SEMMELWEIS EGYETEM

DOKTORI ISKOLA

Ph.D. értekezések

3390.

TÖRÖK DÓRA

Experimentális és klinikai farmakológia

című program

Programvezető: Dr. Szökő Éva, egyetemi tanár

Témavezető: Dr. Juhász Gabriella, egyetemi tanár

Dr. Petschner Péter, egyetemi adjunktus

Approaches of genetic signal detection in polygenic, multifactorial disorders

Ph.D. Thesis

Dóra Török

Semmelweis University Doctoral School
Pharmaceutical Sciences Division

Supervisors: Gabriella Juhász, MD, D.Sc
Péter Petschner, PharmD, Ph.D
Official reviewers: Eszter Ari, Ph.D habil, Andrea Szabó-Vereczkei, Ph.D

Head of the Complex Examination Committee: Zoltán Benyó, MD, D.Sc
Members of the Complex Examination Committee: Zoltán Zádori, MD, Ph.D
Eszter Ducza, PharmD, Ph.D

Budapest, 2025

Table of contents

List of abbreviations

BOLD - blood oxygen level-dependent

CGRP - calcitonin gene-related peptide

CTS - childhood trauma screener

ELS - early life stress

fMRI - functional magnetic resonance imaging

GCTA - genome-wide complex trait analysis

GSEA - gene set enrichment analysis

GWAS - genome-wide association study

HDM - human default model

HPA - hypothalamic–pituitary–adrenal

h2 - heritability estimate

IL-6 - interleukin 6

kbp - kilobase pair

LD - linkage disequilibrium

LEG - leading edge genes

MAF - minor allele frequency

MDD - major depressive disorder

MHC - major histocompatibility complex

mPFC - medial prefrontal cortex

NAc - nucleus accumbens

NAD+ - nicotinamide-adenine dinucleotide

NMN - nicotinamide mononucleotide

PRS - polygenic risk scores

QC - quality control

SD - standard deviation

SE - standard error

SIRT1 - nicotinamide-adenine dinucleotide-dependent protein deacetylase sirtuin-1

SNP - single-nucleotide polymorphism

TNF-α - tumor necrosis factor alpha

1. Introduction

Common migraine and depression are complex disorders characterized by polygenic and multifactorial background (1, 2). Common migraine (henceforth migraine), unlike familial hemiplegic migraine, which follows Mendelian inheritance (3), refers to more prevalent forms of migraine, including migraine with and without aura, affecting approximately 14-15% of the global population (4), with a higher prevalence in women than men (5). Similarly, depression, defined as a mood disorder marked by constant sadness, loss of interest, and cognitive and physical symptoms, affects more than 300 million people worldwide (6), also showing a marked female predominance (7). The diagnostic criteria for major depressive disorder (MDD) include the presence of at least one of two core symptoms, depressed mood or anhedonia, persisting for at least two weeks, accompanied by at least four additional symptoms (8). The polygenic nature of these disorders means that their risk is influenced by the combined effects of numerous common genetic variants, particularly single-nucleotide polymorphisms (SNPs). In addition to their complex genetic architecture, both migraine and depression are shaped by different environmental and lifestyle factors (9–11). This multifactorial nature highlights the need for integrative statistical frameworks to better understand the mechanisms underlying these disorders. In addition to their complex, polygenic nature, depression and migraine are highly comorbid disorders. Comorbidity refers to the co-occurrence of two or more medical conditions within the same individual, occurring not by chance, but potentially indicating a shared biological background (12). For example, migraine and depression frequently co-occur, and both disorders are often comorbid with other conditions such as other psychiatric conditions, anxiety, and cardiovascular diseases (13–15).

Recognizing the significant genetic contribution to these disorders, much research has been conducted to investigate underlying genetic architecture. While genome-wide association studies (GWAS) have advanced our understanding of the genetic basis of complex traits, they still leave important questions unanswered (16). A persistent gap remains between SNP-based heritability estimates and those derived from twin studies.

In migraine research, large-scale meta-analyses, Gormley et al. (17), Hautakangas et al. (18), have estimated SNP-heritability at 14.6% and 11.2%, respectively, whereas twin studies suggest heritability as high as 36-48% (19). Compared to migraine, depression has been more extensively studied in terms of its genetic architecture, yet substantial gaps remain. Twin studies estimate the heritability of major depressive disorder (MDD) at 37% (20), while the largest GWAS to date reported SNP-based heritability at just 8.4% (21), reflecting a notable discrepancy between common variant effects and the broader genetic background. This discrepancy, referred to as missing heritability (22), may reflect the influence of rare variants, copy number variations, epistasis, environmental factors, or complex interactions between genes and the environment, including epigenetic effects. Although some of these elements can be partially addressed through GWAS or advanced methods, many remain unresolved.

One unexplored factor contributing to this gap could be the definition of the phenotype itself (23). To enhance the statistical power of GWAS, large sample sizes are often prioritized, sometimes without a precise definition of the disease or phenotype under investigation. Variations in how depression and migraine are defined, ranging from self-reported symptoms to clinical diagnosis, can introduce noise into genetic analyses. Moreover, comorbidities present at the time of diagnosis may further blur the underlying genetic signal. Previous work from our group identified a network of migraine and depression comorbidities (15). In a follow-up study using machine learning approaches, we found preliminary evidence suggesting that the presence of diagnosed depression may influence the strength and direction of genetic associations in migraine (24). Similar findings in the literature report altered heritability estimates in comorbid populations (25). Although these observations are still preliminary, they raise the possibility that focusing on individuals for whom there is only one diagnosed condition could help refine the phenotype and uncover more specific genetic contributors. Such an approach could also improve target identification for treatments, particularly given that clinical trials often exclude patients with multiple diagnoses.

Another pitfall of large-powered GWASs is that some disease-related genes may not meet the strict genome-wide significance threshold and thus, remain undetected (26).

Since genes participate in complex biological pathways, taking into account prior knowledge, like gene expression findings from animal models, can improve our ability to uncover key genetic contributors to disease (27).

Transcriptomic studies in animal models have proposed several mechanisms for migraine (28); however, these studies have shown limited results, rarely identifying genes that were consistently significant after multiple testing correction. In depression research, interest in gene expression profiling is also increasing, as it could offer valuable insight into the molecular underpinnings of major depressive disorder (29–33). Despite the availability of numerous antidepressant treatments, most of which act on monoaminergic systems such as serotonin and norepinephrine (34, 35), these pathways explain only a fraction of the disorder's heterogeneity and therapeutic response (36). Integrating transcriptomic data into depression research can therefore help to uncover additional pathways and targets beyond monoamines, potentially advancing both diagnostic precision and therapeutic development (37, 38).

One promising approach to enhance understanding of complex disorders is the use of targeted analysis strategies informed by transcriptomic data. Unlike hypothesis-free GWAS, which investigate the genome without prior assumptions, targeted strategies focus on predefined gene sets or pathways supported by biological evidence. Transcriptomic profiling in animal models exposed to disease-relevant conditions can reveal genes and pathways that are differentially expressed and potentially mechanistically involved. These gene sets can then be tested in human genomic data to assess their association with the disorder, thereby increasing biological interpretability. This integrative approach bridges experimental findings and human genetics, allowing for the prioritization of genes that may not reach genome-wide significance in GWAS but are still functionally relevant. In depression and migraine research, such strategies are particularly valuable given the modest SNP-based heritability and the complexity of the underlying biology.

One such pathway that has emerged as a candidate for targeted investigation in depression is the NAD+/SIRT1 signaling pathway (39).

Nicotinamide-adenine dinucleotide (NAD+)-dependent deacetylase sirtuin-1 (SIRT1) has a central role in diverse physiological processes, including stress response, energy metabolism, and neuroplasticity (39–41). Despite its biological relevance, genetic and transcriptomic studies investigating the NAD+/SIRT1 pathway in humans remain underrepresented. In the transcriptomic study of Morató et al. (42), they demonstrated that early life stress (ELS) in male rodents led to increased adiposity and reduced sociability in adulthood, central behavioral features relevant to depression. These effects were mediated by ELS-induced disruption of the stress-fat-brain axis, in which a downregulation of SIRT1 in the nucleus accumbens (NAc) and consequent impairment of the NAD+/SIRT1 signaling pathway played a central mechanistic role. Notably, treatment with nicotinamide mononucleotide (NMN), which boosts NAD+ levels and compensates for reduced SIRT1 activity, was able to reverse these behavioral changes.

Given that social withdrawal is a symptom of depression, these findings suggest that ELS may predispose individuals to depression via dysregulation of the NAD+/SIRT1 pathway. However, the sex-specific and ELS-dependent effects of this pathway, its relationship with obesity, and its influence on NAc function remain largely unexplored in humans. This limits our ability to leverage components of the NAD+/SIRT1 pathway as biomarkers or therapeutic targets in depression. Taken together, both improved phenotype characterization and genetic analyses guided by prior knowledge hold the potential to refine and enhance genetic findings in polygenic disorders such as migraine and depression. Accordingly, the following research questions were formulated.

2. Objectives

The objective was to answer the following research questions:

Question 1

Does a narrower migraine phenotype by excluding comorbidities improve the discovery of polygenic signals and lead to a deeper understanding of migraine pathomechanism?

Question 2

Does a human transcriptomic findings-driven genomic analysis provide novel insights into migraine pathophysiology?

Question 3

Do rodent model-derived transcriptomic findings allow for genomic signal detection in the NAD+/SIRT1 pathway in humans for stress-induced depression, thereby linking early environmental adversity and obesity to the development of depression?

3. Methods

3.1 Study 1: polygenic signal detection of migraine-first diagnosis

3.1.1. UK Biobank cohort

Data from the UK Biobank (application number 71718) were used. Participants aged 40-69 were invited and recruited using NHS patient registers (43). Approval for the study was obtained from the National Research Ethics Service Committee North West-Haydock (44). All participants gave written informed consent, and all procedures were conducted in accordance with the Declaration of Helsinki. The data collection methodology has been previously described by Bycroft et al. (45).

3.1.2. Genomic quality control

Quality control (QC) procedures were applied, including a minor allele frequency (MAF) threshold of 0.01, stepwise filtering for SNP and individual missingness (thresholds of 0.1, 0.05, and 0.01), exclusion of variants failing Hardy-Weinberg equilibrium ($p \geq 1\times10^{-5}$), and linkage disequilibrium (LD) pruning using an R^2 cutoff of 0.2. Multiallelic variants were excluded. In addition, variants with an imputation "info" score below 0.5 were also removed. These QC steps were conducted prior to sex check (discrepancies between reported and genetically inferred sex were excluded), identification of heterozygosity outliers, and principal component analysis (PCA). After excluding variants below minor allele frequency of 0.01 and variants on chromosome X, 6,077,313 SNPs remained. We followed the QC protocol outlined in Eszlari et al. (46).

3.1.3. Phenotype definition

Migraine-first patients were identified based on G43 diagnosis (UKB data-field 131052: date of first reported migraine, with or without aura). We restricted the cohort to individuals for whom G43 was the first recorded medical diagnosis in their lifetime. Healthy controls were defined as individuals with no diagnosed medical conditions at the age when migraine-first patients typically received their diagnosis (mean age 22 ± 12), and with no history of migraine at any point in their lifetime.

Data filtering resulted in 6,307 migraineurs and 200,571 controls. Restricting to individuals who had genetic data available, we had 6,139 migraineurs and 193,790 healthy controls for further analysis. To characterize the population, we made descriptive statistics with R (v4.1.2).

3.1.4. Genome-wide association analysis

Genome-wide association analyses were conducted using PLINK2 (47). The model included the following covariates: sex, age (UKB fields 31 and 21003), 10 principal components of the genome, and the genotyping chip information (UKB Axiom Array or UK BiLEVE Axiom Array).

3.1.5. Heritability estimation

SNP-based heritability refers to the proportion of variation in a trait or disease that can be attributed to the additive effects of common SNPs across the genome. SNP effects are estimated through genome-wide association studies, and the resulting summary statistics are then used in downstream analyses to estimate heritability (h^2). SNP-based heritability was estimated using SumHer from the LDAK software (v5.2) (48), following the authors' recommendation, using the Human Default Model (HDM).

Estimates were calculated for 1) all SNPs, 2) all SNPs excluding those that reached genome-wide and suggestive significance, and 3) a subset of SNPs restricted to the HapMap3 panel. The HapMap3 SNP set consists of common, well-imputed variants that are widely used in genetic studies due to their high quality, reliability (49). The major histocompatibility complex (MHC) region on chromosome 6 (25-35 Mb) was excluded from heritability estimations because of its highly complex linkage disequilibrium (LD) structure and extreme genetic variability (50).

These characteristics can influence heritability estimates and complicate downstream analyses, so this region is typically omitted in genetic analyses to improve accuracy and interpretability. Heritability values were transformed to the liability scale using a sample-based migraine prevalence of 3% and population prevalence of 16%, as done in previous study of Gormley et al. (17).

3.1.6. Validation of heritability estimations

To verify the robustness of our results, we conducted several validation analyses:

1) To demonstrate the importance of stringent control selection and to ensure that the elevated heritability estimate was not because of the limited statistical power, we re-ran the analysis using an alternative control group composed of all participants without any migraine diagnosis (N = 327,986). 2) We assessed the influence of case-control imbalance by repeating the analysis three times, each time randomly removing one-third of the healthy controls to simulate a higher proportion of cases in the sample (i.e., increased ascertainment). This approach allowed us to evaluate the stability and robustness of heritability estimates under different case-control ratios, as heritability on the liability scale can be sensitive to the assumed or observed prevalence of migraine in the population.

3) To prove whether phenotype definition is crucial, we tested a broader migraine phenotype by including all individuals ever diagnosed with migraine (G43), regardless of first diagnosis onset, N_{cases} = 17,679; $N_{controls}$ = 316,446. 4) We tested the effect of different heritability models: genome-wide complex trait analysis (GCTA), Human Default Model (HDM), LDAK model, and BLD-LDAK model, detailed in publications (48, 51), and explained in the Supplementary table S6 of Torok et al. (52).

3.1.7. Post-GWAS analyses

The goal of post-GWAS analyses is to interpret GWAS findings in the context of genomic architecture and biological pathways, which can help uncover potential mechanisms underlying migraine. Risk loci identification was carried out using FUMA (53), which involves identifying genomic regions (loci) that are significantly associated with migraine. We defined index SNPs as those that are LD-independent ($r^2 < 0.1$), meaning they likely represent distinct association signals. Around each index SNP, we constructed a risk locus by including all nearby variants in high LD ($r^2 > 0.6$) within a 250 kilobase pair (kbp) window. This approach helps to capture the broader genetic signal around a key variant, rather than focusing on individual SNPs.

Index SNPs were defined consistent with approaches from large migraine meta-analyses of Gormley et al. (17) and Hautakangas et al. (18). Significance thresholds for SNPs were set at $p < 10^{-8}$ for genome-wide significance, conventional genome-wide significance ($p < 5 \times 10^{-8}$) and $10^{-8} < p < 10^{-5}$ for suggestive associations.

Gene-set enrichment and tissue-specific enrichment analysis were conducted using MAGMA (54), a tool that maps associated SNPs to genes and tests whether certain biological pathways or functional gene sets are overrepresented among the associated genes. Gene-set enrichment aids in identifying biological processes or molecular functions potentially involved in migraine. In addition, tissue-specific enrichment analysis was performed to determine whether the genes associated with migraine are differently expressed in specific tissues or cell types. This analysis leverages gene expression data GTEx (55) to assess whether the distribution of associated genes is non-random across different tissues, offering insights into where in the body, such as the brain or vascular tissues, genetic risk factors may exert their effects.

3.2. Study 2: validation of transcriptomic findings through genetic analysis

3.2.1. UK Biobank cohort

We applied the same criteria to define migraine as a first diagnosis. To ensure consistency between the genetic and transcriptomic analysis, we included smoking, allergy, and information regarding vitamin A and vitamin A retinol equivalent intake as covariates. (56). Smoking status was classified using data on current (field ID: 1239) and past (field ID: 1249) smoking behavior. Asthma status was determined using self-reported medical conditions related to asthma or hayfever/eczema/allergic rhinitis (field ID: 6152). Intake of vitamin A and vitamin A retinol equivalents was estimated from dietary data and supplement use (field IDs: 26061, 20084). Sex and age were extracted from field IDs 31 and 21003, respectively.

3.2.2. Hungarian migraine cohort

Research involving the Hungarian migraine cohort was approved by the Medical Research Council's local ethics committee (reference numbers: 23609-1/2011-EKU, 23421-1/2015-EKU, 31304-1/2014-EKU, 014946-003/2016/OTIG, OGYÉI/49553/2017) and followed the guidelines of the Helsinki Declaration.

Data collection included information on migraine diagnosis, smoking status, age, sex, history of allergies, pregnancy and breastfeeding, serious acute or chronic illnesses (past or present), and regular medication use. The diagnosis of episodic migraine without aura (MO) was made by a specialist neurologist according to the International Classification of Headache Disorders, 3rd edition (beta version) (57).

Controls were defined as individuals without any diagnosed diseases. Data from the entire cohort (N_{cases} = 172, $N_{controls}$ =117) was used in the genetic analyses.

A similar filtering process was applied to the Hungarian migraine cohort's genetic dataset to ensure consistency with the UK Biobank pipeline. Descriptive statistics were done in R (version 4.1.2).

3.2.3. Genetic analysis of the Leading Edge Genes (LEG) and vitamin A pathway genes

In the transcriptomic study of Petschner et al. (56), gene set enrichment analysis (GSEA) comparing migraine cases to controls identified 88 significantly enriched pathways.

To pinpoint the genes most relevant to these replicated, pathway-level findings, they extracted the 69 Leading Edge Genes (LEGs) - the key genes driving the enrichment signals. For the genetic validation analyses, we focused on SNPs (N = 13,650) located within these LEGs and their surrounding regions (±10 kbp), as these areas may contain regulatory elements that influence gene expression.

Gene boundaries were defined based on the hg19 Known Gene track from the UCSC Genome Browser (58). SNPs falling within these boundaries or the neighboring 10 kbp regions were identified using ANNOVAR (59), a functional annotation tool for genetic variants. The complete list of selected SNPs and their corresponding genes is provided in Petschner et al. - Supplementary material 2, and Supplementary material 4 (56).

Vitamin A pathway genes were manually selected based on protein names mentioned in previous studies (60, 61). The corresponding human gene names were identified using GeneCards (62), and SNPs (N = 13,635) were extracted using the same approach as for the Leading Edge Genes (LEGs). Total list of the selected SNPs and genes can be found in Petschner et al. - Supplementary Material 4 (56).

Logistic regression analyses of the selected SNPs were conducted for migraine using PLINK2, including covariates such as sex, age, the top 10 genetic principal components, smoking status, and allergy history in both the UK Biobank and Hungarian migraine cohort.

For the UK Biobank analyses, genotype array type was also included as a covariate. In addition, for SNPs related to the vitamin A pathway, estimated intake of vitamin A and vitamin A retinol equivalents were included as covariates in the regression analyses.
We performed t-test with R (version 4.1.2) to compare the vitamin A and vitamin A retinol equivalent intakes in migraineurs and controls.

3.3. Study 3: genetic risk of NAD+/SIRT1 pathway in depression

3.3.1. UK Biobank cohort

Data from the UK Biobank (application ID: 1602) was used in the third study, including a total of 334,248 participants.
Depressive symptoms were assessed by a composite score derived from multiple UK Biobank data fields 2050, 2060, 2070, and 2080, in line with prior methods described by Hullam et al (63).
Early life stress, considering data fields 20487, 20488, 20489, 20490, 20491, was evaluated using the Childhood Trauma Screener of Walker et al (64). Body fat percentage was measured through impedance-based body composition measurement (field id: 23099). Age (field id: 21003) and sex (field id: 31) were used as covariates in statistical analyses, except in sex-stratified models.

3.3.2. Hungarian cohort

Control individuals of the Hungarian migraine cohort (N=102), excluding migraineurs, were used in genetic and functional connectivity analyses. Depressive symptoms were assessed using the Zung Self-Rating Depression Scale (65), and early life stress (ELS) was measured with the Childhood Trauma Questionnaire (66).
Descriptive statistics were done in R (version 4.1.2), describing both cohorts.

3.3.3. NAD+/SIRT1 pathway

Variants within 20 genes (*PPARGC1A*, *TFAM*, *NDUFS4*, *SHD*, *UQCRB*, *COX6A2*, *ACO2*, *CS*, *MFN1*, *MFN2*, *BNIP3*, *CAT*, *GPX4*, *HSPD1*, *CLPP*, *BDNF*, *NPY*, *HCRTR2*, *KCNMB2*, *SIRT1*) of the NAD+/SIRT1 pathway, as defined in Morató et al. (42), were selected, including a 10,000 kbp flanking region for each gene. SNP annotation was conducted using the biomaRt (67) R package and LDlink (68) database resources.

Following quality control, 3,827 SNPs in the UKB and 1,885 SNPs in the Hungarian dataset were retained for analysis.

3.3.4. Genome-wide association analysis on depression score

The UK Biobank cohort was divided into discovery and target subsets to ensure independent datasets for the estimation of polygenic risk scores (**Figure 1**).

Individuals with available ELS data (N = 105,654) constituted the target group, while the remaining participants (N = 228,595) formed the discovery sample. GWAS was conducted on depression score in the discovery sample using PLINK2 (47), with covariates including age, sex, 10 genetic principal components, and genotyping chip information.

3.3.5. Polygenic risk score estimation

Calculating polygenic risk scores requires both a discovery sample and a target sample. The discovery sample is used to identify SNP-trait associations and estimate effect sizes through genome-wide association studies, while the target sample is used to calculate individual-level risk scores based on these effect sizes. This separation is essential to avoid overfitting and to ensure that the PRS reflects true genetic risk rather than sample-specific noise. The method requires harmonization between the discovery and target datasets, which is crucial when the data originates from different sources. This harmonization process includes aligning SNP identifiers, filtering based on minor allele frequency, excluding ambiguous strand SNPs (A/T or C/G), and managing missing data. As part of our PRS workflow, we carried out parameter estimation procedures. Estimation of the posterior effect sizes depends on two key parameters: heritability (h^2) and the proportion of causal variants. These parameters were inferred using linkage disequilibrium score regression, which incorporates SNP effect sizes from GWAS summary statistics, prior assumptions about genetic architecture, and LD structure. In our study, LD information was calculated specifically within the genomic regions corresponding to the NAD+/SIRT1 pathway genes under investigation.

