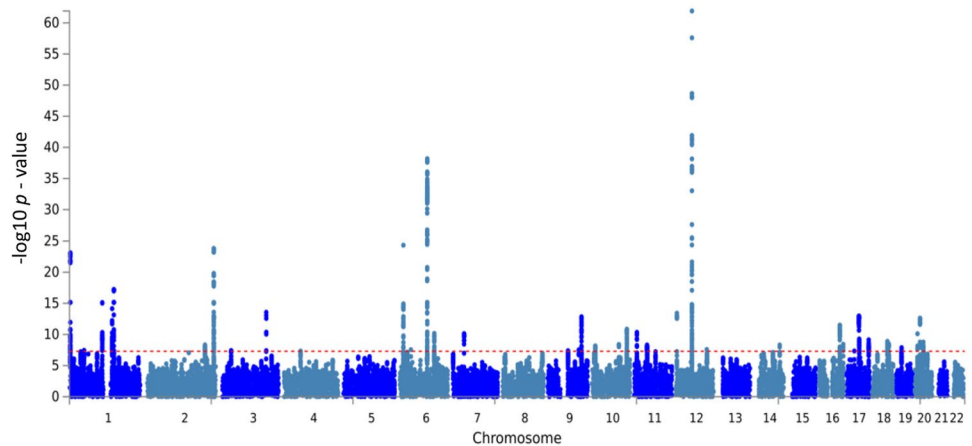
The output generated from metaUSAT was uploaded to Functional Mapping and Annotation (FUMA v1.3.6b) for SNP annotation (Watanabe et al. 2017). We used the 1000 Genome Phase 3 reference panel by default and other default values adapted by FUMA in terms of defining lead SNPs and risk loci. FUMA also generates a Manhattan plot

and a corresponding Q-Q plot for the meta-analysis result (<https://fuma.ctglab.nl/>). FUMA uses “maximum distance of linkage disequilibrium (LD) blocks to merge” (default value = 250 kb) to determine the number of associated loci and the r^2 value (default value $r^2 \geq 0.6$ to be considered as non-independent) to determine the number of independent significant SNPs. In addition, the gene-based association analysis and the gene-set analysis were performed with Multi-marker Analysis of Genomic Annotation (MAGMA v1.08), which was integrated in FUMA (de Leeuw et al. 2015). In gene-based association analysis, summary statistics of SNPs were aggregated to the level of whole genes, testing the joint association of all SNPs in the gene with the phenotype. In other words, all the SNPs were mapped to 19,436 protein coding genes if the SNPs are located within genes. In gene-set analysis, individual genes were aggregated to groups of genes sharing certain biological, functional or other characteristics. This was done to provide insight into the involvement of specific biological pathways or cellular functions in the genetic aetiology of a phenotype. A total of 10,894 gene sets were tested and a competitive test model was applied. Tissue expression analysis was obtained from the Genotype-Tissue Expression (GTEx) project (<https://www.gtexportal.org/home/>) which was also integrated in FUMA. In the tissue expression analysis, average gene-expression per tissue type was used as gene covariate to test positive relationships between gene expression in a specific tissue type and genetic associations. In addition, regional plots of the suggested new loci were generated by LocusZoom (<http://locuszoom.org/>).

Results

There were 8,500,802 common SNPs from both cohorts analysed by the metaUSAT software. FUMA reported 38 independent genetic loci across autosomal chromosomes with the *LDL Receptor related Protein 1 (LRP1)*—*Signal Transducer and Activator of Transcription 6 (STAT6)*—*Short chain Dehydrogenase/Reductase family 9C member 7 (SDR9C7)* region in chromosome 12q13.3 being the most significantly associated locus with a leading p value of 1.24×10^{-62} for rs11172113. The Four And A Half LIM Domains 5 (*FHL5*)—UFM1 Specific Ligase 1 (*UFL1*) locus in chromosome 6q16.1 was the second most significantly associated, with a p value of 6.57×10^{-39} for rs9486715. A Manhattan plot showing these loci is shown in Fig. 1. A corresponding Q-Q plot is included as shown in Supplementary Fig. 1. Among the 38 identified loci, there were 2228 SNPs that demonstrated an association with genome-wide significance, with p value $< 5 \times 10^{-8}$. Among these SNPs, 113 SNPs were considered as independent associations

Fig. 1 The Manhattan plot of the GWAS meta-analysis on headaches ($N=397,385$). The dashed red line indicates the cut-off p value of 5×10^{-8}



($r^2 < 0.6$ with any SNP within the 2228 SNPs) (Supplementary Tables 1 and 2).

Table 1 summarises the information relating to the 38 associated loci. Among these loci, 25 loci had previously been reported by Gormley et al. (2016), and nine further loci had been separately identified by Meng et al. (2018). Four of the 38 loci were newly suggested. The *One Cut homeobox 2* (*ONECUT2*) gene locus (18q21.31) was the strongest signal among these four new loci, associated with a p value of 1.29×10^{-9} for rs673939. Table 2 summarises the details of the four newly suggested loci. Regional plots of these four new loci are included in Fig. 2.

The gene-based association study identified 51 genes (cut-off p value = $0.05/19,436 = 2.57 \times 10^{-6}$) that were associated with headaches, with the *PR/SET domain 16* (*PRDM16*) gene showing the strongest association (p value of 7.10×10^{-15}). All significantly associated genes are summarised in Supplementary Table 3.

The gene-set analysis found that no specific pathway was significantly associated with headaches after Bonferroni correction (cut-off p value = $0.05/10,894 = 4.6 \times 10^{-6}$). The top 10 pathways are included in the Supplementary Table 4.

Two types of tissue analysis were performed. The tissue expression analysis on 30 general tissues revealed that both brain tissues and vascular tissues are potentially involved in the disease mechanisms (Fig. 3). The tissue expression analysis on 54 specific tissues also found 11 brain tissues with significant association ($p < 0.05/54 = 9.26 \times 10^{-4}$) (Fig. 4).

We also compared the 28 loci suggested by Meng et al. (2018) with the 38 loci suggested by this study. We noticed that 24 loci reported by Meng et al. (2018) still exist in the current study while four loci dropped out. Fourteen loci showed up in the current study which were not identified by Meng et al. (2018) (Table 3, Supplementary Tables 5 and 6).

With the specific permission obtained from the 23andMe, we have included the summary results of the GWAS on self-reported migraine using the 23andMe data in a supplementary file. We also provided a loci comparison table among

the four studies (the current study, Gormley et al. 2016; Meng et al. 2018 and 23andMe) (Supplementary Table 7).

Discussion

We performed a GWAS meta-analysis study on two highly genetically correlated phenotypes based on summary statistics from two large GWAS: self-reported headache and self-reported migraine (genetic correlation value = 0.72, $p = 1.66 \times 10^{-68}$, standard error = 0.04). This analysis identified 38 loci associated with headaches, of which 34 had been previously identified (Gormley et al. 2016; Meng et al. 2018) and four were newly suggested loci.

GWAS on complex traits have achieved great success in the past decade (Mills and Rahal 2019). Furthermore, GWAS meta-analysis on same phenotypes from multi-centers and multi-cohorts also improve statistical power to identify genetic variants which otherwise cannot be detected by a single cohort study (Evangelou and Ioannidis 2013). However, the number of cohorts in a genetic consortium will reach a bottleneck when most of the existing cohorts are already included. It becomes more difficult to include extra cohorts into the consortium to achieve a higher study power. Meanwhile, it might also be challenging to fund and allocate resources to increase sample size of cohorts in a genetic consortium. This has led to the development and use of statistical methods that leverage other aspects of a study to increase detection power. Software developed for joint meta-analysis of GWAS on existing correlated phenotypes can improve power with minimum additional resource requirement, particularly as it is now a routine requirement for GWAS summary statistics to be shared publicly after publication (Guo and Wu 2019). Specific software created for this purpose includes Software for Correlated Phenotype Analysis (META-SCOPA), Canonical Correlation Analysis (metaCCA), and Multi-Trait Analysis of GWAS (MTAG) (Cichonska et al. 2016; Mägi et al. 2017; Turley et al. 2018).

