Gene-based pleiotropy across migraine with aura and migraine without aura patient groups

Zhao, Huiying

2016-06

Zhao , H , Eising , E , de Vries , B , Vijfhuizen , L S , Anttila , V , Winsvold , B S , Kurth , T , Stefansson , H , Kallela , K M , Malik , R , Stam , A H , Ikram , M A , Ligthart , L , Freilinger , T , Alexander , M , Mueller-Myhsok , B , Schreiber , S , Meitinger , T , Aromas , A , Eriksson , J G , Boomsma , D I , van Duijn , C M , Zwart , J-A , Quaye , L , Kubisch , C , Dichgans , M , Wessman , M , Stefansson , K , Chasman , D I , Palotie , A , Martin , N G , Montgomery , G W , Ferrari , M D , Terwindt , G M , van den Maagdenberg , A M J M , Nyholt , D R & Int Headache Genetics Consortium 2016 , ' Gene-based pleiotropy across migraine with aura and migraine without aura patient groups ' , Cephalalgia , vol. 36 , no. 7 , pp. 648-657 . https://doi.org/10.1177/03331

http://hdl.handle.net/10138/224081
https://doi.org/10.1177/0333102415591497

publishedVersion

Original Article

Cephalalgia
An International Journal of Headache

International Headache Society

Cephalalgia
2016, Vol. 36(7) 648–657
© International Headache Society 2015
Reprints and permissions:
sagepub.co.uk/journalsPermissions.nav
DOI: 10.1177/0333102415591497
cep.sagepub.com
SAGE

Gene-based pleiotropy across migraine with aura and migraine without aura patient groups

Huiying Zhao[1,2], Else Eising[3], Boukje de Vries[3], Lisanne S Vijfhuizen[3]; for the International Headache Genetics Consortium: Verneri Anttila[4–7], Bendik S Winsvold[8], Tobias Kurth[9–11], Hreinn Stefansson[12], Mikko Kallela[13], Rainer Malik[14], Anine H Stam[15], M Arfan Ikram[16–18], Lannie Ligthart[19,20], Tobias Freilinger[14,21], Michael Alexander[22,23], Bertram Müller-Myhsok[24,25], Stefan Schreiber[26,27], Thomas Meitinger[28,29], Arpo Aromas[30], Johan G Eriksson[30–34], Dorret I Boomsma[19], Cornelia M van Duijn[16], John-Anker Zwart[8], Lydia Quaye[35], Christian Kubisch[36], Martin Dichgans[14,25], Maija Wessman[4,31], Kari Stefansson[12,37], Daniel I Chasman[1], Aarno Palotie[4,5,7,38–40], Nicholas G Martin[2], Grant W Montgomery[2], Michel D Ferrari[15], Gisela M Terwindt[15], Arn M J M van den Maagdenberg[3,15] and Dale R Nyholt[1,2]

---

[1] Institute of Health and Biomedical Innovation, Queensland University of Technology, Australia
[2] QIMR Berghofer Medical Research Institute, Brisbane, Australia
[3] Department of Human Genetics, Leiden University Medical Centre, The Netherlands
[4] Analytical and Translational Genetics Unit, Department of Medicine, Massachusetts General Hospital, USA
[5] Program in Medical and Population Genetics, Broad Institute of MIT and Harvard, USA
[6] Harvard Medical School, USA
[7] Stanley Center for Psychiatric Research, Broad Institute of Harvard and MIT, USA
[8] FORMI and Department of Neurology, Oslo University Hospital and University of Oslo, Norway
[9] Institut National de la Santé et de la Recherche Médicale (INSERM) Research Center for Epidemiology and Biostatistics (U897) – Team Neuroepidemiology, France
[10] University of Bordeaux, France
[11] Department of Medicine, Division of Preventive Medicine, Brigham and Women's Hospital, Harvard Medical School, USA
[12] deCODE Genetics, Iceland
[13] Department of Neurology, Helsinki University Central Hospital, Finland
[14] Institute for Stroke and Dementia Research, Klinikum der Universität München, Ludwig-Maximilians-Universität, Germany
[15] Department of Neurology, Leiden University Medical Centre, The Netherlands
[16] Department of Epidemiology, Erasmus University Medical Centre, The Netherlands
[17] Department of Radiology, Erasmus University Medical Centre, The Netherlands
[18] Department of Neurology, Erasmus University Medical Centre, The Netherlands
[19] Department of Biological Psychology, VU University, The Netherlands
[20] EMGO+ Institute for Health and Care Research, VU University Medical Centre, The Netherlands
[21] Department of Neurology and Epileptology and Hertie-Institute for Clinical Brain Research, University of Tuebingen, Germany
[22] Department of Genomics, Life & Brain Center, University of Bonn, Germany
[23] Institute of Human Genetics, University of Bonn, Germany
[24] Max Planck Institute of Psychiatry, Germany
[25] Munich Cluster for Systems Neurology (SyNergy), Germany
[26] Institute of Clinical Molecular Biology, Christian Albrechts University, Germany
[27] Department of Internal Medicine I, Christian Albrechts University, Germany
[28] Institute of Human Genetics, Helmholtz Center Munich, Germany
[29] Institute of Human Genetics, Klinikum Rechts der Isar, Technische Universität München, Germany
[30] National Institute for Health and Welfare, Finland
[31] Institute of Genetics, Folkhälsan Research Center, Finland
[32] Department of General Practice, Helsinki University Central Hospital, Finland
[33] Vaasa Central Hospital, Finland

Corresponding author:
Dale R Nyholt, Institute of Health and Biomedical Innovation, Queensland University of Technology, GPO Box 2434, Brisbane QLD 4001, Australia.
Email: d.nyholt@qut.edu.au

Abstract

Introduction: It is unclear whether patients diagnosed according to International Classification of Headache Disorders criteria for migraine with aura (MA) and migraine without aura (MO) experience distinct disorders or whether their migraine subtypes are genetically related.