PRS calculations were performed on the target dataset and on the Hungarian cohort using LDpred2 (69), incorporating GWAS-derived beta values for NAD+/SIRT1 pathway variants. PRS was calculated for the full target sample and separately by sex.

3.3.6. Interaction between PRS and ELS

Linear regression models assessed the combined effect of ELS and genetic risk on depression score.

A baseline model (R^2_{null}) included PRS and ELS as main effects, with additional covariates: age, sex, except in sex-stratified models, population structure components, and genotyping array. The full model ($R^2_{PRS \times ELS}$) included an interaction term between PRS and ELS. Analyses were performed using R version 4.1.2.

The difference in explained variance (ΔR^2) between the two models was calculated as shown in *Equation 1*:

$$\Delta R^2 = R^2_{PRS \times ELS} - R^2_{null}$$

Equation 1 - Variance change by the interaction of the PRS and ELS

In this formula, ΔR^2 represents the variance change, explained by the interaction between polygenic risk score (PRS) and early life stress (ELS). $R^2_{PRS \times ELS}$ refers to the variance explained by the full model, which includes covariates as well as the main effects and their interaction. R^2_{null} represents the variance explained by a model containing only the covariates and the main effects of PRS and ELS.

We calculated the explained variance change in percentage (ΔR^2_{perc}) with the following formula, explained in *Equation 2*:

$$\Delta R^2_{perc} = \Delta R^2 / R^2_{null} * 100$$

Equation 2 - Variance change in percentage

3.3.7. Mediation analysis

Mediation analysis can test specifically whether the influence of the PRS × ELS interaction on depressive symptoms could be partially explained by its impact on body fat, which in turn may contribute to depression score. This was assessed in the total sample as well as separately in male and female subgroups. The models included covariates such as age, sex (only in the full sample), top 10 genetic principal components, genotype array, and the main effects of both PRS and ELS. For each model, we estimated the direct effect (denoted as c), the indirect effect (a × b), and the total effect (c + a × b) (**Figure 1**).

Following the mediation framework proposed by Baron and Kenny (70), we tested three sequential hypotheses: (1) that the independent variable (in this case, PRS × ELS) significantly predicts the outcome (depression score), (2) that it also predicts the mediator (body fat percentage), and (3) that the mediator significantly predicts the outcome.
We used "lavaan" (71) R package (version 0.6-12) for mediation analysis.

3.3.8. Resting-state functional connectivity analyses

Resting-state functional magnetic resonance imaging (fMRI) is used to assess spontaneous brain activity by detecting blood oxygen level-dependent (BOLD) signal fluctuations. When the BOLD signal time courses of different brain regions fluctuate in synchrony, it suggests potential functional connectivity, which is typically measured by calculating Pearson correlations and their conversion into Z-scores using Fisher's transformation.
The fMRI data were collected using two types of 3T scanners: a Philips Achieva (used for 40 participants) and a Siemens MAGNETOM Prisma (used for 62 participants). Structural brain images were obtained using a T1-weighted 3D turbo field echo (TFE) sequence on the Philips system and a 3D MPRAGE sequence on the Siemens system. Functional data were recorded with T2*-weighted echo-planar imaging (EPI) sequences. For the Philips scanner, the settings included a repetition time (TR) of 2,500 ms, echo time (TE) of 30 ms, a 240 mm^2 field of view (FOV), and 3 × 3 × 3 mm voxel resolution. On the Siemens scanner, data were collected with a TR of 2,220 ms, TE of 30 ms, FOV of 222 mm^2, and the same spatial resolution. To create seed-based connectivity maps, the average signal from predefined seed regions was correlated voxel-wise with the rest of the brain, and the results were normalized to Z-scores. Imaging data preprocessing steps were the same for both scanners as detailed in the publication of Gecse et al. (72).
For the resting-state functional connectivity analysis, data from a Hungarian cohort of 102 healthy individuals (58 females and 44 males) were involved (**Figure 1**). Each participant underwent a resting-state fMRI scan and provided both biological and genetic samples. Seed-based connectivity analysis was carried out using the nucleus accumbens (NAc) as the seed region, defined by MNI coordinates (x = 10, y = 12, z = -8) with a 4 mm radius (73). The resulting individual connectivity maps were used for further analysis using the Statistical Parametric Mapping software (SPM12) within the Matlab 2016 environment.

To account for potential confounding variables, all analyses included adjustments for different scanner types, sex, age, and motion. For visualization purposes, MRIcroGL (http://www.mccauslandcenter.sc.edu/mricrogl/) and R software (version 4.1.2) were employed.

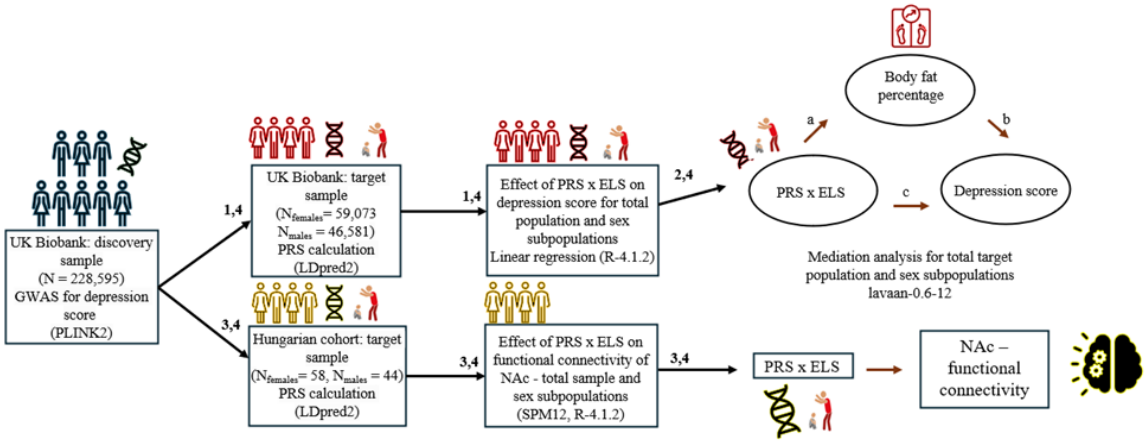

Figure 1. **Methodological overview of Study 3.**

4. Results

4.1. Study 1: polygenic signal detection of migraine-first diagnosis

4.1.1. Descriptive statistics of migraine-first population

For GWAS and heritability analyses, we use data of migraine-first patients (6,139 (3%)) and healthy controls (193,790 (97%)) (**Table 1**). Among migraine-first patients, the large majority were females (4,688 (76%)) compared to males (1,451 males, 24%), while in the healthy control group, the sex ratio was balanced, consisting of 97,078 females (50%) and 96,712 males (50%). The average age for migraine-first diagnosis was 22±12 years.

Table 1. Descriptive statistics of migraine-first patients and healthy controls.

Migraine-first patients and healthy controls		
Age	Mean	58.1743
	SD	±7.4741
Sex	Females	101766 (51%)
	Males	98163 (49%)
Migraine-first cases - healthy controls	Cases	6139 (3%)
	Controls	193790 (97%)
Migraine-first patients - age	Mean	56.0664
	SD	±7.7387
Migraine-first patients - sex	Females	4688 (76%)
	Males	1451 (24%)
Healthy controls - sex	Females	97078 (50%)
	Males	96712 (50%)
Healthy controls - age	Mean	58.2410
	SD	±7.4559
Age when migraine was diagnosed (in migraine-first patients)	Mean	22.2105
	SD	±12.2849

Note: SD denotes standard deviation.

4.1.2. Heritability estimation of migraine-first diagnosis

SNP-based heritability of migraine among migraine-first patients with Human Default Model showed high heritability: 19.37% (± 0.019 SD) on liability scale, when considering all SNPs (**Figure 2**, **Table 2**).

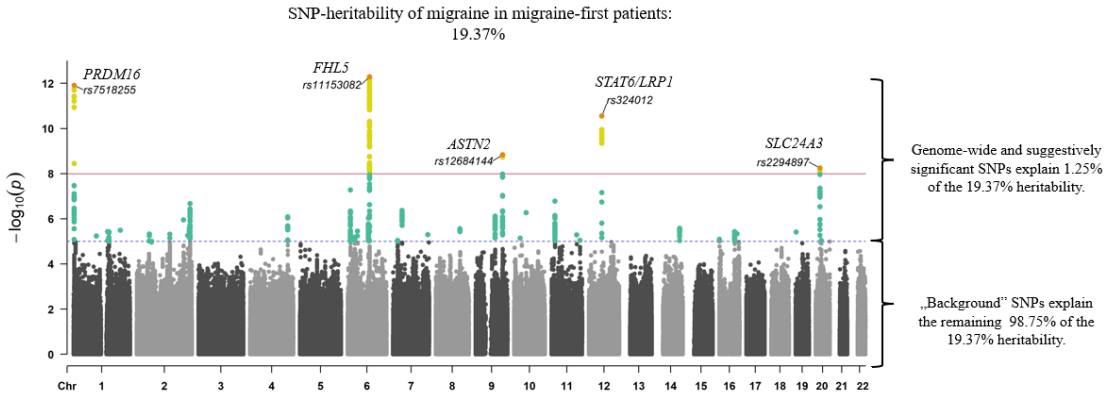

Figure 2. **Manhattan plot shows results of genome-wide association analysis and heritability estimation performed on migraine-first diagnosis**. On the x-axis, all chromosomes are represented, and all dots denote a single-nucleotide polymorphism. On the y-axis, the negative 10 base logarithmic transformation of the p-values is denoted. Genome-wide significance ($p < 10^{-8}$) is indicated by the red line, while the blue dashed line represents the suggestive significance threshold ($p < 10^{-5}$). SNPs shown in yellow are genome-wide significant, and those in green are suggestively significant. SNPs passing these thresholds are shown in yellow (genome-wide significant) and green (suggestively significant), together accounting for 1.25% of the total heritability. In contrast, the remaining SNPs, represented in black and grey, represent background variants that contribute the vast majority (98.75%) of heritability.

After removing both genome-wide ($p < 10^{-8}$) and suggestively ($p < 10^{-5}$) significant variants, heritability remained relatively high at 18.13% (± 0.019 SD), suggesting that much of the genetic contribution arises from the broader, underlying polygenic background. When limiting the analysis to HapMap3 SNPs, consistent with the approaches of Gormley et al. (17) and Hautakangas et al. (18), the heritability estimate increased to 21.31% (± 0.019 SD) (**Table 2**), representing a 45.96% and 90.27% rise compared to the estimates reported in the respective studies.

Table 2. Results of heritability estimation according to the involved SNP sets.

Migraine-first cases and healthy controls	Heritability on liability scale	**SD ±**
All SNPs	19.37%	0.019
All SNPs with GWS and suggestively significant SNPs excluded	18.13%	0.019
SNPs restricted to the HapMap3 SNP-set	21.31%	0.019

Note: SD - standard deviation.

4.1.3. Validation of heritability estimation

4.1.3.1. Validation 1: Migraine-first cases and all controls

We re-estimated SNP-based heritability using migraine-first cases and a broader control group comprising all individuals without a migraine diagnosis, thereby increasing the number of controls (**Table 3**).

Table 3. Summary statistics of migraine-first patients and all individuals as controls.

Migraine-first patients and all controls		
Age	Mean	56.8738
	SD	±7.9919
Sex	Females	179427 (54%)
	Males	154698 (46%)
Migraine-first patients - all controls	Cases	6139 (2%)
	Controls	327986 (98%)
Migraine-first patients - sex	Females	4688 (76%)
	Males	1451 (24%)
All controls - sex	Females	174739 (53%)
	Males	153247 (47%)
Migraine-first patients - age	Mean	56.0665
	SD	±7.7387
All controls - age	Mean	56.8889
	SD	±7.9958

Note: SD - standard deviation.

While increasing the number of controls led to greater statistical power, it resulted in a lower heritability estimate (h^2 = 17.39%, SD = 0.016), reflecting a 10.22% decrease in liability scale.

4.1.3.2. Validation 2: Migraine-first cases and balanced controls

To assess the impact of control group sample size and the increased case-to-control ratio, we randomly subsampled one-third of the control individuals and repeated this procedure three times. (**Table 4**).

Table 4. Descriptive statistics of the validation subsamples.

Subsample 1		
Age	Mean	58.1508
	SD	±7.4857
Sex	Females	69375 (51%)
	Males	65958 (49%)
Migraine-first patients - randomly selected healthy controls	Cases	6139 (5%)
	Controls	129194 (95%)
Subsample 2		
Age	Mean	58.1678
	SD	±7.4784
Sex	Females	69431 (51%)
	Males	65902 (49%)
Migraine-first patients - randomly selected healthy controls	Cases	6139 (5%)
	Controls	129194 (95%)
Subsample 3		
Age	Mean	58.1271
	SD	±7.4866
Sex	Females	69368 (51%)
	Males	65965 (49%)
Migraine-first patients -	Cases	6139 (5%)

randomly selected healthy controls	Controls	129194 (95%)

Note: SD - standard deviation.

Using more balanced samples with a greater proportion of migraine patients resulted in a lower SNP-based heritability estimate (**Table 5**), h^2_{mean} = 18.45%, SD = ± 0.017 , a 13.42% reduction on liability scale, likely due to diminished statistical power, though the heritability remained considerably higher than in earlier studies (17, 18).

Table 5. **Results of the second validation analyses.**

Migraine-first patients and randomly selected healthy controls - subsample 1	Heritability on liability scale	SD ±
SNPs restricted for the HapMap3 SNP-set	18.54%	0.017
Migraine-first patients and randomly selected healthy controls - subsample 2	Heritability on liability scale	SD ±
SNPs restricted for the HapMap3 SNP-set	18.18%	0.017
Migraine-first patients and randomly selected healthy controls - subsample 3	Heritability on liability scale	SD ±
SNPs restricted for the HapMap3 SNP-set	18.63%	0.017
Mean of the heritability of the 3 subsamples (h^2_{mean})	18.45%	0.017

Note: SNP-based heritability estimates on the liability scale for migraine-first patients and three independent subsamples of randomly selected healthy controls. Heritability was estimated using SNPs restricted to the HapMap3 panel. The table reports heritability values and corresponding standard deviations (SD) for each subsample, along with the mean heritability across the three datasets.

4.1.3.3. Validation 3: Heritability of migraine diagnosis without filtering for first onset

We applied the same analysis steps using a broader phenotype definition that included all individuals diagnosed with migraine (G43 diagnosis), without restricting to first-onset

cases (17,679) and all controls (316,446) (**Table 6**). Age of diagnosis was higher, 37 years, compared to migraine-first disease onset, 22 years.

Table 6. Descriptive statistics for the total UK Biobank population, including all migraineurs and all control participants.

All patients diagnosed with G43 and all controls		
Age	Mean	56.8738
	SD	±7.9919
Sex	Females	179427 (54%)
	Males	154698 (46%)
G43 migraine - all controls	Cases	17679 (5%)
	Controls	316446 (95%)
G43 migraine patients - age	Mean	55.5428
	SD	±7.8920
G43 migraine patients - sex	Females	13228 (75%)
	Males	4451 (25%)
All controls - sex	Females	166199 (53%)
	Males	150247 (47%)
All controls - age	Mean	56.9482
	SD	±7.9910
Age when migraine was diagnosed (for G43 migraine patients)	Mean	37.1125
	SD	±18.9314

Note: SD - standard deviation.

The heritability estimate for G43 migraine diagnosis without restricting to first-onset cases was lower, h^2 = 12.92%, SD = ±0.007, representing a 39.32% reduction on the liability scale.

4.1.3.4. Validation 4: Effect of the heritability model

The impact of the GCTA heritability model was small and yielded comparable results to the default heritability model ($h^2 = 23.89\%$, SD = ±0.021). Results of all different heritability estimations can be found in Supplementary Tables S5, S7, and S8 of Torok et al. (52).

4.1.4. Interpreting the biological significance of migraine-first GWAS findings

We identified 9 independent SNPs that survived genome-wide significance correction threshold ($p < 10^{-8}$) and one that survived a conventional ($p < 5 \times 10^{-8}$) threshold.

Among these, eight variants were associated with a risk effect (odds ratio > 1), and two were associated with a protective effect (odds ratio < 1) and were located in 5 loci: *PRDM16*, *FHL5*, *ASTN2*, *STAT6/LRP1* (**Table 7**).

Table 7. Genome-wide significant SNP-level results.

rsID	p-value	OR	gene
rs11153082	5.18×10^{-13}	1.16	*FHL5*
rs7518255	1.23×10^{-12}	1.18	*PRDM16*
rs324012	2.82×10^{-11}	0.87	*STAT6/LRP1*
rs11172113	2.84×10^{-11}	0.87	*LRP1*
rs3798293	4.85×10^{-11}	1.17	*FHL5*
rs4839683	2.23×10^{-10}	1.16	*FHL5*
rs12684144	1.47×10^{-9}	1.15	*ASTN2*
rs2651899	3.55×10^{-9}	1.13	*PRDM16*
rs2294897	5.66×10^{-9}	1.15	*SLC24A3*
rs2455136	3.48×10^{-8}	1.14	*PRDM16*

Note: rsID is the unique identifier of the single-nucleotide polymorphism, p-value is the significance of the association analysis, OR is the odds ratio, the effect of the SNP, and gene is the annotated gene in which the found SNP is located.

Among the risk loci identified by GWS SNPs, all (100%) were previously reported in earlier meta- and GWAS analyses. Using the suggestive significance threshold, 30.56% of the loci were replicated across all studies (**Table 8**). The full list of the 54 identified suggestively significant SNPs can be found in the Supplementary Table S10 of Torok et al. (52).

Table 8. **Overlapping SNPs and risk loci.**

GWS SNPs				
	Current study	Gormley et al.	Hautakangas et al.	Overlap in all three
GWS SNPs	10	44(1)	170(-)	-
GWS risk loci	5	38(5)	123(5)	5
Suggestive and GWS SNPs				
	Current study	Gormley et al.	Hautakangas et al.	Overlap in all three
GWS + suggestive SNPs	64	44(5)	170(-)	-
GWS risk loci	36	38(11)	123(11)	10

Note: Numbers in parentheses indicate overlaps with our current findings. For the study by Hautakangas et al. (18), only data on risk loci were available; (-) denotes data not available.

Among the suggestively significant risk loci (**Table 9**) *ADAMTSL4*, *MEF2D*, *TRPM8*, *PHACTR1*, *SUGCT/C7orf10*, *ITPK1* were replicated in all three studies. The identified *CALCB* was only found in the study of Hautakangas et al. (18).

Table 9. Genome-wide and suggestively significant risk loci ($p < 10^{-5}$) identified by genome-wide association analysis comparing migraine-first patients and healthy controls.

Chr	Start	End	Independent significant SNPs	P-value	Gene
6	12768218	12948388	rs9349379, rs9381500, rs12202891	5.35×10^{-8}	*PHACTR1*
11	14980848	15121130	rs7396909, rs7394786	1.68×10^{-7}	*CALCB*
2	234804509	234867513	rs12988953, rs6709005, rs2302153, rs6760630	2.14×10^{-7}	*TRPM8*

7	40360982	40477363	rs12532479	4.31 x 10^{-7}	*SUGCT/ C7orf10*
10	53156442	53222958	rs113137506	5.41 x 10^{-7}	*PRKG1*
9	86342917	86683120	rs7045525, rs117722586	7.66 x 10^{-7}	*LOC101927575*
4	166062170	166128741	rs73009672, rs112524929	8.23 x 10^{-7}	-
6	91892908	91986228	rs9359932	9.49 x 10^{-7}	-
2	205842583	206337543	rs189372686	1.13 x 10^{-6}	*PARD3B*
2	231896204	231914845	rs6854	2.59 x 10^{-6}	*C2orf72*
14	93591673	93596315	rs11160100, rs7148352, rs11624776	2.64 x 10^{-6}	*ITPK1*
8	106594719	106643913	rs62527241	2.77 x 10^{-6}	*ZFPM2*
1	206685627	206685627	rs944770	3.23 x 10^{-6}	*RASSF5*
6	43455719	43633417	rs140119541	3.52 x 10^{-6}	*POLH, POLR1C*
16	68555187	68680651	rs12921494	3.66 x 10^{-6}	*ZFP90*
1	150204405	150513711	rs6693567, rs141292640	3.71 x 10^{-6}	*LOC124904415, RPRD2, ADAMTSL4*
1	156406381	156474929	rs1342442	3.8 x 10^{-6}	*MEF2D*
19	3042734	3042734	rs11084993	3.9 x 10^{-6}	*TLE2*
16	79207139	79207139	rs79101720	4.66 x 10^{-6}	*WWOX*
2	55564300	56101058	rs115740972	4.73 x 10^{-6}	*PPP4R3B*
2	145755449	145969331	rs79120566	4.81 x 10^{-6}	*LOC100505498*
7	153143250	153158288	rs117637348	5.09 x 10^{-6}	-
11	111109258	111111729	rs7119658	5.16 x 10^{-6}	-
2	224890196	224890196	rs151027228	5.51 x 10^{-6}	*SERPINE2*
1	100998783	101149093	rs115756886	5.85 x 10^{-6}	*LOC124904231*
10	26677166	26688963	rs111936115	7.28 x	-

				10^{-6}	
16	1056547	1076660	rs2432301	8.19 x 10^{-6}	-
11	125283916	125456805	rs11220077	8.99 x 10^{-6}	*FEZ1*
20	25205031	25758201	rs3787078	9.02 x 10^{-6}	*PYGB*
7	20656054	20724826	rs10266137	9.15 x 10^{-6}	*ABCB5*
2	66228401	66228539	rs13429325	9.88 x 10^{-6}	*LINC02934*

Note: Chr: chromosome; start/end: genomic coordinates of the risk loci; independent significant SNPs: significant SNPs located in risk loci, found in GWAS analysis; gene: annotated genes associated with the loci.