Table 1 The 38 loci generated by the GWAS meta-analysis study on self-reported headache and self-reported migraine

Rank	Gene	Lead SNP	Chr	SNP position	p-value (migraine)	p-value (headache)	Z-value (migraine)	Z-value (headache)	p-value (meta)	Reported by Gormley et al. (2016) and the locus rank
1	<i>LRP1-STAT6-SDR9C7</i>	rs11172113	12	57,527,283	4.57×10^{-23}	4.92×10^{-47}	-9.87	-14.41	1.24×10^{-62}	Yes (1)
2	<i>FHL5-UFL1</i>	rs9486715	6	97,059,769	8.69×10^{-13}	2.58×10^{-31}	7.16	11.64	6.57×10^{-39}	Yes (3)
3	<i>PHACTR1</i>	rs9349379	6	12,903,957	5.20×10^{-14}	3.59×10^{-15}	-7.51	-7.87	4.62×10^{-25}	Yes (6)
4	<i>TRPM8-HJURP</i>	rs2362290	2	234,825,369	8.89×10^{-13}	6.25×10^{-16}	-7.11	-8.08	1.49×10^{-24}	Yes (5)
5	<i>PRDM16</i>	rs10218452	1	3,075,597	1.49×10^{-18}	4.55×10^{-9}	8.83	5.86	7.70×10^{-24}	Yes (2)
6	<i>MEF2D</i>	rs2282286	1	156,452,870	5.78×10^{-10}	2.34×10^{-11}	6.21	6.68	6.00×10^{-18}	Yes (7)
7	<i>Intergenic (Near TSPAN2-NGF)</i>	rs12134493	1	115,677,946	1.03×10^{-9}	1.93×10^{-10}	6.14	6.37	7.48×10^{-16}	Yes (4)
8	<i>Intergenic (Near ADAMTSL4-ECM1)</i>	rs6693567	1	150,510,660	8.11×10^{-6}	3.77×10^{-12}	4.47	6.95	7.16×10^{-15}	Yes (28)
9	<i>Intergenic (Near GPRI49)</i>	rs34097149	3	154,263,175	6.93×10^{-7}	1.33×10^{-10}	-4.89	-6.42	2.80×10^{-14}	Yes (19)
10	<i>Intergenic (Near FGF6)</i>	rs10774231	12	4,515,374	4.18×10^{-8}	4.55×10^{-9}	-5.49	-5.86	3.82×10^{-14}	Yes (9)
11	<i>LINC02210-CRHR1-MAPT</i>	rs117368197	17	43,715,924	7.12×10^{-1}	2.36×10^{-14}	-0.37	7.63	1.08×10^{-13}	No, reported by Meng et al. (2018)
12	<i>ASTN2</i>	rs10759844	9	119,249,326	3.65×10^{-6}	2.24×10^{-10}	-4.62	-6.34	1.44×10^{-13}	Yes (13)
13	<i>SLC24A3</i>	rs4814864	20	19,469,817	2.36×10^{-12}	8.88×10^{-4}	7.03	3.32	2.39×10^{-13}	Yes (8)
14	<i>CFDP1</i>	rs11149826	16	75,435,140	8.82×10^{-7}	2.85×10^{-8}	-4.92	-5.55	3.20×10^{-12}	Yes (16)
15	<i>PLEKHAI (ARMS2-HTRA1)</i>	rs78438709	10	124,201,071	9.19×10^{-3}	8.57×10^{-12}	-2.59	-6.83	1.41×10^{-11}	Yes (35)
16	<i>MRVII</i>	rs4909945	11	10,673,739	3.93×10^{-5}	5.13×10^{-9}	-4.11	-5.84	4.58×10^{-11}	Yes (14)
17	<i>Intergenic (Near GJAI)</i>	rs9490318	6	121,860,207	2.27×10^{-3}	1.72×10^{-10}	3.06	6.38	6.84×10^{-11}	Yes (25)
18	<i>SUGCT</i>	rs77410344	7	40,410,924	8.45×10^{-8}	4.84×10^{-6}	5.39	4.57	7.44×10^{-11}	Yes (10)
19	<i>Intergenic (Near JAG1)</i>	rs6040095	20	10,680,221	7.98×10^{-8}	4.67×10^{-6}	5.38	4.58	7.78×10^{-11}	Yes (20)
20	<i>RNF213</i>	rs12943001	17	78,238,645	7.93×10^{-8}	5.37×10^{-5}	-5.36	-4.04	7.35×10^{-10}	Yes (17)
21	<i>ONECUT2</i>	rs673939	18	55,153,266	4.88×10^{-5}	1.63×10^{-7}	-4.05	-5.24	1.29×10^{-9}	No (new locus)
22	<i>NOLAL</i>	rs159058	20	31,108,108	3.51×10^{-5}	3.18×10^{-7}	-4.14	-5.11	1.68×10^{-9}	No, reported by Meng et al. (2018)
23	<i>Intergenic (Near ZCCHC2)</i>	rs4941139	18	60,162,791	7.73×10^{-1}	2.22×10^{-10}	0.29	6.35	3.38×10^{-9}	No, reported by Meng et al. (2018)
24	<i>Intergenic (Near ZCCHC14)</i>	rs8052831	16	87,578,039	1.15×10^{-8}	2.86×10^{-3}	5.72	2.98	3.80×10^{-9}	Yes (22)
25	<i>PLCE1</i>	rs3891783	10	96,015,793	3.85×10^{-7}	7.53×10^{-5}	-5.07	-3.96	4.10×10^{-9}	Yes (11)
26	<i>CARF</i>	rs72928613	2	203,839,628	2.58×10^{-5}	1.23×10^{-6}	-4.19	-4.85	4.80×10^{-9}	Yes (34)
27	<i>Intergenic (Near ITPK1)</i>	rs28540738	14	93,591,673	3.56×10^{-6}	9.60×10^{-6}	-4.63	-4.43	4.95×10^{-9}	Yes (27)
28	<i>CHRM4</i>	rs2067482	11	46,406,767	9.78×10^{-4}	3.93×10^{-8}	-3.29	-5.49	4.99×10^{-9}	No, reported by Meng et al. (2018)
29	<i>CAMK1D</i>	rs10752269	10	12,692,902	4.20×10^{-1}	4.87×10^{-10}	0.81	6.22	7.11×10^{-9}	No, reported by Meng et al. (2018)
30	<i>MAU2</i>	rs34858588	19	19,457,235	3.53×10^{-4}	3.40×10^{-7}	3.59	5.10	1.33×10^{-8}	No (new locus)
31	<i>MYO1H</i>	rs6606710	12	109,848,903	3.86×10^{-2}	8.44×10^{-9}	2.07	5.76	2.60×10^{-8}	No, reported by Meng et al. (2018)

Table 1 (continued)

Rank	Gene	Lead SNP	Chr	SNP position	p-value (migraine)	p-value (headache)	Z-value (migraine)	Z-value (headache)	p-value (meta) Reported by Gormley et al. (2016) and the locus rank
32	<i>Intergenic (Near KCNK17)</i>	rs72854120	6	39,248,533	6.22×10^{-3}	3.99×10^{-8}	- 2.68	- 5.49	2.81×10^{-8} No (new locus)
33	<i>ZNF462</i>	rs2134063	9	109,695,139	1.60×10^{-3}	1.85×10^{-7}	3.16	5.21	2.98×10^{-8} No (new locus)
34	<i>Intergenic (Near CDKN2C)</i>	rs7555006	1	51,480,258	9.56×10^{-2}	5.87×10^{-9}	1.67	5.82	3.59×10^{-8} No, reported by Meng et al. (2018)
35	<i>LOC101927995 (Near TGFBR2)</i>	rs6791480	3	30,480,559	7.70×10^{-5}	4.60×10^{-6}	3.96	4.58	3.81×10^{-8} Yes (26)
36	<i>TJP2</i>	rs7850547	9	71,747,208	5.24×10^{-4}	7.50×10^{-7}	- 3.47	- 4.95	4.09×10^{-8} No, reported by Meng et al. (2018)
37	<i>NUFIP2</i>	rs8614	17	27,588,806	2.36×10^{-1}	4.25×10^{-9}	1.19	5.87	4.56×10^{-8} No, reported by Meng et al. (2018)
38	<i>Intergenic (Near REST-SPINK2)</i>	rs781669	4	57,819,794	4.59×10^{-3}	1.09×10^{-7}	2.83	5.31	4.67×10^{-8} Yes (21)

