

## RESEARCH ARTICLE

## A replication study separates polymorphisms behind migraine with and without depression

Peter Petschner<sup>1,2,3\*</sup>, Daniel Baksa<sup>2,4</sup>, Gabor Hullam<sup>5</sup>, Dora Torok<sup>2</sup>, Andras Millinghoffer<sup>3,5,6</sup>, J. F. William Deakin<sup>7</sup>, Gyorgy Bagdy<sup>2,3,6</sup>, Gabriella Juhasz<sup>2,3,4</sup>

**1** Bioinformatics Center, Institute for Chemical Research, Kyoto University, Gokasho, Uji, Kyoto, Japan, **2** Department of Pharmacodynamics, Faculty of Pharmacy, Semmelweis University, Budapest, Hungary, **3** MTA-SE Neuropsychopharmacology and Neurochemistry Research Group, Hungarian Academy of Sciences, Semmelweis University, Budapest, Hungary, **4** SE-NAP2 Genetic Brain Imaging Migraine Research Group, Hungarian Brain Research Program, Semmelweis University, Budapest, Hungary, **5** Department of Measurement and Information Systems, Budapest University of Technology and Economics, Budapest, Hungary, **6** NAP-2-SE New Antidepressant Target Research Group, Semmelweis University, Budapest, Hungary, **7** Neuroscience and Psychiatry Unit, Division of Neuroscience and Experimental Psychology, The University of Manchester and Manchester Academic Health Sciences Centre, Manchester, United Kingdom

☞ These authors contributed equally to this work.

\* [petschnerp@yahoo.com](mailto:petschnerp@yahoo.com)



## OPEN ACCESS

**Citation:** Petschner P, Baksa D, Hullam G, Torok D, Millinghoffer A, Deakin JFW, et al. (2021) A replication study separates polymorphisms behind migraine with and without depression. PLoS ONE 16(12): e0261477. <https://doi.org/10.1371/journal.pone.0261477>

**Editor:** Badri Padhukasahasram, Illumina Inc, UNITED STATES

**Received:** May 23, 2021

**Accepted:** December 3, 2021

**Published:** December 31, 2021

**Copyright:** © 2021 Petschner et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** Our study was approved by Scientific and Research Ethics Committee of the Medical Research Council (Budapest, Hungary) and North Manchester Local Research Ethics Committee, Manchester. At the time the DNA samples were collected, our participant consent form did not include a paragraph regarding open access data availability. Since our participants were not in a position to give or refuse permission for public data sharing, the authors are not allowed to share individual data publicly (all group-level, summarized data are

## Abstract

The largest migraine genome-wide association study identified 38 candidate loci. In this study we assessed whether these results replicate on a gene level in our European cohort and whether effects are altered by lifetime depression. We tested SNPs of the loci and their vicinity with or without interaction with depression in regression models. Advanced analysis methods such as Bayesian relevance analysis and a neural network based classifier were used to confirm findings. Main effects were found for rs2455107 of *PRDM16* (OR = 1.304, p = 0.007) and five intergenic polymorphisms in 1p31.1 region: two of them showed risk effect (OR = 1.277, p = 0.003 for both rs11209657 and rs6686879), while the other three variants were protective factors (OR = 0.4956, p = 0.006 for both rs12090642 and rs72948266; OR = 0.4756, p = 0.005 for rs77864828). Additionally, 26 polymorphisms within *ADGRL2*, 2 in *REST*, 1 in *HPSE2* and 33 mostly intergenic SNPs from 1p31.1 showed interaction effects. Among clumped results representing these significant regions, only rs11163394 of *ADGRL2* showed a protective effect (OR = 0.607, p = 0.002), all other variants were risk factors (rs1043215 of *REST* with the strongest effect: OR = 6.596, p = 0.003). Bayesian relevance analysis confirmed the relevance of intergenic rs6660757 and rs12128399 (p31.1), rs1043215 (*REST*), rs1889974 (*HPSE2*) and rs11163394 (*ADGRL2*) from depression interaction results, and the moderate relevance of rs77864828 and rs2455107 of *PRDM16* from main effect analysis. Both main and interaction effect SNPs could enhance predictive power with the neural network based classifier. In summary, we replicated p31.1, *PRDM16*, *REST*, *HPSE2* and *ADGRL2* genes with classic genetic and advanced DNA analysis methods. While the p31.1 region and *PRDM16* are worthy of further investigations in migraine in general, *REST*, *HPSE2* and *ADGRL2* may be prime candidates behind migraine pathophysiology in patients with comorbid depression.

available in the manuscript and in Supplementary Information). However, reasonable data access requests can be made to the corresponding author ([petschnerp@yahoo.com](mailto:petschnerp@yahoo.com)) or to Dr. Xenia Gonda ([gonda.xenia@med.semmelweis-univ.hu](mailto:gonda.xenia@med.semmelweis-univ.hu)) at the Department of Psychiatry and Psychotherapy, Semmelweis University, a key member of the NewMood (New Molecules in Mood Disorders, 2004-2009) research program who took part in data collection and can be considered as a completely independent data manager regarding this paper.

**Funding:** Our study was part of NewMood (New Molecules in Mood Disorders, 2004-2009) research program funded by the European Union and was supported by the Hungarian Brain Research Program (KTIA\_13\_NAP-A-II/14, 2017-1.2.1-NKP-2017-00002); by the National Development Agency (KTIA\_NAP\_13-1-2013-0001); by the Sixth Framework Program of the European Union, NewMood (Grant No. LSHM-CT-2004-503474); by the Hungarian Academy of Sciences, Semmelweis University and the Hungarian Brain Research Program (MTA-SE-NAP B Genetic Brain Imaging Migraine Research Group, Grant KTIA\_NAP\_13-2-2015-0001); by the National Institute for Health Research Manchester Biomedical Research Centre; by the Hungarian Academy of Sciences (MTA-SE Neuropsychopharmacology and Neurochemistry Research Group); by the National Research, Development and Innovation Office, Hungary (2019-2.1.7-ERA-NET-2020-00005), under the frame of ERA PerMed (ERAPERMED2019-108); by the New National Excellence Program of Ministry of Human Capacities (UNKP-17-4-I-SE-8, UNKP-19-4-BME-344); by the New National Excellence Program of the Ministry for Innovation and Technology from the source of the National Research, Development and Innovation Fund (UNKP-20-3-II-SE-51, UNKP-20-5-BME-92, UNKP-21-5-BME-362, UNKP-21-4-I-SE-15); by the BME-Biotechnology FIKP grant of EMMI (BME FIKP-BIO); by the Thematic Excellence Programme (Tématerületi Kiválósági Program, 2020-4.1.1.-TKP2020) of the Ministry for Innovation and Technology in Hungary, within the framework of the Neurology and Translational Biotechnology thematic programmes of the Semmelweis University; by OTKA 119866; and by the NRDI Fund (BME NC TKP2020, BME IE-BIO TKP2020, Artificial Intelligence National Laboratory Programme) based on the charter of bolster issued by the NRDI Office under the auspices of the Ministry for Innovation and Technology. GH is a recipient of the Janos Bolyai Research Scholarship of the Hungarian Academy of Sciences. PP is an

## Introduction

Migraine is a serious neurological disorder characterized by recurrent headache, typically with unilateral pulsating pain aggravated by physical activity, nausea, photo- and/or phonophobia. Worldwide on average 11% of adults are affected, with women three times more often [1]. Migraineurs suffer severe interference with daily activities due to the disease that also manifests in high direct (e.g. healthcare service utilization) and indirect (including reduced productivity) costs [2–4]. Estimated annual expenses of migraine in Europe are around €27 billion [5] and the disease is responsible for 25 million lost school- and working days in the UK alone [1]. All these data motivate investigations of the pathomechanisms of the disease, which are still not fully understood.

The pathophysiology of common migraine involves multiple factors including genetic elements. Genetic heritability of the disease is around 40–50% [6] and several hypothesis-driven candidate gene studies reported associations, e.g. with polymorphisms of *HTR1A*, *HTR1B*, *HTR2A*, *MAOA*, *SLC6A4*, *COMT*, and *CNR1* genes, though their etiological roles are not certain [7–10]. Since migraine is polygenic in nature, genome-wide association studies (GWAS) that investigate polymorphisms associated with the disease along the entire genome, can point to novel candidates for further testing. A recent, seminal GWAS meta-analysis identified 38 loci associated with migraine [11]. After initial identification of such variants replication studies are necessary to verify and refine findings.

In addition to genetic risk, comorbid conditions may also play a role in the pathophysiology of migraine including psychiatric disorders, like depression [6, 12–15]. A recent study on 117,392 subjects from the UK Biobank cohort further supported a comorbid relationship between depression and migraine [16]. Depression also has a polygenic background and genetic association and twin studies suggest shared genetic risk factors between migraine and depression [17–20]. Based on GWAS results the genetic correlation ( $r_G$ ) between migraine and depression is between 0.25–0.32 [18, 21], however, migraine with and without comorbid depression might also develop on partially different genetic background [6] and a causal relationship is also possible [17]. Investigating monozygotic twins discordant for depression showed that only the twin suffering from depression has increased risk for migraine, indicating similar conclusions [17].

All these findings indicate that migraine has complex relationships involving several factors. Standard methods, however, are often inadequate to explore complex relationships such as nonlinear interactions. Advanced analysis tools utilizing machine learning methods may reveal additional aspects of variables and relationships. One such aspect is their relevance with respect to the disease. More precisely, which variables are strongly relevant, i.e. belong to the minimal Markov blanket of the target variable, which is migraine in this case. These variables represent the sufficient set of variables to uniquely define the value of the target variable in a graph representing dependencies [22]. We applied a method to determine strong relevance, the so called Bayesian network based Bayesian multilevel analysis of relevance, or Bayesian relevance analysis in short. This method was recently applied to analyze relevance of comorbidities, environmental factors and genetic variants [16, 23, 24].

The other important property of a variable is its predictive power. That is, how accurately a given variable predicts the target variable. Predictive power can be measured by using the variables as features in a classifier, such as a neural network, whose prediction accuracy is then assessed. Based on the above, both methods can complement traditional methods, such as linear or logistic regressions.

In light of the above, the aim of the present study was threefold: 1) to replicate the results of the largest migraine GWAS [11] by conducting a candidate gene study involving all

international research fellow of Japan Society for the Promotion of Science (Postdoctoral Fellowships for Research in Japan, standard program, P20809). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing interests:** J. F. William Deakin has share options in P1vital and Gyorgy Bagdy is a member of the Board of directors at Gedeon Richter. The other authors declare no competing interests exist.

polymorphisms within the 10,000 base pair (bp) vicinity of proposed genes/loci based on our independent samples from Budapest and Manchester using logistic regression models; 2) to investigate whether polymorphisms within the already proposed relevant genes/loci show interactions with lifetime depression with regression models; 3) to validate the identified single nucleotide polymorphisms (SNPs) with two machine learning algorithms, a Bayesian relevance analysis and a neural network based classifier to confirm the relevance, and the predictive power of the identified factors, respectively.