We were able to explicitly replicate the significant tissue-specific enrichment of our risk loci, using suggestive significance SNP threshold ($p < 10^{-5}$), in tibial artery tissue ($p = 0.0001927$), consistent with findings reported by both Gormley et al. (17) and Hautakangas et al. (18). Enrichment in other tissues remained non-significant.

Pathway-level analysis based on risk loci defined by genome-wide significant (GWS) SNPs did not identify any significant gene sets. However, analysis using suggestively significant SNPs revealed enrichment in seven KEGG pathways (MsigDB C2 sets), including retinol metabolism and steroid hormone biosynthesis. Additionally, analysis of MsigDB C3 sets identified significant transcription factor-related enrichments involving STAT1 and PEA3. Detailed results can be found in Supplementary Figures S2, S3 of Torok et al. (52).

4.2. Study 2: validation of transcriptomic findings through genetic analyses in migraine

4.2.1. UK Biobank cohort descriptors

In these analyses, apart from migraine-first diagnosis, sex, and age, we included smoking status, allergy, vitamin A intake, and retinol equivalent intake (**Table 10**).

Table 10. **Sample characteristics of the UK Biobank cohort included in the genetic analyses of LEGs and vitamin A-related pathway's SNP-level investigations**.

Age	Mean	58.1743
	SE	0.0167
	Range	39-72
Smoking	*Cases*	*Controls*
	102976	66403
Allergy	Cases	*Controls*
	25321	144469
Migraine	*Cases*	*Controls*
	6139	193790
Sex	*Males*	*Females*
	101766	98163
Vitamin A retinol equivalents intake	Mean	978.9086
	SE	3.8709
	Range (μg)	0-41446
Vitamin A intake	Mean	482.1076
	SE	3.4461
	Range (μg)	0-38784

Note: SE - standard error; μg - micrograms.

4.2.2. Migraine transcriptomic cohort description

Mean age of the Hungarian migraine transcriptomic study population (27.5) was smaller than in the UK Biobank (58.2), and the migraine case-control ratio was more balanced (**Table 11**).

Table 11. Sample characteristics of the migraine transcriptomic cohort included in the genetic analyses of LEGs and SNP-level analyses related to vitamin A.

Age	Mean	27.5051
	SE	0.3868
	Range	19-49
Smoking	Mean	4.6329
	SE	0.0511
	Range	1-5
Sex	*Males*	*Females*
	83	206
Migraine	*Cases*	*Controls*
	172	117
Allergy	*Cases*	*Controls*
	106	177

Note: SE - standard error.

4.2.3. Findings from the Leading Edge Gene (LEG) genetic analyses

No results remained significant after applying Bonferroni correction for multiple testing ($p \leq 3.66 \times 10^{-6}$; 0.05/13,650). However, numerous variants showed nominal significance ($p \leq 0.05$), including those in *CBR3* ($N_{snp} = 46$), *CORIN* ($N_{snp} = 14$), *EFNB2* ($N_{snp} = 5$), *KCNMA1* ($N_{snp} = 90$), and *TEK* ($N_{snp} = 6$) in the models including covariates sex, age, allergy, and smoking. Detailed results for each SNPs can be found in Petschner et al. - Supplementary Table 10 (56).

4.2.4. Results from the analyses of genes involved in vitamin A pathway

The result of the t-test showed that the vitamin A intake was significantly different in migraineurs compared to controls. In case of vitamin A retinol equivalent intake, we did not find a significant difference (**Table 12**).

Table 12. Results of the comparison of vitamin A and retinol equivalent intake between the migraine and control groups.

	Vitamin A intake mean	t-test p-value	Vitamin A retinol equivalent intake mean	t-test p-value
migraineurs	426.7946	0.00021	946.5640	0.06491
controls	484.0020		980.0163	

4.2.4.1. UK Biobank cohort

Eleven variants in the *LRP1* gene, rs4759276, rs4759275, rs1385526, rs1799737, rs1466535, rs10876964, rs4759045, rs10876965, rs11172113, rs4367982, and rs4759277, were significantly associated with migraine both for controlling for vitamin A and vitamin A retinol equivalent intake, with p-values (**Table 13**) exceeding the multiple testing correction threshold of 3.66×10^{-6} (0.05/13,635).

Table 13. Significant SNPs of vitamin A pathway genes

vitamin A intake		
ID	OR	P
rs4759276	0.84	1.46×10^{-8}
rs4759275	0.84	1.63×10^{-8}
rs1385526	0.84	6.29×10^{-8}
rs1799737	0.84	7.85×10^{-8}
rs1466535	0.84	8.63×10^{-8}
rs10876964	0.84	8.66×10^{-8}
rs4759045	0.84	9.23×10^{-8}
rs10876965	0.84	9.45×10^{-8}
rs11172113	0.85	1.22×10^{-7}
rs4367982	0.85	1.30×10^{-7}
rs4759277	0.85	1.57×10^{-7}
vitamin A retinol equivalent intake		
ID	OR	P
rs4759276	0.84	1.45×10^{-8}
rs4759275	0.84	1.62×10^{-8}
rs1385526	0.84	6.21×10^{-8}
rs1799737	0.84	7.79×10^{-8}

rs1466535	0.84	8.53×10^{-8}
rs10876964	0.84	8.56×10^{-8}
rs4759045	0.84	9.13×10^{-8}
rs10876965	0.84	9.35×10^{-8}
rs11172113	0.85	1.20×10^{-7}
rs4367982	0.85	1.28×10^{-7}
rs4759277	0.85	1.55×10^{-7}

Note: ID - unique identifier of SNPs, OR - odds ratio, P - p-value. We did correction for 1) age, sex, genotype array, smoking, allergy, and vitamin A intake, and 2) age, sex, genotype array, smoking, allergy, and vitamin A retinol equivalent intake.

4.2.4.2. Hungarian migraine cohort

Seven *LRP1* SNPs that were significant in the UK Biobank and present in Hungarian cohort showed nominal significance ($p \leq 0.05$) for migraine status, regardless of the applied corrections (**Table 14**).

Table 14. **Nominally significant SNPs of vitamin A pathway genes, correcting for 1) age, sex, smoking, allergy.**

ID	OR	P
rs10876964	0.60	0.0176
rs10876965	0.60	0.0176
rs1466535	0.60	0.0176
rs4759045	0.60	0.0176
rs4367982	0.61	0.0187
rs1799737	0.63	0.0287
rs4759277	0.63	0.0304

Note: ID - unique identifier of SNPs, OR - odds ratio, P - p-value.

Across all analyses for UK Biobank and Hungarian cohort, the minor alleles were associated with a protective effect (OR < 1) against migraine, suggesting increased risk for carriers of the major allele.

4.3. Study 3: genetic risk of NAD+/SIRT1 pathway in depression

4.3.1. Descriptive statistics

4.3.1.1. UK Biobank cohort

Discovery sample included 228,594 individuals with genetic data, age, genotype information, and depression score (**Table 15**).

Table 15. Mean and standard error of the age and depression score in the UKB discovery sample.

UK Biobank discovery sample				
		Total	Female (N=120392)	Male (N=108202)
Age	Mean	57.1867	57.1429	57.2355
	SE	±0.0170	±0.0230	±0.0251
Depression score	Mean	5.7041	5.8352	5.5581
	SE	±0.0046	±0.0065	±0.0065

As a target sample, as was explained in the methods (**Figure 1**), we used independent individuals from UK Biobank (**Table 16**), including data of 105,654 individuals. In our analyses, we utilized the same depression score, early life stress, body fat percentage, and PRS for NAD+/SIRT1 pathway.

Table 16. Descriptive statistics of the target sample.

UK Biobank target sample				
		Total	Female (N=59073)	Male (N=46581)
Age	Mean	56.1947	55.6724	56.8571
	SE	±0.0236	±0.0313	±0.0357
Depression score (range: 4-16)	Mean	5.3883	5.5133	5.2297
	SE	±0.0057	±0.0079	±0.0082

Early life stress (ELS) (range: 0-20)	Mean	1.6872	1.7740	1.5772
	SE	±0.0072	±0.0104	±0.0097
Pathway PRS	Mean	-0.0006	-0.0005	-0.0006
	SE	$\pm 1.5461 \times 10^{-5}$	$\pm 2.0727 \times 10^{-5}$	$\pm 2.3213 \times 10^{-5}$
Body fat percentage	Mean	30.7378	35.6352	24.5118
	SE	±0.0260	±0.0282	±0.0265

Note: Mean values and corresponding standard errors of age, depression score, early life stress (ELS), body fat percentage, and pathway-specific polygenic risk score in the UK Biobank target sample.

4.3.1.2. Hungarian cohort

In this study, we used the same cohort as used in study 2, but restricted only to control individuals, and considered depression score, early life stress, and calculated polygenic genetic risk for NAD+/SIRT1 pathway (**Table 17**).

Table 17. **Hungarian cohort description**

Hungarian cohort:				
		Total population	Females (N= 58)	Males (N=44)
Age	Mean	26.0294	26.1035	25.9318
	SE	±0.5132	±0.7873	±0.5919
Depression score (range: 22-55)	Mean	33.4608	34.1724	32.5227
	SE	±0.5808	±0.8341	±0.7649
Early life stress (ELS) (range: 25-90)	Mean	33.9902	33.00	35.2955
	SE	±1.0109	±1.0426	±1.8960

Pathway PRS	Mean	0.0024	0.0026	0.0020
	SE	±0.0004	±0.0006	±0.0007

Note: Mean and standard error of depression score, early life stress (ELS), and pathway-specific polygenic risk score in the Hungarian cohort.

4.3.2. Interacting effect of PRS and early life stress on depression score

Significant p-values for the interaction effect ($p \leq 0.016$; 0.05/3) were observed in the total sample and the male subpopulation. For the male subpopulation, we identified the largest percentage change in explained variance (ΔR^2_{perc}) (**Table 18**).

Table 18. **Interacting effect of PRS and early life stress.**

Effect of PRS x ELS on depression score			
	Total population	*Females*	*Males*
Explained variance change (ΔR^2_{perc})	0.0992%	0.0143%	0.5037%
P-value of PRS x ELS	**0.0074***	0.4623	**0.0002***
Effect size and SE of PRS x ELS	1.2615(±0.4712)	0.4392(±0.5976)	2.9275(±0.7768)

Note: Percentage changes in explained variance (ΔR^2_{perc}) for depression score resulting from the interaction between early life stress (ELS) and genetic risk associated with NAD+/SIRT1 pathway genes, analyzed in the total sample and in subpopulations. P-values reflect the significance of the interaction term between PRS and ELS; bold and * indicate values that remain significant after multiple testing correction ($p \leq 0.0167$; 0.05/3). Effect size refers to the beta coefficient, with the corresponding standard error (SE) shown in parentheses. PRS: polygenic risk score; ELS: early life stress.

4.3.3. Mediating role of body fat percentage in the interaction between PRS and ELS

In this model, consistent with the regression analyses, the interaction between PRS and ELS had a significant direct effect (c) on depression ($p \leq 0.05$) in both the total sample and the male subpopulation. Body fat percentage also demonstrated a significant direct effect (b) on depression score across all groups ($p \leq 0.05$). However, no significant indirect effects (a x b) via body fat percentage were found (**Table 19**), indicating that body fat does not mediate the interaction effect. A visual summary of the mediation analyses is provided in **Figure 1.**

Table 19. Results of mediation analyses in the total population and subgroups.

Total population			
	Beta	SE	P
(a) effect of interaction on body fat percentage	0.0111	±1.6434	0.9946
(b) effect of body fat percentage on depression score	0.0248	±0.0009	**$< 2.2 \times 10^{-16}$***
(a x b) effect of interaction through body fat percentage on depression score	0.0003	±0.0408	0.9946
(c) effect of interaction on depression score	1.1768	±0.4738	**0.0130***
(a x b + c) total effect of interaction and body fat percentage on depression score	1.1771	±0.4756	**0.0133***
Female subpopulation			
	Beta	SE	P
(a) effect of interaction on body fat percentage	-2.5926	±2.1697	0.2321

(b) effect of body fat percentage on depression score	0.0269	±0.0011	**< 2.2 x 10^{-16}***
(a x b) effect of interaction through body fat percentage on depression	-0.0697	±0.0584	0.2327
(c) effect of interaction on depression score	0.4167	±0.6000	0.4874
(a x b + c) total effect of interaction and body fat percentage on depression score	0.3469	±0.6028	0.5649
Male subpopulation			
	Beta	SE	P
(a) effect of interaction on body fat percentage	4.8282	±2.5347	0.0568
(b) effect of body fat percentage on depression score	0.0209	±0.0014	**< 2.2 x 10^{-16}***
(a x b) effect of interaction through body fat percentage on depression score	0.1011	±0.0535	0.0589
(c) effect of interaction on depression score	2.7691	±0.7827	**0.0004***
(a x b + c) total effect of interaction and body fat percentage on depression score	2.8703	±0.7845	**0.0003***

Note: Beta values represent effect sizes, with SE indicating the corresponding standard error. P-values: Significant results ($p \leq 0.05$) are indicated in bold and marked with an asterisk (*). PRS: polygenic risk score; ELS: early life stress.

4.3.4. Sex-specific association between NAc resting-state functional connectivity and the interaction of PRS and ELS in the Hungarian cohort

No significant association was found between NAc resting-state functional connectivity and the interaction between NAD+/SIRT1 pathway PRS and ELS in the overall sample. However, a sex-specific pattern emerged. Males exhibited stronger connectivity between the NAc and both the middle frontal gyrus (Peak T = 3.967; MNI coordinates: 44, 30, 30) and the triangular part of the inferior frontal gyrus (Peak T = 4.142; MNI: 42, 22, 30) compared to females (pFWE = 0.0139, cluster size k = 159; **Figure 3**).

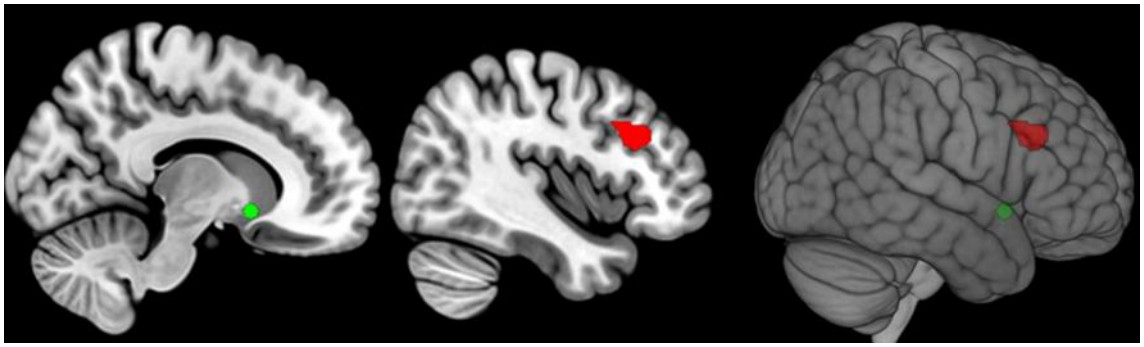

Figure 3. Brain regions showing significant connectivity with the nucleus accumbens (NAc) in males, associated with the interaction between NAD+/SIRT1 pathway genetic risk and early life stress. Green: NAc seed region (MNI coordinates: x = 10, y = 12, z = –8; radius = 4 mm). Red: Middle frontal gyrus and triangular part of the inferior frontal gyrus. Significance threshold: cluster-level pFWE < 0.05, MNI: Montreal Neurological Institute space.

In males, increased interaction scores between PRS and ELS were associated with enhanced functional connectivity between the NAc and the frontal regions (β = 0.3859, SE = 0.1513, R^2 = 0.1135). In contrast, females showed the opposite trend, where higher interaction scores were linked to weakened connectivity (β = -0.3850, SE = 0.1104, R^2 = 0.1637) (**Table 20 and Figure 4**).

Table 20. Shows the sex-specific association between resting-state functional connectivity of the NAc (Z-scores) and the interaction between NAD+/SIRT1 pathway PRS and early life stress (ELS).

Linear regression results				
	Beta	SE	P	R^2
Females	-0.3850	±0.1104	0.0010	0.1637
Males	0.3859	±0.1513	0.0145	0.1135

Note: beta - effect size; SE - standard error; P - p-value of the association; R^2- coefficient of determination.

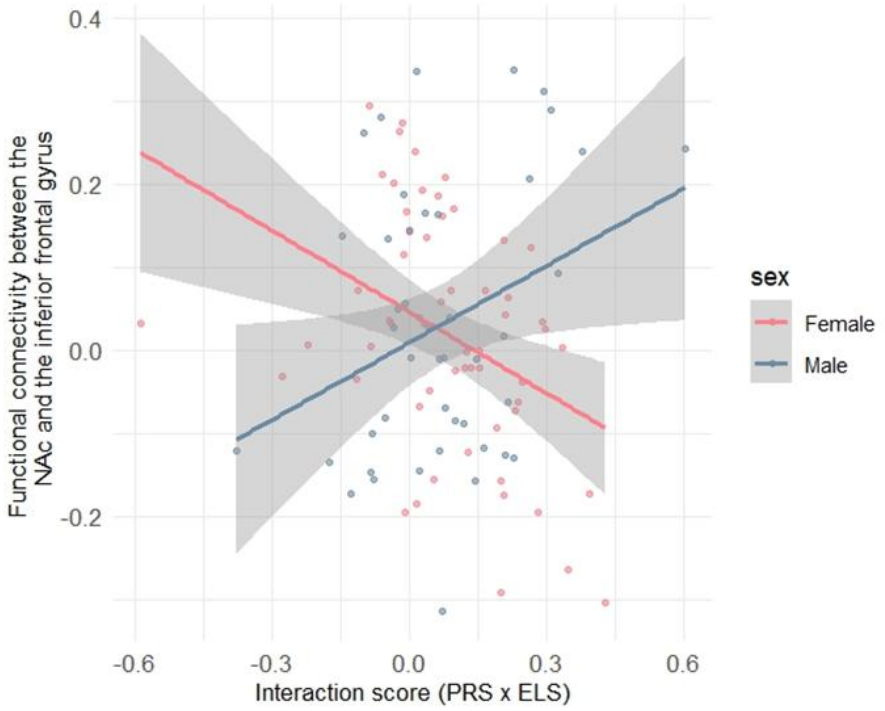

Figure 4. Sex-specific functional connectivity between the nucleus accumbens and the inferior frontal gyrus is modulated by the interaction between genetic risk in the NAD+/SIRT1 pathway and early life stress. The X-axis shows the interaction between PRS and ELS, while the Y-axis displays the Z-score of functional connectivity between the NAc and the inferior frontal gyrus. Shaded regions indicate standard errors (SE).

5. Discussion

Question 1

Previous large-scale meta-analyses have identified 38 and 123 genome-wide significant loci using cohorts of 59,674 cases with 316,078 controls (17) and 102,084 cases with 771,257 controls (18), reaching p-values as low as 5.6×10^{-49} and 10^{-90}, respectively. These efforts also reported SNP-based heritability estimates of 14.6% and 11.2%, underscoring migraine's polygenic nature.

In contrast, by focusing on a more homogeneous migraine-first subgroup, we identified a 45.96% and 90.27% increase in SNP-based heritability, reaching 19.37%, despite our considerably smaller sample size. Notably, only 1.25% of this heritability was accounted for by genome-wide significant and suggestive hits (**Figure 2**), reinforcing the polygenic architecture of migraine and suggesting that the observed increase in heritability largely arises from widespread genetic effects of non-genome-wide significant variants. Furthermore, all genome-wide significant loci we detected (5/5) overlapped with previously reported findings, and 31% (11/36) (**Table 8**) of our suggestive associations also aligned with previous results.

The identified genes are biologically relevant and consistent with prior literature: RDM16 is involved in thermoregulation, adult neurogenesis, and cerebrospinal fluid regulation (74, 75), FHL5 plays roles in vascular remodeling and gene regulation (76). ASTN2 is associated with neuronal migration and psychiatric conditions such as schizophrenia (77), LRP1 participates in vitamin A transport, immune modulation, lipid metabolism, and various vascular diseases (Alzheimer's disease, aortic and coronary artery conditions) (78, 79), SLC24A3 is linked to sodium/calcium exchange and intracellular calcium balance (24, 80).

Among our findings reaching suggestive significance, we identified *CALCB*, a gene encoding a calcitonin gene-related peptide (CGRP). Notably, among previous large-scale meta-analyses, only the study by Hautakangas et al. (18) detected this gene, likely due to their extensive sample size. Remarkably, our smaller-scale study was also able to capture this signal, which we assign to the use of a well-defined and refined migraine phenotype. The relevance of this finding is emphasized by the growing body of literature linking CGRP to migraine pathophysiology (81, 82).

In recent years, migraine treatment has advanced significantly through the development of therapies targeting CGRP or its receptor (83). These include small-molecule receptor antagonists (gepants) for acute migraine and monoclonal antibodies (Eptinezumab, Fremanezumab, Galcanezumab, and Erenumab) for migraine prevention. Unlike earlier treatments, these CGRP-targeted drugs are designed specifically to modulate the trigeminal pain system, offering greater specificity and minimal side effects (83).
Our findings, therefore, underscore the importance of phenotype refinement focusing on comorbidity exclusion in genetic studies, as it enables the detection of biologically meaningful signals, even in studies with small sample sizes.