Chr: chromosome

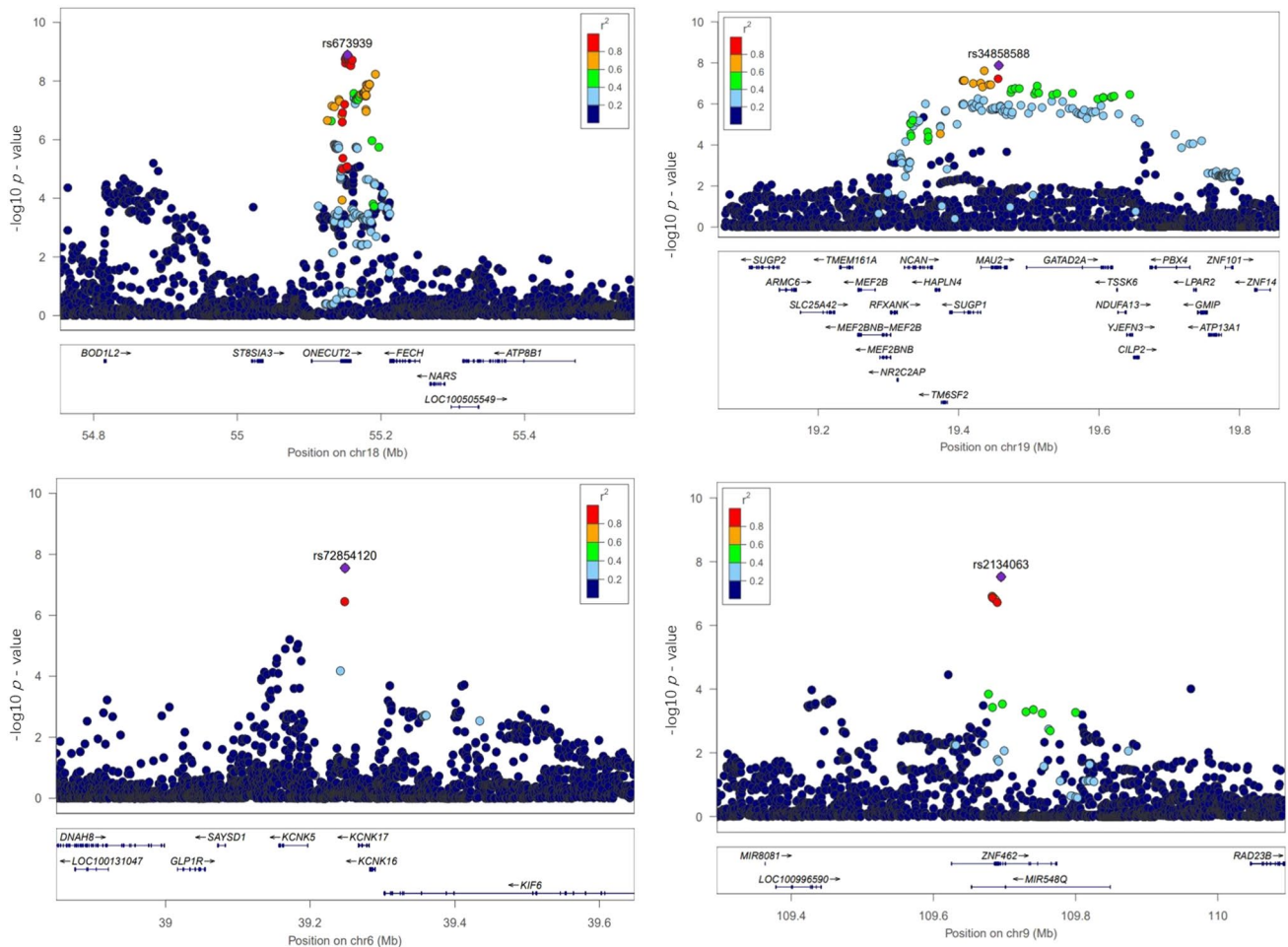
The Z values (ratio of effect size to standard error) stand for the specific SNP effect contribution from each cohort
The locus rank reported in final column is based on the study by Gormley et al. (2016)

In this study, we used a recently developed software called metaUSAT, whose properties have been illustrated using simulated data, and the Type 2 Diabetes Genetic Exploration by Next-generation sequencing in multi-Ethnic Samples (T2D-GENES) and the METabolic Syndrome In Men (METSIM) datasets (Fuchsberger et al. 2016; Stancáková et al. 2009). The software does not generate an overall Beta or standard error for each SNP but it calculates a Z value representing the contribution of each SNP from each cohort or each phenotype. As with Beta values, if Z values are both positive or both negative, then the direction of the SNP effect on the traits is the same. In Table 2, we see that the most strongly associated SNPs in the top nine loci all have the same direction of effect. However, the lead SNP in the *Long Intergenic Non-Protein Coding RNA 2210 (LINC02210)—Corticotropin Releasing Hormone Receptor 1 (CRHR1)—Microtubule Associated Protein Tau (MAPT)* locus (ranked 10th), although significantly associated, showed a different direction of effect in the two datasets. This might indicate that the role of this locus might be different in these two phenotypes. It is possible that a locus could contribute to non-migraine type headaches while contributing minimally to migraine. However, this assumption definitely needs further lab evidence. Comparing the Z-values could be a novel way to differentiate the genetic impact of certain SNPs in genetically correlated yet different phenotypes (Ray and Chatterjee 2020).

Consistent with previous studies, the most significantly associated locus in the *LRP1-STAT6-SDR9C7* region was the strongest locus identified in the meta-analysis ($p = 1.24 \times 10^{-62}$ for rs11172113) (Gormley et al. 2016; Meng et al. 2018). This locus, ranging from 57,244,168 to 57,629,608 in chromosome 12, contained 122 SNPs associated with genome-wide significance, among the 166 SNPs in the output dataset. The *LRP1* gene has been well established as a migraine gene (Anttila et al. 2010, 2013). One theory about its possible link with migraine is that the LRP1 protein interacts with the glutamate receptors on neurons while the pathophysiology of migraine has been suggested to be related with the glutamate homeostasis (Andreou and Goadsby 2009). The gene-based association study revealed that the *PRDM16* gene was the most significantly associated gene, followed by *CRHR1*, *MAPT*, and *KAT8 regulatory NSL complex subunit 1 (KANSL1)* (Supplementary Table 3). Through the tissue expression analysis, both brain and vascular tissues were indicated as being involved in the mechanisms of headaches. Gormley et al. (2016) found that vascular factors played a main role in migraine, while in our UK Biobank study, we found that neural tissues were major factors in self-reported headache. We, therefore, deduce that for other types of headaches, such as tension-type headache which produces most headaches in the general population, the role of neural tissue is likely to be greater than that of

Table 2 The summary statistics of the four newly suggested loci of headaches

New Locus (Chr)	Lead SNP	Effective allele	23andMe (migraine)		UK Biobank (headache)		Joint meta-analysis p value
			Beta (SE)	p value	Beta (SE)	p value	
<i>ONECUT2</i> (18q21.31)	rs673939	C	-0.04 (0.01)	4.88×10^{-5}	-0.0078 (0.001)	1.63×10^{-7}	1.29×10^{-9}
<i>MAU2</i> (19p13.11)	rs34858588	G	0.067 (0.019)	3.53×10^{-4}	0.013 (0.003)	3.40×10^{-7}	1.33×10^{-8}
Intergenic (Near <i>KCNK17</i> , 6p21.2)	rs72854120	C	-0.23 (0.08)	6.22×10^{-3}	-0.047 (0.009)	3.99×10^{-8}	2.81×10^{-8}
<i>ZNF462</i> (9q31.2)	rs2134063	G	0.04 (0.01)	1.60×10^{-3}	0.0099 (0.002)	1.85×10^{-7}	2.98×10^{-8}

**Fig. 2** The regional plots of the four new loci. Up left: the *ONECUT2* region; Up right: the *MAU2* region; Bottom left: the Intergenic region (Near *KCNK17*, 6p21.2); Bottom right: the *ZNF462* region

vascular factors. The cortex has demonstrated the strongest link with headache in the tissue expression analysis. It has been reported that migraine is associated with the changes in cortex functions (Barbanti et al. 2020; Charles and Brennan 2010).