In summary, analyses presented in this paper validate polymorphisms in p31.1 region of chromosome 1, *PRDM16*, *ADGRL2*, *REST* and *HPSE2* genes in migraine. Furthermore, results allowed the separation of variants that play a role in migraine accompanied by depression, like *ADGRL2*, *REST* and *HPSE2*, from those that play a general role in migraine, like *PRDM16*. Finally, we also show, how different machine learning methods can be applied to analyze and evaluate effects of genetic polymorphisms interacting with non-genetic factors.

## Materials and methods

To achieve our objectives, first we conducted standard genetic association analyses of polymorphisms within the genes/loci identified by Gormley et al. [11] and their 10,000 base pair (bp) vicinity in two European population genetic samples (from Budapest and Manchester) and their combination. Subsequently, we performed interaction analyses with lifetime depression using the same samples. We used a less stringent significance criterion ( $p < 0.05$  in all three samples, for a justification see below), however we also applied two machine learning algorithms, a Bayesian relevance analysis method and a neural network classifier to confirm findings.

## Subjects

Around 2700 participants (aged between 18–60) were recruited through advertisements and general practices from Greater Manchester, United Kingdom and Budapest, Hungary. Ethics Committees in both cities approved the study (Scientific and Research Ethics Committee of the Medical Research Council, Budapest, Hungary, ad.225/KO/2005.; ad.323-60/2005-1018EKU and ad.226/KO/2005.; ad.323-61/2005-1018 EKU; North Manchester Local Research Ethics Committee, Manchester, UK REC reference number: 05/Q1406/26), and all our procedures were in accordance with the Declaration of Helsinki. Written informed consent was provided by all participants. We included subjects of European white origin (based on self-reported data) who provided DNA sample. Our total subject number was  $n = 1815$  (in Budapest:  $n = 839$ ; in Manchester:  $n = 975$ ) after screening for available data on sex, age, ethnicity, migraine and lifetime depression status. For exact subject numbers used in each analysis, see S3–S8 Tables. Recruitment strategies and response rates were specified earlier [25].

## Phenotype

English and Hungarian versions of brief standard questionnaires were used in the study. Our validated background questionnaire [26] collected data about sex, age, ethnicity, sociodemographic status, personal and family psychiatric history. Information about lifetime depression (DEPR) also originated from this questionnaire. Lifetime depression was defined as a positive answer to reported lifetime depression in the personal psychiatric history of the background questionnaire. Furthermore, this was validated in a smaller subset of the original sample plus an independent sample by face-to-face diagnostic interviews as described previously [26].

The ID-Migraine questionnaire, a validated migraine-screening tool [27] was used to measure migraine. This questionnaire consists of 3 items measuring main symptoms of migraine

in the last 3 months: photophobia, nausea and disability. Migraine (ID\_MIGR) was defined as 2 or 3 YES answers to the questions about symptoms, and was validated earlier [27].

### Genotyping, quality control and imputation

A genetic saliva sample kit was used to collect DNA data—extracted from buccal mucosa cells [25]. Genotyping was implemented using Illumina's CoreExom PsychChip yielding 573,141 variants with genomic position according to the Genome Reference Consortium Human 37 build (GRCh37/hg19). Details of quality control and imputation were published earlier [28] and were based on an established protocol [29]. Briefly, after restricting the measured variants to biallelic and autosomal polymorphisms, imputation was made using SHAPEIT ([https://mathgen.stats.ox.ac.uk/genetics\\_software/shapeit/shapeit.html](https://mathgen.stats.ox.ac.uk/genetics_software/shapeit/shapeit.html)) and IMPUTE2 ([http://mathgen.stats.ox.ac.uk/impute/impute\\_v2.html](http://mathgen.stats.ox.ac.uk/impute/impute_v2.html)) softwares from the reference data. Quality control (QC) was performed for the different populations (Budapest, Manchester and total sample) separately and started with exclusion of multiallelic variants and variants with an 'info' and 'certainty' score of less than 0.5 and 0.7, respectively. Afterwards, the following thresholds were used for filtering: a minor allele frequency (MAF) of 0.01; iteratively 0.1, 0.05, and 0.01 missingness; a Hardy-Weinberg equilibrium test p-value  $\geq 1 \times 10^{-5}$ ; 0.2 of the squared correlation coefficient ( $R^2$ ) value for LD pruning; and an identical-by-descent ( $\pi$ ) value  $\leq 0.1875$ . Principal components were also calculated to control for genetic heterogeneity. Individuals with uncertain gender (discrepancy between chromosome composition and questionnaire data), or outliers by heterozygosity were also excluded. SNPs were then filtered for the candidate genes/loci reported by Gormley et al. [11] and their 10,000 bp vicinity yielding 36579 candidate polymorphisms in 1815 individuals as a final dataset for subsequent analysis. Further details about quality control steps can be seen in [S1 Appendix](#).

### Statistical analyses

**Plink analysis.** We used PLINK v1.07 analysis program (<https://zzz.bwh.harvard.edu/plink/>), the standard in genetic testing, to calculate Hardy-Weinberg equilibrium for individual SNPs (not on genome-wide level), and to run logistic regression analysis. SNPs yielding a Hardy-Weinberg p-value of less than 0.05 were excluded. In all regression analyses two-tailed tests were used.

ID\_MIGR was our categorical outcome variable. The following covariates were added to every regression analysis: age, sex and the calculated first 10 principal components of the whole genetic dataset to control for population substructure (i.e. hidden subgroups as defined by their genetic compositions). First, we tested the main effects of the selected single nucleotide polymorphisms on migraine. After main effects analysis DEPR was added as an interacting variable to test SNP x DEPR interaction on migraine, representing the biological phenomenon that a SNP acts on migraine differently in those individuals who reported lifetime depression than in those who did not. We performed our analyses in our subsamples (Budapest and Manchester separately) and in the total sample (see [S1](#) and [S2](#) Tables for population descriptors).

In case of multiple genetic association tests, Bonferroni correction is often used to control multiple hypothesis testing [30–32]. In our study, however, we applied a more permissive significance criterion. The reasons were the followings: 1) Bonferroni correction is often criticized for its stringency on the field of genomics [33, 34] 2) most corrections assume independent tests, however, this is not the case in our analysis due to substantial linkage between polymorphisms (e.g. 1p31.1 region is a large haploblock (see Gormley et al. [11], supplementary materials)); 3) in contrast to exploratory, hypothesis-free GWASs, we wanted to confirm and extend robust findings proven by Gormley et al. [11], which allows for less

stringent thresholds [35]; 4) recent modelling demonstrated that application of more permissive p-value thresholds could enhance the detection of true positive signals, which is more important in our replication study [36]; 5) it is well-known that both common migraine and depression are polygenic disorders with very small expected effect sizes for individual SNPs.

Therefore, we used permissive thresholds for p-values and SNPs were considered statistically significant according to the following criteria: 1) significant effect in both subsamples with  $p < 0.05$  and 2) replication of these results in the total sample with the same effect allele and effect direction,  $p < 0.05$ . For comparison, if we only applied the criterion of a 0.05 significance in the total sample, altogether 1286 polymorphisms were significant. By restricting these SNPs to those showing a significant effect in both subsamples and also in the total sample we obtained only six (see [Results](#)). Note, however, that none of the SNPs remained significant using Bonferroni correction either in the main or interaction effect analyses (see also [Limitations](#)).

We also tested, whether any of the polymorphisms that were found significant with respect to migraine in main effect or interaction analyses, have an effect on lifetime depression. Logistic regression analysis was performed with the same covariates as for migraine, in the total sample. The significance threshold was the same as above.

For genotype and phenotype frequencies see [S1](#) and [S2](#) Tables. Standard methods like logistic/linear regression, however, are often inadequate to explore complex relationships such as nonlinear interactions. Advanced analysis tools utilizing machine learning methods may enable a more detailed analysis of relationships, especially in a multivariate case. Previously, we applied Bayesian relevance analysis in order to create a Bayesian direct multimorbidity map of depression and its comorbidities, investigating the direct and non-direct relationships of depression [16]; to characterize the relevance of relationships between depression and several environmental and lifestyle factors [23].

Another aspect of variables in addition to their relevance is their predictive power, which measures how accurately do individual or multiple variables in a model predict a target variable. In the present scenario this can be stated as: how accurately do genetic variables (SNPs) and the presence (or absence) of depression (as a comorbid disease) predict the migraine status of patients. For this purpose, we utilize a neural network classifier which is a universal approximator, capable of learning multivariate nonlinear relationships given adequate sample size. The application of such methods for the analysis of genetic associations is not as frequent as the presence of complex relationships. One of the goals of this study is to show that a detailed analysis requires additional tools besides the standard methods.

**Bayesian relevance analysis.** As a post-hoc test, we applied a Bayesian systems-based method called Bayesian relevance analysis [22], which relies on Bayesian model averaging [37]. The method considers possible variable dependency models fitting the data using a Markov chain Monte Carlo method to perform a random walk in the space of possible dependency models that are directed acyclic graphs [38], and assesses the probability of models and model properties [39]. The number of 'burn-in' steps were 1,000,000 followed by 2,000,000 additional steps starting from a randomly initialized structure, using 4 chains. The structure prior was set to uniform, and a Cooper-Herskovits prior (of the Bayesian Dirichlet prior family) was utilized as the parameter prior. Possible directed acyclic graph structures were limited such that each node could have at most 5 incoming edges (parents).

Variables that are directly connected to a selected target variable or those that form an interaction pattern with the target variable (please, note, this is not the same interaction, like the DEPR x SNP interaction parameter used above) in the majority of possible dependency models, can be considered strongly relevant [40].

We have to note that the Bayesian relevance analysis method tests all the SNPs jointly in a multivariate model and automatically corrects for multiple hypothesis testing, however, the method cannot correct automatically for a strong correlation of SNPs. Therefore, we used PLINK's built-in clumping method to select important variants from relevant blocks in high LD. For clumping, the only relevant criterion was that from each block with an  $R^2$  higher than 0.6 only one SNP can be selected (for clumping results see [S9 Table](#)). Please, note, that throughout the manuscript relevance values represent relevance of the given variant with respect to migraine.

For additional information on Bayesian relevance analysis and strong relevance, see [S2 Appendix](#).

**Analysis of predictive power.** We also utilized a neural network based classifier to assess the predictive power of relevant SNPs as a multivariate model. We created a neural network in the Tensorflow framework [41] which consisted of a fully connected input layer and a hidden layer using the most commonly used rectified linear unit (relu) activation function, and an output layer. Based on multiple experiments with various architectures, we found that more than 10 neurons per layer did not improve classification results. Therefore, we opted to use 10 neurons in both the input and hidden layers. Weights and biases were randomly initialized according to the uniform distribution. Since the target variable was binary, we utilized binary-cross entropy as a loss function. The ADAM method [42] was used as an optimizer with a batch size of 20 samples, and 50 as the number of learning epochs.