At pathway level, enrichment of STAT1-, PEA3-, and retinol metabolism-related gene sets highlights potential roles of immune and vitamin A-related processes in migraine (84–86). Moreover, tissue-specific enrichment in the tibial artery supports the vascular hypothesis of migraine pathophysiology (87).
These findings are in line with previous literature but offer an important refinement: while meta-analyses using broader, heterogeneous disease definitions have greater statistical power to detect genome-wide significant loci, they may underrepresent the polygenic contribution underlying migraine. Our data suggest that the genetic signal captured by top hits represents only a small fraction of overall heritability, and that variants under the genome-wide significance threshold collectively account for a much larger portion of the genetic architecture. Thus, taking into account and interpreting top hits alone may provide an incomplete understanding of migraine genetics.
From a clinical perspective, three implications arise: 1) More precise genetic risk estimation may be achievable using better-defined subgroups such as migraine-first patients. 2) The search for novel therapeutic targets may benefit from focusing on biological pathways enriched in this polygenic background, rather than only genome-wide significant loci. 3) Given that a major part of migraine heritability reflects small-effect variants, multi-target therapeutic strategies may have greater promise than single-target approaches.
In summary, our findings indicate that patients diagnosed with migraine-first represent a genetically informative subgroup, characterized by higher SNP-based heritability and consistent replication of known risk loci.

The majority of this heritability remains unexplained by genome-wide significant variants, supporting the need to consider the broader polygenic background.

Question 2

Genomic analysis of the LEGs, selected based on our transcriptomic study of interictal migraine without aura patients, did not show significant results. We showed that top leading edge gene, *CYP26B1*, downregulation is not explained by genetic polymorphisms, but rather reflects dysregulated retinoic acid receptor signaling, driven by reduced retinoic acid availability linked to disease-associated variants in the *LRP1* gene.

In our previous genome-wide association analysis (Study 1), *LRP1* emerged as one of the top replicated risk loci among migraine-first patients. In line with these findings, Study 2 provided evidence for the importance of *LRP1* independently of the different vitamin A retinol equivalent intake of migraineurs, reinforcing its central role through pathway-level genetic analyses.

Although candidate gene-level associations across differentially expressed genes did not yield genome-wide significant SNPs in migraine-first patients, suggesting limited direct overlap between transcriptomic changes and individual common variants, even in this well-defined migraine subgroup, *LRP1* provided replicable evidence. Specifically, we identified eleven polymorphisms within the *LRP1* gene that remained significant after correction for multiple testing (**Table 13**) and seven replicated investigating an independent population of Hungarian migraine cohort (**Table 14**). This highlights a potential functional role of LRP1 in retinoid signaling dysregulation, which may be behind migraine.

Mechanistically, LRP1 facilitates the hepatic uptake of retinyl ester-containing chylomicron, a key step in storing and redistributing vitamin A derivatives to peripheral tissues (61). Moreover, it mediates the internalization of serum amyloid A-retinol complexes in immune cells within the intestinal tract (79). These dual functions suggest that LRP1 influences both immune cell vitamin A access and peripheral retinoic acid availability. Disturbance of these processes, due to disease-associated SNPs in LRP1, may also be involved in reductions in retinoic acid levels, impair retinoic acid receptor signaling, and suppress downstream targets such as CYP26B (88).

Taken together, these results provide context for the transcriptomic findings in the form of *LRP1* as a genetic marker of dysregulated retinoid signaling in migraine. This not only aligns with its known vascular and immune-related roles but also proposes novel perspectives on migraine pathophysiology.

The convergence of findings across genome-wide and pathway-level analyses highlights the importance of extending traditional GWAS approaches. By examining biologically coherent pathways, such as vitamin A metabolism pathway, we can better characterize the underlying mechanisms in migraine and identify candidate targets for precision treatment, even if in naive, systemic analyses these genes would otherwise remain insignificant.

Question 3

This study's results present new evidence that the genetic risk of NAD+/SIRT1 signaling pathway in interaction with early life stress (ELS) influences depressive symptoms in human males. This gene-environment interaction appears to impact depressive symptom scores independently of body fat percentage. Additionally, we found that this interaction alters the functional connectivity of the nucleus accumbens, a key brain region in reward processing, with connectivity patterns differing between males and females. The results extend previous findings from animal models to humans and highlight the NAD+/SIRT1 pathway in interaction with ELS behind depression.

The observed interaction between NAD+/SIRT1 RPS and ELS underscores the role of early environmental exposures and genetic risk of depression. This aligns with prior human studies reporting decreased SIRT1 expression in individuals exposed to childhood adversity, correlating with depression severity (89). Another finding in a rodent model has shown that early-life stress downregulates SIRT1 expression in the brain and in peripheral blood, which is linked to adult depressive-like behavior, and pharmacological treatment reversed this process along depressive symptomatology (90). Our current findings, in line with the literature, provide important translational insights by suggesting that there is an interplay between ELS and NAD+/SIRT1 pathway genetic risk in humans. Though the variance explained by the PRS × ELS interaction is modest (0.5037% in males, Table 18) is relatively high for a pathway-level result, especially when compared

to the typical genome-wide PRS effect sizes, which generally explain about 2-12% of variance on the liability scale (91, 92).

Our results suggest that the observed effects are not limited to SIRT1 but extend to other components of the NAD+/SIRT1 pathway. In the pathway, many genes encode mitochondrial proteins. This supports emerging literature pointing to mitochondrial dysfunction, particularly in nuclear-encoded mitochondrial genes, as a contributor to depression (93).

Given the link between mitochondrial processes and cognitive symptoms of depression, future research could explore whether cognitive deficits in depression are modulated by genetic variation in this pathway.

Contrary to findings from rodent studies where ELS was associated with increased adiposity and SIRT1 alterations, our results suggest that in humans, body fat percentage does not mediate the effect of NAD+/SIRT1 genetic risk on depressive symptoms. It is important to consider that our UK Biobank cohort has a mean age of approximately 57 years (**Table 15**), and the timing and duration of ELS in this population vary widely, in contrast to the animal models. This difference may have influenced the observed lack of mediation by body fat percentage in our sample.

Several genes within this pathway encode proteins involved in oxidative stress responses and metabolic regulation as well (*NPY*, *PPARGC1A*, *CAT*, *GPX4*, and *HSPD1*) (62). *NPY*, the gene for neuropeptide Y, is particularly remarkable, as it has been implicated in both major depressive disorder through interaction with ELS (94) and in obesity and energy metabolism (95). Despite evidence from original findings of Morató et al. (42) and established associations between high body mass index and depression (96, 97), our data did not support a mediating role of adiposity in the interaction between NAD+/SIRT1 genetic risk and ELS in depression.

The NAc, central component of reward processing (98), showed altered functional connectivity with middle frontal gyrus and inferior frontal gyrus in males exposed to ELS with high NAD+/SIRT1 PRS (**Table 20**, **Figure 4**). These regions are located in the dorsolateral prefrontal cortex. Early life stress has been linked to increased medial

prefrontal cortex (mPFC) connectivity, which is associated with higher levels of symptoms of depression and anxiety (99).

In addition, the prefrontal cortex regulates reward processing by exerting top-down control over the striatum (100); thus, increased connectivity between the NAc and middle frontal gyrus (part of the dorsolateral PFC) may reflect reduced hedonic capacity following ELS. This aligns with recent studies linking elevated NAc-PFC connectivity to melancholic depression (101) and anhedonia (102). Our results are in line with this evidence and suggest that the interaction of NAD+/SIRT1 PRS and ELS may influence reward-related brain circuits in a sex-dependent manner, contributing to depressive symptoms in males.

Our findings reveal a male-specific effect of the interaction of NAD+/SIRT1 PRS and ELS on depressive symptoms and functional connectivity. This sex difference is consistent with prior work, including rodent studies of Morató et al. (42), where ELS-induced increased fat mass and reduced sociability were only observed in males. Depression is more prevalent in women during reproductive years (7), and our findings emphasize that the biological mechanisms underlying depression may differ by sex. Rodent studies often rely exclusively on male animals due to the complexity of female reproductive cycles. Even when females are included, the effects observed in males are frequently absent (103). Notably, this study has successfully translated findings from male rodents to human males. Our findings underscore that meaningful translation in the form of genomic identification is possible when sex differences are properly considered, emphasizing the importance of including both sexes in animal research and recognizing sex as a critical factor in cross-species studies of depression.

In the study by Morató et al. (42), treatment with nicotinamide mononucleotide reversed behavioral effects of early life stress in adult mice. Similarly, resveratrol, an activator of SIRT1 (39), has shown antidepressant effects in animal models (104). Our findings highlight the relevance of the interaction of NAD+/SIRT1 pathway and ELS in depression, in males, and help define a subgroup that may benefit most from such targeted interventions.

Limitations

A key limitation of Study 1 is that the available migraine diagnosis (ICD code G43 - migraine) does not distinguish between migraine with or without aura. Future studies should aim to differentiate between these subtypes, ideally focusing on individuals with a first diagnosis of migraine without comorbidities, in order to gain a more precise understanding of the distinct genetic backgrounds associated with each type.

In Study 2, the main limitation when comparing transcriptomic and genetic analyses findings is that genetic association analyses reflect tissue-independent associations. Another limitation is that we were unable to measure retinoic acid (vitamin A) levels in our small Hungarian migraine cohort, likely due to degradation by the time of analysis.

The following limitations should be acknowledged when interpreting the results of Study 3. First, the division of the UK Biobank data into discovery and target samples was based on the availability of early life stress data. First, using the same cohort data for both discovery and validation may raise concerns about representativeness. However, since participation in the online follow-up and filling out the Childhood Trauma Screener assessing early life stress occurred randomly, the risk of non-representative bias is likely reduced.

Second, the genetic coverage of the NAD+/SIRT1 pathway in our analyses was incomplete. This was primarily due to standard genomic QC procedures and the unavailability of mitochondrial DNA data in the UK Biobank and in the Hungarian sample. While this represents a limitation, it is worth noting that mitochondrial function is influenced by nuclear-encoded genes, and common variants in these genes have been shown to affect depressive phenotypes with modest but less compensable effects (93). Therefore, it is unlikely that inclusion of mitochondrial variants would change the observed patterns; rather, it could further strengthen the associations found. Nonetheless, future research incorporating mitochondrial genomic data could offer a complete understanding of the role of mitochondrial function in ELS-related depression.

6. Conclusion

This PhD thesis demonstrates that enhancing genetic signal detection in polygenic disorders is achievable through the integration of two strategies: clear phenotype definition and biologically grounded, mainly transcriptomics-driven pathway analysis. In the context of migraine, refining the phenotype by focusing on migraine-first individuals without comorbidity led to improved heritability estimates and clearer identification of genetic loci. When combined with transcriptomic findings-based pathway-level focus, particularly on the vitamin A pathway, these analyses enabled more targeted, replicable findings. Translating findings from an animal study, calculating pathway-level polygenic risk score, and testing its interaction with early environmental adversity, it became possible to identify interacting genetic influences on both depressive symptoms and on functional connectivity levels.

These findings emphasize the importance of phenotype definition and hypothesis-driven statistical analyses in genetics. Rather than seeking general associations in broad, noisy samples, this work encourages targeted, biologically informed analyses, whether to explore migraine pathomechanism or to translate neurobiological insights from animal models into human contexts. This thesis contributes to a growing framework for more interpretable and biologically meaningful genetic research in various disorders.

7. Summary

The work presents novel strategies to improve genetic signal detection in polygenic disorders by combining refined phenotype definitions with biologically grounded, pathway-level analyses. Using migraine and depression phenotypes, it highlights how hypothesis-driven insights improve discovery and interpretability in genetics.

In migraine, focusing on individuals without comorbidities at disease onset, those whose first diagnosis was migraine, increased SNP-based heritability to 19.37%, compared to the previously reported 14.6%. This refined group showed stronger genetic signals, replicated known migraine loci (*PRDM16*, *FHL5*, *ASTN2*, *STAT6*/*LRP1*, *SLC24A3*), and identified suggestive associations in *CALCB*, a key CGRP coding gene.

Integrating findings of transcriptomic study in migraine, focusing on the vitamin A pathway, provided additional biological context. Although most differentially expressed genes showed no significant association, *LRP1* emerged as a key gene: eleven SNPs were significantly associated with migraine, and seven replicated in an independent cohort, supporting its involvement in pathophysiology.

The pathway-based approach was then extended to depression. Based on rodent findings, NAD+/SIRT1 pathway polygenic risk scores were calculated and tested in interaction with early life stress. In males, NAD+/SIRT1 genetic risk in interaction with early life stress explained variance from depressive symptoms at 0.5037% with p-value of 0.0002. Body fat percentage did not mediate this effect. Resting-state fMRI revealed that these individuals also showed altered connectivity between the nucleus accumbens and prefrontal cortex regions (females, beta = -0.3850, p-value = 0.0010, males, beta = 0.3859, p-value = 0.0145), involved in reward and anhedonia. Notably, the increased connectivity observed in males aligns with patterns previously reported in melancholic depression.

Together, these findings show that precise phenotype selection and pathway-focused genetic analyses can improve signal detection and biological relevance. This work supports a shift from broad, noisy samples toward targeted, replicable frameworks in genetics.

8. References

1. Wessman M, Terwindt GM, Kaunisto MA, Palotie A, Ophoff RA. Migraine: a complex genetic disorder. Lancet Neurol. 2007 Jun 1;6(6):521–32.

2. Goldberg D. The heterogeneity of "major depression". World Psychiatry. 2011;10(3):226–8.

3. Pietrobon D. Familial Hemiplegic Migraine. Neurotherapeutics. 2007 Apr 1;4(2):274–84.

4. Steiner TJ, Stovner LJ. Global epidemiology of migraine and its implications for public health and health policy. Nat Rev Neurol. 2023 Feb;19(2):109–17.

5. Burch RC, Buse DC, Lipton RB. Migraine: Epidemiology, Burden, and Comorbidity. Neurol Clin. 2019;37(4):631–49.

6. World Health O. Depression and other common mental disorders: global health estimates. Geneva: World Health Organization; 2017 2017. Contract No.: WHO/MSD/MER/2017.2.

7. Otte C, Gold SM, Penninx BW, Pariante CM, Etkin A, Fava M, et al. Major depressive disorder. Nat Rev Dis Primer. 2016 Sep 15;2(1):1–20.

8. American Psychiatric Association D-TF. Diagnostic and statistical manual of mental disorders: DSM-5™, 5th ed. Arlington, VA, US: American Psychiatric Publishing, Inc.; 2013. xliv, 947-xliv, p.

9. Gałecka M, Bliźniewska-Kowalska K, Maes M, Su KP, Gałecki P. Update on the neurodevelopmental theory of depression: is there any 'unconscious code'? Pharmacol Rep PR. 2020 Dec 31;

10. Lopresti AL, Hood SD, Drummond PD. A review of lifestyle factors that contribute to important pathways associated with major depression: Diet, sleep and exercise. J Affect Disord. 2013 May 15;148(1):12–27.

11. Martinelli D, Pocora MM, De Icco R, Putortì A, Tassorelli C. Triggers of migraine: where do we stand? Curr Opin Neurol. 2022 Jun 1;35(3):360–6.

12. Feinstein AR. The pre-therapeutic classification of co-morbidity in chronic disease. J Chronic Dis. 1970 Dec 1;23(7):455–68.

13. Low NCP, Merikangas KR. The Comorbidity of Migraine. CNS Spectr. 2003 Jun;8(6):433–44.

14. Radat F, Swendsen J. Psychiatric Comorbidity in Migraine: A Review. Cephalalgia. 2005 Mar 1;25(3):165–78.

15. 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 Comput Biol. 2017 Jun 23;13(6):e1005487.

16. Visscher PM, Brown MA, McCarthy MI, Yang J. Five Years of GWAS Discovery. Am J Hum Genet. 2012 Jan 13;90(1):7–24.

17. 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. Nat Genet. 2016 Aug;48(8):856–66.

18. Hautakangas H, Winsvold BS, Ruotsalainen SE, Bjornsdottir G, Harder AVE, Kogelman LJA, et al. Genome-wide analysis of 102,084 migraine cases identifies 123 risk loci and subtype-specific risk alleles. Nat Genet. 2022 Feb;54(2):152–60.

19. Olofsson IA. Migraine heritability and beyond: A scoping review of twin studies. Headache. 2024 Sep;64(8):1049–58.

20. Sullivan PF, Neale MC, Kendler KS. Genetic epidemiology of major depression: review and meta-analysis. Am J Psychiatry. 2000 Oct;157(10):1552–62.

21. Adams MJ, Streit F, Meng X, Awasthi S, Adey BN, Choi KW, et al. Trans-ancestry genome-wide study of depression identifies 697 associations implicating cell types and pharmacotherapies. Cell. 2025 Feb 6;188(3):640-652.e9.

22. Maher B. Personal genomes: The case of the missing heritability. Nature. 2008 Nov 1;456(7218):18–21.

23. Cai N, Revez JA, Adams MJ, Andlauer TFM, Breen G, Byrne EM, et al. Minimal phenotyping yields genome-wide association signals of low specificity for major depression. Nat Genet. 2020 Apr;52(4):437–47.

24. Petschner P, Baksa D, Hullam G, Torok D, Millinghoffer A, Deakin JFW, et al. A replication study separates polymorphisms behind migraine with and without depression. Padhukasahasram B, editor. PLOS ONE. 2021 Dec 31;16(12):e0261477.

25. Baksa D, Gonda X, Juhasz G. Why are migraineurs more depressed? A review of the factors contributing to the comorbidity of migraine and depression. Neuropsychopharmacol Hung Magy Pszichofarmakologiai Egyesulet Lapja Off J Hung Assoc Psychopharmacol. 2017 Mar;19(1):37–44.

26. Cantor RM, Lange K, Sinsheimer JS. Prioritizing GWAS Results: A Review of Statistical Methods and Recommendations for Their Application. Am J Hum Genet. 2010 Jan 8;86(1):6–22.

27. Wang K, Li M, Hakonarson H. Analysing biological pathways in genome-wide association studies. Nat Rev Genet. 2010 Dec;11(12):843–54.

28. Charles A. The pathophysiology of migraine: implications for clinical management. Lancet Neurol. 2018 Feb 1;17(2):174–82.

29. Kogelman LJ, Falkenberg K, Halldorsson GH, Poulsen LU, Worm J, Ingason A, et al. Comparing migraine with and without aura to healthy controls using RNA sequencing. Cephalalgia Int J Headache. 2019 Oct;39(11):1435–44.

30. Kogelman LJA, Falkenberg K, Buil A, Erola P, Courraud J, Laursen SS, et al. Changes in the gene expression profile during spontaneous migraine attacks. Sci Rep. 2021 Apr 15;11(1):8294.

31. Aczél T, Körtési T, Kun J, Urbán P, Bauer W, Herczeg R, et al. Identification of disease- and headache-specific mediators and pathways in migraine using blood transcriptomic and metabolomic analysis. J Headache Pain. 2021 Oct 6;22(1):117.

32. Gerring ZF, Powell JE, Montgomery GW, Nyholt DR. Genome-wide analysis of blood gene expression in migraine implicates immune-inflammatory pathways. Cephalalgia Int J Headache. 2018 Feb;38(2):292–303.

33. Perry CJ, Blake P, Buettner C, Papavassiliou E, Schain AJ, Bhasin MK, et al. Upregulation of inflammatory gene transcripts in periosteum of chronic migraineurs: Implications for extracranial origin of headache. Ann Neurol. 2016 Jun;79(6):1000–13.

34. Bremshey S, Groß J, Renken K, Masseck OA. The role of serotonin in depression—A historical roundup and future directions. J Neurochem. 2024;168(9):1751–79.

35. Le GH, Wong S, Lu A, Vasudeva S, Gill H, Badulescu S, et al. Electroencephalography (EEG) spectral signatures of selective serotonin reuptake inhibitors (SSRIs), selective norepinephrine reuptake inhibitors (SNRIs) and vortioxetine in major depressive disorder: A systematic review. J Affect Disord. 2025 Jan 1;368:798–819.

36. Nestler EJ, Barrot M, DiLeone RJ, Eisch AJ, Gold SJ, Monteggia LM. Neurobiology of depression. Neuron. 2002 Mar 28;34(1):13–25.

37. Hicks EM, Seah C, Cote A, Marchese S, Brennand KJ, Nestler EJ, et al. Integrating genetics and transcriptomics to study major depressive disorder: a conceptual framework, bioinformatic approaches, and recent findings. Transl Psychiatry. 2023 Apr 19;13(1):129.

38. Mariani N, Cattane N, Pariante C, Cattaneo A. Gene expression studies in Depression development and treatment: an overview of the underlying molecular mechanisms and biological processes to identify biomarkers. Transl Psychiatry. 2021 Jun 8;11(1):354.

39. Alageel A, Tomasi J, Tersigni C, Brietzke E, Zuckerman H, Subramaniapillai M, et al. Evidence supporting a mechanistic role of sirtuins in mood and metabolic disorders. Prog Neuropsychopharmacol Biol Psychiatry. 2018 Aug 30;86:95–101.

40. Lu G, Li J, Zhang H, Zhao X, Yan LJ, Yang X. Role and Possible Mechanisms of Sirt1 in Depression. Oxid Med Cell Longev. 2018;2018:8596903.

41. Xu J, Jackson CW, Khoury N, Escobar I, Perez-Pinzon MA. Brain SIRT1 Mediates Metabolic Homeostasis and Neuroprotection. Front Endocrinol. 2018;9:702.

42. Morató L, Astori S, Zalachoras I, Rodrigues J, Ghosal S, Huang W, et al. eNAMPT actions through nucleus accumbens NAD+/SIRT1 link increased adiposity with sociability deficits programmed by peripuberty stress. Sci Adv. 2022 Mar 4;8(9):eabj9109.

43. Smith DJ, Nicholl BI, Cullen B, Martin D, Ul-Haq Z, Evans J, et al. Prevalence and characteristics of probable major depression and bipolar disorder within UK biobank: cross-sectional study of 172,751 participants. PloS One. 2013;8(11):e75362.

44. Nagel M, Watanabe K, Stringer S, Posthuma D, van der Sluis S. Item-level analyses reveal genetic heterogeneity in neuroticism. Nat Commun. 2018 Mar 2;9(1):905.

45. Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, et al. The UK Biobank resource with deep phenotyping and genomic data. Nature. 2018 Oct;562(7726):203–9.

46. Eszlari N, Bruncsics B, Millinghoffer A, Hullam G, Petschner P, Gonda X, et al. Biology of Perseverative Negative Thinking: The Role of Timing and Folate Intake. Nutrients. 2021 Dec 8;13(12):4396.

47. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. GigaScience. 2015 Dec 1;4(1):s13742-015-0047–8.

48. Speed D, Balding DJ. SumHer better estimates the SNP heritability of complex traits from summary statistics. Nat Genet. 2019 Feb;51(2):277–84.

49. Altshuler DM, Gibbs RA, Peltonen L, Altshuler DM, Gibbs RA, Peltonen L, et al. Integrating common and rare genetic variation in diverse human populations. Nature. 2010 Sep;467(7311):52–8.

50. Kennedy AE, Ozbek U, Dorak MT. What has GWAS done for HLA and disease associations? Int J Immunogenet. 2017;44(5):195–211.

51. Speed D, Holmes J, Balding DJ. Evaluating and improving heritability models using summary statistics. Nat Genet. 2020 Apr;52(4):458–62.

52. Torok D, Petschner P, Baksa D, Juhasz G. Improved polygenic risk prediction in migraine-first patients. J Headache Pain. 2024 Sep 27;25(1):161.

53. Watanabe K, Taskesen E, van Bochoven A, Posthuma D. Functional mapping and annotation of genetic associations with FUMA. Nat Commun. 2017 Nov 28;8(1):1826.

54. de Leeuw CA, Mooij JM, Heskes T, Posthuma D. MAGMA: Generalized Gene-Set Analysis of GWAS Data. PLoS Comput Biol. 2015 Apr 17;11(4):e1004219.

55. Watanabe K, Umićević Mirkov M, de Leeuw CA, van den Heuvel MP, Posthuma D. Genetic mapping of cell type specificity for complex traits. Nat Commun. 2019 Jul 19;10(1):3222.

56. Petschner P, Kumar S, Nguyen DA, Torok D, Gal Z, Baksa D, et al. The interictal transcriptomic map of migraine without aura. J Headache Pain. 2025 May 12;26(1):109.

57. Headache Classification Committee of the International Headache Society (IHS). The International Classification of Headache Disorders, 3rd edition (beta version). Cephalalgia Int J Headache. 2013 Jul;33(9):629–808.

58. Karolchik D, Baertsch R, Diekhans M, Furey TS, Hinrichs A, Lu YT, et al. The UCSC Genome Browser Database. Nucleic Acids Res. 2003 Jan 1;31(1):51–4.

59. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res. 2010 Sep;38(16):e164.

60. Carazo A, Macáková K, Matoušová K, Krčmová LK, Protti M, Mladěnka P. Vitamin A Update: Forms, Sources, Kinetics, Detection, Function, Deficiency, Therapeutic Use and Toxicity. Nutrients. 2021 May 18;13(5):1703.

61. D'Ambrosio DN, Clugston RD, Blaner WS. Vitamin A metabolism: an update. Nutrients. 2011;3(1):63–103.

62. Stelzer G, Rosen N, Plaschkes I, Zimmerman S, Twik M, Fishilevich S, et al. The GeneCards Suite: From Gene Data Mining to Disease Genome Sequence Analyses. Curr Protoc Bioinforma. 2016 Jun 20;54:1.30.1-1.30.33.

63. Hullam G, Antal P, Petschner P, Gonda X, Bagdy G, Deakin B, et al. The UKB envirome of depression: from interactions to synergistic effects. Sci Rep. 2019 Jul 5;9(1):1–19.

64. Walker EA, Gelfand A, Katon WJ, Koss MP, Von Korff M, Bernstein D, et al. Adult health status of women with histories of childhood abuse and neglect. Am J Med. 1999 Oct;107(4):332–9.

65. Zung WW. A SELF-RATING DEPRESSION SCALE. Arch Gen Psychiatry. 1965;12:63–70.

66. Bernstein DP, Fink L, Handelsman L, Foote J, Lovejoy M, Wenzel K, et al. Initial reliability and validity of a new retrospective measure of child abuse and neglect. Am J Psychiatry. 1994;151(8):1132–6.

67. Durinck S, Moreau Y, Kasprzyk A, Davis S, De Moor B, Brazma A, et al. BioMart and Bioconductor: a powerful link between biological databases and microarray data analysis. Bioinforma Oxf Engl. 2005;21(16):3439–40.

68. Machiela MJ, Chanock SJ. LDlink: a web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants. Bioinforma Oxf Engl. 2015;31(21):3555–7.

69. Privé F, Arbel J, Vilhjálmsson BJ. LDpred2: better, faster, stronger. Bioinformatics. 2020 0 1;36(22–23):5424–31.

70. Baron RM, Kenny DA. The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. J Pers Soc Psychol. 1986;51:1173–82.

71. Rosseel Y. lavaan: An R Package for Structural Equation Modeling. J Stat Softw. 2012 May 24;48:1–36.

72. Gecse K, Baksa D, Dobos D, Aranyi CS, Galambos A, Kocsel N, et al. Sex Differences of Periaqueductal Grey Matter Functional Connectivity in Migraine. Front Pain Res Lausanne Switz. 2021;2:767162.

73. Haynos A, Hall L, Lavender J, Peterson C, Crow S, Klimes-Dougan B, et al. Resting state functional connectivity of networks associated with reward and habit in anorexia nervosa. Hum Brain Mapp. 2018 Sep 25;40.

74. 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 Dev. 2017 Jun 1;31(11):1134–46.

75. Ishibashi J, Seale P. Functions of Prdm16 in thermogenic fat cells. Temp Multidiscip Biomed J. 2015 Mar 17;2(1):65–72.

76. Wong D, Auguste G, Lino Cardenas CL, Turner AW, Chen Y, Song Y, et al. FHL5 Controls Vascular Disease-Associated Gene Programs in Smooth Muscle Cells. Circ Res. 2023 Apr 28;132(9):1144–61.

77. Hayashi Y, Okumura H, Arioka Y, Kushima I, Mori D, Lo T, et al. Analysis of human neuronal cells carrying ASTN2 deletion associated with psychiatric disorders. Transl Psychiatry. 2024 Jun 3;14(1):1–12.

78. Pirruccello JP, Rämö JT, Hoan Choi S, Chaffin MD, Kany S, Nekoui M, et al. The Genetic Determinants of Aortic Distension. J Am Coll Cardiol. 2023 Apr 11;81(14):1320–35.

79. Bang YJ, Hu Z, Li Y, Gattu S, Ruhn KA, Raj P, et al. Serum amyloid A delivers retinol to intestinal myeloid cells to promote adaptive immunity. Science. 2021 Sep 17;373(6561):eabf9232.

80. Albury CL, Stuart S, Haupt LM, Griffiths LR. Ion channelopathies and migraine pathogenesis. Mol Genet Genomics. 2017 Aug 1;292(4):729–39.

81. Scuteri D, Corasaniti MT, Tonin P, Nicotera P, Bagetta G. Role of CGRP pathway polymorphisms in migraine: a systematic review and impact on CGRP mAbs migraine therapy. J Headache Pain. 2021 Jul 30;22(1):87.

82. Ho TW, Edvinsson L, Goadsby PJ. CGRP and its receptors provide new insights into migraine pathophysiology. Nat Rev Neurol. 2010 Oct;6(10):573–82.

83. Edvinsson L, Haanes KA, Warfvinge K, Krause DN. CGRP as the target of new migraine therapies — successful translation from bench to clinic. Nat Rev Neurol. 2018 Jun 1;14(6):338–50.

84. Hu X, Li J, Fu M, Zhao X, Wang W. The JAK/STAT signaling pathway: from bench to clinic. Signal Transduct Target Ther. 2021 Nov 26;6(1):1–33.

85. McCabe KL, McGuire C, Reh TA. Pea3 expression is regulated by FGF signaling in developing retina. Dev Dyn. 2006;235(2):327–35.

86. Qi T, Qu Q, Li G, Wang J, Zhu H, Yang Z, et al. Function and regulation of the PEA3 subfamily of ETS transcription factors in cancer. Am J Cancer Res. 2020 Oct 1;10(10):3083–105.

87. Parsons AA, Strijbos PJ. The neuronal versus vascular hypothesis of migraine and cortical spreading depression. Curr Opin Pharmacol. 2003 Feb 1;3(1):73–7.

88. Osanai M, Takasawa A, Takasawa K, Kyuno D, Ono Y, Magara K. Retinoic acid metabolism in cancer: potential feasibility of retinoic acid metabolism blocking therapy. Med Mol Morphol. 2023 Mar 1;56(1):1–10.

89. Lo Iacono L, Bussone S, Andolina D, Tambelli R, Troisi A, Carola V. Dissecting major depression: The role of blood biomarkers and adverse childhood experiences in distinguishing clinical subgroups. J Affect Disord. 2020 Nov 1;276:351–60.

90. Lo Iacono L, Visco-Comandini F, Valzania A, Viscomi MT, Coviello M, Giampà A, et al. Adversity in childhood and depression: linked through SIRT1. Transl Psychiatry. 2015 Sep 1;5:e629.

91. Pardiñas AF, Holmans P, Pocklington AJ, Escott-Price V, Ripke S, Carrera N, et al. Common schizophrenia alleles are enriched in mutation-intolerant genes and in regions under strong background selection. Nat Genet. 2018 Mar;50(3):381–9.

92. Stahl EA, Breen G, Forstner AJ, McQuillin A, Ripke S, Trubetskoy V, et al. Genome-wide association study identifies 30 loci associated with bipolar disorder. Nat Genet. 2019 May;51(5):793–803.

93. Petschner P, Gonda X, Baksa D, Eszlari N, Trivaks M, Juhasz G, et al. Genes Linking Mitochondrial Function, Cognitive Impairment and Depression are Associated with Endophenotypes Serving Precision Medicine. Neuroscience. 2018 Feb 1;370:207–17.

94. Maul S, Giegling I, Fabbri C, Corponi F, Serretti A, Rujescu D. Genetics of resilience: Implications from genome-wide association studies and candidate genes of the stress response system in posttraumatic stress disorder and depression. Am J Med Genet B Neuropsychiatr Genet. 2020;183(2):77–94.

95. Zhang W, Cline MA, Gilbert ER. Hypothalamus-adipose tissue crosstalk: neuropeptide Y and the regulation of energy metabolism. Nutr Metab. 2014 Jun 10;11(1):27.

96. Yuan L, Su Y, Zhao J, Cho M, Wang D, Yuan L, et al. Investigating the shared genetic architecture between obesity and depression: a large-scale genomewide cross-trait analysis. Front Endocrinol. 2025 May 8;16:1578944.

97. Selman A, Dai J, Driskill J, Reddy AP, Reddy PH. Depression and obesity: Focus on factors and mechanistic links. Biochim Biophys Acta BBA - Mol Basis Dis. 2025 Jan 1;1871(1):167561.

98. Nestler EJ, Carlezon WA. The Mesolimbic Dopamine Reward Circuit in Depression. Biol Psychiatry. 2006 Jun 15;59(12):1151–9.

99. Hanson JL, Knodt AR, Brigidi BD, Hariri AR. Heightened connectivity between the ventral striatum and medial prefrontal cortex as a biomarker for stress-related psychopathology: understanding interactive effects of early and more recent stress. Psychol Med. 2018;48(11):1835–43.

100. Ferenczi EA, Zalocusky KA, Liston C, Grosenick L, Warden MR, Amatya D, et al. Prefrontal cortical regulation of brainwide circuit dynamics and reward-related behavior. Science. 2016;351(6268):aac9698.

101. Chen Z, Ou Y, Liu F, Li H, Li P, Xie G, et al. Increased brain nucleus accumbens functional connectivity in melancholic depression. Neuropharmacology. 2024 Feb 1;243:109798.

102. Hu Y, Zhao C, Zhao H, Qiao J. Abnormal functional connectivity of the nucleus accumbens subregions mediates the association between anhedonia and major depressive disorder. BMC Psychiatry. 2023 Apr 21;23(1):282.

103. Gururajan A, Reif A, Cryan JF, Slattery DA. The future of rodent models in depression research. Nat Rev Neurosci. 2019 Nov;20(11):686–701.

104. Moore A, Beidler J, Hong M. Resveratrol and Depression in Animal Models: A Systematic Review of the Biological Mechanisms. Molecules. 2018 Aug 30;23(9):2197.

9. Bibliography of the candidate's publications

Publications related to the PhD thesis

- **Torok D**, Petschner P, Baksa D, Juhasz G. Improved polygenic risk prediction in migraine-first patients. J Headache Pain. 2024 Sep 27;25(1):161. doi: 10.1186/s10194-024-01870-8. PMID: 39333847; PMCID: PMC11438044. IF: 7,9
- Petschner P, Kumar S, Nguyen DA, **Torok D**, Gal Z, Baksa D, Gecse K, Kokonyei G, Mamitsuka H, Juhasz G. The interictal transcriptomic map of migraine without aura. J Headache Pain. 2025 May 12;26(1):109. doi: 10.1186/s10194-025-02033-z. PMID: 40350427; PMCID: PMC12067696. IF: 7,9

Publications not related to the PhD thesis

- **Torok D**, Hegedus K, Varga H, Gonda X, Bagdy G, Petschner P. Neuroinflammation and Mood: Dissecting the Role of Eotaxin-1 in Depression. Neuropsychopharmacol Hung. 2025 Jun;27(2):115-120. English. PMID: 40608294.
- Erdelyi-Hamza B, **Torok D**, Krause S, Eszlari N, Bagdy G, Juhasz G, Gonda X. Disentangling nature and nurture: Exploring the genetic background of depressive symptoms in the absence of recent stress exposure using a GWAS approach. J Affect Disord. 2025 Jun 18;388:119731. doi: 10.1016/j.jad.2025.119731. Epub ahead of print. PMID: 40541832. IF: 4,9
- Kristof Z, Szabo D, Sperlagh B, **Torok D**, Gonda X. From Childhood Woes to Adult Blues: Unmasking the Role of Early Traumas, P2X7 Receptor, and Neuroinflammation in Anxiety and Depression. Int J Mol Sci. 2025 May 14;26(10):4687. doi: 10.3390/ijms26104687. PMID: 40429831; PMCID: PMC12111330. IF: 4,9
- Németh A, Gecse K, **Török D**, Baksa D, Dobos D, Aranyi CS, Emri M, Bagdy G, Juhász G. Hypothalamic connectivity strength is decreasing with polygenic risk in migraine without aura patients. Eur J Pharm Sci. 2025 Jul 1;210:107123. doi: 10.1016/j.ejps.2025.107123. Epub 2025 May 10. PMID: 40354987. IF: 4,7

- Lan A, **Torok D**, Einat H, Bagdy G, Juhasz G, Gonda X. Chronotypes are significantly associated with emotional well-being in a general population cohort in Hungary. Sci Rep. 2025 May 8;15(1):16054. doi: 10.1038/s41598-025-00893-8. PMID: 40341219; PMCID: PMC12062299. IF: 3,9
- Kumar S, Petschner P, Gecse K, **Torok D**, Juhasz G. Acute neuroendocrine challenge elicits enhanced cortisol response and parallel transcriptomic changes in patients with migraine. Pain Rep. 2025 May 1;10(3):e1254. doi: 10.1097/PR9.0000000000001254. PMID: 40322023; PMCID: PMC12047896. IF: 3,1
- Baksa D, Eszlari N, **Torok D**, Hullam G, Bagdy G, Juhasz G. Evening Chronotype Associates With Worse Physical and Mental Health and Headache-Related Disability Among Migraine Patients. J Sleep Res. 2025 Apr 10:e70066. doi: 10.1111/jsr.70066. Epub ahead of print. PMID: 40205879. IF: 3,9
- Krause S, **Torok D**, Bagdy G, Juhasz G, Gonda X. Genome-wide by trait interaction analyses with neuroticism reveal chronic pain-associated depression as a distinct genetic subtype. Transl Psychiatry. 2025 Mar 29;15(1):108. doi: 10.1038/s41398-025-03331-5. PMID: 40157903; PMCID: PMC11954882. IF: 6,2
- Gal Z, **Torok D**, Gonda X, Eszlari N, Anderson IM, Deakin B, Petschner P, Juhasz G, Bagdy G. New Evidence for the Role of the Blood-Brain Barrier and Inflammation in Stress-Associated Depression: A Gene-Environment Analysis Covering 19,296 Genes in 109,360 Humans. Int J Mol Sci. 2024 Oct 21;25(20):11332. doi: 10.3390/ijms252011332. PMID: 39457114; PMCID: PMC11508422. IF: 4,9
- Eszlari N, Hullam G, Gal Z, **Torok D**, Nagy T, Millinghoffer A, Baksa D, Gonda X, Antal P, Bagdy G, Juhasz G. Olfactory genes affect major depression in highly educated, emotionally stable, lean women: a bridge between animal models and precision medicine. Transl Psychiatry. 2024 Apr 8;14(1):182. doi: 10.1038/s41398-024-02867-2. PMID: 38589364; PMCID: PMC11002013. IF: 6,2
- Pothorszki D, Koncz S, **Török D**, Papp N, Bagdy G. Unique Effects of (R)-Ketamine Compared to (S)-Ketamine on EEG Theta Power in Rats.

Pharmaceuticals (Basel). 2024 Feb 1;17(2):194. doi: 10.3390/ph17020194. PMID: 38399409; PMCID: PMC10893209. IF: 4,8

- Kristof Z, Gal Z, **Torok D**, Eszlari N, Sutori S, Sperlagh B, Anderson IM, Deakin B, Bagdy G, Juhasz G, Gonda X. Embers of the Past: Early Childhood Traumas Interact with Variation in *P2RX7* Gene Implicated in Neuroinflammation on Markers of Current Suicide Risk. Int J Mol Sci. 2024 Jan 10;25(2):865. doi: 10.3390/ijms25020865. PMID: 38255938; PMCID: PMC10815854. IF: 4,9
- Kristof Z, Gal Z, **Torok D**, Eszlari N, Sutori S, Erdelyi-Hamza B, Petschner P, Sperlagh B, Anderson IM, Deakin JFW, Bagdy G, Juhasz G, Gonda X. Variation along P2RX7 interacts with early traumas on severity of anxiety suggesting a role for neuroinflammation. Sci Rep. 2023 May 12;13(1):7757. doi: 10.1038/s41598-023-34781-w. PMID: 37173368; PMCID: PMC10182087. IF: 3,8
- Gal Z, **Torok D**, Gonda X, Eszlari N, Anderson IM, Deakin B, Juhasz G, Bagdy G, Petschner P. Inflammation and Blood-Brain Barrier in Depression: Interaction of CLDN5 and IL6 Gene Variants in Stress-Induced Depression. Int J Neuropsychopharmacol. 2023 Mar 22;26(3):189-197. doi: 10.1093/ijnp/pyac079. PMID: 36472886; PMCID: PMC10032304. IF: 4,5
- Petschner P, Baksa D, Hullam G, **Torok D**, Millinghoffer A, Deakin JFW, Bagdy G, Juhasz G. A replication study separates polymorphisms behind migraine with and without depression. PLoS One. 2021 Dec 31;16(12):e0261477. doi: 10.1371/journal.pone.0261477. PMID: 34972135; PMCID: PMC8719675. IF: 3,752
- Ujma PP, Eszlári N, Millinghoffer A, Bruncsics B, **Török D**, Petschner P, Antal P, Deakin B, Breen G, Bagdy G, Juhász G. Genetic effects on educational attainment in Hungary. Brain Behav. 2022 Jan;12(1):e2430. doi: 10.1002/brb3.2430. Epub 2021 Nov 29. PMID: 34843176; PMCID: PMC8785634. IF: 3,1
- Bokor J, Sutori S, **Torok D**, Gal Z, Eszlari N, Gyorik D, Baksa D, Petschner P, Serafini G, Pompili M, Anderson IM, Deakin B, Bagdy G, Juhasz G, Gonda X. Inflamed Mind: Multiple Genetic Variants of *IL6* Influence Suicide Risk Phenotypes in Interaction With Early and Recent Adversities in a Linkage Disequilibrium-Based Clumping Analysis. Front Psychiatry. 2021 Oct

29;12:746206. doi: 10.3389/fpsyt.2021.746206. PMID: 34777050; PMCID: PMC8585756. IF: 5,435

- Gyorik D, Eszlari N, Gal Z, **Torok D**, Baksa D, Kristof Z, Sutori S, Petschner P, Juhasz G, Bagdy G, Gonda X. Every Night and Every Morn: Effect of Variation in *CLOCK* Gene on Depression Depends on Exposure to Early and Recent Stress. Front Psychiatry. 2021 Aug 26;12:687487. doi: 10.3389/fpsyt.2021.687487. PMID: 34512413; PMCID: PMC8428175. IF: 5,435
- Kristof Z, Eszlari N, Sutori S, Gal Z, **Torok D**, Baksa D, Petschner P, Sperlagh B, Anderson IM, Deakin JFW, Juhasz G, Bagdy G, Gonda X. P2RX7 gene variation mediates the effect of childhood adversity and recent stress on the severity of depressive symptoms. PLoS One. 2021 Jun 10;16(6):e0252766. doi: 10.1371/journal.pone.0252766. PMID: 34111150; PMCID: PMC8191953. IF: 3,752
- Gonda X, Eszlari N, **Torok D**, Gal Z, Bokor J, Millinghoffer A, Baksa D, Petschner P, Antal P, Breen G, Juhasz G, Bagdy G. Genetic underpinnings of affective temperaments: a pilot GWAS investigation identifies a new genome-wide significant SNP for anxious temperament in ADGRB3 gene. Transl Psychiatry. 2021 Jun 1;11(1):337. doi: 10.1038/s41398-021-01436-1. PMID: 34075027; PMCID: PMC8169753. IF: 7,989
- Bokor J, Krause S, **Torok D**, Eszlari N, Sutori S, Gal Z, Petschner P, Anderson IM, Deakin B, Bagdy G, Juhasz G, Gonda X. "Out, out, brief candle! Life's but a walking shadow": *5-HTTLPR* Is Associated With Current Suicidal Ideation but Not With Previous Suicide Attempts and Interacts With Recent Relationship Problems. Front Psychiatry. 2020 Jun 25;11:567. doi: 10.3389/fpsyt.2020.00567. PMID: 32670107; PMCID: PMC7331851. IF: 4,157

∑110,12

9. Acknowledgements

I would like to express my deepest gratitude to my supervisors, whose guidance, encouragement, and expertise have been important to every stage of this thesis. I am also thankful to all colleagues at the Department of Pharmacodynamics for their support and assistance. Special thanks go to Sándor Krause, Xénia Gonda, Nóra Eszlári, Kinga Gecse, and Dóra Pothorszki; from you, I gained not only scientific knowledge but also lasting friendships. I am deeply grateful to my hosts, Doug Speed and Hiroshi Mamitsuka, for broadening both my scientific horizons and my life experience. Finally, and most importantly, I thank my family, friends, and my cats; without you, this thesis would not have been possible.