In this study, we suggested four new loci which have not previously been reported to be associated with headaches. The *ONECUT2* gene region was the most strongly

associated among these four loci. The Z values from the headache study ($Z = -5.24$) and the migraine study ($Z = -4.05$) were in the same direction. The original p values of the top SNP (rs673939) in this region were found to be 9.00×10^{-8} by Meng et al. (2018) and 4.88×10^{-5} in the 23andMe migraine dataset (Gormley et al. 2016). The *ONECUT2* gene, also termed *OC-2*, is a newly discovered member of the ONECUT transcription factor family

Fig. 3 Tissue expression results on 30 specific tissue types by GTEx in the FUMA. The dashed line shows the cut-off p value for significance with Bonferroni adjustment for multiple hypothesis testing

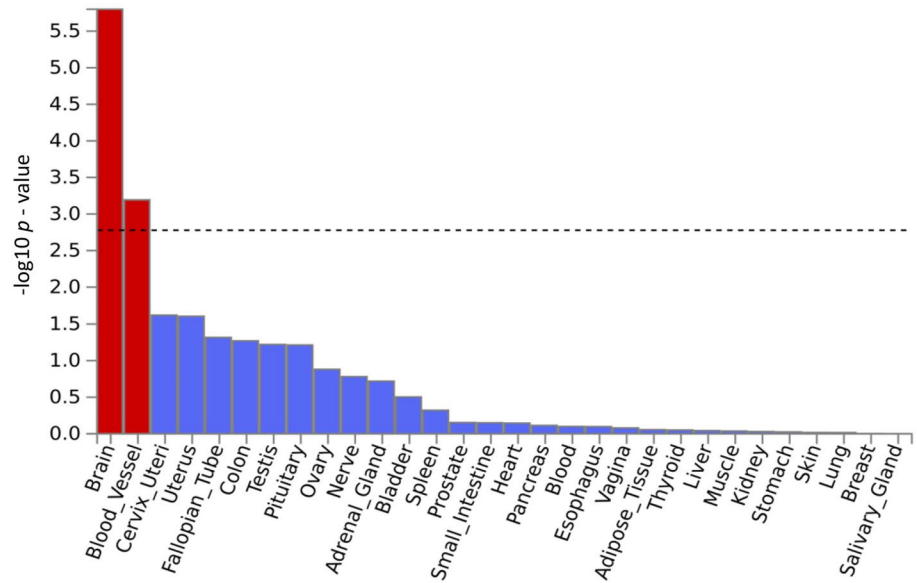
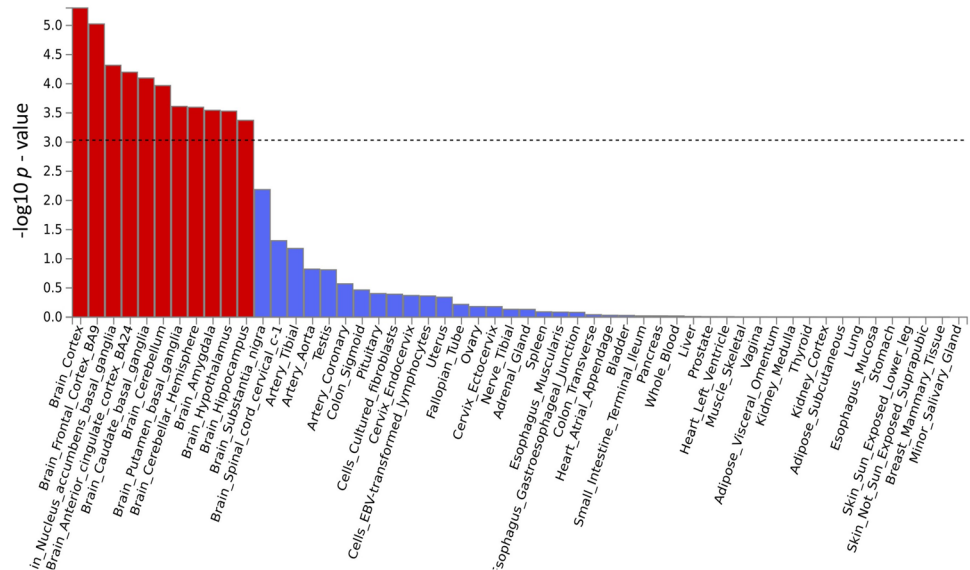


Fig. 4 Tissue expression results on 53 specific tissue types by GTEx in the FUMA. The dashed line shows the cut-off p value for significance with Bonferroni adjustment for multiple hypothesis testing



(Jacquemin et al. 1999). *ONECUT2* can widely regulate the protein expression associated with cell proliferation, migration, adhesion, and differentiation, thus being involved in the regulation of the development of an organism (Yu et al. 2020). It has been well reported for its associations with multiple cancers. Although we do not know why it is statistically associated with headaches, the gene is expressed in the brain (<https://www.ncbi.nlm.nih.gov/gene/9480>). It is not uncommon that SNPs can be associated with multiple phenotypes which seem completely unrelated (Solovieff et al. 2013). The *MAU2* sister chromatid cohesion factor (*MAU2*) is a protein-coding gene, which plays an important role when cohesions (chromosome-associated multi-subunit protein complex) try to bind to DNA to carry out a large

spectrum of chromatin-related functions, including sister chromatid cohesion, DNA repair, transcriptional regulation, and three-dimensional organization of chromatin (Zhu and Wang 2019). Mutations of *MAU2* have been linked with a rare disorder of Cornelia de Lange Syndrome (Parenti et al. 2020). The *Potassium two pore domain channel subfamily K member 17 (KCNK17)* is the nearest gene to the leading SNP of rs72854120 in the third new locus. Variants of this gene have been reported to be associated with ischaemic stroke, cerebral hemorrhage, and arrhythmia (Friedrich et al. 2014; He et al. 2014). The protein products of *Zinc Finger protein 462 (ZNF462)* have shown important roles in embryonic development in animal models (Cosemans et al. 2018). Variants of this gene have been reported to contribute to

Table 3 Four loci reported by Meng et al. (2018) while dropped out in the current study

Loci reported by Meng et al	Gene	Chromosome	Lead SNP In Meng et al	SNP position	Still a locus in the current study?
1	<i>LRP1-STAT6-SDR9C7</i>	12	rs11172113	57,527,283	Yes
2	<i>FHL5-UFL1</i>	6	rs9486715	97,059,769	Yes
3	<i>TRPM8-HJURP</i>	2	rs2362290	234,825,369	Yes
4	<i>PHACTR1</i>	6	rs9349379	12,903,957	Yes
5	<i>LINC02210-CRHRI-MAPT</i>	17	rs77804065	43,810,896	Yes
6	<i>Intergenic</i>	1	rs12740679	150,262,270	Yes
7	<i>Intergenic</i>	10	rs78438709	124,201,071	Yes
8	<i>MEF2D</i>	1	rs1050316	156,434,703	Yes
9	<i>ASTN2</i>	9	rs17220352	119,248,059	Yes
10	<i>Intergenic</i>	3	rs34097149	154,263,175	Yes
11	<i>Intergenic</i>	6	rs9490318	121,860,207	Yes
12	<i>Intergenic</i>	1	rs12134493	115,677,946	Yes
13	<i>Intergenic</i>	18	rs4941139	60,162,791	Yes
14	<i>CAMK1D</i>	10	rs2895526	12,726,061	Yes
15	<i>PRDM16</i>	1	rs56304645	3,085,186	Yes
16	<i>NUFIP2</i>	17	rs8614	27,588,806	Yes
17	<i>Intergenic</i>	12	rs10774231	4,515,374	Yes
18	<i>MRVII</i>	11	rs4909945	10,673,739	Yes
19	<i>BTN2A2</i>	6	rs2072806	26,385,093	No
20	<i>Intergenic</i>	1	rs7555006	51,480,258	Yes
21	<i>MYO1H</i>	12	rs6606710	109,848,903	Yes
22	<i>IFT81</i>	12	rs7300001	110,581,731	No
23	<i>NOLAL</i>	20	rs1555132	31,046,567	Yes
24	<i>CFDP1</i>	16	rs1011121	75,325,933	Yes
25	<i>PTBP2</i>	1	rs3748784	97,187,174	No
26	<i>FXN</i>	9	rs4596713	71,699,216	Yes
27	<i>ATG13</i>	11	rs56349329	46,695,483	Yes
28	<i>MACF1</i>	1	rs2036465	39,575,982	No

craniofacial and neurodevelopmental abnormalities (Weiss et al. 2017). It is worth noting that the Z values of each leading SNP in the four new loci were all in the same direction in each of the two cohorts. We also noted that Meng et al. (2018) suggested 28 loci associated with self-reported headache; with 24 of these loci still found in the current study while four dropped out (Table 3), which means we found 14 newly suggested loci when performing the meta-analysis with the 23andMe data in the current study. One reason for dropping the four previously suggested loci might be that the studies are based on the genetically correlated samples but are examining different phenotypes. Of note is that the *p* values of the dropped loci in Meng et al. (2018) could be considered as of marginal GWAS significance, and further work needs to be done to explore their relevance.