Predictive power was determined according to a weighted accuracy score involving sensitivity and specificity measures taking into account the case-control ratio. All measures were computed using a cross-validation framework of  $k = 10$  partitions and  $n = 10$  repetitions, i.e. the data was split into  $k = 10$  partitions and for each classifier learning phase 9 partitions were used as training data and 1 partition as testing data, cycling through all possible combinations. This process starting with a partitioning was repeated  $n = 10$  times. In other words, the average of each measure computed  $n \times k$  (i.e. 100) times was used for the final evaluation. The same clumping method was used to filter SNPs prior to predictive power analysis as in the case of relevance analysis.

## Functional prediction

Functional characterization of the significant SNPs was performed with Functional Prediction tool, FuncPred (<https://snpinfo.niehs.nih.gov/snpinfo/snpfunc.html>) and with SNPnexus (<https://www.snp-nexus.org/v4/>). SNPnexus provided an opportunity to get a more detailed description of functionally relevant SNPs.

## Results

### Logistic regression results for main effect analysis

Among the 36579 polymorphisms only 6 showed significant ( $p < 0.05$ ) main effects for migraine that replicated in both subsamples (Budapest, Manchester) and in the total sample. From the six polymorphisms 5 (rs11209657, rs6686879, rs77864828, rs12090642, rs72948266) were located in the intergenic regions on the short arm of the first chromosome (1p31.1). Among them, based on the odds ratios (ORs) of the individual polymorphisms, 3 (rs77864828, rs12090642, rs72948266) represented protective factors ( $OR < 1$ ) and 2 (rs11209657, rs6686879) could be considered as risk factors for migraine ( $OR > 1$ ). The only polymorphism that could be associated with a gene was rs2455107 in the intronic region of *PRDM16* showing an OR higher than 1. For detailed results, see [Table 1](#) and [S3–S5 Tables](#). None of the above mentioned polymorphisms showed main effects with respect to lifetime depression.

**Table 1. Significant genetic polymorphisms in main effects logistic regression analysis for migraine in the subsamples (Budapest, Manchester) and the total sample.**

Variant name	Function prediction based on UCSC	Chromosome	Reference allele	Effect Allele	Associated gene / localization	Results of logistic regression models in main effect analyses					
						Budapest		Manchester		Total sample	
						p-value	OR	p-value	OR	p-value	OR
rs2455107	intron variant	1	A	C	<i>PRDM16</i>	0.0442	1.384	0.0463	1.288	0.0072	1.304
rs11209657	intergenic	1	G	A	chr1.p31.1	0.0477	1.298	0.0342	1.255	0.003	1.277
rs6686879	intergenic	1	G	A	chr1.p31.1	0.0477	1.298	0.0342	1.255	0.003	1.277
rs77864828	intergenic	1	C	T	chr1.p31.1	0.0295	0.3546	0.0478	0.526	0.0048	0.4756
rs12090642	intergenic	1	T	C	chr1.p31.1	0.0499	0.4215	0.0394	0.5128	0.0064	0.4956
rs72948266	intergenic	1	A	G	chr1.p31.1	0.0499	0.4215	0.0394	0.5128	0.0064	0.4956

**Table 1** shows significant polymorphisms, their respective functions based on the UCSC Genome Browser, chromosome localization, reference and effect allele, their associated gene or more exact genomic region, and the p-values and odds ratios (ORs) from the main effects logistic regression with age, sex, and 10 principal components as covariates. Note, that only one polymorphism could be associated with a gene, namely, *PRDM16* and all other variants are intergenic from a large LD block (p31.1) on the first chromosome. From the 38 investigated loci only these replicated according to our criteria.

<https://doi.org/10.1371/journal.pone.0261477.t001>

### Logistic regression results for interaction effect analysis

51 polymorphisms showed a replicated ( $p < 0.05$  in both subsamples and total sample) effect in interaction with lifetime depression for migraine, from which 26 could be associated with *ADGRL2* (also known as *LPHN2*) on chromosome 1, with rs7412827 being a downstream gene variant, while the others were intron variants. Only 2 of them (rs11163394, rs3790895) could be considered protective. We also identified 22 intergenic risk polymorphisms on chromosome 1 in interaction with lifetime depression. One variant (rs1043215) in the 3' untranslated region and another downstream variant (rs143167654) of *REST* (also known as *NRSF*) showed surprisingly high ORs for migraine in the interaction analyses [total sample OR: 6.596 (95% CI: 1.90–22.86), Manchester OR: 5.769 (95% CI: 1.07–31.04), Budapest OR: 10.5 (95% CI: 1.37–80.38) for both SNPs]. Rs1889974 of *HPSE2*, also showed significant effects. Clumped results representing the regions were collected in **Table 2**, for full results see **S6–S8** Tables. None of the polymorphisms showed main effects for lifetime depression.

**Table 2. Clumped results representing significant regions showing interaction with lifetime depression for migraine by logistic regression analysis.**

Variant name	Function prediction based on UCSC	Chromosome	Reference allele	Effect allele	Associated gene / localization	Results of logistic regression models in interaction effect analyses					
						Budapest		Manchester		Total sample	
						p-value	OR	p-value	OR	p-value	OR
rs11163394	intron variant	1	G	A	<i>ADGRL2</i>	0.0491	0.5598	0.0015	0.4946	0.002	0.6079
rs6598982	intergenic	1	T	C	chr1.p31.1	0.0194	1.998	0.0024	2	0.0002	1.837
rs12128399	intergenic	1	G	T	chr1.p31.1	0.0116	2.268	0.0045	2.159	0.0004	1.923
rs12129408	intergenic	1	A	G	chr1.p31.1	0.0313	1.899	0.0274	1.67	0.02093	1.462
rs6660757	intergenic	1	A	C	chr1.p31.1	0.019	1.919	0.01	1.767	0.0001841	1.833
rs1043215	3' UTR variant	4	G	A	<i>REST</i>	0.0235	10.5	0.0412	5.769	0.002939	6.596
rs1889974	intron variant	10	G	A	<i>HPSE2</i>	0.019	1.985	0.0265	1.647	0.002424	1.635

**Table 2** shows significant polymorphisms, their respective functions based on UCSC Genome Browser, chromosome localization, reference and effect allele, their associated gene or more exact genomic region, and the p-values and odds ratios (ORs) in interaction with lifetime depression for migraine using logistic regression with age, sex, and 10 principal components as covariates.

<https://doi.org/10.1371/journal.pone.0261477.t002>

**Table 3. Posterior probability of relevance with respect to migraine for SNPs with main effects.**

Variant name	Relevance	Chromosome	Associated gene
rs77864828	0.421	1	intergenic
rs2455107	0.191	1	<i>PRDM16</i>
rs11209657	0.138	1	intergenic

**Table 3** shows posterior probabilities of main effect SNPs using a Bayesian relevance analysis. A higher posterior probability indicates higher relevance with respect to migraine and thus, serves as a post-hoc test for the already significant polymorphisms identified by regression models. Interestingly, the intergenic polymorphism rs77864828 shows a larger posterior probability, than the only gene-associated polymorphism rs2455107, *PRDM16*. Note that this method investigates possible multivariate models, i.e. it tests the SNPs jointly.

<https://doi.org/10.1371/journal.pone.0261477.t003>

### Bayesian relevance analysis

To determine which of the identified polymorphisms are the most relevant for migraine we also ran a Bayesian relevance analysis investigating multivariate dependency structures on our total sample including sex, age, and a population descriptor (Budapest or Manchester). Significant SNPs from main effect analyses were checked for linkage and only the most representative SNPs were included in the Bayesian analysis (rs2455107, rs77864828, and rs11209657). Results for representative SNPs with main effects are presented in [Table 3](#).

To distinguish between SNPs that showed significant effects in an interaction model, we performed Bayesian relevance analysis separately in depressed and non-depressed subjects on a clumped SNP set and calculated differences between relevance scores (i.e. posterior probabilities) (see [Table 4](#)). Four out of these polymorphisms showed at least a magnitude larger probability of relevance in depressed subjects than in non-depressed subjects and with the exception of rs11163394 all had an OR larger than 1 in regression models. The largest difference in relevance was found in the case of intergenic rs12128399 from the 1p31.1 region with low probability of relevance in non-depressed and high probability of relevance in depressed subjects on migraine. Rs12128399 was followed by rs6660757 in the p31.1 region, rs1889974 in *HPSE2*, rs1043215 in *REST* and rs11163394 in *ADGRL2* sorted by the relevance difference between depressed and non-depressed individuals.

**Table 4. Posterior probability of relevance of SNPs with respect to migraine in interaction with lifetime depression.**

Variant name	Relevance in non-depressed subjects	Relevance in depressed subjects	Difference in relevance between depressed and non-depressed subjects	Chromosome	Associated gene
rs12128399	0.025	0.929	0.904	1	Intergenic
rs6660757	0.210	0.997	0.787	1	Intergenic
rs1889974	0.047	0.704	0.657	10	<i>HPSE2</i>
rs1043215	0.290	0.907	0.617	4	<i>REST</i>
rs11163394	0.014	0.311	0.298	1	<i>ADGRL2</i>
rs6598982	0.013	0.059	0.047	1	Intergenic
rs12129408	0.0056	0.0062	0.0006	1	Intergenic

In [Table 4](#), a higher posterior probability value indicates higher relevance with respect to migraine, shown separately for non-depressed and depressed subjects. A difference between relevance values of a SNP indicates that the SNP plays different roles in non-depressed and depressed subjects and, thus, confirms gene-disease interaction. With the exception of the intergenic rs12129408 where the difference is negligible, all of the polymorphisms show larger relevance in depressed individuals, suggesting that these polymorphisms are more likely to contribute to migraine in depressed subjects. Please, note that the model tests all SNPs at the same time. The low performance of rs12129408 is a result of multivariant effects discussed in the Limitations section.

<https://doi.org/10.1371/journal.pone.0261477.t004>



**Table 5. Predictive power of multivariate models compared to a baseline containing age, sex and population variables.**

	Models	Accuracy	
		Score	Difference
M0	Age, sex, population	0.565	n/a
M1	Age, sex, population, rs2455107, rs11209657, rs77864828	0.664	17.56%
M2	Age, sex, population, lifetime depression	0.709	25.45%
M3	Age, sex, population, lifetime depression, rs11163394, rs6598982, rs12128399, rs12129408, rs6660757, rs1043215, rs1889974	0.767	35.76%

**Table 5** shows the predictive score calculated by a neural network based classifier and the relative difference from the base model using only age, sex and population to predict migraine. The predictive score is a weighted average of sensitivity and specificity measures computed using the total sample. All scores are compared to the score of the baseline model M0. In the M1 model, the significant main effect polymorphisms could enhance predictive power, nevertheless, the addition of the lifetime depression variable without any genetic variants (M2) achieved better performance. The further addition of the interaction polymorphisms to M2 (M3) yielded somewhat better results, showing that lifetime depression in itself is one of the best predictors of migraine, and probably reflecting the large effect size difference between lifetime depression and genetic variants.