The work was funded by the Hungarian National Research, Development, and Innovation Office 2019-2.1.7-ERA-NET-2020, (ERAPERMED2019-108); the Hungarian National Research, Development, and Innovation Office (K 143391, K 139330, PD 146014, and PD 134449 grants); the Hungarian Brain Research Programs (2017-1.2.1-NKP-2017-00002; KTIA_NAP_13-1-2013-000; KTIA_NAP_13-2-2015-0001; NAP2022-I-4/2022); TKP2021-EGA funding scheme (2020-4.1.1.-TKP2020; TKP2021-EGA-25 and TKP2021-EGA-02). Peter Petschner was an international research fellow of the Japan Society for the Promotion of Science (Postdoctoral Fellowships for Research in Japan, standard program, P20809). Dóra Török was a recipient of EKÖP-2024-68, Kinga Gecse was a recipient of EKÖP-2024-164.

The Journal of Headache and Pain

BRIEF REPORT Open Access

Improved polygenic risk prediction in migraine-first patients

Dora Torok[1,2†], Peter Petschner[1,2,3†], Daniel Baksa[1,2,4] and Gabriella Juhasz[1,2*]

Abstract

Background Recent meta-analyses estimated 14.6% and 11.2% SNP-based heritability of migraine, compared to twin-heritability estimates of 30–60%. This study aimed to investigate heritability estimates in "migraine-first" individuals, patients for whom G43 (migraine with or without aura) was their first medical diagnosis in their lifetime.

Findings Using data from the UK Biobank ($N = 199{,}929$), genome-wide association studies (GWAS) were conducted on 6,139 migraine-first patients and 193,790 healthy controls. SNP-based heritability was estimated using SumHer, yielding 19.37% (±0.019) for all SNPs and 21.31% (±0.019) for HapMap3 variants, substantially surpassing previous estimates. Key risk loci included *PRDM16*, *FHL5*, *ASTN2*, *STAT6/LRP1*, and *SLC24A3*, and pathway analyses highlighted retinol metabolism and steroid hormone biosynthesis as important pathways in these patients.

Conclusions The findings underscore that excluding comorbidities at onset time can enhance heritability estimates and genetic signal detection, significantly reducing the extent of "missing heritability" in migraine.

Keywords Migraine, Genetics, Missing heritability, Disease onset

Introduction

Previous large meta-analyses [1, 2] (henceforth, Gormley et al., Hautakangas et al.) estimated 14.6% and 11.2% single nucleotide polymorphism (SNP)-based heritability in migraine (with or without aura), respectively. In contrast, twin-heritability estimates of the disorder range from 30 to 60% [3]. Differences between SNP-based and twin-heritability estimates are considered "missing heritability", and are often attributed to various factors, like rare variants, environmental factors, or their epigenetic interactions. Some of these factors could be compensated during analyses.

A prominent example can be known comorbidities upon onset. We have previously identified the comorbidity network of migraine. Among others, a factor in this comorbidity map of the disorder was major depression [4]. A later study from our lab using machine learning methods provided substantial evidence that already diagnosed, comorbid major depression can influence relevance and effect size of genomic hits for migraine [5]. Another study reported altered heritability in the presence of comorbid depression [6]. Taken together, findings suggest, given the above are true for other comorbidities too, that migraine-first diagnosis may have a better characterizable genomic background. Thus, to investigate the biological background of the migraine-first diagnosis

† Dora Torok and Peter Petschner indicates shared first authorship.

*Correspondence:
Gabriella Juhasz
juhasz.gabriella@semmelweis.hu
[1]Department of Pharmacodynamics, Faculty of Pharmaceutical Sciences, Semmelweis University, Nagyvarad ter 4., Budapest 1096, Hungary
[2]NAP3.0-SE Neuropsychopharmacology Research Group, Hungarian Brain Research Program, Semmelweis University, Budapest, Hungary
[3]Bioinformatics Center, Institute of Chemical Research, Kyoto University, Gokasho, Uji, Kyoto, Japan
[4]Department of Personality and Clinical Psychology, Institute of Psychology, Faculty of Humanities and Social Sciences, Pazmany Peter Catholic University, Budapest, Hungary

is important to improve our understanding of migraine pathophysiology and to enhance migraine-specific drug target prediction, particularly as clinical drug trials mainly focus on migraine patients without other comorbidities.

Accurate migraine definition may also improve GWAS results and heritability estimates. Meta-analyses (Gormley et al., and Hautakangas et al.), to include as many individuals as possible and increase statistical power, often use data from self-reported questionnaires to identify participants with migraine. A clinical diagnosis or more homogeneous patient subgroups can provide more homogeneous samples and reduce noise for genomic methods.

In light of the above, our aim was to define a migraine-first population, i.e. migraine patients diagnosed with the disorder (G43, migraine with or without aura) without apparent comorbidities at the time of the diagnosis (henceforth: migraine-first patients) and run genome-wide association study and heritability estimation comparing them to age-matched controls without diagnosed diseases. Our goal was to test if such an approach is capable to discover relevant top hits or increase background polygenic signal in migraine.

Methods

Data used in our analysis were part of the UK Biobank (application number 71718). Genomic quality control steps and filtering were performed as described in the supplementary information of Eszlari et al. [7], including filtering for minor allele frequency, Hardy-Weinberg equilibrium, missingness, info and certainty filters, linkage disequilibrium pruning, kinship check, sex check, and heterozygosity outlier detection, leaving 334,125 participants with valid genomic data.

Migraine-first patients (N=6,139) were identified by the date of their migraine diagnosis (UKB data-field 131052 - Date G43 first reported [migraine]), with the criterion that migraine was their initial diagnosis. Healthy controls (N=193,790) were defined as individuals who had no recorded diseases at the same age (22±12 years, mean±standard deviation) when the migraine-first patients received their diagnosis and had no migraine diagnosis in their lifetime. Descriptive statistics were calculated with R (v 4.1.2) and can be found in Supplementary Table S1.

Genome-wide association analyses were performed with PLINK2 [8] on migraine-first patients and healthy controls. After quality control steps and exclusion of SNPs on chromosome X, 6,077,313 SNPs remained. The following covariates were included in the analyses: sex, age (UKB data-field 31 and 21003), first 10 principal components of the genome and genotype array (UKB Axiom Array and Affymetrix UK BiLEVE Axiom Array).

Evaluation of the GWAS result (risk loci identification, gene-set enrichment analysis) was performed with FUMA [9] and MAGMA [10]. Index variants were defined as LD-independent variants (r^2<0.1) and loci containing all variants in high linkage disequilibrium (r^2>0.6) within 250 kb distance to the index variants. (The definitions were the same as in Gormley et al. and Hautakangas et al.). We used two thresholds for lead SNP identification: a genome-wide significance (GWS) value of $p<10^{-8}$, and a suggestive significance of $10^{-8} < p<10^{-5}$.

SNP-based heritability was estimated with SumHer (part of the software package LDAK, version 5.2.) [11], following the authors' recommendation, using a one-parameter heritability model - The Human Default Model (specified by adding these parameters: --ignore weights YES and --power −0.25) for (1) all SNPs, (2) all SNPs with genome-wide (GWS) and suggestive significant SNPs excluded, and (3) SNPs restricted for the HapMap3 SNP-set. MHC region (chr 6, 25–35 Mb) was excluded from FUMA and MAGMA analyses and all heritability estimations. The estimated heritability was converted to liability scale using 0.03 "ascertainment" (reflecting that migraine proportion was 0.03 in our sample) and "prevalence" was set to 0.16 (as in Hautakangas et al.). During analyses no genomic inflation correction was used to be able to compare results with Gormley et al. and Hautakangas et al., and follow the guidelines of the authors of SumHer. In addition, such a correction is recommended only for low quality data with population structure and relatedness, which we controlled due to the standard quality control pipeline employed.

For validation purposes, we also tested (1) our control group definition by performing the same analysis steps with controls without restrictions (henceforth: all participants without migraine diagnosis in their lifetime, ($N_{\text{all\_controls}}$ = 327,986, Supplementary Table S2), (2) the influence of the small proportion of migraine patients in our population (ascertainment) by running and averaging three analyses after excluding a third of randomly selected healthy controls to increase the "ascertainment" parameter (Supplementary Table S3), (3) our phenotype definition by conducting the same analysis steps for all migraine diagnosed patients (G43) without filtering for first onset ($N_{\text{G43\_diagnosis}}$ = 17,679, $N_{\text{all\_controls}}$ = 316,446, Supplementary Table S4), (4) the effect of the heritability model [11–13] namely, genome-wide complex trait analysis (GCTA) using the following parameters: --ignore weights YES and --power 1; Human Default Model (HD) with --ignore weights YES and --power −0.25; LDAK model with --weights<weights for each variant>and --power −0.25; and BLD-LDAK model with --annotation-number 65, --annotation-prefix<baseline annotation categories>, --ignore-weights YES and --power

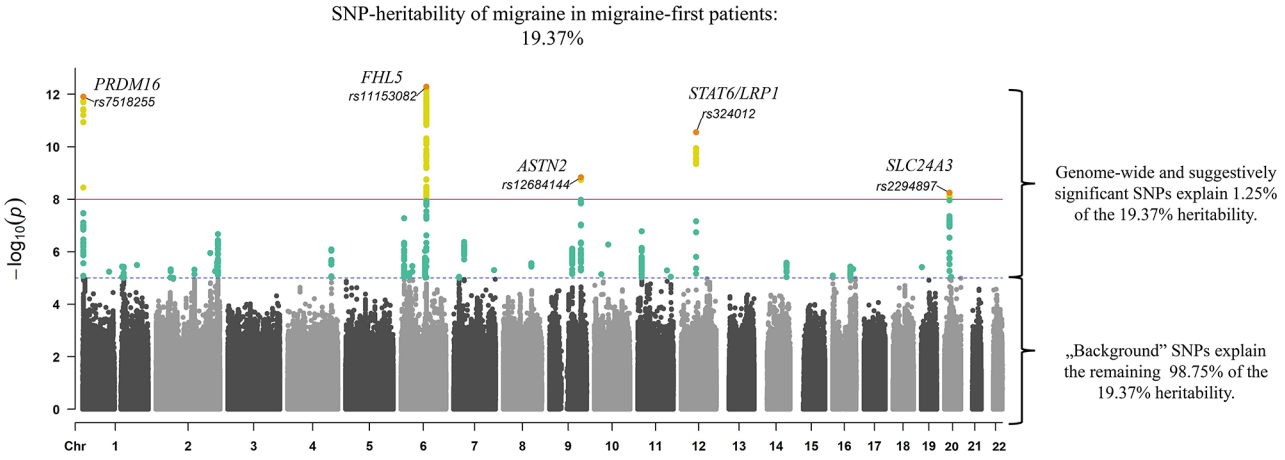

Fig. 1 Summary of the results: genome-wide significant loci and SNP-heritability of migraine-first patients and healthy controls

−0.25. All model assumptions are described in Supplementary Table S6.

In addition, to assess whether including variants with a lower MAF (MAF≥0.001) could enhance heritability estimation, we expanded our analysis to include these rare variants (MAF extended analysis).

Results

Figure 1 shows the five genome-wide significant risk loci identified both in our GWAS, in the study of Gorlmey et al., and Hautakangas et al. The red line denotes genome-wide significance, and the blue dashed line denotes the suggestive significance threshold. Yellow dots are genome-wide significant, and green dots are suggestively significant SNPs. The colored SNPs denote the combination of genome-wide- and suggestively significant SNPs (explaining 1.25% of the total heritability). In contrast, black and grey dots denote the "background" SNPs, explaining a large majority of the total heritability (98.75%). Migraine-first patients: initial filtering was based on G43 migraine (with or without aura) diagnosis and narrowed down to migraine-first patients for whom G43 (migraine with or without aura) was their first medical diagnosis in their lifetime. Healthy controls: individuals who had no recorded diseases at the same age when the migraine-first patients received their diagnosis and had no migraine diagnosis in their lifetime.

SumHer estimated SNP-based heritability was 19.37% (±0.019, SD) for migraine in migraine-first patients using all SNPs (Fig. 1). Excluding genome-wide and suggestively significant hits still resulted in a SNP-based heritability of 18.13% (±0.019, SD), indicating that the majority of the signal stems from the background polygenic signal. Restricting our SNP set to HapMap3 variants, in accordance with Gormley et al. and Hautakangas et al., gave a SNP-based heritability of 21.31% (±0.019, SD). This is a 45.96% and 90.27% increase compared to the study of Gormley et al. and Hautakangas et al., respectively. All heritability results can be found in Supplementary Table S5.

Table 1 Genome-wide and suggestively significant SNPs and risk loci

GWS SNPs				
	Current study	Gormley et al.	Hautakangas et al.	Overlap in all three
GWS SNPs	10	44(1)	170(n/a)	-
GWS risk loci	5	38(5)	123(5)	5
GWS + suggestive SNPs				
	Current study	Gormley et al.	Hautakangas et al.	Overlap in all three
GWS + suggestive SNPs	64	44(5)	170(n/a)	-
GWS + suggestive risk loci	36	38(11)	123(11)	11

Table 1 shows the overlapping SNPs and loci based on the results of Gormely et al., and Hautakangas et al., and our current study. Numbers in brackets denote the overlap with our current study. In case of Hautakangas et al., only risk loci data was available. GWS: genome-wide significant ($p<10^{-8}$), GWS + suggestive: significance threshold is $p<10^{-5}$, n/a -not available.

Full list of GWS and suggestively significant SNPs can be found in Supplementary Tables S9-S10. Among risk loci calculated using GWS SNPs (Supplementary Table S11), 100% of our present results were found in the previous meta-analyses. With suggestive significance threshold (to avoid a reduction in potential risk loci caused by reduced power in the current study, Supplementary Table S12), 30.56% of risk loci replicated in all studies (see Table 1). Prominent genes associated with GWS and suggestive significant risk loci were identified as: *PRDM16*, *FHL5*, *ASTN2*, *STAT6/LRP1*, *SLC24A3*, and the former five genes plus *ADAMTSL4*, *MEF2D*, *TRPM8*, *PHACTR1*, *SUGCT/C7orf10*, *ITPK1*, respectively.

We could explicitly replicate significant tissue-specific enrichment of our risk loci in tibial artery tissues (Supplementary Table S13), similarly to both Gormley et al. and Hautakangas et al. Other tissue-specific enrichment remained non-significant.

Using risk loci based on GWS SNPs, pathway-level analysis yielded no significant gene sets, while employing suggestive significant SNPs gave seven KEGG pathways (MsigDB C2 sets), including e.g., retinol metabolism and steroid hormone biosynthesis, among others (Supplementary Figure S2). Enrichment of MsigDB C3 sets resulted in STAT1 and PEA3 significant transcription factor-related sets (Supplementary Figure S3).

Validation efforts showed (Supplementary Table S5) that (1) although we observed a higher statistical power with more significant GWS hits increasing the number of controls, the heritability estimate was lower (h^2=17.39%, a 10.22% reduction on liability scale), (2) employing more balanced samples with higher proportion of migraine patients compared to controls decreased SNP-based heritability (h^2_{mean} = 18.45%, a 13.42% reduction on liability scale), probably due to the reduced statistical power, but still remained substantially higher than in previous studies, (3) heritability of G43 migraine diagnosis without filtering for first onset was low (h^2=12.92%, a 39.32% reduction on liability scale) (4) the impact of GCTA heritability model was small and yielded comparable results (h^2=23.89%) to HD. Results for all other heritability estimation analysis can be found in Supplementary Tables S5, S7 and S8. Results regarding MAF extended analyses showed that extending our analyses to include rarer variants can enhance the explained heritability, albeit, average contribution of the rarer variants was in general smaller compared to more common ones (Supplementary Table S7, Supplementary Fig. 1).

All in all, validation efforts indicated that control group definition, ascertainment parameter and heritability model could not explain the elevated SNP-based heritability in migraine-first patients, while the diagnosis alone without considering first onset on the UKB population yielded comparable heritability to Gormley et al. Although there was a modest enhancement in heritability estimation, the overall contribution of variants with lower MAF appears to be low.

Discussion

In our study, substantially elevated SNP-based heritability and replicable top hits in risk loci of *PRDM16*, *FHL5*, *ASTN2*, STAT6/*LRP1*, *SLC24A3* genes were found among migraine-first patients (Fig. 1).

Previous meta-analyses identified 38 and 123 risk loci with significance values reaching 5.6×10^{-49} and 10^{-90} in 59,674 cases and 316,078 controls, and 102,084 cases and 771,257 controls in the studies of Gormley et al. and Hautakangas et al., respectively. The two meta-analyses also indicated a 14.6% and 11.2% SNP-based heritability of migraine.

Using migraine-first patients, we found a 45.96% and 90.27% increase in SNP-based heritability with a much smaller sample size. From our estimated 19.37% heritability, top hits (at a GWS and suggestive significance level) explained only 1.25% (Fig. 1), which is well in line with the polygenic structure of migraine and suggests that an elevated background polygenic signal is primarily responsible for the observed increase in SNP-based heritability. In addition, 100% of our risk loci overlapped with results of previous meta-analyses in case of GWS- (5/5, Table 1) and 31% in case of suggestive significant hits (11/36, Table 1), albeit, the most significant risk locus obtained a significance of only 10^{-13}. The identified risk loci are related to (1) thermoregulation, adult neurogenesis, and cerebrospinal fluid maintenance (*PRDM16*); (2) vascular remodeling and gene programs (*FHL5*); (3) neuronal migration and schizophrenia (*ASTN2*); (4) vitamin A transport into immune cells, Alzheimer's disease, lipid homeostasis, various vascular disorders, spontaneous coronary artery dissection [14, 15], in cervical artery dissection [16], in abdominal aortic aneurysm [17, 18], in descending aorta strain and, in descending aortic diameter [19] (*LRP1*); (5) sodium-calcium change and intracellular calcium homeostasis (*SLC24A3)* (www.genecards.org, https://www.ebi.ac.uk/gwas/) [5, 20]. Among the pathways, STAT1-, PEA3- and retinol metabolism-related sets implicate immunologic and vitamin A-bound processes behind migraine. Tissue-specific enrichment in tibial artery confirmed previous assumptions about vascular components of the disease, which is in line with LRP1-associated processes. In sum, significant findings were consistent with previous studies.

Although large meta-analyses with less well-defined phenotypes possess large statistical power and can identify top hits in migraine with high significance, they seem to substantially underestimate the background polygenic signal based on the current results. This phenomenon may distort the picture of the real underlying genetic architecture of migraine, overemphasizing top hits and underemphasizing the underlying background polygenic signal. In accordance, results also show that these top hits do contribute, but a larger part of the genomic signal stems from non-significant hits. Given these considerations, polygenic risk scores that estimate such polygenicity and could be useful in disease and prognosis prediction may be calculated more accurately from better-defined, more homogenous patient subgroups, like our migraine-first group.

All in all, our study provided evidence that migraine patients who had their migraine diagnosis first have higher heritability compared to a general migraine

population. Additionally, this migraine-first subgroup also showed association with the same genetic risk factors that were established based on large heterogeneous cohorts, but the top hits could only explain a small proportion of this heritability. Three clinical implications emerge: (1) optimized and more precise estimation of the polygenic risk of migraine based on migraine-first patients can pave the way for improved prediction of migraine risk, (2) key research attempts to find novel migraine medications could be improved by further investigating biological mechanisms related to the polygenic background of patients with a migraine-first diagnosis, (3) and finally, due to the fact that migraine heritability mainly explained by polygenicity, novel migraine treatment strategies probably benefit from a multi-target approach.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s10194-024-01870-8.

Supplementary Material 1

Acknowledgements
This work uses data provided by patients and collected by the NHS as part of their care and support. Copyright © (2019), NHS England. Re-used with the permission of the UK Biobank (Application Number 71718). All rights reserved.

Author contributions
Dora Torok: conceptualization, analysis, methodology, writing – original draft, writing – review and editing, Peter Petschner: conceptualization, analysis, methodology, writing – original draft, writing – review and editing.Dora Torok and Peter Petschner contributed equally to this work. Daniel Baksa: conceptualization, writing – original draft, writing – review and editing, Gabriella Juhasz: funding acquisition, conceptualization, supervision, writing – original draft, writing – review and editing.

Funding
The study was supported by the Hungarian Brain Research Program (Grant: KTIA_13_NAPA-II/14; KTIA_NAP_13-1-2013- 0001; KTIA_NAP_13-2- 2015-0001; 2017 – 1.2.1-NKP-2017-00002); Hungarian Brain Research Program 3.0 (NAP2022-I-4/2022); by the Ministry of Innovation and Technology of Hungary from the National Research, Development and Innovation Fund, financed under the TKP2021-EGA-25; by the Hungarian National Research, Development, and Innovation Office Grant: K 143391, and by the ÚNKP-23-4-II-SE-2 New National Excellence Program of the Ministry for Culture and Innovation from the source of the National Research, Development and Innovation Fund. Peter Petschner was an international research fellow of the Japan Society for the Promotion of Science (Postdoctoral Fellowships for Research in Japan, standard program, P20809).

Data availability
No datasets were generated or analysed during the current study.