Aim: Using a novel gene-based (statistical) approach, we aimed to identify individual genes and pathways associated both with MA and MO.

Methods: Gene-based tests were performed using genome-wide association summary statistic results from the most recent International Headache Genetics Consortium study comparing 4505 MA cases with 34,813 controls and 4038 MO cases with 40,294 controls. After accounting for non-independence of gene-based test results, we examined the significance of the proportion of shared genes associated with MA and MO.

Results: We found a significant overlap in genes associated with MA and MO. Of the total 1514 genes with a nominally significant gene-based p value ($p_{\text{gene-based}} \leq 0.05$) in the MA subgroup, 107 also produced $p_{\text{gene-based}} \leq 0.05$ in the MO subgroup. The proportion of overlapping genes is almost double the empirically derived null expectation, producing significant evidence of gene-based overlap (pleiotropy) ($p_{\text{binomial-test}} = 1.5 \times 10^{-4}$). Combining results across MA and MO, six genes produced genome-wide significant gene-based p values. Four of these genes (*TRPM8, UFL1*, *FHL5* and *LRP1*) were located in close proximity to previously reported genome-wide significant SNPs for migraine, while two genes, *TARBP2* and *NPFF* separated by just 259 bp on chromosome 12q13.13, represent a novel risk locus. The genes overlapping in both migraine types were enriched for functions related to inflammation, the cardiovascular system and connective tissue.

Conclusions: Our results provide novel insight into the likely genes and biological mechanisms that underlie both MA and MO, and when combined with previous data, highlight the neuropeptide FF-amide peptide encoding gene (*NPFF*) as a novel candidate risk gene for both types of migraine.

Keywords

Migraine, aura, genome-wide, association, gene-based, pleiotropy

Date received: 28 February 2015; revised: 11 May 2015; accepted: 24 May 2015

Introduction

Two common forms of migraine exist that are clinically distinguished by the presence of aura symptoms prior to the headache phase and hence are called migraine with aura (MA) and migraine without aura (MO) (1). The International Classification of Headache Disorders (ICHD) specifies MO patients have only migraine attacks without aura, whereas MA patients are characterized by their migraine attacks with aura, but can also have attacks without aura (1). The migraine aura is likely caused by cortical spreading depression (CSD), a slowly self-propagating wave of neuronal and glial depolarization in the cortex (2). The headache is caused by activation of the trigeminovascular system, which leads to abnormal processing of pain signals in the brain stem and subsequent activation of higher-order brain centers giving the sensation of pain (3). CSD has been shown to activate this pain pathway in experimental animal studies, although such proof is essentially lacking in patients (4). Whereas the biological mechanisms in the early phase of an attack seem to differ between MA and MO, they may be similar in the headache phase.

Supportive evidence for converging headache-generating mechanisms in both migraine types comes from the fact that some drugs have equal efficacy in treating both types of migraines (5). Moreover, several studies have reported that MA and MO frequently coexist within the same family (6,7). A large population-based study has reported high co-occurrence of MA and MO, and found that 13% of patients with active migraine (at least one migraine attack in the past year) have both MA and MO, which is higher than expected from their individual prevalence (8). Thus, the frequent

---

[34] Department of General Practice and Primary Health Care, University of Helsinki, Finland

[35] Department of Twin Research and Genetic Epidemiology, King's College London, UK

[36] Institute of Human Genetics, University Medical Center Hamburg-Eppendorf, Germany

[37] School of Medicine, University of Iceland, Iceland

[38] Institute for Molecular Medicine Finland (FIMM), University of Helsinki, Finland

[39] Psychiatric & Neurodevelopmental Genetics Unit, Department of Psychiatry, Massachusetts General Hospital, USA

[40] Department of Neurology, Massachusetts General Hospital, USA

H.Z. and E.E. contributed equally to this work.

A.M.J.M.v.d.M. and D.R.N. jointly directed this work.

co-occurrence of the two disorders within families and individuals suggest MA and MO share—at least to some extent—the same biological mechanisms. In support of this idea, latent class and genetic analysis of migraine symptom data in large Australian (9) and Dutch (10) twin samples indicated the existence of a continuum of severity, with MA being more severe but not etiologically distinct from MO.