<https://doi.org/10.1371/journal.pone.0261477.t005>

### Predictive power analysis

We also investigated the predictive performance of three multivariate models related to either the main effects or the interactions of SNPs with respect to migraine (see [Table 5](#)). All models were compared to a baseline model (M0) consisting of age, sex and population. In this analysis, the population variable was selected instead of the 10 principal components for two reasons: 1) previously, we found that the first principal component is congruous with the population variable, and the addition of further principal components to the analysis does not change significantly the lambda-value of genomic inflation factor [43]; 2) this way, the results of the two machine learning algorithms are more comparable since the Bayesian relevance analysis can only handle categorical variables. The first model (M1) contained M0 and the SNPs with main effects with respect to migraine, the second model (M2) M0 and lifetime depression, and the third model (M3) those SNPs that were found relevant with respect to migraine in interaction with lifetime depression and can be considered as representative SNPs of their respective locality.

Main effect SNPs could enhance predictive power by 17.56% when compared to the M0 model. Inclusion of lifetime depression enhanced the predictive power by 25.45% when compared to M0. Inclusion of the representative interaction SNPs in addition to lifetime depression, slightly improved predictions, resulting in a better score by 35.76% than that of M0 (and a 10.31% better score if compared to M2).

To confirm results of the predictive power analysis, we have investigated logistic regression models with the same terms used in models M0-M3 described in the predictive power analysis section; migraine served as the target variable. Specifically, M0: sex, age, population; M1: M0 + rs2455107, rs11209657, rs77864828; M2: M0 + DEPR; M3: M2 + rs12129408, rs6598982, rs11163394, rs12128399, rs1889974, rs1043215, rs6660757. Note that in case of M3 the interaction terms (formed between a SNP and depression) are modelled explicitly, e.g. DEPR: rs12129408.

Results are shown in [S13 Table](#). In the case of M1, all 3 SNPs showing a main effect can be considered as significant terms of the logistic regression model. Furthermore, the depression descriptor variable in model M2 is highly significant (p-value: 1.39E-12), which further confirms that depression can be considered a relevant factor in the etiology of migraine. Model

M3 contains the main effects and interactions terms of SNP found significant in interaction with depression. Among the 7 interaction terms only 3 are significant (DEPR:rs6660757, p-value: 6.75E-05; DEPR:rs1043215, p-value: 0.00275; rs1889974, p-value: 0.0096).

Due to the nature of the multivariate model, it is possible that given the effect of the most significant terms, other terms might appear insignificant. Therefore, we investigated each SNP individually in the sense that we formed variations of the M3 model such that only a single SNP (main effect and interaction with depression) was admitted to the model besides age, sex, population and depression. [S14 Table](#) shows the significance of interaction terms of such models. Results indicate that in each case the interaction term is significant to some extent even for those SNPs whose interaction term was not significant in the complex M3 model.

In addition, we compared the performance of the models using the Akaike information criterion (AIC). [S15 Table](#) shows the AIC score for each of the models. AIC is an estimator of prediction error, a lower value is considered better, furthermore it includes a penalty term for model complexity. Results confirm the outcome of predictive power analysis. M0 has the highest score (AIC: 2033.7) outperformed by M1 (AIC:2017.5) containing additional SNPs with main effects with respect to migraine. Model M2 containing the depression descriptor (AIC: 1984.9) is even better than M1 due to the significant effect of depression and the decreased complexity with respect to M1 (i.e. there are fewer variables in the model). Finally, M3 has to lowest score (AIC: 1954.3), indicating that the interaction of depression and the selected SNPs may play an important role in predicting migraine susceptibility.

## Functional prediction results

Functional prediction results can be seen in [S10–S12 Tables](#). Among the significant polymorphisms of the main effect analysis, several showed indicative scores for functional effect. For example, rs6686879 showed high score based on gene annotation, while rs12090642 could have an effect as a transcription start site or as an eQTL variant (both SNPs are situated in the 1p31.1 region). Among the hits of interaction analysis, all variants showed scores suggesting functional effect, especially rs1043215 from *REST* gene.

## Discussion

Our analyses confirmed main effects of *PRDM16* and 1p31.1 SNPs on migraine in two independent European cohorts. Furthermore, this study also provides evidence that SNPs of the previously identified migraine risk genes *REST*, *ADGRL2*, *HPSE2* and 1p31.1 show lifetime depression-dependent associations with the disease. The neural network-based classifier proved the predictive value of both main and interaction effect SNPs on migraine, while Bayesian relevance analysis showed large relevance difference of the interaction SNPs between depressed and non-depressed migraineurs.

## Main effects

In 2016 Gormley and colleagues identified 38 credible sets of SNPs that may have a genetic impact on migraine [11]. To replicate the findings, we extracted 36579 polymorphisms (from them 25161 belonged to 1p31.1, a single, huge LD-block according to Gormley and colleagues [11]) that were in the vicinity of these loci. Our main effect analysis for migraine demonstrated only 6 nominally significant SNPs, and among them only one could be associated with a gene, the intron region of *PRDM16*: rs2455107. This polymorphism showed previously no associations, whatsoever, with migraine or any other diseases. At the same time, different variants of *PRDM16* associated with migraine in several earlier studies [11].

Rs2651899 of *PRDM16* showed associations in the Women's Genome Health Study (WGHS) with migraine compared to non-migraineurs and also with migraine compared to non-migraine type headache [44]. A replication study on a Chinese population supported a role for rs2651899 as a risk factor in migraine without aura regardless of gender [45] and its role was also demonstrated in two North Indian, a Chinese Han, a Spanish, a Swedish, and a Pakistani population [46–51]. A recent meta-analysis also demonstrated an association between rs2651899 and migraine risk [52]. Risk allele carriers at rs2651899 also associated with larger response to serotonin 1B/1D receptor agonist antimigraine medications [53]. In contrast, Gormley et al. [11] suggested that rs10218452 and rs12135062 of *PRDM16* are responsible for main effects with respect to migraine.

Yet, we know little about the gene itself. *PRDM16* is implicated in brown adipose tissue differentiation that is a major contributor to thermoregulation and is capable for heat production [54, 55] (also according to the Kyoto Encyclopedia of Genes and Genomes (KEGG): [https://www.kegg.jp/kegg-bin/search\\_pathway\\_text?map=&keyword=PRDM16&mode=1&viewImage=true](https://www.kegg.jp/kegg-bin/search_pathway_text?map=&keyword=PRDM16&mode=1&viewImage=true)). A case study reported association between body temperature and headache intensity in migraineurs with aura [56] and a hypothesis by Horvath suggests a causal role for altered thermoregulation in migraine [57]. From another perspective, *PRDM16* was shown to be necessary for adult neurogenesis (also see the AmiGO 2 database (Gene Ontology): [http://amigo.geneontology.org/amigo/gene\\_product/UniProtKB:Q9HAZ2](http://amigo.geneontology.org/amigo/gene_product/UniProtKB:Q9HAZ2)) and ependymal cell differentiation in the adult mouse brain after deletion of *PRDM16* with the Nestin-Cre method. In the latter study, the few remaining ciliated ependymal cells were abnormal and could not secure the proper flow of the cerebrospinal fluid (CSF) [58]. In chronic migraine, concentrations of several biomarkers are altered in CSF, e.g. glutamate or calcitonin-gene related peptide to name a few [59] that may connect altered ependymal cell function with the disease.

However, these links remain strongly hypothetical. Nonetheless, previous and current results indicate that polymorphisms of *PRDM16* may be consistent and relevant candidates for migraine susceptibility. The redundancy of the polymorphisms on the gene level, or allelic heterogeneity, suggests that altered gene functions or expression may be more important than individual variants within *PRDM16*.

**Main effects—Bayesian relevance analysis and analysis of predictive power.** Our post-hoc Bayesian relevance analysis showed the following order of relevance: 1) rs77864828 of 1p31.1 region; 2) rs2455107 of *PRDM16* and 3) rs11209657 of 1p31.1 region. Furthermore, inclusion of these 3 representative SNPs in the neural network classifier enhanced predictive power by 17.56% which is large considering that individual genetic variants have usually very small effects and the inclusion of lifetime depression could enhance predictive power within the same magnitude (25.45%). Gormley et al. [11] identified the 1p31.1 region as a risk locus and we can confirm these findings. Recent studies show that intronic and intergenic regions are responsible for harboring vast numbers of non-coding RNAs modulating gene transcription [60, 61]. These results point out that the fine modulation of transcription might be responsible for common migraine instead of large impact mutations.

In summary, main effect analyses including regression models and neural network based predictive power calculations indicated that rs2455107 of *PRDM16* and intergenic rs11209657, rs77864828 of 1p31.1 are potential candidate polymorphisms behind migraine. Other loci proposed by Gormley et al. [11] could not be replicated in our sample.

## Interaction effects

Interestingly, none of the polymorphisms that were significant in interaction effect analysis with lifetime depression showed significance in main effect results. Previous studies already

suggested a substantial difference between migraine with and without depression [6, 17]. A twin study demonstrated that migraine and anxious depression may be in a causal relationship, though the direction of such causality remained undetermined [17]. The authors implied that migraine with and without depression can be regarded as different phenotypes and refuted the idea of a shared etiology [17]. In further support, according to polygenic risk scores, migraine with depression showed higher resemblance to depression than to migraine [6]. Our study adds further evidence for a substantial difference since there were candidate SNPs elevating the risk of migraine with comorbid lifetime depression that were not involved in mediating migraine in general. These results implicate that migraine with and without lifetime depression may develop on different genetic background, and that large GWASs might find SNPs relevant in specific subgroups of the disorder. In the future, these SNPs can be used for personalized risk/phenotype prediction (like in [62, 63]). The regression analysis identified 51 SNPs from three genes. We discuss them in the context of the results from the two machine learning methods.

#### **Interaction effects—Bayesian relevance analysis and analysis of predictive power.**

Post-hoc analysis showed that among the identified SNPs the two most relevant ones were rs12128399 and rs6660757 of p31.1, both are intergenic and their exact role is at present unknown.