Notably, we had particular advantages in this study that would be important to consider if applying these methods to other phenotypes and or samples. One was that the two phenotypes we chose (self-reported headache and migraine) are

highly genetically correlated. However, the ability of a study to detect more new variants would be reduced if the two phenotypes are not so highly correlated. Second, our two cohorts were both of mainly European descent and with minimum sample overlap; therefore, we avoided some negative impact (such as increasing type I and type II errors) which could be caused by these factors in the study. There are limitations associated with this approach, being novel in its application. For example, although we successfully addressed phenotypes and datasets which are highly genetically correlated, there are insufficient published studies to allow us to determine the strength of correlation which is required to allow this approach for future studies. This will require our approach to be replicated with other phenotypes and datasets, followed by formal statistical appraisal of the results. Similarly, although it was not directly relevant in our study, consideration will need to be given to applying this approach when there is sample overlapping. It is worth mentioning that the genetic correlation between diagnosed

cluster headache phenotype from a mixed UK and Swedish cohort and self-reported headache from the UK Biobank was 0.50 (O'Connor et al. 2021). At the time we conducted this analysis, the four novel loci we suggested were previously unknown. During the preparation stage of this paper, a large GWAS meta-analysis on migraine has been published suggesting 123 migraine-related loci (Hautakangas et al. 2022). Among our four newly suggested loci, three loci have been reported (*ONECUT2*, *MAU2* and *ZNF462*). Although that study was specifically addressing migraine, rather than the more general headache phenotype that we addressed, this overlap between the new loci suggested in the two studies both helps to confirm the findings of Hautakangas et al. (2022), and the success of our methodological approach. The loci near *KCNK17* on chromosome 6p21.2 was not clearly reported and its nearest loci in the paper was *potassium two pore domain channel subfamily K member 5 (KCNK5)* which is 60 kb away from *KCNK17*. A recent study on rare variants of migraine showed that significant cis-expression quantitative trait loci (eQTL) in the polycomb response elements (regulatory sites that mediate the silencing of homeotic and other genes) mapped to the *KCNK17* (Techlo et al. 2020).

It is also important to note that the control definitions in the two GWAS datasets were different. The controls used in the UK Biobank self-reported headache phenotype reported no pain within previous month, while the controls used in the 23andMe self-reported migraine phenotype could have had pain in body sites other than the head. This mean that genes identified in the UK Biobank cohort may not be specific to headache, but could be more generally associated with pain. Please also note that self-reported headache or migraine phenotypes are different from clinical ascertained phenotypes. Although our results on self-reported phenotypes will provide reference values to other researchers, there could be potential biases because of this and the results should be interpreted with caution in a clinical setting.

Conclusion

In summary, our study suggested four new genetic loci which are associated with self-reported headaches and/or migraine, and shed further light on their potential mechanisms. Further research could attempt a meta-analysis study on GWAS of different types of primary headaches (on the condition that they are reasonably genetically correlated) to identify further genetic components.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s43657-022-00078-7>.

Acknowledgements The current study was conducted under approved UK Biobank data application number 4844. We would like to thank all participants of the UK Biobank cohort who have provided necessary

genetic and phenotypic information. We would like to thank the research participants and employees of 23andMe for making this work possible.

23andMe Consortium: The following members of the 23andMe Research Team contributed to this study: Stella Aslibekyan, Adam Auton, Elizabeth Babalola, Robert K. Bell, Jessica Bielenberg, Katarzyna Bryc, Emily Bullis, Daniella Coker, Gabriel Cuellar Partida, Devika Dhamija, Sayantan Das, Sarah L. Elson, Teresa Filshtein, Kipper Fletez-Brant, Pierre Fontanillas, Will Freyman, Anna Faaborg, Shirin T. Fuller, Pooja M. Gandhi, Karl Heilbron, Barry Hicks, Ethan M. Jewett, Katelyn Kukar, Keng-Han Lin, Maya Lowe, Jey C. McCreight, Matthew H. McIntyre, Steven J. Micheletti, Meghan E. Moreno, Joanna L. Mountain, Priyanka Nandakumar, Elizabeth S. Noblin, Jared O'Connell, Yunru Huang, Aaron A. Petrakovitz, Vanessa Lane, Aaron Petrakovitz, Joanne S. Kim, G. David Poznik, Morgan Schumacher, Anjali J. Shastri, Janie F. Shelton, Jingchunzi Shi, Suyash Shringarpure, Vinh Tran, Joyce Y. Tung, Xin Wang, Wei Wang, Catherine H. Weldon, Peter Wilton, Alejandro Hernandez, Corinna Wong, Christophe Toukam Tchakouté.

Authors' contributions WM organised project, drafted the paper and contributed to the analysis. PSR, CN and ALR performed the meta-analysis. HLH contributed to Table 1. QG contributed to data formatting. MJA performed the UK Biobank GWAS analysis. HZ and ZHL contributed to discussion parts. The 23andMe Research Team provided the GWAS summary statistics of the 23andMe cohort. DR supervised the usage of the metaUSAT software and provided comments. LAC and CNAP provided comments to the paper. AMM and BHS organised the project and provided comments.

Funding This study was mainly funded by the Wellcome Trust Strategic Award "Stratifying Resilience and Depression Longitudinally" (STRADL) with Reference Number 104036/Z/14/Z.

Data availability The GWAS meta-analysis summary statistics can be downloaded from <https://app.box.com/s/gm2qkf17hc9w1fc5ymvi fgbxz0httvhs>.

The FUMA results can be viewed from <https://fuma.ctglab.nl/browse/334>.

Data from 23andMe were obtained under a data transfer agreement. Further information about obtaining access to the 23andMe Inc. summary statistics is available from: <https://research.23andme.com/collaborate/>. Any other data relevant to the study that are not included in the article or its supplementary materials are available from the authors upon reasonable request.

Declarations

Conflict of interest The employees of 23andMe/23andMe Research Team hold stock in the company. The other authors declare that they have no conflict of interest.

Ethical approval This study was approved by the Ethics Committee of the University of Dundee, School of Medicine.

Consent to participate Informed consent was obtained by the UK Biobank and 23andMe cohorts from all individual participants included in the study.

Consent to publish All authors have consent for publication.