Declarations

Ethics approval and consent to participate
Ethical approval was given by the National Research Ethics Service Committee North West–Haydock and all procedures were carried out in accordance with the Declaration of Helsinki.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 18 July 2024 / Accepted: 18 September 2024
Published online: 27 September 2024

References

1. Gormley P, Anttila V, Winsvold BS et al (2016) Meta-analysis of 375,000 individuals identifies 38 susceptibility loci for migraine. Nat Genet 48:856–866
2. Hautakangas H, Winsvold BS, Ruotsalainen SE et al (2022) Genome-wide analysis of 102,084 migraine cases identifies 123 risk loci and subtype-specific risk alleles. Nat Genet 54:152–160
3. Sutherland HG, Albury CL, Griffiths LR (2019) Advances in genetics of migraine. J Headache Pain 20:72
4. Marx P, Antal P, Bolgar B et al (2017) Comorbidities in the diseasome are more apparent than real: what bayesian filtering reveals about the comorbidities of depression. PLOS Comput Biol 13:e1005487
5. Petschner P, Baksa D, Hullam G et al (2021) A replication study separates polymorphisms behind migraine with and without depression. PLoS ONE 16:e0261477
6. Baksa D, Gonda X, Juhasz G (2017) Why are migraineurs more depressed? A review of the factors contributing to the comorbidity of migraine and depression. Neuropsychopharmacol Hung Magy Pszichofarmakologiai Egyesulet Lapja off J Hung Assoc Psychopharmacol 19:37–44
7. Eszlari N, Bruncsics B, Millinghoffer A et al (2021) Biology of perseverative negative thinking: the role of timing and Folate Intake. Nutrients 13:4396
8. Chang CC, Chow CC, Tellier LC et al (2015) Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience* ; 4: s13742-015-0047–8
9. Watanabe K, Taskesen E, van Bochoven A et al (2017) Functional mapping and annotation of genetic associations with FUMA. Nat Commun 8:1826
10. de Leeuw CA, Mooij JM, Heskes T et al (2015) MAGMA: generalized gene-set analysis of GWAS Data. PLoS Comput Biol 11:e1004219
11. Speed D, Balding DJ (2019) SumHer better estimates the SNP heritability of complex traits from summary statistics. Nat Genet 51:277–284
12. Speed D, Holmes J, Balding DJ (2020) Evaluating and improving heritability models using summary statistics. Nat Genet 52:458–462
13. Speed D, Kaphle A, Balding DJ (2022) SNP-based heritability and selection analyses: improved models and new results. BioEssays 44:2100170
14. Turley TN, O'Byrne MM, Kosel ML et al (2020) Identification of susceptibility loci for spontaneous coronary artery dissection. JAMA Cardiol 5(8):929–938
15. Adlam D, Berrandou T-E, Georges A et al (2023) Genome-wide association meta-analysis of spontaneous coronary artery dissection identifies risk variants and genes related to artery integrity and tissue-mediated coagulation. Nat Genet 55:964–972
16. Debette S, Kamatani Y, Metso TM et al (2015) Common variation in PHACTR1 is associated with susceptibility to cervical artery dissection. Nat Genet 47:78–83
17. Bown MJ, Jones GT, Harrison SC et al (2011) Abdominal aortic aneurysm is associated with a variant in low-density lipoprotein receptor-related protein 1. Am J Hum Genet 89:619–627
18. Jones GT, Tromp G, Kuivaniemi H et al (2017) Meta-analysis of Genome-Wide Association Studies for Abdominal aortic aneurysm identifies four New Disease-specific risk loci. Circ Res 120:341–353
19. Pirruccello JP, Rämö JT, Choi SH et al (2023) The genetic determinants of aortic distention. J Am Coll Cardiol 81:1320–1335
20. Wong D, Auguste G, Lino Cardenas CL et al (2023) FHL5 controls vascular Disease-Associated Gene Programs in smooth muscle cells. Circ Res 132:1144–1161

Publisher's note

Petschner *et al.*
The Journal of Headache and Pain (2025) 26:109
https://doi.org/10.1186/s10194-025-02033-z

The Journal of Headache and Pain

RESEARCH

Open Access

The interictal transcriptomic map of migraine without aura

Peter Petschner[1,2,3,4,5*†], Sahel Kumar[2,3†], Duc A. Nguyen[1,6], Dora Torok[2,3,4], Zsofia Gal[2,3,4], Daniel Baksa[2,3,4,7], Kinga Gecse[2,3,4], Gyongyi Kokonyei[2,3,4,8], Hiroshi Mamitsuka[1] and Gabriella Juhasz[2,3,4*]

Abstract

Background The present study aimed to deliver a replicable transcriptomic map of migraine without aura (MO) and its comprehensive, genome- and drug discovery focused analysis to identify hypotheses for future research- and clinical attempts.

Methods We recruited 30 controls and 22 MO patients without serious chronic comorbidities/regular medication intake. RNA-sequencing was conducted interictally at two different time points to identify replicable differential gene expression and enriched pathways. Subsequent refining and functional analyses were performed, including: 1) testing additional patient factors, 2) running genetic association analysis on migraine in the UK Biobank (UKB) and our cohort, and 3) predicting drug binding with AutoDock Vina and machine learning to proteins of transcriptomic changes.

Results Difference in CYP26B1 was identified as key alteration in migraine. Gene set enrichment analysis identified 88 replicated, significant, exclusively downregulated core pathways, including metabolic, cardiovascular, and immune system-related gene sets and 69 leading genes, like CORIN. Logistic regression of leading genes' and vitamin A pathway-related polymorphisms identified 11 significant polymorphisms in LRP1. Confirmatory analyses excluded a substantial impact of sex, allergy and different vitamin A/retinol intake. Binding simulations and predictions pointed to potential future drug molecules, like tetrandrine and probucol.

Conclusion The replicable transcriptomic map of MO and functional analyses: 1) identified pathomechanisms related to metabolic, cardiovascular and immune system related processes on a molecular level, 2) reported gene level hits, 3) proposed novel potential etiology, like LRP1-induced decreased retinoic acid signaling, and 4) delivered novel drug candidates for the disorder.

Keywords Migraine without aura, Transcriptomics, Genomics, Drug discovery Molecular pathophysiology

†Peter Petschner and Sahel Kumar these authors contributed equally to the present work.

*Correspondence:
Peter Petschner
petschner.peter@semmelweis.hu
Gabriella Juhasz
juhasz.gabriella@semmelweis.hu
[1] Bioinformatics Center, Institute of Chemical Research, Kyoto University, Uji, Kyoto 611 - 0011, Japan
[2] Department of Pharmacodynamics, Faculty of Pharmaceutical Sciences, Semmelweis University, Budapest 1089, Hungary
[3] Center of Pharmacology and Drug Research & Development, Semmelweis University, Budapest 1089, Hungary
[4] NAP3.0-SE Neuropsychopharmacology Research Group, Hungarian Brain Research Program, Semmelweis University, Budapest 1089, Hungary
[5] Research Unit for Realization of Sustainable Society, Kyoto University, Uji, Kyoto 611 - 0111, Japan
[6] Hanoi University of Science and Technology, Hanoi 100000, Vietnam
[7] Department of Personality and Clinical Psychology, Institute of Psychology, Pazmany Peter Catholic University, Budapest 1088, Hungary
[8] Institute of Psychology, ELTE Eötvös Loránd University, Budapest 1064, Hungary

Introduction

Migraine was responsible for about 5% of the total global years lived with disability in 2019 [1] indicating one of the most significant disease burdens. Numerous mechanisms for the disorder and its symptoms have been proposed based on animal models and targeted interventionist approaches [2], yet systematic investigation of the responsible underlying genetic and transcriptomic changes in humans could only provide modest results. Genomic meta-analyses focusing on single nucleotide polymorphisms (SNPs) were able to identify a moderate number of loci, which were in line with both vascular- and neuronal origins of the disease [3, 4]. On a transcriptomic level, human migraine gene-expression studies, investigating mRNAs, were scarcely published [5–10], despite the method being able to capture direction of changes, identify both primary causes and secondary alterations, highlight novel pathophysiology and deliver potential drug targets. These few published mRNA-sequencing (RNA-Seq) and microarray studies could not determine genes that were consistently replicated and/or remained significant after multiple testing correction between migraine patients and controls. Systematic analyses in RNA-Seq studies on a pathway-level yielded few results too. One of them pointed to a downregulation of immune system-related pathways [8]. Another study implicated, for migraine in general, macromolecular complex, nucleus and protein complex-related genes [5], while a non-systematic analysis suggested enrichment of nominally significant genes in mitochondrial processes and inflammatory response between migraine patients and controls [7]. Despite known differences between migraine with and without aura (MO), only one study addressed specifically the more common MO, and found no significant genes after multiple hypothesis correction and only two significant pathways, the macromolecular and ribonucleoprotein complex pathways [5].

A reason for the inconclusiveness could be the underlying heterogeneity of the investigated samples, influencing gene expression, migraine or both. We hypothesised that rigorous collection of a large sample with clinical diagnosis of MO subtype, strict exclusion criteria for serious acute/chronic illnesses and regular medication intake, and correction for remaining phenotypic variables might yield replicable results and provide a reliable transcriptomic map of the disease. Such a map and its functional analysis could widen knowledge regarding the underlying pathophysiology, connection between MO-transcriptomics and -genetics, boost drug development and provide testable hypotheses for clinical interventions and future studies.

To generate such a reliable interictal transcriptomic map of MO, we collected a large, rigorously phenotyped cohort of MO patients and healthy controls, and performed gene- and pathway-level analyses on raw RNA-Seq data of interictal, whole blood samples taken at two independent time points. Various analysis steps and additional datasets were used: 1) to test if correction for remaining phenotypic variables (smoking, history of allergy) influence gene expression replicability, 2) to obtain a differentially expressed core set of genes and pathways for MO, 3) to identify allergy-/sex-dependent genes and pathways in the core set, 4) to determine potential underlying genetic variants behind transcriptomic findings, 5) to characterise the leading gene-level candidate, *CYP26B1*, and its role in the observed changes, 6) to test binding affinity of current medications to the identified genes and 7) search for novel drugs capable of influencing the gene- and pathway-level targets (Fig. 1).

Methods

Participants

Volunteer recruitment for the migraine cohort took place through advertisements at university, headache clinics, in newspaper articles and online in Hungary. We collected data about diagnosis of MO, smoking status, age, sex, history of allergy, pregnancy, breast feeding, any other serious acute, previous and chronic illnesses and regular medication intake. Additionally, we collected Migraine Disability Assessment (MIDAS) measures, including attack frequency (MIDAS A), headache severity (MIDAS B) [11], and years since MO onset, from the MO cohort using standardized questionnaires and headache diaries. Diagnosis of episodic MO was made by an expert neurologist based on the International Classification of Headache Disorders 3rd edition (beta version) [12]. Subjects were excluded for 1) neurological problems (except MO), 2) psychiatric disorders, 3) serious, acute, previous and chronic illnesses (except history of allergy), 4) regular medication intake, except contraceptives, 5) pregnancy and breastfeeding determined by self-reported questionnaires and a neurologist. Mental health problems were assessed by Mini International Neuropsychiatric Interview (MINI) [13]. On measurement days, blood was taken from cubital veins of a random subset of participants (N= 52, transcriptomic cohort) following the recording of randomly assigned psychological tasks from participants. Sampling and measurements were repeated for a second time with average time difference of around four weeks. Subjects had to confirm that 48 h before sample collection no analgesic, anti-migraine medication and 4 h before no caffeine was consumed. MO samples were excluded if they were not attack- and headache-free 24 h before and after sample collection at both time points. Healthy controls had to be headache free 24 h before sampling.

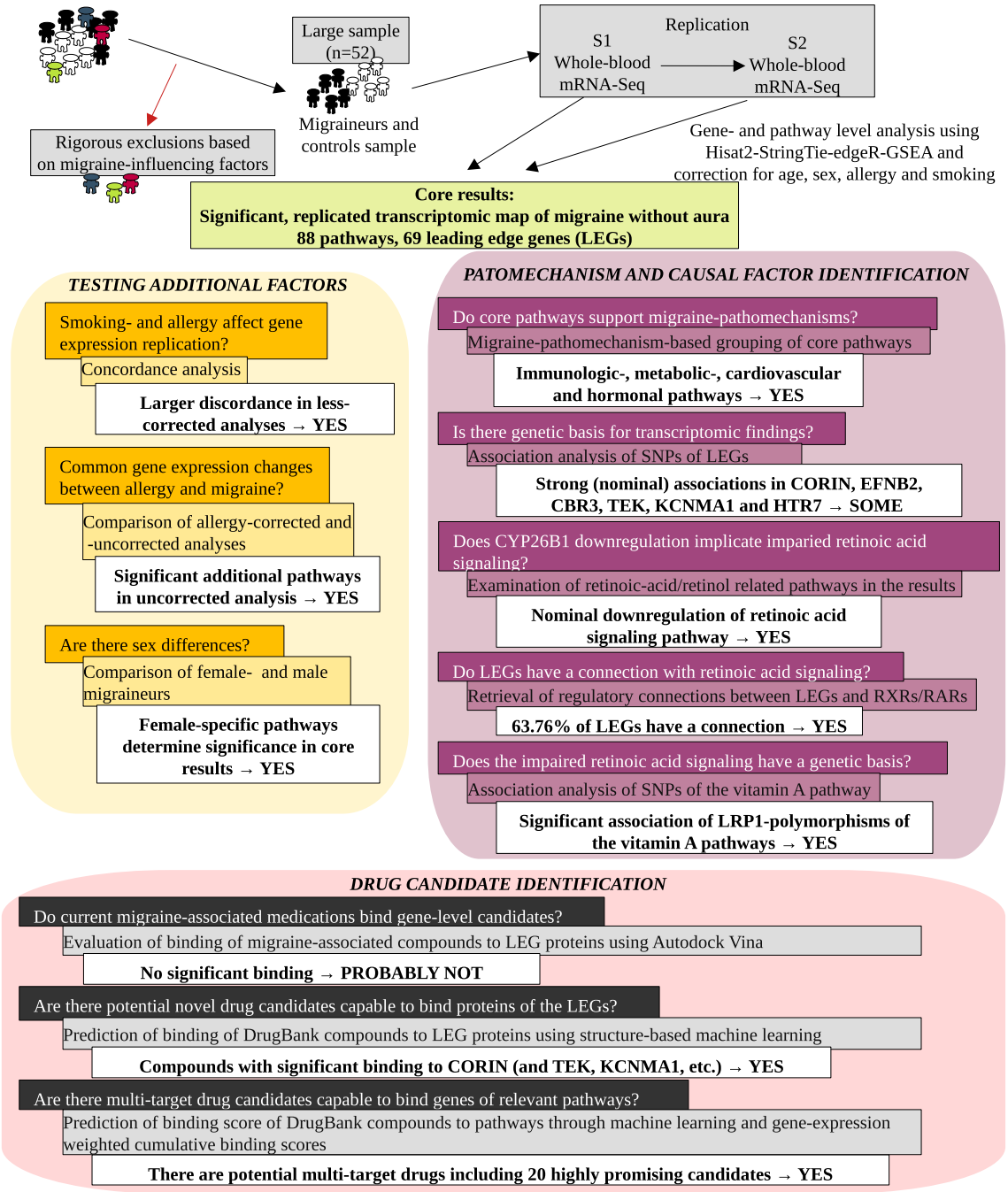

Fig. 1 Workflow and Experimental Design for Migraine Transcriptomic Analysis. Figure shows scientific questions, methodological approaches and key findings and implications of the study. Significance, unless otherwise noted, represents significant values after multiple hypothesis correction. Abbreviations: *LEGs* – Leading Edge Genes, *S1* – Time Point 1, *S2* – Time Point 2. *SNP* – Single Nucleotide Polymorphism

RNA sample preparation, sequencing, quality control and gene-level analysis

For purification and extraction of blood samples PAXgene Blood mRNA kit (Qiagen, Venlo, The Netherlands) was used with QiaCube instrument at the Institute of Psychiatry, Psychology & Neuroscience, King's College London (United Kingdom) per the manufacturers handbook. Samples were processed according to "NEBNext Ultra Directional RNA Library Prep Kit for Illumina" (NEB #E7420S/L) at GenomeScan (Leiden, The Netherlands). Sequencing was performed using Illumina NextSeq 500, with 75 bp single-end sequencing according to

Illumina's and GenomeScan's standard operating procedures and ISO standards with a sequencing depth of 20 M reads per sample.

Remaining adapters (attached during sequencing) were removed with Trimmomatic (v 0.39) [14]. Raw data quality control (QC) was done by FastQC (v0.11.8, [15]), resulting in a quality score > 30 for all reads, meaning a 99.9% accuracy in base calls. Consequently, no sample exclusions were made.

RNA-seq reads were aligned to Ensembl human reference GRCh38.96 with HISAT2 (v2.2.1, [16]), with default settings and spliced alignment options for downstream transcriptome assembly and rna strandness R, providing alignment rates ranging between ~ 91–96%, except one sample with 84%. Subsequently, StringTie (v2.1.4, [17]) was used for transcriptome assembly and quantification.

Gene level analysis to determine differentially expressed genes was done in R (v4.0.3, [18]) with Bioconductor (v3.11, [19]) package edgeR (v3.30.3, [20]) using the quasi-likelihood F-test. P-values were adjusted for multiple testing by the Benjamini–Hochberg false discovery rate (FDR) [21].

All transcriptomic analyses included covariates: 1) age and sex or 2) age (for sex-specific analyses), where age corresponded to the exact age at the time of measurement (at S1 and S2). In analyses considering the impact of vitamin A intake and using the covariate of supplements with certain or potential vitamin A content we used supplement intake as a binary variable.

For gene set enrichment analyses (GSEA, [22]) we used all C5 gene sets representing Gene Ontology (GO) sets. C5 sets were chosen, because: 1) these provide the largest coverage, 2) represent various functional gene sets able to connect a high-level neurologic disorder with gene expression changes [23], and,3) are the most comparable with other studies on the field. Results using the cellular compartment sets are presented only in supplementary materials, as we considered these hard to interpret with respect to migraine.

Gene-set enrichment analysis

GSEA was performed with Bioconductor package fgsea (v. 1.14.0, [24]) using default settings with size of gene sets between 15 and 500. Gene sets were obtained from Molecular Signature Database (MSigDB) v.7.4. Results are published for Gene Ontology molecular functions (GOMF) and biological processes (GOBP) sets in the main text.

A pathway was significant if FDR < = 0.05 at both time points. Leading edge genes (LEGs) of the significant GOBP and GOMF pathways, a subset of the pathway's genes, which appear in the ranked list of genes before the running sum reaches its maximum, were extracted. In short, LEGs were genes driving significant pathway-level enrichment. LEGs were considered important if they have shown a mean absolute logarithm twofold change (Log2 FC) value over 2 (for LEG5) and over 0.5 (for LEG69). Since these genes were important behind the significantly replicated pathways, our criteria were rather as a post-hoc filtering than a standalone statistical evaluation. The thresholds were selected based on literature for 0.5 [25], and conventions in transcriptomic analyses, which usually consider a log2 FC > = 2 a threshold for substantially altered genes.

Sign concordance analysis

Gene expression values (Supplementary material 2) were filtered for the common set of genes present at all time points and corrections investigated, with the filtered set consisting of 10,765 genes. Log2 FC signs were considered concordant if a given gene had either positive or negative log2 FC at both time points for a given comparison, discordant otherwise. Similar analyses were conducted after filtering for an absolute log2 FC value of 0.001 and 0.05 (Supplementary material 3) to assess genes with increasingly non-near zero expression change.

UK Biobank participants

Migraine phenotype was determined based on cases, who had migraine as a first medical diagnosis in UK Biobank (UKB) (project ID: 71718). We used field ID: 131052, G43 to determine first reported migraine (resulting in 6,307 migraineurs and 200,571 controls). This phenotype is considered to be a more genetically determined migraine phenotype [26]. Smoking status was defined based on current (field ID: 1239) and past (field ID: 1249) smoking habits. Information related to asthma or hayfever/eczema/allergic rhinitis from medical conditions (field ID: 6152) were used to determine asthma status and estimated food nutrients and vitamin and/or mineral supplement use was utilized for vitamin A- and vitamin A retinol equivalent intake (field IDs: 26061, 20084). We used field IDs 31, 21003 to determine sex and age of participants in UKB. Control individuals were participants who, at the mean age of the migraine group (+ -SD) had no diagnosed disease.

Migraine cohort

Data from whole migraine cohort including those that were not sampled for transcriptomic analysis (N= 289) were included in the genetic analyses. Phenotypic variables were determined similarly for the subset that underwent transcriptomic analysis, resulting in 172 migraine cases and 117 controls.

Selection of Single Nucleotide Polymorphisms for Leading Edge Genes for genetic analyses

SNPs that belong to the LEGs and their 10 kbp vicinity were retrieved using gene boundaries according to hg19 known gene gene track of UCSC and ANNOVAR [27]. The final list of SNPs and genes can be found in Supplementary material 4.

Vitamin A pathway genes and selection of Single Nucleotide Polymorphisms for genetic analyses

Vitamin A pathway genes were manually selected from [28] and [29], based on protein names mentioned in the texts, human gene names retrieved using GeneCards [30], and corresponding SNPs extracted like for LEGs (Supplementary material 4).

Plink2 genetic analyses

Logistic regression of the above SNPs was performed for migraine with Plink2 using covariates, sex, age, top 10 principal components of the genome, smoking, allergy in both UKB and our cohort. Additionally, genotype array was included in the UKB analyses. For vitamin A pathway SNPs, retinol equivalent- and vitamin A intake was also used in the UKB analyses (Supplementary material 1). All genetic analyses used Bonferroni-correction (threshold of significance/number of tests performed) as this is standard on the genomic field [3, 4].