The third polymorphism, rs1889974, based on relevance difference between depressed and non-depressed migraineurs, is located in the intron of *HPSE2*, encoding the heparanase-2 (Hpa2) enzyme [64]. Hpa2 was identified as a novel member of the heparanase family [64]. Heparanases are endoglycosidases cleaving the heparan sulfate (HS), which leads to the remodelling of the extracellular matrix [64], and is connected to tumor metastasis, angiogenesis and inflammation, but Hpa2 does not exhibit this HS-degrading activity and is not a subject of proteolytic processing [65]. Its wild type (Hpa2c) has an inhibitory effect on heparanase enzymatic activity, likely through Hpa2c's high affinity to HS and heparin [65]. Our result with *HPSE2* might support the vascular theory of migraine. A prothrombotic tendency has been suggested in migraine pathogenesis by the connection between migraine and cardio- and cerebrovascular ischemic events, although related studies show conflicting data [66–68] and the literature of anticoagulants' efficacy in migraine patients also shows mixed results [66]. Depression has been connected to cardiovascular diseases too [69, 70] and the 'vascular depression' hypothesis also links cerebrovascular disease to geriatric depressive syndromes [71, 72]. In addition, *HPSE2* associated with the depression-related trait, neuroticism [73]. These results may represent a lead connecting the gene with migraine and depression.

The fourth largest relevance difference was found in case of rs1043215, which is located within the *REST* gene, an important regulator of gene expression with the ability to bind RE1 (also known as NRSF) in the promoter region of more than 2000 genes [74–76]. In non-neuronal cells *REST* is responsible for the suppression of neuron specific gene expression during embryogenesis [77]. In adult neuronal cells *REST* is expressed at low levels and is induced by different types of neuronal activation [78]. Among the regulated genes in neuronal cell lines were synapsin, synaptophysin, Crh, *BDNF*, and the serotonin 1A receptor [74, 79, 80], which are highly relevant for depression and antidepressant response [81]. In fact, aberrant *REST* transcriptional regulation was shown in patients with major depressive disorder in current depressive state, but not in remission [79]. While it remained so far uninvestigated with respect to migraine directly, *REST* was also involved in neuropathic pain symptoms [82]. Furthermore, decreased *REST* levels associated with increased neuronal excitability and cortical activity that may be directly related to elevated pain sensations [83]. The two identified polymorphisms of *REST* showed large ORs for migraine. The relevant SNP in the Bayesian analysis, rs1043215 in the 3' untranslated region was uninvestigated previously. However, our

functional prediction analysis showed that this SNP is functional with great probability, and it was detected as a highly conserved SNP suggesting an evolutionary importance for this variant. The above discussed empirical evidences and the present study delineate *REST* polymorphisms, especially rs1043215, as possible candidates for further studies in migraine with depression.

The next SNP according to relevance difference, rs11163394, belongs to an intronic region of *ADGRL2* from 1p31.1 encoding latrophilin 2, an adhesion G-protein coupled receptor [84]. Latrophilin receptors are involved in the maintenance of synapses and *ADGRL2* in particular is involved in the intracellular release of Ca-ions [85]. While the latrophilin 1 and 3 receptors were proposed as contributors to the regulation of mood [85, 86], *ADGRL2* remains hitherto less investigated. The reason was probably the relatively low expression of *ADGRL2* compared to other latrophilins, within the brain [87], however, it cannot be excluded that *ADGRL2* is also involved in the regulation of mood. Elevated excitability of tissues via calcium-channelopathies were already proposed in migraine [88]. In certain cases, familial hemiplegic migraine is caused by mutations in the calcium voltage-gated channel subunit alpha1 A (*CACNA1A*), and it was also proposed that physiologic cation concentrations may be altered in other forms of the disease [88, 89]. Our study suggests that via variation in the *ADGRL2* gene, intracellular calcium levels might be elevated, resulting in a hyperexcitability similar to that seen in the familial form of the disease. While it may be premature to draw conclusions about the role of *ADGRL2* in migraine with accompanying depression, the 26 different polymorphisms associated with *ADGRL2* show that the gene might represent an important novel target of migraine research.

The last two variants based on relevance difference were rs6598982 and rs12129408 from the 1p31.1 region. Our functional prediction analysis showed high scores for rs6598982 in some categories suggesting high probability for being an eQTL, a trait-associated and inherited disease-associated variant. Furthermore, it has to be noted that non-coding *LINC01360* and *RN7SKP19* pseudogene in the region associated with depression and anxiety disorder in GWASs [90–93]. From the perspective of the predictive power, inclusion of these interacting SNPs in the model slightly improved its prediction value compared to adding only lifetime depression (10.31% gain).

In summary, interaction analysis indicated that intergenic rs12128399, rs6660757 and variants in *HPSE2*, *REST* and from the 1p31.1 region (including *ADGRL2*) represent lifetime depression-dependent genetic factors for migraine.

## Limitations

Our study has to be interpreted in light of the following limitations. First, our sample is small compared to the original study. The lack of large statistical power may explain why we couldn't replicate many loci from the Gormley study. To remain rigorous, our significance criterion was to replicate results in both subsamples and the total sample with nominal significance instead of the Bonferroni correction used in large genetic studies. We applied two machine learning methods leading to further validation and characterization of results: the Bayesian relevance analysis validated relevance of interaction results, while the predictive power analysis confirmed both main effect and interaction polymorphisms. The lack of Bonferroni correction remains a major limitation of the study. We already discussed why we choose more permissive p-value thresholds at the relevant section of Methods.

Second, we also have to note that the two subsamples differ in some characteristics (see S1 and S2 Tables), but exactly for this very reason, we believe that those SNPs that were able to replicate are indeed important.

Third, we measured migraine and lifetime depression with questionnaires. The migraine questionnaire was already validated against a clinical sample as demonstrated by Lipton and colleagues [27, 94]. The latter is also true for the depression questionnaire, which was validated by us, as indicated in the methods section [95].

Fourth, the polymorphism rs12129408 showed risk OR ( $> 1$ ) in the PLINK analysis, but the difference between relevance in depressed and non-depressed individuals was negligible. This can be explained by the different aims of the methods. The PLINK analysis assesses effect sizes and based on these estimations calculates significance. At the same time, Bayesian relevance analysis addresses whether a given polymorphism has a direct effect on the outcome variable, but does not state whether this effect is protective or not.

Fifth, the minor allele frequency of the *REST* gene polymorphism was low in our sample that might explain the large ORs seen in the analyses (see [S2 Table](#)). Therefore, these results should be interpreted with caution.

Finally, predictive power analysis could confirm the effect of main effect SNPs, and also the effect of interaction SNPs to some extent, while Bayesian relevance analysis validated interaction effects only. It has to be noted, that these two methods answer different questions about the roles of these SNPs (predictive power and direct effects, respectively) and, therefore, it is not surprising that different results were obtained.

## Conclusions

The elaborate analysis of 38 risk loci for migraine already proposed by the GWAS of Gormley et al. [11] could identify only 57 nominally significant variants from altogether four genes/regions (*PRDM16*, *REST*, *ADGRL2*, *HPSE2*). Main effect analysis indicated that intergenic polymorphism rs77864828, and with smaller relevance, intron variant rs2455107 in *PRDM16* represent genetic factors for common migraine. In contrast, intergenic rs6660757 and rs12128399, rs1889974 in *HPSE2*, rs11163394 in *ADGRL2* and rs1043215 of *REST* may only serve as biomarkers in migraine accompanied by lifetime depression. In addition, while most of the above genes could be connected to known migraine pathophysiology, the 1p31.1 region harboring significant intergenic SNPs also assumes a role for non-coding RNAs behind common migraine.

The relative scarcity of significantly validated variants emphasize the need of validation even for large GWAS results especially through the inclusion of interaction factors. It also reveals that even significantly replicated genetic variants may be involved in different disease subtypes, in this particular case, migraine without and with accompanying lifetime depression.

## Supporting information

**S1 Fig. Genomic location of the significant SNPs in *PRDM16* gene.**  
(PDF)

**S2 Fig. Genomic location of the significant SNPs from 1.p31.1 region.**  
(PDF)

**S3 Fig. Genomic location of the significant SNPs near *REST* gene.**  
(PDF)

**S4 Fig. Genomic location of the significant SNPs in *HPSE2* gene.**  
(PDF)

**S1 Table. Summary statistics.**  
(PDF)

**S2 Table. Minor allele frequencies.**

(PDF)

**S3 Table. Results for main effect term in Budapest subsample.**

(PDF)

**S4 Table. Results for main effect term in Manchester subsample.**

(PDF)

**S5 Table. Results for main effect term in total sample.**

(PDF)

**S6 Table. Results for interaction term in Budapest subsample.**

(PDF)

**S7 Table. Results for interaction term in Manchester subsample.**

(PDF)

**S8 Table. Results for interaction term in total sample.**

(PDF)

**S9 Table. Formatted results of the clumping procedure as performed in Plink on the total sample.**

(PDF)

**S10 Table. Results of the functional characterization with GWAVA non-coding scoring algorithm.**

(PDF)

**S11 Table. Results of the functional characterization with DeepSEA non-coding scoring algorithm.**

(PDF)

**S12 Table. Results of the functional characterization with FunSeq2 non-coding scoring algorithms.**

(PDF)

**S13 Table. Multivariate logistic regression models corresponding to models M0-M3 used for assessing predictive power.**

(PDF)

**S14 Table. Significance of interaction terms in variations of logistic regression model M3 each containing only one relevant SNP.**

(PDF)

**S15 Table. Evaluation of logistic regression models M0-M3 using Akaike information criterion (AIC).**

(PDF)

**S1 Appendix. Quality control (QC) steps.**

(PDF)

**S2 Appendix. Additional information on methods.**

(PDF)

## Author Contributions

**Conceptualization:** Peter Petschner, Daniel Baksa, Gabriella Juhasz.

**Data curation:** Peter Petschner, Daniel Baksa, Gabor Hullam, Dora Torok, Andras Millinghoffer.

**Formal analysis:** Peter Petschner, Daniel Baksa, Gabor Hullam, Dora Torok, Andras Millinghoffer.

**Funding acquisition:** J. F. William Deakin, Gyorgy Bagdy, Gabriella Juhasz.

**Investigation:** J. F. William Deakin, Gyorgy Bagdy, Gabriella Juhasz.

**Methodology:** Peter Petschner, Daniel Baksa, Gabor Hullam, Dora Torok.

**Supervision:** Gabriella Juhasz.

**Visualization:** Dora Torok.

**Writing – original draft:** Peter Petschner, Daniel Baksa, Gabor Hullam, Dora Torok, Andras Millinghoffer, J. F. William Deakin, Gyorgy Bagdy, Gabriella Juhasz.

**Writing – review & editing:** Peter Petschner, Daniel Baksa, Gabor Hullam, Dora Torok, Andras Millinghoffer, J. F. William Deakin, Gyorgy Bagdy, Gabriella Juhasz.