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 licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence 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 licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Andreou AP, Goadsby PJ (2009) Therapeutic potential of novel glutamate receptor antagonists in migraine. *Expert Opin Investig Drugs* 18(6):789–803. <https://doi.org/10.1517/13543780902913792>
- Anttila V, Stefansson H, Kallela M, Todt U, Terwindt GM, Calafato MS, Nyholt DR, Dimas AS, Freilinger T, Müller-Myhsok B, Artto V, Inouye M, Alakurtti K, Kaunisto MA, Hämäläinen E, de Vries B, Stam AH, Weller CM, Heinze A, Heinze-Kuhn K, Goebel I, Borck G, Göbel H, Steinberg S, Wolf C, Björnsson A, Gudmundsson G, Kirchmann M, Hauge A, Werge T, Schoenen J, Eriksson JG, Hagen K, Stovner L, Wichmann HE, Meitinger T, Alexander M, Moebus S, Schreiber S, Aulchenko YS, Breteler MM, Uitterlinden AG, Hofman A, van Duijn CM, Tikka-Kleemola P, Vepsäläinen S, Lucae S, Tozzi F, Muglia P, Barrett J, Kaprio J, Färkkilä M, Peltonen L, Stefansson K, Zwart JA, Ferrari MD, Olesen J, Daly M, Wessman M, van den Maagdenberg AM, Dichgans M, Kubisch C, Dermitzakis ET, Frants RR, Palotie A (2010) Genome-wide association study of migraine implicates a common susceptibility variant on 8q22.1. *Nat Genet* 42(10):869–873. <https://doi.org/10.1038/ng.652>
- Anttila V, Winsvold BS, Gormley P, Kurth T, Bettella F, McMahon G, Kallela M, Malik R, de Vries B, Terwindt G, Medland SE, Todt U, McArdle WL, Quaye L, Koiranen M, Ikram MA, Lehtimäki T, Stam AH, Ligthart L, Wedenoja J, Dunham I, Neale BM, Palta P, Hamalainen E, Schürks M, Rose LM, Buring JE, Ridker PM, Steinberg S, Stefansson H, Jakobsson F, Lawlor DA, Evans DM, Ring SM, Färkkilä M, Artto V, Kaunisto MA, Freilinger T, Schoenen J, Frants RR, Pelzer N, Weller CM, Zielman R, Heath AC, Madden PAF, Montgomery GW, Martin NG, Borck G, Göbel H, Heinze A, Heinze-Kuhn K, Williams FMK, Hartikainen AL, Pouta A, van den Ende J, Uitterlinden AG, Hofman A, Amin N, Hottenga JJ, Vink JM, Heikkilä K, Alexander M, Muller-Myhsok B, Schreiber S, Meitinger T, Wichmann HE, Aromaa A, Eriksson JG, Traynor B, Tratzumi D, Rossin E, Lage K, Jacobs SBR, Gibbs JR, Birney E, Kaprio J, Penninx BW, Boomsma DI, van Duijn C, Raitakari O, Jarvelin MR, Zwart JA, Cherkas L, Strachan DP, Kubisch C, Ferrari MD, van den Maagdenberg A, Dichgans M, Wessman M, Smith GD, Stefansson K, Daly MJ, Nyholt DR, Chasman D, Palotie A (2013) Genome-wide meta-analysis identifies new susceptibility loci for migraine. *Nat Genet* 45(8):912–917. <https://doi.org/10.1038/ng.2676>
- Barbanti P, Brighina F, Egeo G, Di Stefano V, Silvestro M, Russo A (2020) Migraine as a cortical brain disorder. *Headache* 60(9):2103–2114. <https://doi.org/10.1111/head.13935>
- Boardman HF, Thomas E, Croft PR, Millson DS (2003) Epidemiology of headache in an English district. *Cephalalgia* 23(2):129–137. <https://doi.org/10.1046/j.1468-2982.2003.00468.x>
- Charles A, Brennan KC (2010) The neurobiology of migraine. *Handb Clin Neurol* 97:99–108. [https://doi.org/10.1016/s0072-9752\(10\)97007-3](https://doi.org/10.1016/s0072-9752(10)97007-3)
- Chasman DI, Schürks M, Anttila V, de Vries B, Schminke U, Launer LJ, Terwindt GM, van den Maagdenberg AM, Fendrich K, Völzke H, Ernst F, Griffiths LR, Buring JE, Kallela M, Freilinger T, Kubisch C, Ridker PM, Palotie A, Ferrari MD, Hoffmann W, Zee RY, Kurth T (2011) Genome-wide association study reveals three susceptibility loci for common migraine in the general population. *Nat Genet* 43(7):695–698. <https://doi.org/10.1038/ng.856>
- Cichonska A, Rousu J, Marttinen P, Kangas AJ, Soinen P, Lehtimäki T, Raitakari OT, Jarvelin MR, Salomaa V, Ala-Korpela M, Ripatti S, Pirinen M (2016) metaCCA: summary statistics-based multivariate meta-analysis of genome-wide association studies using canonical correlation analysis. *Bioinformatics* 32(13):1981–1989. <https://doi.org/10.1093/bioinformatics/btw052>
- Cosemans N, Vandenhove L, Maljaars J, Van Esch H, Devriendt K, Baldwin A, Fryns JP, Noens I, Peeters H (2018) ZNF462 and KLF12 are disrupted by a de novo translocation in a patient with syndromic intellectual disability and autism spectrum disorder. *Eur J Med Genet* 61(7):376–383. <https://doi.org/10.1016/j.ejmg.2018.02.002>
- de Leeuw CA, Mooij JM, Heskes T, Posthuma D (2015) MAGMA: generalized gene-set analysis of GWAS data. *PLoS Comput Biol* 11(4):e1004219. <https://doi.org/10.1371/journal.pcbi.1004219>
- Evangelou E, Ioannidis JP (2013) Meta-analysis methods for genome-wide association studies and beyond. *Nat Rev Genet* 14(6):379–389. <https://doi.org/10.1038/nrg3472>
- Freilinger T, Anttila V, de Vries B, Malik R, Kallela M, Terwindt GM, Pozo-Rosich P, Winsvold B, Nyholt DR, van Oosterhout WP, Artto V, Todt U, Hämäläinen E, Fernández-Morales J, Louter MA, Kaunisto MA, Schoenen J, Raitakari O, Lehtimäki T, Vila-Pueyo M, Göbel H, Wichmann E, Sintas C, Uitterlinden AG, Hofman A, Rivadeneira F, Heinze A, Tronvik E, van Duijn CM, Kaprio J, Cormand B, Wessman M, Frants RR, Meitinger T, Müller-Myhsok B, Zwart JA, Färkkilä M, Macaya A, Ferrari MD, Kubisch C, Palotie A, Dichgans M, van den Maagdenberg AM (2012) Genome-wide association analysis identifies susceptibility loci for migraine without aura. *Nat Genet* 44(7):777–782. <https://doi.org/10.1038/ng.2307>
- Friedrich C, Rinné S, Zumhagen S, Kiper AK, Silbernagel N, Netter MF, Stallmeyer B, Schulze-Bahr E, Decher N (2014) Gain-of-function mutation in TASK-4 channels and severe cardiac conduction disorder. *EMBO Mol Med* 6(7):937–951. <https://doi.org/10.15252/emmm.201303783>
- Fuchsberger C, Flannick J, Teslovich TM, Mahajan A, Agarwala V, Gaulton KJ, Ma C, Fontanillas P, Moutsianas L, McCarthy DJ, Rivas MA, Perry JRB, Sim X, Blackwell TW, Robertson NR, Rayner NW, Cingolani P, Locke AE, Tajas JF, Highland HM, Dupuis J, Chines PS, Lindgren CM, Hartl C, Jackson AU, Chen H, Huyghe JR, van de Bunt M, Pearson RD, Kumar A, Müller-Nurasyid M, Grarup N, Stringham HM, Gamazon ER, Lee J, Chen Y, Scott RA, Below JE, Chen P, Huang J, Go MJ, Stitzel ML, Pasko D, Parker SCJ, Varga TV, Green T, Beer NL, Day-Williams AG, Ferreira T, Fingerlin T, Horikoshi M, Hu C, Huh I, Ikram MK, Kim BJ, Kim Y, Kim YJ, Kwon MS, Lee J, Lee S, Lin KH, Maxwell TJ, Nagai Y, Wang X, Welch RP, Yoon J, Zhang W, Barzilay N, Voight BF, Han BG, Jenkinson CP, Kuulasmaa T, Kuusisto J, Manning A, Ng MCY, Palmer ND, Balkau B, Stančáková A, Abboud HE, Boeing H, Giedraitis V, Prabhakaran D, Gottesman O, Scott J, Carey J, Kwan P, Grant G, Smith JD, Neale BM, Purcell S, Butterworth AS, Howson JMM, Lee HM, Lu Y, Kwak SH, Zhao W, Danesh J, Lam VKL, Park KS, Saleheen D, So WY, Tam CHT, Afzal U, Aguilar D, Arya R, Aung T, Chan E, Navarro C, Cheng CY, Palli D, Correa A, Curran JE, Rybin D, Farook VS, Fowler SP, Freedman BI, Griswold M, Hale DE, Hicks PJ, Khor CC, Kumar S, Lehne B, Thuillier D, Lim WY, Liu J, van der Schouw YT, Loh M, Musani SK, Puppala S, Scott WR, Yengo L, Tan ST, Taylor HA Jr, Thameem F, Wilson G Sr, Wong TY, Njølstad PR, Levy JC, Mangino M, Bonnycastle LL, Schwarzmayr T, Fadista J, Surdulescu GL, Herder C, Groves CJ, Wieland T,