Genetics studies are starting to discover genes and loci associated with migraine. A recent genome-wide association (GWA) meta-analysis of 23,285 migraine cases and 95,425 controls of European ancestry that was conducted by the International Headache Genetics Consortium (IHGC) identified 12 independent single nucleotide polymorphism (SNP) loci (rs2651899, rs10915437, rs12134493, rs2274316, rs7577262, rs6790925, rs9349379, rs13208321, rs4379368, rs10504861, rs6478241, rs11172113) significantly associated with migraine (p values $< 5 \times 10^{-8}$) (11). Many genes in or near these loci are involved in neuronal function, thus supporting a role for neuronal signaling in migraine etiology as had previously been shown for monogenic familial hemiplegic migraine, a subtype of MA (12). In addition, SNPs at 134 independent loci showed suggestive association with migraine (p values $< 1 \times 10^{-5}$) (11). When considering only the genome-wide significant hits (11,13–15), GWA studies have been more successful using case samples satisfying criteria for only MO, with six loci (rs2274316, rs7577262, rs6790925, rs9349379, rs6478241, rs11172113) identified (13) compared to only one locus (rs1835740) using case samples satisfying criteria for only MA (15). However, in the GWA meta-analysis of 23,285 migraine cases and 95,425 controls (11), nine of the 12 genome-wide significant loci—which comprise five of the six loci previously identified using only MO cases (13)—produced the strongest evidence for association using case samples satisfying MA and/or MO; the remaining three loci (rs9349379, rs13208321, rs10504861) were most strongly associated in the only MO subgroup (16). A lack of detailed clinical data precluded a more in-depth analysis of migraine subtypes.

We previously compared the effects (odds ratios, OR) of the top SNPs from the 12 genome-wide significant loci in the MA and only MO subgroups, and showed that four SNPs (rs10915437, rs9349379, rs13208321, rs10504861) had a significant heterogeneous effect size (p value $= 4.4 \times 10^{-3}$, 3.2×10^{-4}, 4.9×10^{-2}, 4.5×10^{-3}, respectively). However, for all 12 SNPs, the allele associated with increasing risk was the same across these subgroups and an analysis of $\sim$23,000 independent SNPs found the majority of genome-wide SNP effects to be in the same direction across the MA and MO subgroups (16).

To more thoroughly assess the genetic overlap between MA and MO (i.e. beyond inspection of the genome-wide significant loci and individual SNP effects), we performed gene-based tests of association using GWA data for MA and MO from the IHGC meta-analysis. We evaluated the gene-level genetic relationship between MA and MO patients by testing whether the number of overlapping (pleiotropic) genes—i.e. genes with nominally significant gene-based association p values for both MA and MO—was more than expected by chance. The overlapping genes were also used in pathway and network analyses, to identify potential canonical pathways, biological functions, and molecular networks underlying both MA and MO.

Methods

Ethics statement

For all study cohorts, participation was based on informed consent. Each study was approved by local research ethics boards in the country where the study cohort was collected. See the original publication of the IHGC GWA meta-analysis for full details of ethics and consent procedures in each study cohort (11).

Cohorts and sample collection

The IHGC GWA meta-analysis used SNP marker data from 23,285 cases with migraine and 95,425 controls of European ancestry from 18 GWA studies, comprising five clinic-based studies, mainly compared to population-matched control samples with unknown migraine status, as well as 13 population-based cohorts (11). For a subset of the 18 total GWA studies comprising the original GWA meta-analysis, sufficiently detailed phenotype information was available to allow subclassification into either of the two migraine subtypes, MA (seven cohorts: total 4505 MA cases versus 34,813 controls) or MO (four cohorts: total 4038 MO cases versus 40,294 controls) (Table S1). For more detailed descriptions of the cohorts, please see the original report (11).

GWA data

GWA summary statistic results for MA and MO were obtained from seven studies with case samples satisfying criteria for MA, and four studies with case samples satisfying criteria for only MO from the 2013 IHGC GWA meta-analysis (11). The selected MA and MO studies do not contain overlapping individuals and are of similar total size.

Genome-wide SNP genotyping was performed independently in each cohort with the use of various standard genotyping technologies, and imputed for each study with reference to HapMap release 21 or 22

CEU-phased genotypes. Each study contributed summary statistic data from an association analysis performed using a frequentist additive model based on estimated SNP allelic dosages, adjusting for gender. SNPs were filtered on a per-study level based on inclusion criteria of minor allele frequency (MAF) $> 0.1\%$ and imputation quality measures of $I_A > 0.6$ (IMPUTE 2) or $r^2 > 0.3$ (MaCH). The 1,680,313 "consensus" SNPs analyzed in the original IHGC GWA meta-analysis (11) were used in the present study. GWA summary statistic results for the seven MA studies were meta-analyzed in a fixed-effect model using GWAMA (17). GWA summary statistic results for the 4 MO studies were analogously meta-analyzed using GWAMA (17). The resulting MA and MO GWA meta-analysis results were subsequently used in gene-based tests for association.

Gene-based association test

We obtained RefSeq gene information (hg19) from the UCSC genome browser (downloaded 20 March 2014). Overlapping isoforms of the same gene were combined to form a single full-length version of the gene. Isoforms that did not overlap were left as duplicates of that gene. For the duplicated genes, the gene with the lowest downstream gene-based p value was retained. This led to 24,383 unique genes. The IHGC SNP positions were converted from hg18 to hg19 using the liftOver utility (https://genome.ucsc.edu/cgi-bin/hgLiftOver), and were assigned to genes if they mapped to between 15 kb 5' of the transcription start site (TSS) and 15 kb 3' of the transcription end site (TES). The 15-kb gene boundary extension was chosen based on the observation that 90% of SNPs affecting expression quantitative trait loci (eQTLs) were within this proximity (18). Gene-based association tests were performed using the GATES test (19) implemented in the Fast ASsociation Tests (FAST) package (20). GATES performs the gene-based test by adjusting the observed p value of the most significant SNP located in a gene, by the effective number of independent SNPs tested across the gene.