## References

1. World Health Organization. Atlas of headache disorders and resources in the world 2011. Geneva, WHO. 2011.
2. Foster SA, Chen CC, Ding Y, Mason O, McGuinness CB, Morrow P, et al. Economic burden and risk factors of migraine disease progression in the US: a retrospective analysis of a commercial payer database. *Journal of medical economics*. 2020; 23(11):1356–64. Epub 2020/08/28. <https://doi.org/10.1080/13696998.2020.1814790> PMID: 32845189.
3. Irimia P, Garrido-Cumbrera M, Santos-Lasaosa S, Braçe O, Colomina I, Blanch C, et al. Estimating the savings associated with a migraine-free life: results from the Spanish Atlas. *European journal of neurology*. 2020; 27(12):2616–24. Epub 2020/07/10. <https://doi.org/10.1111/ene.14431> PMID: 32643853.
4. Osumili B, McCrone P, Cousins S, Ridsdale L. The Economic Cost of Patients With Migraine Headache Referred to Specialist Clinics. *Headache*. 2018; 58(2):287–94. Epub 2017/11/03. <https://doi.org/10.1111/head.13210> PMID: 29094342.
5. Bloudek LM, Stokes M, Buse DC, Wilcox TK, Lipton RB, Goadsby PJ, et al. Cost of healthcare for patients with migraine in five European countries: results from the International Burden of Migraine Study (IBMS). *The journal of headache and pain*. 2012; 13(5):361–78. Epub 2012/05/31. <https://doi.org/10.1007/s10194-012-0460-7> PMID: 22644214.
6. Ligthart L, Hottenga JJ, Lewis CM, Farmer AE, Craig IW, Breen G, et al. Genetic risk score analysis indicates migraine with and without comorbid depression are genetically different disorders. *Human genetics*. 2014; 133(2):173–86. Epub 2013/10/02. <https://doi.org/10.1007/s00439-013-1370-8> PMID: 24081561.
7. Juhasz G, Lazary J, Chase D, Pegg E, Downey D, Toth ZG, et al. Variations in the cannabinoid receptor 1 gene predispose to migraine. *Neuroscience letters*. 2009; 461(2):116–20. Epub 2009/06/23. <https://doi.org/10.1016/j.neulet.2009.06.021> PMID: 19539700.
8. Juhasz G, Zsombok T, Laszik A, Gonda X, Sotonyi P, Faludi G, et al. Association analysis of 5-HTTLPR variants, 5-HT2a receptor gene 102T/C polymorphism and migraine. *Journal of neurogenetics*. 2003; 17(2–3):231–40. Epub 2003/12/12. PMID: 14668201.
9. Juhasz G, Zsombok T, Laszik A, Jakus R, Faludi G, Sotonyi P, et al. Despite the general correlation of the serotonin transporter gene regulatory region polymorphism (5-HTTLPR) and platelet serotonin concentration, lower platelet serotonin concentration in migraine patients is independent of the 5-HTTLPR variants. *Neuroscience letters*. 2003; 350(1):56–60. Epub 2003/09/10. [https://doi.org/10.1016/s0304-3940\(03\)00834-6](https://doi.org/10.1016/s0304-3940(03)00834-6) PMID: 12962916.
10. Kondratieva N, Azimova J, Skorobogatikh K, Sergeev A, Naumova E, Kokaeva Z, et al. Biomarkers of migraine: Part 1—Genetic markers. *Journal of the neurological sciences*. 2016; 369:63–76. Epub 2016/09/23. <https://doi.org/10.1016/j.jns.2016.08.008> PMID: 27653868.



11. Gormley P, Anttila V, Winsvold BS, Palta P, Esko T, Pers TH, et al. Meta-analysis of 375,000 individuals identifies 38 susceptibility loci for migraine. *Nature genetics*. 2016; 48(8):856–66. Epub 2016/06/21. <https://doi.org/10.1038/ng.3598> PMID: 27322543.
12. Breslau N, Lipton RB, Stewart WF, Schultz LR, Welch KM. Comorbidity of migraine and depression: investigating potential etiology and prognosis. *Neurology*. 2003; 60(8):1308–12. Epub 2003/04/23. <https://doi.org/10.1212/01.wnl.0000058907.41080.54> PMID: 12707434.
13. Chen MH, Pan TL, Lin WC, Huang KL, Hsu JW, Li CT, et al. Bidirectional association between migraine and depression among probands and unaffected siblings: A nationwide population-based study. *Journal of affective disorders*. 2021; 279:687–91. Epub 2020/11/16. <https://doi.org/10.1016/j.jad.2020.10.056> PMID: 33190120.
14. Lipton RB, Seng EK, Chu MK, Reed ML, Fanning KM, Adams AM, et al. The Effect of Psychiatric Comorbidities on Headache-Related Disability in Migraine: Results From the Chronic Migraine Epidemiology and Outcomes (CaMEO) Study. *Headache*. 2020; 60(8):1683–96. Epub 2021/01/16. <https://doi.org/10.1111/head.13914> PMID: 33448374.
15. Minen MT, Begasse De Dhaem O, Kroon Van Diest A, Powers S, Schwedt TJ, Lipton R, et al. Migraine and its psychiatric comorbidities. *Journal of neurology, neurosurgery, and psychiatry*. 2016; 87(7):741–9. Epub 2016/01/07. <https://doi.org/10.1136/jnnp-2015-312233> PMID: 26733600.
16. Marx P, Antal P, Bolgar B, Bagdy G, Deakin B, Juhasz G. Comorbidities in the diseasome are more apparent than real: What Bayesian filtering reveals about the comorbidities of depression. *PLoS computational biology*. 2017; 13(6):e1005487. Epub 2017/06/24. <https://doi.org/10.1371/journal.pcbi.1005487> PMID: 28644851 Bristol-Myers Squibb, AstraZeneca, Eli Lilly, Schering Plough, Janssen-Cilag, and Servier (all fees are paid to the University of Manchester to reimburse them for the time taken); he also has share options in P1vital. The other authors report no conflict of interest.
17. Ligthart L, Nyholt DR, Penninx BW, Boomsma DI. The shared genetics of migraine and anxious depression. *Headache*. 2010; 50(10):1549–60. Epub 2010/06/18. <https://doi.org/10.1111/j.1526-4610.2010.01705.x> PMID: 20553331.
18. Anttila V, Bulik-Sullivan B, Finucane HK, Walters RK, Bras J, Duncan L, et al. Analysis of shared heritability in common disorders of the brain. *Science (New York, NY)*. 2018; 360(6395). Epub 2018/06/23. <https://doi.org/10.1126/science.aap8757> PMID: 29930110.
19. de Boer I, van den Maagdenberg A, Terwindt GM. Advance in genetics of migraine. *Current opinion in neurology*. 2019; 32(3):413–21. Epub 2019/03/19. <https://doi.org/10.1097/WCO.0000000000000687> PMID: 30883436.
20. Schur EA, Noonan C, Buchwald D, Goldberg J, Afari N. A twin study of depression and migraine: evidence for a shared genetic vulnerability. *Headache*. 2009; 49(10):1493–502. Epub 2009/05/15. <https://doi.org/10.1111/j.1526-4610.2009.01425.x> PMID: 19438739.
21. Yang Y, Zhao H, Boomsma DI, Ligthart L, Belin AC, Smith GD, et al. Molecular genetic overlap between migraine and major depressive disorder. *European journal of human genetics: EJHG*. 2018; 26(8):1202–16. Epub 2018/07/12. <https://doi.org/10.1038/s41431-018-0150-2> PMID: 29995844.
22. Antal P, Millinghoffer A, Hullám G, Szalai C, Falus A. A Bayesian View of Challenges in Feature Selection: Feature Aggregation, Multiple Targets, Redundancy and Interaction. In: Yvan S, Huan L, Iñaki I, Louis W, Yves Van de P, editors. *Proceedings of the Workshop on New Challenges for Feature Selection in Data Mining and Knowledge Discovery at ECML/PKDD 2008; Proceedings of Machine Learning Research: PMLR %J Proceedings of Machine Learning Research*; 2008. p. 74–89.
23. Hullam G, Antal P, Petschner P, Gonda X, Bagdy G, Deakin B, et al. The UKB envirome of depression: from interactions to synergistic effects. *Scientific reports*. 2019; 9(1):9723. Epub 2019/07/07. <https://doi.org/10.1038/s41598-019-46001-5> PMID: 31278308 for Bristol-Myers Squibb, AstraZeneca, Eli Lilly, Schering Plough, Janssen-Cilag, and Servier (all fees are paid to the University of Manchester to reimburse them for the time taken); he also has share options in P1vital. The other authors report no conflict of interest.
24. Gonda X, Hullam G, Antal P, Eszlari N, Petschner P, Hökfelt TG, et al. Significance of risk polymorphisms for depression depends on stress exposure. *Scientific reports*. 2018; 8(1):3946. Epub 2018/03/04. <https://doi.org/10.1038/s41598-018-22221-z> PMID: 29500446 Bristol-Myers Squibb, AstraZeneca, Eli Lilly, Schering Plough, Janssen-Cilag and Servier (all fees are paid to the University of Manchester to reimburse them for the time taken); he also has share options in P1vital. I.M.A. received consultancy fees from Alkermes, Lundbeck-Otsuka Ltd., Janssen-Cilag Ltd and fees for speaking from Lundbeck. The remaining authors declare no conflict of interest.
25. Juhasz G, Chase D, Pegg E, Downey D, Toth ZG, Stones K, et al. CNR1 gene is associated with high neuroticism and low agreeableness and interacts with recent negative life events to predict current depressive symptoms. *Neuropsychopharmacology: official publication of the American College of Neuropsychopharmacology*. 2009; 34(8):2019–27. Epub 2009/02/27. <https://doi.org/10.1038/npp.2009.19> PMID: 19242408.