- Bork-Jensen J, Brandslund I, Christensen C, Koistinen HA, Doney ASF, Kinnunen L, Esko T, Farmer AJ, Hakaste L, Hodgkiss D, Kravic J, Lyssenko V, Hollensted M, Jørgensen ME, Jørgensen T, Ladenvall C, Justesen JM, Käräjämäki A, Kriebel J, Rathmann W, Lannfelt L, Lauritzen T, Narisu N, Linneberg A, Melander O, Milani L, Neville M, Orho-Melander M, Qi L, Qi Q, Roden M, Rolandsson O, Swift A, Rosengren AH, Stirrups K, Wood AR, Mihailov E, Blancher C, Carneiro MO, Maguire J, Poplin R, Shakir K, Fennell T, DePristo M, de Angelis MH, Deloukas P, Gjesing AP, Jun G, Nilsson P, Murphy J, Onofrio R, Thorand B, Hansen T, Meisinger C, Hu FB, Isomaa B, Karpe F, Liang L, Peters A, Huth C, O'Rahilly SP, Palmer CNA, Pedersen O, Rauramaa R, Tuomilehto J, Salomaa V, Watanabe RM, Syvänen AC, Bergman RN, Bharadwaj D, Bottinger EP, Cho YS, Chandak GR, Chan JCN, Chia KS, Daly MJ, Ebrahim SB, Langenberg C, Elliott P, Jablonski KA, Lehman DM, Jia W, Ma RCW, Pollin TI, Sandhu M, Tandon N, Froguel P, Barroso I, Teo YY, Zeggini E, Loos RJF, Small KS, Ried JS, DeFronzo RA, Grallert H, Glaser B, Metspalu A, Wareham NJ, Walker M, Banks E, Gieger C, Ingelsson E, Im HK, Illig T, Franks PW, Buck G, Trakalo J, Buck D, Prokopenko I, Mägi R, Lind L, Farjoun Y, Owen KR, Gloyn AL, Strauch K, Tuomi T, Kooner JS, Lee JY, Park T, Donnelly P, Morris AD, Hattersley AT, Bowden DW, Collins FS, Atzmon G, Chambers JC, Spector TD, Laakso M, Strom TM, Bell GI, Blangero J, Duggirala R, Tai ES, McVean G, Hanis CL, Wilson JG, Seielstad M, Frayling TM, Meigs JB, Cox NJ, Sladek R, Lander ES, Gabriel S, Burt NP, Mohlke KL, Meitinger T, Groop L, Abecasis G, Florez JC, Scott LJ, Morris AP, Kang HM, Boehnke M, Altshuler D, McCarthy MI (2016) The genetic architecture of type 2 diabetes. *Nature* 536(7614):41–47. <https://doi.org/10.1038/nature18642>
- Fuller G, Kaye C (2007) Headaches. *BMJ* 334(7587):254–256. <https://doi.org/10.1136/bmj.39090.652847.DE>
- GBD 2019 Diseases and Injuries Collaborators (2020) Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. *Lancet* 396(10258):1204–1222. [https://doi.org/10.1016/s0140-6736\(20\)30925-9](https://doi.org/10.1016/s0140-6736(20)30925-9)
- Gormley P, Anttila V, Winsvold BS, Palta P, Esko T, Pers TH, Farh KH, Cuenca-Leon E, Muona M, Furlotte NA, Kurth T, Ingason A, McMahon G, Ligthart L, Terwindt GM, Kallela M, Freilinger TM, Ran C, Gordon SG, Stam AH, Steinberg S, Borck G, Koiranen M, Quaye L, Adams HH, Lehtimäki T, Sarin AP, Wedenoja J, Hinds DA, Buring JE, Schürks M, Ridker PM, Hrafnisdóttir MG, Stefánsson H, Ring SM, Hottenga JJ, Penninx BW, Färkkilä M, Artto V, Kaunisto M, Vepsäläinen S, Malik R, Heath AC, Madden PA, Martin NG, Montgomery GW, Kurki MI, Kals M, Mägi R, Pärn K, Hämäläinen E, Huang H, Byrnes AE, Franke L, Huang J, Stergiakouli E, Lee PH, Sandor C, Webber C, Cader Z, Muller-Miyhok B, Schreiber S, Meitinger T, Eriksson JG, Salomaa V, Heikkilä K, Loehrer E, Uitterlinden AG, Hofman A, van Duijn CM, Cherkas L, Pedersen LM, Stubhaug A, Nielsen CS, Männikkö M, Mihailov E, Milani L, Göbel H, Esserlind AL, Christensen AF, Hansen TF, Werge T, Kaprio J, Aromaa AJ, Raitakari O, Ikram MA, Spector T, Järvelin MR, Metspalu A, Kubisch C, Strachan DP, Ferrari MD, Belin AC, Dichgans M, Wessman M, van den Maagdenberg AM, Zwart JA, Boomsma DI, Smith GD, Stefánsson K, Eriksson N, Daly MJ, Neale BM, Olesen J, Chasman DI, Nyholt DR, Palotie A (2016) Meta-analysis of 375,000 individuals identifies 38 susceptibility loci for migraine. *Nat Genet* 48(8):856–866. <https://doi.org/10.1038/ng.3598>
- Guo B, Wu B (2019) Powerful and efficient SNP-set association tests across multiple phenotypes using GWAS summary data. *Bioinformatics* 35(8):1366–1372. <https://doi.org/10.1093/bioinformatics/bty811>
- Hautakangas H, Winsvold BS, Ruotsalainen SE, Björnsdóttir G, Harder AVE, Kogelman LJA, Thomas LF, Noordam R, Benner C, Gormley P, Artto V, Banasik K, Björnsdóttir A, Boomsma DI, Brumpton BM, Burgdorf KS, Buring JE, Chalmer MA, de Boer I, Dichgans M, Erikstrup C, Färkkilä M, Garbrielsen ME, Ghanbari M, Hagen K, Häppölä P, Hottenga J-J, Hrafnisdóttir MG, Hveem K, Johnsen MB, Kähönen M, Kristoffersen ES, Kurth T, Lehtimäki T, Ligthart L, Magnusson SH, Malik R, Pedersen OB, Pelzer N, Penninx BWJH, Ran C, Ridker PM, Rosendaal FR, Sigurdardóttir GR, Skogholt AH, Sveinsson OA, Thorgeirsson TE, Ullum H, Vijfhuizen LS, Widén E, van Dijk KW, de Boer I, van den Maagdenberg AMJM, Aromaa A, Belin AC, Freilinger T, Ikram MA, Järvelin M-R, Raitakari OT, Terwindt GM, Kallela M, Wessman M, Olesen J, Chasman DI, Nyholt DR, Stefánsson H, Stefánsson K, van den Maagdenberg AMJM, Hansen TF, Ripatti S, Zwart J-A, Palotie A, Pirinen M, Headacheer HA, International Headache Genetics C, Danish Blood Donor Study Genomic C (2022) Genome-wide analysis of 102,084 migraine cases identifies 123 risk loci and subtype-specific risk alleles. *Nat Genetics* 54(2):152–160. <https://doi.org/10.1038/s41588-021-00990-0>
- He L, Ma Q, Wang Y, Liu X, Yuan Y, Zhang Y, Ou W, Liu L, Tan X, Wang X (2014) Association of variants in KCNK17 gene with ischemic stroke and cerebral hemorrhage in a Chinese population. *J Stroke Cerebrovasc Dis* 23(9):2322–2327. <https://doi.org/10.1016/j.jstrokecerebrovasdis.2014.04.029>
- Headache Classification Committee of the International Headache Society (IHS) (2018) The International Classification of Headache Disorders. *Cephalalgia* 38(1):1–211. <https://doi.org/10.1177/0333102417738202>
- Hill WD, Marioni RE, Maghazian O, Ritchie SJ, Hagenaars SP, McIntosh AM, Gale CR, Davies G, Deary IJ (2019) A combined analysis of genetically correlated traits identifies 187 loci and a role for neurogenesis and myelination in intelligence. *Mol Psychiatry* 24(2):169–181. <https://doi.org/10.1038/s41380-017-0001-5>
- Jacquemin P, Lannoy VJ, Rousseau GG, Lemaigre FP (1999) OC-2, a novel mammalian member of the ONECUT class of homeodomain transcription factors whose function in liver partially overlaps with that of hepatocyte nuclear factor-6. *J Biol Chem* 274(5):2665–2671. <https://doi.org/10.1074/jbc.274.5.2665>
- Ligthart L, de Vries B, Smith AV, Ikram MA, Amin N, Hottenga JJ, Koelewijn SC, Kattenberg VM, de Moor MH, Janssens AC, Aulchenko YS, Oostra BA, de Geus EJ, Smit JH, Zitman FG, Uitterlinden AG, Hofman A, Willemsen G, Nyholt DR, Montgomery GW, Terwindt GM, Gudnason V, Penninx BW, Breteler M, Ferrari MD, Launer LJ, van Duijn CM, van den Maagdenberg AM, Boomsma DI (2011) Meta-analysis of genome-wide association for migraine in six population-based European cohorts. *Eur J Hum Genet* 19(8):901–907. <https://doi.org/10.1038/ejhg.2011.48>
- Mägi R, Suleimanov YV, Clarke GM, Kaakinen M, Fischer K, Prokopenko I, Morris AP (2017) SCOPA and META-SCOPA: software for the analysis and aggregation of genome-wide association studies of multiple correlated phenotypes. *BMC Bioinform* 18(1):25. <https://doi.org/10.1186/s12859-016-1437-3>
- Masotti M, Guo B, Wu B (2019) Pleiotropy informed adaptive association test of multiple traits using genome-wide association study summary data. *Biometrics* 75(4):1076–1085. <https://doi.org/10.1111/biom.13076>
- Meng W, Adams MJ, Hebert HL, Deary IJ, McIntosh AM, Smith BH (2018) A genome-wide association study finds genetic associations with broadly-defined headache in UK Biobank (N=223,773). *EBioMedicine* 28:180–186. <https://doi.org/10.1016/j.ebiom.2018.01.023>
- Meng W, Adams MJ, Reel P, Rajendrakumar A, Huang Y, Deary IJ, Palmer CNA, McIntosh AM, Smith BH (2020) Genetic correlations between pain phenotypes and depression and neuroticism. *Eur J Hum Genet* 28(3):358–366. <https://doi.org/10.1038/s41431-019-0530-2>