Given the possibility that gene-based results of neighboring genes may in fact be correlated because the most significant SNPs are the same or in strong linkage disequilibrium (LD), we estimated the effective number of independent genes (gene-based tests) by examining the LD between the most significant SNPs within each gene. This calculation was performed using the Genetic type I Error Calculator (GEC) (21). GEC is a method addressing the multiple-testing issue with dependent SNPs. In short, GEC first divides the input SNPs into LD blocks, and assumes LD blocks as independent ($r^2 < 0.1$), then by examining the eigenvalues obtained from spectral decomposition of the LD correlation matrix of SNPs, GEC estimates the effective number of independent SNPs in the blocks.

Overlapping genes and statistic tests

After the gene-based test, we generated gene sets with gene-based association p values less than three nominally significant thresholds (p values<0.01, <0.05, or <0.1) in the MA and MO datasets. For each gene set, the effective number of independent genes was calculated. Subsequently, we regarded the MA dataset as the "discovery" set and the MO dataset as the "target" set to test whether the proportion of overlapping genes between MA and MO for each of the thresholds is more than expected by chance. Here, the observed number of overlapping genes is defined as the effective number of genes with p values less than the threshold in both the discovery (MA) and target (MO) sets. The observed proportion of overlapping genes is the observed effective number of overlapping genes divided by the effective number of genes with a p value less than the threshold in the discovery set. The expected proportion of overlapping genes is the effective number of genes with a p value less than the threshold in the target set divided by the total effective number of genes in the target set. The statistical significance of whether the number of overlapping genes was more than expected by chance was calculated using exact binomial statistical tests.

Pathway analysis

The pathway analysis was designed to discover biological mechanisms that are shared between MA and MO. Here, we used Ingenuity Pathway Analysis (IPA) software, which is widely used to identify pathways enriched in a gene-set based on biological interactions and functional annotations (22,23). The significance of the identified pathways was evaluated by Fisher's exact statistical test. The pathway data sources include "Canonical pathways," "Biological Functions" and "Networks." "Canonical pathways" contain well-characterized metabolic and cell signaling pathways from specific journal articles, review articles, text books, and KEGG LIGAND. "Biological functions" are used to explore the enriched functions and annotated diseases of input genes. "Networks" represent diagrams of known protein-protein interactions, and are generated based on the input data.

We performed the IPA analysis using genes with a p value < 0.1 in both the MA and MO GWA datasets. The results can potentially be biased owing to non-independent (neighboring) genes. For example, one pathway may include multiple genes tagged with the same top SNP or SNPs that are in strong LD. We, therefore,

Table 1. Effective number of genes in MA and MO.

	Total genes		p value < 0.1			p value < 0.05			p value < 0.01		
Dataset	Raw	Effective[c]	Raw	Effective[c]	Proportion[d]	Raw	Effective[c]	Proportion[d]	Raw	Effective[c]	Proportion[d]
MA[a]	21,116	14,395	2502	1694	0.1177	1514	1023	0.0711	424	282	0.0196
MO[b]	21,116	14,485	2154	1498	0.1034	1185	838	0.0579	321	237	0.0164

[a]Migraine with aura. [b]Migraine without aura. [c]Effective number of genes. [d]Proportion of total effective number of genes.

Table 2. Number of overlapping genes and binomial test results under three p value thresholds of gene-based association.

		Overlapping genes		Proportion of overlap		Binominal test
Discovery	Target	Raw	Effective	Expected	Observed	p value
p value < 0.01						
MA[a]	MO[b]	9	8	0.0164	0.0284	9.55×10^{-2}
p value<0.05						
MA[a]	MO[b]	107	83	0.0579	0.0811	1.50×10^{-4}
p value < 0.1						
MA[a]	MO[b]	271	209	0.1034	0.1234	4.64×10^{-3}

[a]Migraine with aura. [b]Migraine without aura.

checked the IPA results to ensure such bias did not contribute to the identified associated pathways.

Given that significant canonical pathways or biological functions may include redundant genes and gene-gene interactions, we explored their relationships by constructing customized networks. First, the genes overlapping between MA and MO were input into IPA. Subsequently, the focus genes (i.e. the subset of overlapping genes) involved in the significant canonical pathways/biological functions were again input into IPA but now to build functional networks for the significant canonical pathways/biological functions. As per IPA recommendations, biological functions were defined as significant if their p values were < 0.01, and canonical pathways were considered significant if their p values were < 0.05.

Results

Effective number of genes for MA and MO

For both MA and MO, we estimated the effective number of genes by inputting the most significant SNPs of all genes in our GEC analysis. Since we assigned the SNPs to genes, the total raw numbers of genes (21,116) are the same for the MA and MO datasets (Table 1). The slight difference in the effective number of total genes in the MA dataset (14,395) versus the MO dataset (14,485) results from differences in LD structure between the respective sets of most significant SNPs in each gene.

Gene-level genetic relationships of MA and MO

As shown in Table 2, significant overlaps were observed for genes with p values <0.05 or <0.1 in both MA and MO, which indicates that a significant gene-level genetic relationship exists between MA and MO. Since the observed effective number of genes was higher than expected, we assumed that these genes have an increased probability to be truly associated with both disorders.