26. Juhasz G, Dunham JS, McKie S, Thomas E, Downey D, Chase D, et al. The CREB1-BDNF-NTRK2 pathway in depression: multiple gene-cognition-environment interactions. *Biological psychiatry*. 2011; 69(8):762–71. Epub 2011/01/11. <https://doi.org/10.1016/j.biopsych.2010.11.019> PMID: 21215389.
27. Lipton RB, Dodick D, Sadovsky R, Kolodner K, Endicott J, Hettiarachchi J, et al. A self-administered screener for migraine in primary care: The ID Migraine validation study. *Neurology*. 2003; 61(3):375–82. Epub 2003/08/13. <https://doi.org/10.1212/01.wnl.0000078940.53438.83> PMID: 12913201.
28. Eszlari N, Millinghoffer A, Petschner P, Gonda X, Baksa D, Pulay AJ, et al. Genome-wide association analysis reveals KCTD12 and miR-383-binding genes in the background of rumination. *Translational psychiatry*. 2019; 9(1):119. Epub 2019/03/20. <https://doi.org/10.1038/s41398-019-0454-1> PMID: 30886212 and speaking engagements (all fees are paid to the University of Manchester to reimburse them for the time taken) for AstraZeneca, Autifony, Bristol-Myers Squibb, Eli Lilly, Janssen-Cilag, P1vital, Schering Plough, and Servier. The remaining authors declare no conflict of interest.
29. Coleman JR, Euesden J, Patel H, Folarin AA, Newhouse S, Breen G. Quality control, imputation and analysis of genome-wide genotyping data from the Illumina HumanCoreExome microarray. *Briefings in functional genomics*. 2016; 15(4):298–304. Epub 2015/10/08. <https://doi.org/10.1093/bfpg/elv037> PMID: 26443613.
30. Dudbridge F, Gusnanto A. Estimation of significance thresholds for genomewide association scans. *Genetic epidemiology*. 2008; 32(3):227–34. Epub 2008/02/27. <https://doi.org/10.1002/gepi.20297> PMID: 18300295.
31. Pe'er I, Yelensky R, Altshuler D, Daly MJ. Estimation of the multiple testing burden for genomewide association studies of nearly all common variants. *Genetic epidemiology*. 2008; 32(4):381–5. Epub 2008/03/19. <https://doi.org/10.1002/gepi.20303> PMID: 18348202.
32. Risch N, Merikangas K. The future of genetic studies of complex human diseases. *Science (New York, NY)*. 1996; 273(5281):1516–7. Epub 1996/09/13. <https://doi.org/10.1126/science.273.5281.1516> PMID: 8801636.
33. Conneely KN, Boehnke M. So many correlated tests, so little time! Rapid adjustment of P values for multiple correlated tests. *American journal of human genetics*. 2007; 81(6):1158–68. Epub 2007/10/30. <https://doi.org/10.1086/522036> PMID: 17966093.
34. van den Oord EJ. Controlling false discoveries in genetic studies. *American journal of medical genetics Part B, Neuropsychiatric genetics: the official publication of the International Society of Psychiatric Genetics*. 2008; 147b(5):637–44. Epub 2007/12/20. <https://doi.org/10.1002/ajmg.b.30650> PMID: 18092307.
35. Baker M. Statisticians issue warning over misuse of P values. *Nature*. 2016; 531(7593):151. Epub 2016/03/11. <https://doi.org/10.1038/nature.2016.19503> PMID: 26961635.
36. Chen Z, Boehnke M, Wen X, Mukherjee B. Revisiting the genome-wide significance threshold for common variant GWAS. *G3 (Bethesda, Md)*. 2021; 11(2). Epub 2021/02/16. <https://doi.org/10.1093/g3journal/jkaa056> PMID: 33585870.
37. Hoeting JA, Madigan D, Raftery AE, Volinsky CT. Bayesian Model Averaging: A Tutorial. *Statistical Science*. 1999; 14(4):382–401.
38. Friedman N, Koller D. Being Bayesian about network structure: A Bayesian approach to structure discovery in Bayesian networks. *Machine Learning*. 2003; 50(1–2):95–125. <https://doi.org/10.1023/A:1020249912095>
39. Cooper GF, Herskovits E. A Bayesian method for the induction of probabilistic networks from data. *Machine Learning*. 1992; 9(1–3):309–47. <https://doi.org/10.1007/BF00994110>
40. Hullam G, Juhasz G, Deakin B, Antal P, editors. Structural and parametric uncertainties in full Bayesian and graphical lasso based approaches: Beyond edge weights in psychological networks. 2017 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB); 2017 23–25 Aug. 2017.
41. Abadi M, Barham P, Chen J, Chen Z, Davis A, Dean J, et al. TensorFlow: a system for large-scale machine learning. *Proceedings of the 12th USENIX conference on Operating Systems Design and Implementation*; Savannah, GA, USA: USENIX Association; 2016. p. 265–83.
42. Diederik P. Kingma JB. Adam: A Method for Stochastic Optimization. arXiv:1412.6980. 2014.
43. Juhasz G. Endofenotípusok alkalmazása a depresszió patomechanizmusának megismerésében [Doctor of the Hungarian Academy of Sciences thesis]: Faculty of Pharmacy, Semmelweis University; 2018.
44. Chasman DI, Schürks M, Anttila V, de Vries B, Schminke U, Launer LJ, et al. Genome-wide association study reveals three susceptibility loci for common migraine in the general population. *Nature genetics*. 2011; 43(7):695–8. Epub 2011/06/15. <https://doi.org/10.1038/ng.856> PMID: 2166692.

45. An XK, Ma QL, Lin Q, Zhang XR, Lu CX, Qu HL. PRDM16 rs2651899 variant is a risk factor for Chinese common migraine patients. *Headache*. 2013; 53(10):1595–601. Epub 2013/09/12. <https://doi.org/10.1111/head.12212> PMID: 24021092.
46. Fan X, Wang J, Fan W, Chen L, Gui B, Tan G, et al. Replication of migraine GWAS susceptibility loci in Chinese Han population. *Headache*. 2014; 54(4):709–15. Epub 2014/03/29. <https://doi.org/10.1111/head.12329> PMID: 24666033.
47. Ghosh J, Pradhan S, Mittal B. Genome-wide-associated variants in migraine susceptibility: a replication study from North India. *Headache*. 2013; 53(10):1583–94. Epub 2013/11/26. <https://doi.org/10.1111/head.12240> PMID: 24266335.
48. Ran C, Graae L, Magnusson PKE, Pedersen NL, Olson L, Belin AC. A replication study of GWAS findings in migraine identifies association in a Swedish case–control sample. *BMC Medical Genetics*. 2014; 15(1):38. <https://doi.org/10.1186/1471-2350-15-38> PMID: 24674449
49. Sintas C, Fernández-Morales J, Vila-Pueyo M, Narberhaus B, Arenas C, Pozo-Rosich P, et al. Replication study of previous migraine genome-wide association study findings in a Spanish sample of migraine with aura. *Cephalalgia: an international journal of headache*. 2015; 35(9):776–82. Epub 2014/11/13. <https://doi.org/10.1177/0333102414557841> PMID: 25388962.
50. Kaur S, Ali A, Ahmad U, Pandey AK, Singh B. rs2651899 variant is associated with risk for migraine without aura from North Indian population. *Molecular biology reports*. 2019; 46(1):1247–55. Epub 2019/01/13. <https://doi.org/10.1007/s11033-019-04593-1> PMID: 30635810.
51. Zafar R, Saleem T, Sheikh N, Maqbool H, Mukhtar M, Abbasi MH. PRDM16, LRP1 and TRPM8 genetic polymorphisms are risk factor for Pakistani migraine patients. *Saudi Journal of Biological Sciences*. 2021. <https://doi.org/10.1016/j.sjbs.2021.06.028> PMID: 34588893
52. Lee HH, Chen CC, Ong JR, Lin YF, Lee FP, Hu CJ, et al. Association of rs2651899 Polymorphism in the Positive Regulatory Domain 16 and Common Migraine Subtypes: A Meta-Analysis. *Headache*. 2020; 60(1):71–80. Epub 2019/09/27. <https://doi.org/10.1111/head.13670> PMID: 31557325.
53. Christensen AF, Esserlind AL, Werge T, Stefánsson H, Stefánsson K, Olesen J. The influence of genetic constitution on migraine drug responses. *Cephalalgia: an international journal of headache*. 2016; 36(7):624–39. Epub 2015/10/28. <https://doi.org/10.1177/0333102415610874> PMID: 26502740.
54. Fenzl A, Kiefer FW. Brown adipose tissue and thermogenesis. *Hormone molecular biology and clinical investigation*. 2014; 19(1):25–37. Epub 2014/11/13. <https://doi.org/10.1515/hmbci-2014-0022> PMID: 25390014.
55. Seale P, Bjork B, Yang W, Kajimura S, Chin S, Kuang S, et al. PRDM16 controls a brown fat/skeletal muscle switch. *Nature*. 2008; 454(7207):961–7. Epub 2008/08/23. <https://doi.org/10.1038/nature07182> PMID: 18719582.
56. Ordás CM, Cuadrado ML, Rodríguez-Cambrón AB, Casas-Limón J, del Prado N, Porta-Etessam J. Increase in body temperature during migraine attacks. *Pain medicine (Malden, Mass)*. 2013; 14(8):1260–4. Epub 2013/05/29. <https://doi.org/10.1111/pme.12145> PMID: 23710707.
57. Horváth C. Alterations in brain temperatures as a possible cause of migraine headache. *Medical hypotheses*. 2014; 82(5):529–34. Epub 2014/03/04. <https://doi.org/10.1016/j.mehy.2014.02.004> PMID: 24581675.
58. Shimada IS, Acar M, Burgess RJ, Zhao Z, Morrison SJ. Prdm16 is required for the maintenance of neural stem cells in the postnatal forebrain and their differentiation into ependymal cells. *Genes & development*. 2017; 31(11):1134–46. Epub 2017/07/13. <https://doi.org/10.1101/gad.291773.116> PMID: 28698301.
59. van Dongen RM, Zielman R, Noga M, Dekkers OM, Hankemeier T, van den Maagdenberg AM, et al. Migraine biomarkers in cerebrospinal fluid: A systematic review and meta-analysis. *Cephalalgia: an international journal of headache*. 2017; 37(1):49–63. Epub 2016/02/19. <https://doi.org/10.1177/0333102415625614> PMID: 26888294.
60. Giral H, Landmesser U, Kratzer A. Into the Wild: GWAS Exploration of Non-coding RNAs. *Front Cardiovasc Med*. 2018; 5(181):181. Epub 2019/01/09. <https://doi.org/10.3389/fcvm.2018.00181> PMID: 30619888.
61. Mattick JS, Gagen MJ. The evolution of controlled multitasked gene networks: the role of introns and other noncoding RNAs in the development of complex organisms. *Molecular biology and evolution*. 2001; 18(9):1611–30. Epub 2001/08/16. <https://doi.org/10.1093/oxfordjournals.molbev.a003951> PMID: 11504843.
62. Rajput DS, Basha SM, Xin Q, Gadekallu TR, Kaluri R, Lakshmana K, et al. Providing diagnosis on diabetes using cloud computing environment to the people living in rural areas of India. *Journal of Ambient Intelligence and Humanized Computing*. 2021. <https://doi.org/10.1007/s12652-021-03154-4>
63. Roqueiro D, Witteveen MJ, Anttila V, Terwindt GM, van den Maagdenberg AM, Borgwardt K. In silico phenotyping via co-training for improved phenotype prediction from genotype. *Bioinformatics (Oxford,*