- Mills MC, Rahal C (2019) A scientometric review of genome-wide association studies. *Commun Biol* 2:9. <https://doi.org/10.1038/s42003-018-0261-x>
- O'Connor E, Fourier C, Ran C, Sivakumar P, Liesecke F, Southgate L, Harder AVE, Vijfhuizen LS, Yip J, Giffin N, Silver N, Ahmed F, Hostettler IC, Davies B, Cader MZ, Simpson BS, Sullivan R, Efthymiou S, Adebimpe J, Quinn O, Campbell C, Cavalleri GL, Vikelis M, Kelderman T, Paemeleire K, Kilbride E, Grangeon L, Lagrata S, Danno D, Trembath R, Wood NW, Kockum I, Winsvold BS, Steinberg A, Sjöstrand C, Waldenlind E, Vandrovцова J, Houlden H, Matharu M, Belin AC (2021) Genome-wide association study identifies risk loci for cluster headache. *Ann Neurol* 90(2):193–202. <https://doi.org/10.1002/ana.26150>
- Parenti I, Diab F, Gil SR, Mulugeta E, Casa V, Berutti R, Brouwer RWW, Dupé V, Eckhold J, Graf E, Puisac B, Ramos F, Schwarzmayer T, Gines MM, van Staveren T (2020) MAU2 and NIPBL Variants Impair the Heterodimerization of the Cohesin Loader Subunits and Cause Cornelia de Lange Syndrome. *Cell Rep* 31(7):107647. <https://doi.org/10.1016/j.celrep.2020.107647>
- Ray D, Boehnke M (2018) Methods for meta-analysis of multiple traits using GWAS summary statistics. *Genet Epidemiol* 42(2):134–145. <https://doi.org/10.1002/gepi.22105>
- Ray D, Chatterjee N (2020) A powerful method for pleiotropic analysis under composite null hypothesis identifies novel shared loci between Type 2 Diabetes and Prostate Cancer. *PLoS Genet* 16(12):e1009218. <https://doi.org/10.1371/journal.pgen.1009218>
- Riesco N, Cernuda-Morollón E, Pascual J (2017) Neuropeptides as a Marker for Chronic Headache. *Curr Pain Headache Rep* 21(4):18. <https://doi.org/10.1007/s11916-017-0618-8>
- Solovieff N, Cotsapas C, Lee PH, Purcell SM, Smoller JW (2013) Pleiotropy in complex traits: challenges and strategies. *Nat Rev Genet* 14(7):483–495. <https://doi.org/10.1038/nrg3461>
- Stancáková A, Javorský M, Kuulasmaa T, Haffner SM, Kuusisto J, Laakso M (2009) Changes in insulin sensitivity and insulin release in relation to glycemia and glucose tolerance in 6,414 Finnish men. *Diabetes* 58(5):1212–1221. <https://doi.org/10.2337/db08-1607>
- Stearns FW (2010) One hundred years of pleiotropy: a retrospective. *Genetics* 186(3):767–773. <https://doi.org/10.1534/genetics.110.122549>
- Steiner TJ, Scher AI, Stewart WF, Kolodner K, Liberman J, Lipton RB (2003) The prevalence and disability burden of adult migraine in England and their relationships to age, gender and ethnicity. *Cephalalgia* 23(7):519–527. <https://doi.org/10.1046/j.1468-2982.2003.00568.x>
- Steiner TJ, Stovner LJ, Jensen R, Uluduz D, Katsarava Z (2020) Migraine remains second among the world's causes of disability, and first among young women: findings from GBD2019. *J Headache Pain* 21(1):137. <https://doi.org/10.1186/s10194-020-01208-0>
- Stovner L, Hagen K, Jensen R, Katsarava Z, Lipton R, Scher A, Steiner T, Zwart JA (2007) The global burden of headache: a documentation of headache prevalence and disability worldwide. *Cephalalgia* 27(3):193–210. <https://doi.org/10.1111/j.1468-2982.2007.01288.x>
- Techlo TR, Rasmussen AH, Møller PL, Böttcher M, Winther S, Davidsen OB, Olofsson IA, Chalmer MA, Kogelman LJA, Nyegaard M, Olesen J, Hansen TF (2020) Familial analysis reveals rare risk variants for migraine in regulatory regions. *Neurogenetics* 21(3):149–157. <https://doi.org/10.1007/s10048-020-00606-5>
- Turley P, Walters RK, Maghzian O, Okbay A, Lee JJ, Fontana MA, Nguyen-Viet TA, Wedow R, Zacher M, Furlotte NA, Magnusson P, Oskarsson S, Johannesson M, Visscher PM, Laibson D, Cesarini D, Neale BM, Benjamin DJ (2018) Multi-trait analysis of genome-wide association summary statistics using MTAG. *Nat Genet* 50(2):229–237. <https://doi.org/10.1038/s41588-017-0009-4>
- Watanabe K, Taskesen E, van Bochoven A, Posthuma D (2017) Functional mapping and annotation of genetic associations with FUMA. *Nat Commun* 8(1):1826. <https://doi.org/10.1038/s41467-017-01261-5>
- Weiss K, Wigby K, Fannemel M, Henderson LB, Beck N, Ghali N, Study DDD, Anderlid BM, Lundin J, Hamosh A, Jones MC, Ghedia S, Muenke M, Kruszka P (2017) Haploinsufficiency of ZNF462 is associated with craniofacial anomalies, corpus callosum dysgenesis, ptosis, and developmental delay. *Eur J Hum Genet* 25(8):946–951. <https://doi.org/10.1038/ejhg.2017.86>
- Yu J, Li D, Jiang H (2020) Emerging role of ONECUT2 in tumors. *Oncol Lett* 20(6):328. <https://doi.org/10.3892/ol.2020.12192>
- Zhu Z, Wang X (2019) Roles of cohesin in chromosome architecture and gene expression. *Semin Cell Dev Biol* 90:187–193. <https://doi.org/10.1016/j.semcdb.2018.08.004>
- Zhu X, Feng T, Tayo BO, Liang J, Young JH, Franceschini N, Smith JA, Yanek LR, Sun YV, Edwards TL, Chen W, Nalls M, Fox E, Sale M, Bottinger E, Rotimi C, Liu Y, McKnight B, Liu K, Arnett DK, Chakravati A, Cooper RS, Redline S (2015) Meta-analysis of correlated traits via summary statistics from GWASs with an application in hypertension. *Am J Hum Genet* 96(1):21–36. <https://doi.org/10.1016/j.ajhg.2014.11.011>