Combined gene-based association across MA and MO

Given the observed gene-based genetic overlap, we combined the evidence of gene-based association across MA and MO using the Fishers' combined p value method. As the effective number of independent gene-based tests ranged from 14,395 (MA) to 14,485 (MO), a gene-based p value $< 3.45 \times 10^{-6}$ (i.e. 0.05/14,485) is required to retain a Type I error rate of 5% and represent genome-wide significant association. Using this threshold value, six genes (*TRPM8, UFL1, FHL5, TARBP2, NPFF* and *LRP1*) were defined as genome-wide significantly associated with migraine (Table 3). The GWA summary statistic results of the top SNPs for these six genes are shown in Table 4. The top significant SNPs for *UFL1* and *FHL5* on chromosome 6q16.1 are in strong LD, and the neighboring genes *NPFF* and *TARBP2* on chromosome 12q13.13 have the same top SNP (rs11170566), hence the most

Table 3. Genes with Fisher's combined *p* value lower than the threshold.

Chromosome	Gene	Start (hg19)	End (hg19)	p_{MA}[c]	p_{MO}[d]	Fisher's combined *p* value
2q37.1	*TRPM8*[a]	234811042	234943166	1.0×10^{-4}	6.9×10^{-6}	1.6×10^{-8}
6q16.1	*UFL1*[a]	96954701	97018151	4.5×10^{-2}	8.4×10^{-7}	7.0×10^{-7}
6q16.1	*FHL5*[a]	96995423	97079512	4.5×10^{-2}	1.7×10^{-7}	1.5×10^{-7}
12q13.13	*TARBP2*[b]	53880376	53915215	6.3×10^{-3}	2.1×10^{-5}	2.2×10^{-6}
12q13.13	*NPFF*[b]	53885473	53916422	4.4×10^{-3}	1.4×10^{-5}	1.1×10^{-6}
12q13.3	*LRP1*[a]	57507281	57622125	3.2×10^{-3}	2.5×10^{-5}	1.4×10^{-6}

[a] Genes close to single nucleotide polymorphisms (SNPs) significantly associated with migraine without aura (MO) in a genome-wide association (GWA) meta-analysis (11).
[b] Genes not close to SNPs significantly associated with MO or total migraine in a GWA meta-analysis (11).
[c] Gene-based *p* value for migraine with aura.
[d] Gene-based *p* value for migraine without aura.

Table 4. Association results for the top SNPs in the six genome-wide significant genes.

Chromosome	Gene	Top SNP (subtype)	EA[a]	NEA[b]	MA[e]		MO[f]	
					OR[c]	*p* value[d]	OR[c]	*p* value[d]
2q37.1	*TRPM8*	rs17863838 (MA[e])	G	A	1.21	3.63×10^{-6}	1.26	9.69×10^{-7}
		rs10187654 (MO[f])	C	T	1.15	7.65×10^{-6}	1.21	9.32×10^{-8}
6q16.1	*UFL1*	rs4598081 (MA[e])	A	G	1.07	1.07×10^{-2}	1.14	7.02×10^{-6}
		rs11153058 (MO[f])	T	C	1.06	8.81×10^{-2}	1.21	2.01×10^{-7}
6q16.1	*FHL5*	rs11759769 (MA[e])	A	G	1.09	3.02×10^{-3}	1.20	4.49×10^{-8}
		rs2983896 (MO[f])	A	G	1.09	5.11×10^{-3}	1.22	9.70×10^{-9}
12q13.13	*TARBP2*	rs11170566 (MA[e], MO[f])	T	C	1.11	2.20×10^{-3}	1.19	7.04×10^{-6}
12q13.13	*NPFF*	rs11170566 (MA[e], MO[f])	T	C	1.11	2.20×10^{-3}	1.19	7.04×10^{-6}
12q13.3	*LRP1*	rs11172113 (MA[e], MO[f])	T	C	1.10	2.06×10^{-4}	1.15	1.57×10^{-6}

[a] Effect allele. [b] Non-effect allele. [c] Odds ratio for effect allele calculated in the GWA data. [d] *p* value of SNP. [e] Migraine with aura. [f] Migraine without aura.
SNPs: single nucleotide polymorphisms; GWA: genome-wide association.

likely causative risk gene at these loci cannot be determined from the gene-based association results alone. Compared to results from the original migraine GWA meta-analysis of all 18 studies (11), in addition to implicating four genes in close proximity to three previously reported genome-wide significant SNPs (*TRPM8* near rs7577262 on 2q37.1; *UFL1* and *FHL5* near rs13208321 on 6q16.1; and *LRP1* near rs11172113 on 12q13.3), this study identified two genes (*NPFF* and *TARBP2*) at a novel candidate migraine risk locus on chromosome 12q13.13, which was not previously implicated in the GWA meta-analysis of all 23,285 migraine cases and 95,425 controls (11).

Pathway analysis on overlapping genes

The significant biological functions, canonical pathways, and networks found in the overlap between MA and MO are shown in Tables S2–S4. The most significant biological function is "chronic inflammatory disorder" (Table S2). Notably, two genes linked to the function "chronic inflammatory disorder": *TRPM8* (2q37.1) and *UFL1* (6q16.1) have Fisher's combined gene-based *p* values surpassing the genome-wide threshold (3.45×10^{-6}). The most significant canonical pathway is "Notch Signaling" (Table S3), which plays important roles in neuronal function and development (24–27). The most significant network is "Cardiovascular Disease, Organismal Injury and Abnormalities, Cardiac Stenosis" (Table S4). Other biological functions and networks that were significantly enriched in the overlap between MA and MO were mostly involved in cardiovascular function, inflammation, development and connective tissue-related functions (Tables S3 and S4).

As shown in Tables S2 and S3, several significant biological functions and canonical pathways share the same candidate genes, which indicate that these functions are related. In order to find the relationship between them, we constructed networks by IPA (described in Methods). The significant canonical pathways and biological functions were combined into five networks without overlapping genes (Table S5).