- England). 2015; 31(12):i303–10. Epub 2015/06/15. <https://doi.org/10.1093/bioinformatics/btv254> PMID: 26072497.
64. McKenzie E, Tyson K, Stamps A, Smith P, Turner P, Barry R, et al. Cloning and expression profiling of Hpa2, a novel mammalian heparanase family member. *Biochemical and biophysical research communications*. 2000; 276(3):1170–7. Epub 2000/10/12. <https://doi.org/10.1006/bbrc.2000.3586> PMID: 11027606.
  65. Levy-Adam F, Feld S, Cohen-Kaplan V, Shteingauz A, Gross M, Arvatz G, et al. Heparanase 2 interacts with heparan sulfate with high affinity and inhibits heparanase activity. *The Journal of biological chemistry*. 2010; 285(36):28010–9. Epub 2010/06/26. <https://doi.org/10.1074/jbc.M110.116384> PMID: 20576607.
  66. Maggioni F, Bruno M, Mainardi F, Lisotto C, Zanchin G. Migraine responsive to warfarin: an update on anticoagulant possible role in migraine prophylaxis. *Neurological sciences: official journal of the Italian Neurological Society and of the Italian Society of Clinical Neurophysiology*. 2012; 33(6):1447–9. Epub 2012/01/10. <https://doi.org/10.1007/s10072-011-0926-4> PMID: 22222938.
  67. Rajan R, Ahluwalia J, Lal V. Prothrombotic states in migraine. *Clinical and applied thrombosis/hemostasis: official journal of the International Academy of Clinical and Applied Thrombosis/Hemostasis*. 2014; 20(8):851–6. Epub 2013/05/03. <https://doi.org/10.1177/1076029613486538> PMID: 23637003.
  68. Tietjen GE, Khubchandani J. Vascular biomarkers in migraine. *Cephalalgia: an international journal of headache*. 2015; 35(2):95–117. Epub 2014/10/05. <https://doi.org/10.1177/0333102414544976> PMID: 25281220.
  69. Hare DL. Depression and cardiovascular disease. *Current opinion in lipidology*. 2021; 32(3):167–74. Epub 2021/04/17. <https://doi.org/10.1097/MOL.0000000000000749> PMID: 33859128.
  70. Hare DL, Toukhsati SR, Johansson P, Jaarsma T. Depression and cardiovascular disease: a clinical review. *European heart journal*. 2014; 35(21):1365–72. Epub 2013/11/28. <https://doi.org/10.1093/eurheartj/ehu462> PMID: 24282187.
  71. Alexopoulos GS, Meyers BS, Young RC, Campbell S, Silbersweig D, Charlson M. 'Vascular depression' hypothesis. *Archives of general psychiatry*. 1997; 54(10):915–22. Epub 1997/10/24. <https://doi.org/10.1001/archpsyc.1997.01830220033006> PMID: 9337771.
  72. Taylor WD, Aizenstein HJ, Alexopoulos GS. The vascular depression hypothesis: mechanisms linking vascular disease with depression. *Molecular psychiatry*. 2013; 18(9):963–74. Epub 2013/02/27. <https://doi.org/10.1038/mp.2013.20> PMID: 23439482.
  73. Okbay A, Baselmans BM, De Neve JE, Turley P, Nivard MG, Fontana MA, et al. Genetic variants associated with subjective well-being, depressive symptoms, and neuroticism identified through genome-wide analyses. *Nature genetics*. 2016; 48(6):624–33. Epub 2016/04/19. <https://doi.org/10.1038/ng.3552> PMID: 27089181.
  74. Bruce AW, Donaldson IJ, Wood IC, Yerbury SA, Sadowski MI, Chapman M, et al. Genome-wide analysis of repressor element 1 silencing transcription factor/neuron-restrictive silencing factor (REST/NRSF) target genes. *Proceedings of the National Academy of Sciences of the United States of America*. 2004; 101(28):10458–63. Epub 2004/07/09. <https://doi.org/10.1073/pnas.0401827101> PMID: 15240883.
  75. Chong JA, Tapia-Ramírez J, Kim S, Toledo-Aral JJ, Zheng Y, Boutros MC, et al. REST: a mammalian silencer protein that restricts sodium channel gene expression to neurons. *Cell*. 1995; 80(6):949–57. Epub 1995/03/24. [https://doi.org/10.1016/0092-8674\(95\)90298-8](https://doi.org/10.1016/0092-8674(95)90298-8) PMID: 7697725.
  76. Schoenherr CJ, Anderson DJ. The neuron-restrictive silencer factor (NRSF): a coordinate repressor of multiple neuron-specific genes. *Science (New York, NY)*. 1995; 267(5202):1360–3. Epub 1995/03/03. <https://doi.org/10.1126/science.7871435> PMID: 7871435.
  77. Chen ZF, Paquette AJ, Anderson DJ. NRSF/REST is required in vivo for repression of multiple neuronal target genes during embryogenesis. *Nature genetics*. 1998; 20(2):136–42. Epub 1998/10/15. <https://doi.org/10.1038/2431> PMID: 9771705.
  78. Warburton A. The REST/NRSF Pathway as a Central Mechanism in CNS Dysfunction [PhD thesis]: University of Liverpool; 2015.
  79. Otsuki K, Uchida S, Wakabayashi Y, Matsubara T, Hobara T, Funato H, et al. Aberrant REST-mediated transcriptional regulation in major depressive disorder. *Journal of psychiatric research*. 2010; 44(6):378–84. Epub 2009/10/23. <https://doi.org/10.1016/j.jpsychires.2009.09.009> PMID: 19846118.
  80. Singh-Taylor A, Molet J, Jiang S, Korosi A, Bolton JL, Noam Y, et al. NRSF-dependent epigenetic mechanisms contribute to programming of stress-sensitive neurons by neonatal experience, promoting resilience. *Molecular psychiatry*. 2018; 23(3):648–57. Epub 2017/01/11. <https://doi.org/10.1038/mp.2016.240> PMID: 28070121.
  81. Tamási V, Petschner P, Adori C, Kirilly E, Ando RD, Tothfalusi L, et al. Transcriptional evidence for the role of chronic venlafaxine treatment in neurotrophic signaling and neuroplasticity including also

- Glutamatergic [corrected]—and insulin-mediated neuronal processes. *PloS one*. 2014; 9(11):e113662. Epub 2014/11/26. <https://doi.org/10.1371/journal.pone.0113662> PMID: 25423262.
82. Uchida H, Ma L, Ueda H. Epigenetic gene silencing underlies C-fiber dysfunctions in neuropathic pain. *The Journal of neuroscience: the official journal of the Society for Neuroscience*. 2010; 30(13):4806–14. Epub 2010/04/02. <https://doi.org/10.1523/JNEUROSCI.5541-09.2010> PMID: 20357131.
  83. Zullo JM, Drake D, Aron L, O'Hern P, Dhamne SC, Davidsohn N, et al. Regulation of lifespan by neural excitation and REST. *Nature*. 2019; 574(7778):359–64. Epub 2019/10/18. <https://doi.org/10.1038/s41586-019-1647-8> PMID: 31619788.
  84. Meza-Aguilar DG, Boucard AA. Latrophilins updated. *Biomolecular concepts*. 2014; 5(6):457–78. Epub 2014/11/28. <https://doi.org/10.1515/bmc-2014-0032> PMID: 25429599.
  85. Woelfle R, D'Aquila AL, Pavlović T, Husić M, Lovejoy DA. Ancient interaction between the teneurin C-terminal associated peptides (TCAP) and latrophilin ligand-receptor coupling: a role in behavior. *Frontiers in neuroscience*. 2015; 9:146. Epub 2015/05/13. <https://doi.org/10.3389/fnins.2015.00146> PMID: 25964737.
  86. Woelfle R, D'Aquila AL, Lovejoy DA. Teneurins, TCAP, and latrophilins: roles in the etiology of mood disorders. *Translational neuroscience*. 2016; 7(1):17–23. Epub 2017/01/27. <https://doi.org/10.1515/tnsci-2016-0004> PMID: 28123817.
  87. Matsushita H, Lelianova VG, Ushkaryov YA. The latrophilin family: multiply spliced G protein-coupled receptors with differential tissue distribution. *FEBS letters*. 1999; 443(3):348–52. Epub 1999/02/20. [https://doi.org/10.1016/s0014-5793\(99\)00005-8](https://doi.org/10.1016/s0014-5793(99)00005-8) PMID: 10025961.
  88. Gargus JJ. Genetic calcium signaling abnormalities in the central nervous system: seizures, migraine, and autism. *Annals of the New York Academy of Sciences*. 2009; 1151:133–56. Epub 2009/01/22. <https://doi.org/10.1111/j.1749-6632.2008.03572.x> PMID: 19154521.
  89. Ophoff RA, Terwindt GM, Vergouwe MN, van Eijk R, Oefner PJ, Hoffman SM, et al. Familial hemiplegic migraine and episodic ataxia type-2 are caused by mutations in the Ca<sup>2+</sup> channel gene CACNL1A4. *Cell*. 1996; 87(3):543–52. Epub 1996/11/01. [https://doi.org/10.1016/s0092-8674\(00\)81373-2](https://doi.org/10.1016/s0092-8674(00)81373-2) PMID: 8898206
  90. Howard DM, Adams MJ, Clarke TK, Hafferty JD, Gibson J, Shirali M, et al. Genome-wide meta-analysis of depression identifies 102 independent variants and highlights the importance of the prefrontal brain regions. *Nature neuroscience*. 2019; 22(3):343–52. Epub 2019/02/06. <https://doi.org/10.1038/s41593-018-0326-7> PMID: 30718901.
  91. Levey DF, Gelernter J, Polimanti R, Zhou H, Cheng Z, Aslan M, et al. Reproducible Genetic Risk Loci for Anxiety: Results From ~200,000 Participants in the Million Veteran Program. *The American journal of psychiatry*. 2020; 177(3):223–32. Epub 2020/01/08. <https://doi.org/10.1176/appi.ajp.2019.19030256> PMID: 31906708.
  92. Liu W, Li W, Cai X, Yang Z, Li H, Su X, et al. Identification of a functional human-unique 351-bp Alu insertion polymorphism associated with major depressive disorder in the 1p31.1 GWAS risk loci. *Neuropsychopharmacology: official publication of the American College of Neuropsychopharmacology*. 2020; 45(7):1196–206. Epub 2020/03/21. <https://doi.org/10.1038/s41386-020-0659-2> PMID: 32193514.
  93. Wray NR, Ripke S, Mattheisen M, Trzaskowski M, Byrne EM, Abdellaoui A, et al. Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *Nature genetics*. 2018; 50(5):668–81. Epub 2018/04/28. <https://doi.org/10.1038/s41588-018-0090-3> PMID: 29700475.
  94. Csépany É, Tóth M, Gyüre T, Magyar M, Bozsik G, Bereczki D, et al. The validation of the Hungarian version of the ID-migraine questionnaire. *The journal of headache and pain*. 2018; 19(1):106. Epub 2018/11/14. <https://doi.org/10.1186/s10194-018-0938-z> PMID: 30419813.
  95. Juhasz G, Eszlari N, Pap D, Gonda X. Cultural differences in the development and characteristics of depression. *Neuropsychopharmacologia Hungarica: a Magyar Pszichofarmakologiai Egyesület lapja = official journal of the Hungarian Association of Psychopharmacology*. 2012; 14(4):259–65. Epub 2012/12/28. PMID: 23269213.