Discussion

This study identified a significant gene-level genetic relationship between migraine cases satisfying ICHD criteria for MA and migraine cases satisfying criteria for only MO by integrating gene-based tests and estimating the effective number of genes using the GWA summary statistic results for both disorders from the recent meta-analysis conducted by the IHGC (11).

Our approach is different from single-SNP-based approaches, e.g. polygenic prediction (28) and SNP effect concordance analysis (SECA) (29), which previously showed a significant SNP-based genetic overlap between MA and MO (16), as it has the ability to identify genes across disorders in the presence of allelic heterogeneity. The latter is of great importance as only 12 out of 271 genes with gene-based p values < 0.1 for both MA and MO were shown to be tagged by the same top significant SNP. The other advantage of our gene-based approach comes from the fact that genes are the predominant functional unit of the human genome and are therefore more closely related to biological mechanisms.

From the combined analysis of MA and MO gene-based p values, six genes surpass our genome-wide significance threshold: *TRPM8, UFL1*, *FHL5*, *TARBP2, NPFF* and *LRP1* on chromosome 2q37.1, 6q16.1, 6q16.1, 12q13.13, 12q13.13 and 12q13.3, respectively. These results support the utility of our gene-based approach, especially given they are based on 14,742 fewer migraine cases and 20,318 fewer controls (~41% smaller effective sample size) compared to the original IHGC GWA study of 23,285 migraine cases and 95,425 controls.

Four of these genes (*TRPM8, UFL1*, *FHL5* and *LRP1*) were located at or in close proximity to previously reported genome-wide significant SNPs. Transient receptor potential cation channel, subfamily M, member 8 (*TRPM8*) is a calcium channel expressed in primary sensory neurons essential for cold sensation. It can also mediate cold-mediated analgesics in chronic pain models (30), suggesting a modulatory role for TRPM8 in pain sensation. TRPM8 was also identified in the vasculature where it plays a role in setting the vascular tone (31), another function through which the channel could play a role in migraine. UFM1-specific ligase 1 (UFL1) is a protein located at the ER membrane. GO molecular functions assigned to *UFL1* are *ligase activity*, *protein binding* and *UFM1 transferase activity*. The biological roles of UFL1 are largely unknown, but the protein has recently been associated with female hormone signaling pathways (32), providing a potential link with female preponderance in migraine. Four and a half LIM domains 5 (FHL5) is a transcriptional activator of cAMP-responsive element modulator that activates transcriptional programs during spermatogenesis; however, a brain-related role is currently not known. Low-density lipoprotein receptor-related protein 1 (LRP1) is an endocytic receptor with many ligands. Many biological functions have been described for LRP1. In the brain, LRP1 associates with and regulates the expression of NMDAR1 and GluR1 at postsynaptic neurons and can thereby regulate neurotransmission (33), providing an apparent link with migraine pathophysiology.

Two genes, *TARBP2* and *NPFF*, represent novel candidate risk genes for migraine. The known molecular functions of TARBP2, TAR (HIV-1) RNA binding protein 2, that is encoded by the *TARBP2* gene, centers on its requirement for formation of the RNA-induced silencing complex. *TARBP2* binds between the bulge and the loop of the HIV-1 TAR RNA regulatory element and activates HIV-1 gene expression in synergy with the viral Tat protein. TARBP2 is also an integral component of the DICER1-containing complex and involved in miRNA processing (34). Together with the gene ontology (GO) molecular function terms assigned to *TARBP2* (*double-stranded RNA binding*, *protein binding*, *siRNA binding*, *miRNA binding*, and *protein homodimerization activity*), there is no obvious functional link between *TARBP2* and migraine risk.

The neuropeptide FF-amide peptide precursor (NPFF), encoded by the *NPFF* gene, has wide-ranging physiologic effects, including the modulation of morphine-induced analgesia, elevation of arterial blood pressure, and increased somatostatin secretion from the pancreas. The GO molecular functions assigned to *NPFF* are *G-protein coupled receptor binding*, *receptor binding*, and *neuropeptide hormone activity*. Notably, neuropeptide FF potentiates and sensitizes acid-sensing ion channels ASIC1 and ASIC3. The ASICs represent proton-gated channels that are able to transport Na^+ and Ca^{2+}. A recent study using whole-cell patch-clamp electrophysiology showed that a decrease in extracellular pH can directly excite primary dural-afferent neurons via the opening of ASICs and produce migraine-related pain behavior, suggesting ASIC inhibitors may represent novel candidates for migraine therapy (35). Indeed, the ASIC inhibitor amiloride was recently shown to block CSD—the neurophysiological correlate of migraine aura—and inhibited trigeminal activation in in vivo in animals, via an ASIC1 mechanism (36). These previous findings suggest that *NPFF* is the more probable migraine risk gene on 12q13.13.

To gain better understanding of the biological mechanisms that are involved in the two migraine subtypes, we also performed pathway analyses on the genes that showed nominal association with both MA and MO. Pathway analysis based on biological functions provided 22 significant diseases/functions (Table S2). Among them, many were related to inflammatory disorders. A causal role for inflammation in migraine pathophysiology has been described previously, as CSD can activate an inflammatory cascade that can reach the meninges (37) where it may cause neurogenic inflammation and subsequent activation of trigeminal neurons and thereby activation of the migraine headache (37,38).

Gene functions and networks related to cardiovascular and connective tissue disorders showed enrichment in the overlap between MA and MO as well (Tables S2, S4 and S5). A possible role for the cardiovascular system in migraine has been suggested by the high coincidence of MA with stroke, several cardiac disorders, and cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy (CADASIL) (39–41). Very few brain-related functions were identified in the genes nominally associated with both migraine subtypes.

Although approximately three-quarters of the MA patients studied may also suffer migraine attacks without aura (15), our analyses are valid with respect to testing the hypothesis of a shared genetic susceptibility between patients experiencing migraine attacks with aura and patients experiencing only migraine attacks without aura. It is unclear what the relative proportion of the observed genetic overlap is due to patients suffering only migraine attacks with aura, and patients suffering migraine attacks both with and without aura. To specifically assess the genetic overlap between patients suffering only migraine attacks with aura and patients suffering only migraine attacks without aura, additional studies comparing results from GWA analysis of MA-only patients, to results from GWA analysis of MO-only patients are required. However, the MA-only GWA samples currently available are too small to deliver sufficient power to provide a valid conclusion regarding the genetic overlap between MA-only patients and the more common MA + MO and MO-only patients groups.

In summary, these results show a significant gene-based overlap (pleiotropy) between migraine cases satisfying criteria for MA and migraine cases satisfying criteria for only MO, further explaining—at least partly—their co-occurrence within the same patient or family. Our results also highlight four genes (*TRPM8*, *UFL1*, *FHL5* and *LRP1*) that had earlier surfaced in GWA studies and two novel genes (*TARBP2* and *NPFF*) as candidate risk genes for migraine, and indicate that inflammatory and cardiovascular processes may be involved in the etiology of both MA and MO.

Article highlights

- A significant overlap in associated genes exists between the primary subtypes migraine with aura (MA) and migraine without aura (MO), further explaining—at least partly—their co-occurrence within the same patient or family.
- Combining gene-based association results across MA and MO confirmed association of four genes at three previously implicated loci, and when combined with previous data, highlight *NPFF* as a novel candidate risk gene for both types of migraine.
- Results from pathway analyses indicate that inflammatory and cardiovascular processes may be involved in the etiology of both MA and MO.

Funding

This work was supported by a grant from the National Health and Medical Research Council (NHMRC) (APP1075175), and by the European Union's Seventh Framework program (2007–2013) under grant agreement no. 602633 (EUROHEADPAIN).

Conflict of interest

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Acknowledgments

Huiying Zhao was supported by a National Health and Medical Research Council (NHMRC) Early Career Fellowship (APP1091816). Dale R. Nyholt was supported by the Australian Research Council (ARC) Future Fellowship (FT0991022) and NHMRC Research Fellowship (APP0613674) Schemes. Grant W. Montgomery was supported by the NHMRC Fellowships Scheme (339446, 619667). See the original publication of the 2013 IHGC migraine GWA meta-analysis (11) for detailed cohort-specific acknowledgments.

References

1. Headache Classification Committee of the International Headache Society. The International Classification of Headache Disorders: 3 rd edition. *Cephalalgia* 2013; 33: 629–808.
2. Lauritzen M, Dreier JP, Fabricius M, et al. Clinical relevance of cortical spreading depression in neurological

disorders: Migraine, malignant stroke, subarachnoid and intracranial hemorrhage, and traumatic brain injury. *J Cereb Blood Flow Metab* 2011; 31: 17–35.

3. Noseda R and Burstein R. Migraine pathophysiology: Anatomy of the trigeminovascular pathway and associated neurological symptoms, cortical spreading depression, sensitization, and modulation of pain. *Pain* 2013; 154(Suppl 1): S44–S53.
4. Ferrari MD, Klever RR, Terwindt GM, et al. Migraine pathophysiology: Lessons from mouse models and human genetics. *Lancet Neurol* 2015; 14: 65–80.
5. Andersson PG, Dahl S, Hansen JH, et al. Prophylactic treatment of classical and non-classical migraine with metoprolol—a comparison with placebo. *Cephalalgia* 1983; 3: 207–212.
6. Mochi M, Sangiorgi S, Cortelli P, et al. Testing models for genetic determination in migraine. *Cephalalgia* 1993; 13: 389–394.
7. Joutel A, Ducros A, Vahedi K, et al. Genetic heterogeneity of familial hemiplegic migraine. *Am J Hum Genet* 1994; 55: 1166–1172.
8. Launer LJ, Terwindt GM and Ferrari MD. The prevalence and characteristics of migraine in a population-based cohort: The GEM study. *Neurology* 1999; 53: 537–542.
9. Nyholt DR, Gillespie NA, Heath AC, et al. Latent class analysis does not support migraine with aura and migraine without aura as separate entities. *Genet Epidemiol* 2004; 26: 231–244.
10. Ligthart L, Boomsma DI, Martin NG, et al. Migraine with aura and migraine without aura are not distinct entities: Further evidence from a large Dutch population study. *Twin Res Hum Genet* 2006; 9: 54–63.
11. Anttila V, Winsvold BS, Gormley P, et al. Genome-wide meta-analysis identifies new susceptibility loci for migraine. *Nat Genet* 2013; 45: 912–917.
12. de Vries B, Frants RR, Ferrari MD, et al. Molecular genetics of migraine. *Hum Genet* 2009; 126: 115–132.
13. Freilinger T, Anttila V, de Vries B, et al. Genome-wide association analysis identifies susceptibility loci for migraine without aura. *Nat Genet* 2012; 44: 777–782.
14. Chasman DI, Schurks M, Anttila V, et al. Genome-wide association study reveals three susceptibility loci for common migraine in the general population. *Nat Genet* 2011; 43: 695–698.
15. Anttila V, Stefansson H, Kallela M, et al. Genome-wide association study of migraine implicates a common susceptibility variant on 8q22.1. *Nat Genet* 2010; 42: 869–873.
16. Nyholt DR, International Headache Genetics Consortium, Anttila V, et al. Concordance of genetic risk across migraine subgroups: Impact on current and future genetic association studies. *Cephalalgia* 2015; 35: 489–499.
17. Magi R and Morris AP. GWAMA: Software for genome-wide association meta-analysis. *BMC bioinformatics* 2010; 11: 288.
18. Pickrell JK, Marioni JC, Pai AA, et al. Understanding mechanisms underlying human gene expression variation with RNA sequencing. *Nature* 2010; 464: 768–772.
19. Li MX, Gui HS, Kwan JS, et al. GATES: A rapid and powerful gene-based association test using extended Simes procedure. *Am J Hum Genet* 2011; 88: 283–293.
20. Chanda P, Huang H, Arking DE, et al. Fast association tests for genes with FAST. *PLoS One* 2013; 8: e68585.
21. Li MX, Yeung JM, Cherny SS, et al. Evaluating the effective numbers of independent tests and significant *p*-value thresholds in commercial genotyping arrays and public imputation reference datasets. *Hum Genet* 2012; 131: 747–756.
22. Bradel-Tretheway BG, Mattiacio JL, Krasnoselsky A, et al. Comprehensive proteomic analysis of influenza virus polymerase complex reveals a novel association with mitochondrial proteins and RNA polymerase accessory factors. *J Virol* 2011; 85: 8569–8581.
23. Belisle SE, Tisoncik JR, Korth MJ, et al. Genomic profiling of tumor necrosis factor alpha (TNF-alpha) receptor and interleukin-1 receptor knockout mice reveals a link between TNF-alpha signaling and increased severity of 1918 pandemic influenza virus infection. *J Virol* 2010; 84: 12576–12588.
24. Gaiano N and Fishell G. The role of notch in promoting glial and neural stem cell fates. *Annu Rev Neurosci* 2002; 25: 471–490.
25. Bolos V, Grego-Bessa J and de la Pompa JL. Notch signaling in development and cancer. *Endocr Rev* 2007; 28: 339–363.
26. Aguirre A, Rubio ME and Gallo V. Notch and EGFR pathway interaction regulates neural stem cell number and self-renewal. *Nature* 2010; 467: 323–327.
27. Hitoshi S, Alexson T, Tropepe V, et al. Notch pathway molecules are essential for the maintenance, but not the generation, of mammalian neural stem cells. *Genes Dev* 2002; 16: 846–858.
28. Purcell SM, Wray NR, Stone JL, et al. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature* 2009; 460: 748–752.
29. Nyholt DR. SECA: SNP effect concordance analysis using genome-wide association summary results. *Bioinformatics* 2014; 30: 2086–2088.
30. Proudfoot CJ, Garry EM, Cottrell DF, et al. Analgesia mediated by the *TRPM8* cold receptor in chronic neuropathic pain. *Curr Biol* 2006; 16: 1591–1605.
31. Johnson CD, Melanaphy D, Purse A, et al. Transient receptor potential melastatin 8 channel involvement in the regulation of vascular tone. *Am J Physiol Heart Circ Physiol* 2009; 296: H1868–H1877.
32. Yoo HM, Kang SH, Kim JY, et al. Modification of ASC1 by UFM1 is crucial for ERalpha transactivation and breast cancer development. *Mol Cell* 2014; 56: 261–274.
33. Liu Q, Trotter J, Zhang J, et al. Neuronal LRP1 knockout in adult mice leads to impaired brain lipid metabolism and progressive, age-dependent synapse loss and neurodegeneration. *J Neurosci* 2010; 30: 17068–17078.
34. Melo SA, Ropero S, Moutinho C, et al. A *TARBP2* mutation in human cancer impairs microRNA processing and *DICER1* function. *Nat Genet* 2009; 41: 365–370.
35. Yan J, Edelmayer RM, Wei X, et al. Dural afferents express acid-sensing ion channels: A role for decreased

meningeal pH in migraine headache. *Pain* 2011; 152: 106–113.

36. Holland PR, Akerman S, Andreou AP, et al. Acid-sensing ion channel 1: A novel therapeutic target for migraine with aura. *Ann Neurol* 2012; 72: 559–563.
37. Karatas H, Erdener SE, Gursoy-Ozdemir Y, et al. Spreading depression triggers headache by activating neuronal Panx1 channels. *Science* 2013; 339: 1092–1095.
38. Waeber C and Moskowitz MA. Migraine as an inflammatory disorder. *Neurology* 2005; 64(10 Suppl): S9–S15.
39. Sacco S and Kurth T. Migraine and the risk for stroke and cardiovascular disease. *Curr Cardiol Rep* 2014; 16: 524.
40. Kurth T, Chabriat H and Bousser MG. Migraine and stroke: A complex association with clinical implications. *Lancet Neurol* 2012; 11: 92–100.
41. Chabriat H, Joutel A, Dichgans M, et al. Cadasil. *Lancet Neurol* 2009; 8: 643–653.