Binding predictions using AutoDock Vina

Sixty-four migraine medications were selected from DrugBank 5.1 [31] by two pharmacists (SK and KG). Antiemetic and non-small molecule drugs (e.g., antibodies, botulinum toxin, etc.) were excluded (Supplementary material 5). For binding simulations, small molecule 3D structures of the remaining 54 drugs were downloaded from PubChem [32]. For AutoDock Vina analyses pdb files containing AlphaFold2 predicted protein structures were downloaded from UniProt [33]. Proteins and ligands were prepared with default settings using prepare_receptor command (except that "-A" flag was used to add hydrogens to structures) and ml_prepare-ligand command, respectively. Basic docking simulations using 0,0,0 as centre and 126,126,126 as size for grid parameters with Vina forcefield using AutoDock Vina (v1.2.3., [34]) were run with "exhaustiveness" set to 32. Based on preliminary runs maximum number of binding modes was left at default. From the simulated binding pair modes, the one with the lowest score (strongest binding) was selected as representative of the drug-protein binding. Docking simulations require normalization [35]. Normalization was performed by randomly sampling 100 proteins for each drug (Supplementary material 6) and Z-scores for the drugs using the mean and standard deviation of random samples were calculated and one-sided p-values derived from these scores (energetically more favourable binding pairs only). Z-test significance values were considered significant if they survived Bonferroni multiple hypothesis correction (p-value $<= 1.3617 \times 10^{-5}$).

Drug-protein binding prediction with machine learning

Genes of the significant pathways and all drugs from the 5.1 version of DrugBank [31] were used. To our knowledge, no immediately deployable comprehensive drug binding prediction pipeline exists for genes, therefore we developed our own (Supplementary material 1). In detail, genes were extracted from all pathways for a relevant comparison, and structures downloaded from the UniProt database [33]. All proteins of a gene were used. HyperAttentionDTI, using a convolutional neural network (CNN) with an attention block and the drug's SMILES strings and the amino acid sequence code for proteins, was utilized to predict drug-binding scores for all proteins [36]. Key hyperparameter settings were: learning rate of 1×10^{-4}, weight decay of 1×10^{-4}, input embeddings of 64. CNN blocks had three stacked 1D-CNN layers with 32, 64, 96 filters, window sizes were 4, 6, 8 for drugs and 4, 6, 12 for proteins, output block was four layers with 1024, 1024, 512 and 2 neurons, respectively. Dropout rate was 0.1 and batch size 32. Predictions were run on a 24 GB Nvidia RTX A5000 GPU.

After obtaining protein-drug prediction for 11,064 drugs from DrugBank, the maximum binding score among all proteins for the gene was associated as the binding score of the drug and the gene. We calculated all drug-gene binding scores and assessed how many times each drug was in the top 10 most binding drugs for every gene from the 9,657 genes. Such drugs are less suitable drug candidates due to their ubiquitous binding, thus drugs with more than 10 times in the top 10 predictions of the 9,657 genes were excluded. Finally, normalisation using the mean and standard deviation of the drug's predicted binding score to all proteins was performed and right-sided p-values were calculated. FDR values were calculated considering binding to all LEGs for a given drug. The values, in sum, show which proteins are bound by the given drug with significantly higher affinity than a hypothetical average protein.

One drug-one target approaches in migraine drugs were criticized for inadequacy to reveal relevant candidates [37]. Thus, a pipeline from the significantly enriched pathways has been developed to obtain pathway level drug candidates. Drug-gene prediction scores for a given gene were multiplied with the mean absolute log2 FC values of a gene from the migraine versus control allergy-corrected comparison, if it had both prediction and gene expression values in all measurements.

To calculate effect on a pathway of a given drug, genes belonging to a pathway weighted with the drugs' normalised binding affinity to an average protein (as in single target case) were aggregated using the L2-norm. The L2 norm gave binding score for each drug on the pathway level and can be considered as a weighted sum of the binding scores with the expression change of the genes in a pathway. Empirical p-values were calculated using the Monte Carlo integral of the ranked p-values list for all pathways and drugs and pathway-gene pairs below the arbitrary significance threshold of 0.0005 are presented.

Results

Descriptive statistics

We assessed clinical characteristics and basic demographic data of the transcriptomic cohort summarized in Table 1 (for additional population characteristics such as sleep quality/chronotype, diet, exercise, smoking status, allergy history, supplement intake, and contraceptive use, see Supplementary material 1).

Smoking- and allergy-correction increase gene expression replicability in migraine

First, we tested if remaining phenotypic factors like history of allergy (henceforth allergy) and smoking status (henceforth smoking), in addition to the standard age and sex variables, could influence replicability of gene expression in general (and if their correction is necessary). Discordant gene expression signs at the two time points (S1 and S2) decreased in the following order of corrections in migraine versus control comparisons: 1) age and sex correction (44.38%), 2) age, sex and allergy (36.01%), 3) age, sex and smoking (21.68%), and 4) age, sex, allergy and smoking (21.39%) using all 10,765 genes. Results were consistent with this pattern if genes with small expression differences between groups were excluded (Supplementary material 2).

Core gene-level expression changes in migraine implicate CYP26B1 gene

Correction for age, sex, smoking and allergy (henceforth the allergy-corrected analysis) left no significant, replicated genes at both time points (no significant genes at S1 and five at S2) comparing MO and controls (Supplementary material 2). Investigation of log2 FC revealed that *CYP26B1* had the largest average log2 FC in the analyses and was downregulated in MO (log2 $FC_{S1} = -5.865$, log2 $FC_{S2} = -5.860$, Supplementary material 2).

Core pathway-level differences in migraine support pathophysiology

GSEA of allergy-corrected expression of the migraine versus control comparison yielded 88 enriched pathways (FDR < = 0.05 at both time points and concordant normalized enrichment score [NES]). All significant, replicated pathways were downregulated (Supplementary material 7). For the list of the top 10 pathways based on their mean NES of S1 and S2, see Table 2.

Grouping of the significant pathways into migraine-related, high-level categories based on the pathway's functional descriptions is visible on Fig. 2 and support known pathophysiology mechanisms including altered amino-acid-, lipid, carbohydrate metabolism, cardiovascular- and hormonal elements and immune system-related processes.

Leading edge analysis identifies 69 genes in migraine

To extract genes with relevance for the replicated, significant core pathway-level results, LEGs – genes driving pathway-level enrichment signals – were extracted from pathways. Among the LEGs, five (LEG5) showed a mean log2 FC value below −2, namely: *CYP26B1*, *CORIN*, *PRTN3*, *CCL23*, *CTSG* and 69 (LEG69) showed a mean log2 FC below −0.5. There were no LEGs with positive log2 FC (Table 3, Supplementary material 2).

Table 1 Demographic and Clinical Characteristics in Control vs. Migraine Groups

Characteristic	Control ($n = 30$)	Migraine ($n = 22$)	Test statistic	*p*-value
Sex (M/F)	14/16	5/17	$\chi^2 = 2.1895$	0.139
Age (S1) [years]	26.4 ± 4.02 (21–37)	26.8 ± 5.13 (20–37)	U = 324	0.92
Age (S2) [years]	26.6 ± 3.99 (21–37)	26.9 ± 5.17 (20–38)	U = 319	0.847
MIDAS A	not applicable	11.71 ± 5.83 (2–21)	not applicable	not applicable
MIDAS B	not applicable	5.57 ± 1.60 (2–9)	not applicable	not applicable
Duration of MO onset [years]	not applicable	12.42 ± 6.31 (2.55–28.23)	not applicable	not applicable

Data for continuous variables (age, MIDAS, duration of migraine onset) are presented as mean ± SD (range). Sex distribution was compared using a chi-square test between MO and controls, and age was compared using a Mann–Whitney U test at both scan time points (S1, S2). MIDAS (A, B) and duration of MO onset were not tested against controls as these were not recorded in the control group. *Abbreviations: SD* – Standard Deviation, *U* – Mann–Whitney U test statistic, χ^2 – Chi-squared statistic, *S1* – Time Point 1, *S2* – Time Point 2, *MIDAS* – Migraine Disability Assessment, *MO* – Migraine without Aura

Table 2 Top 10 Significant Gene Ontology Pathways in Migraine Without Aura

Pathway name	Mean NES	FDR at S1	FDR at S2
Myeloid leukocyte mediated immunity	– 2.66	5.44E- 08	4.3E- 08
Myeloid leukocyte activation	– 2.55	5.44E- 08	4.3E- 08
Cell activation involved in immune response	– 2.41	5.44E- 08	4.3E- 08
Defense response to fungus	– 2.20	1.72E- 03	1.1E- 03
Antimicrobial humoral response	– 2.19	4.98E- 04	1.3E- 03
Regulation of DNA templated transcription in response to stress	– 2.19	5.72E- 04	6.4E- 06
Response to fungus	– 2.18	4.20E- 03	2.5E- 04
Regulation of transcription from rna polymerase II promoter in response to hypoxia	– 2.12	3.53E- 03	1.5E- 04
Regulation of cellular amino acid metabolic process	– 2.02	3.25E- 03	7.2E- 03
Activation of innate immune response	– 2.00	1.08E- 03	2.4E- 04

The table shows top 10 replicated significant Gene Ontology pathways for interictal episodic migraine without aura patients compared to controls based on mean normalised enrichment scores using two time points. For detailed statistics, see Supplementary material **7**. *Abbreviations: FDR* – False Discovery Rate, *NES* – Normalized Enrichment Score, *S1* – Time Point 1, *S2* – Time Point 2

Influence of allergy and sex on core gene expression results

We considered two remaining phenotypic factors, allergy and female/male sex, that may have influenced our findings and their analysis could reveal interesting insights into MO. On one hand, correction for allergy—despite its overall beneficial effect on replicability—may have removed migraine-relevant genes from comparisons due to the high comorbidity and potential shared aetiology with migraine (38). We tested this possibility by comparing results of allergy-corrected and -uncorrected analyses. With the notable exception of 16 non-overlapping LEGs (Fig. 3) the majority of LEG69 and all LEG5 genes remained relevant in driving pathway level enrichments independent from allergy-correction. *CORIN* and *CYP26B1* remained LEGs with largest mean log2 FC, with the latter reaching FDR-significance at both time points in allergy-uncorrected analysis (Supplementary material 2,7,8,9).

Sex-specific analyses confirm pathway-level findings and importance of CYP26B1 and CORIN

Correction for sex in above tests has ensured that the identified core results are important in both sexes. Nonetheless, sex-specific analyses are recommended in migraine research and migraine prevalence is two- to threefold in females compared to males [39]. Therefore, we speculated that analyses comparing the sexes may deliver additional insights about the stability of the found genes and pathways and can point to sex-specific characteristics for future studies.

Comparison of gene expression between female and male migraine patients with correction for control samples in the sexes left no significantly different genes at any time points neither in allergy-corrected nor in allergy-uncorrected analyses (Supplementary material 1), indicating lack of significant sex-specific gene-level differences.

Pathway-level findings indicated that the general comparison of migraine versus controls delivered stable results and pathway-level difference between the sexes stems from larger downregulation in men, and smaller downregulation or slight, non-significant upregulation in women (For full results see Supplementary material 2,7,8,9).

Genetic associations of Leading Edge Genes refine background of expression changes

To test if the observed changes have an underlying genetic cause, we performed logistic regression analyses with age, sex, smoking status and history of allergy as covariates on SNPs of the LEG69 genes for migraine in the UK Biobank. No results survived Bonferroni-type multiple hypothesis correction (p-value $< = 3.66 \times 10^{-6}$, 0.05/13,650). Nominally significant hits were numerous and included variants from *EFNB2*, *CBR3*, *TEK*, *KCNMA1* and *CORIN*. Many SNPs below nominal threshold were related to *SGCD* (Supplementary material 10).

CYP26B1 indicated a role for retinoic acid pathway in migraine

The fact that CYP26B1 is regulated by and metabolises all-trans-retinoic acid (atRA) and our previous results, i.e. 1) being the top gene with largest average log2 FC, 2) having stable, sex- and allergy-independent expression change and 3) showing no underlying SNP hits indicated that retinoic acid-related pathways may be involved in MO. Therefore, we conducted further investigations and

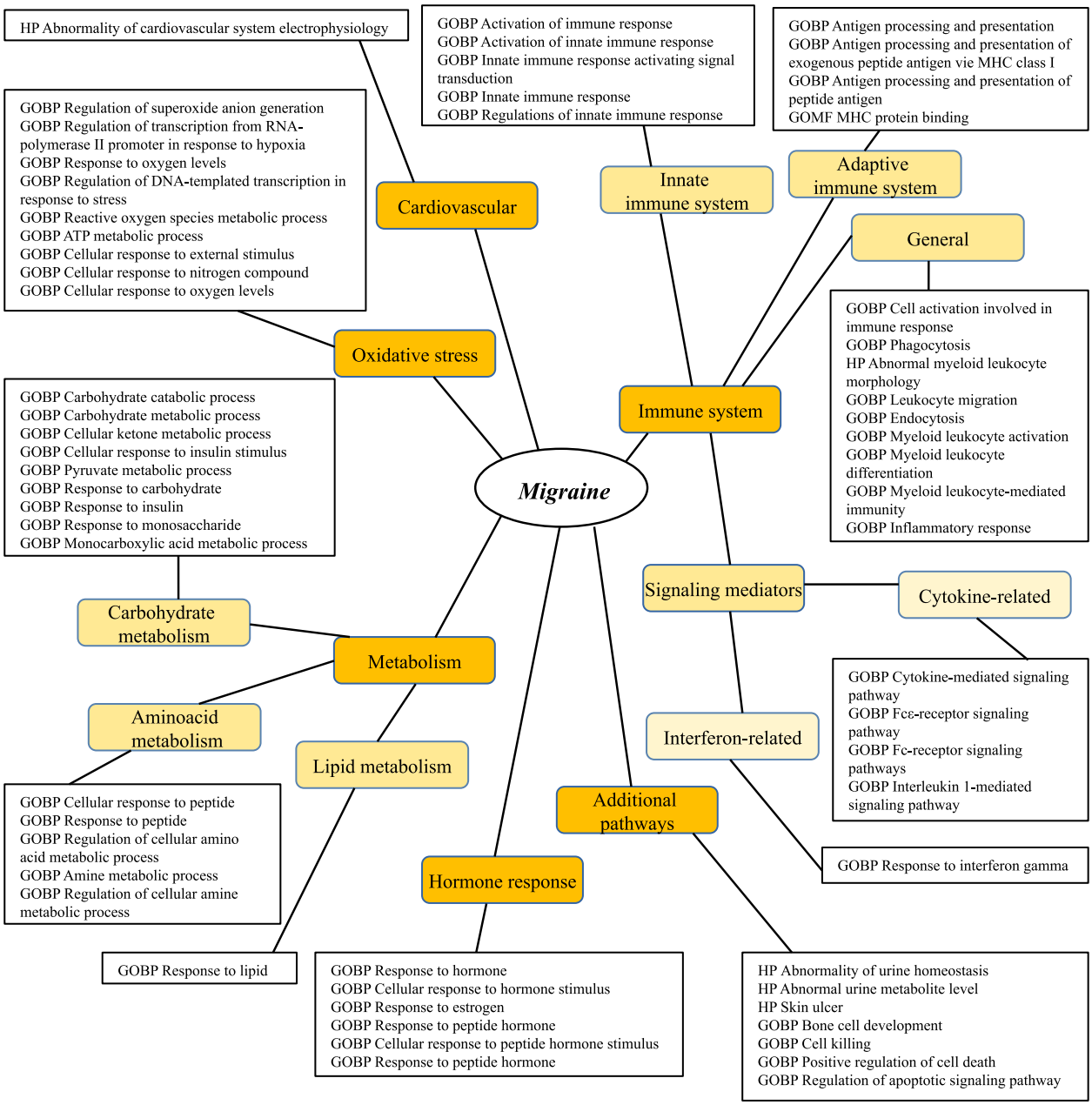

Fig. 2 Functional Categorization of Significant Pathways in Migraine Pathophysiology. Functional categories were selected based on expert knowledge. Pathways were assigned to terms based on the relevance of their name/functional description to the given term. *Abbreviations: HP* – Human Phenotype Ontology, *GOBP* – Gene Ontology Biological Process, *GOMF* – Gene Ontology Molecular Function

Table 3 Leading Edge Genes in Migraine and Their Expression Changesdant normalized enrichme

Gene name	Mean log2 FC	Log2 FC at S1	*p*-value at S1	FDR at S1	Log2 FC at S2	*p*-value at S2	FDR at S2
CYP26B1	− 5.8622	− 5.8646	0.0001	0.2005	− 5.8598	0.0001	0.0585
CORIN	− 4.1653	− 3.5429	0.0015	0.5275	− 4.7877	0.0004	0.1426
PRTN3	− 3.0534	− 3.3806	0.0005	0.2991	− 2.7263	0.0029	0.4484
CCL23	− 2.7878	− 2.7640	0.0386	0.8735	− 2.8115	0.0458	0.8297
CTSG	− 2.6687	− 2.7865	0.0004	0.2936	− 2.5509	0.0009	0.2410

Table shows gene level expression statistics of leading edge genes with absolute mean log2 FC value > = 2 (LEG_5) of allergy-corrected analyses, for detailed values see Supplementary material 2. *Abbreviations: Log2 FC* – Logarithm Two fold Change, *FDR* – False Discovery Rate, *S1* – Time Point 1, *S2* – Time Point 2

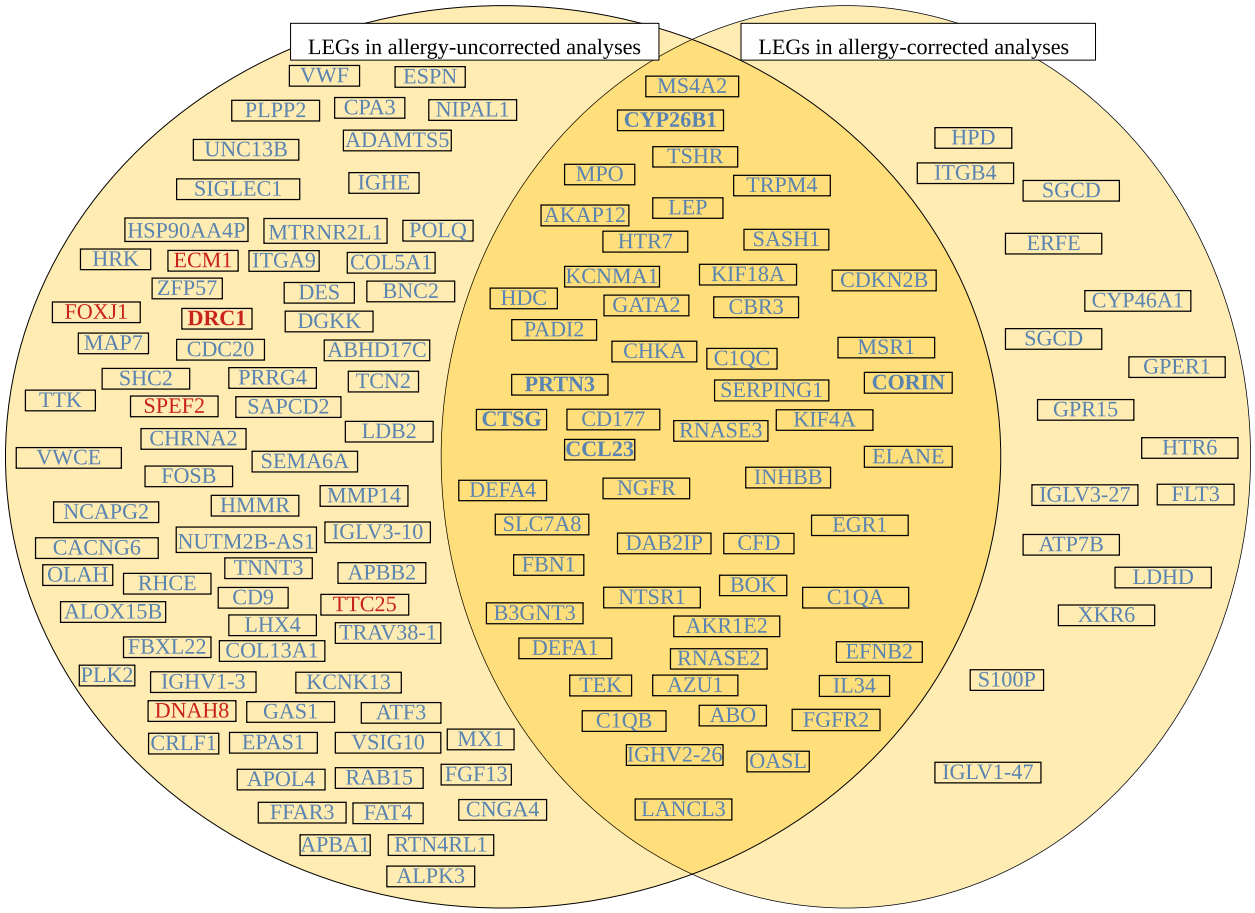

Fig. 3 Distribution and Comparison of Leading Edge Genes Between Allergy-Corrected and Uncorrected Analyses. Figure shows LEGs with absolute mean log2 FC values over 0.5. Bold values, blue and red colours denote log2 FC > = 2, down- and upregulation, respectively. *Log2 FC* – Logarithm Twofold Change, *LEGs* – Leading Edge Genes

first, scrutinized specific related gene sets in pathway-level results.

Allergy-corrected analyses showed a nominal significance for retinoic acid receptor signalling pathway at S2. This result became FDR-significant at S2 in allergy-uncorrected analysis, suggesting a potential involvement of retinoic acid signalling (Supplementary material 9).

Genetic analyses of vitamin A pathway point to connection with LRP1

Logistic regression of vitamin A pathway SNPs for migraine with age, sex, history of allergy, smoking status and vitamin A intake correction was done in the UKB dataset.

Eleven polymorphisms, rs4759276, rs4759275, rs1385526, rs1799737, rs1466535, rs10876964, rs4759045, rs10876965, rs11172113, rs4367982, rs4759277 of the *LRP1* gene showed significant results for migraine with *p*-values below the multiple hypothesis correction threshold of 3.66×10^{-6} (0.05/13635, Supplementary material 11). Analyses correcting for sex, age, smoking, allergy and vitamin A showed that the associations are independent of vitamin A intake, excluding associations due to an underlying correlation of different vitamin A intake with migraine status and highlighting *LRP1* polymorphisms as potential independent causal factors. Seven *LRP1* SNPs that were significant in the UKB and were present in our population showed nominal significance for MO status independent of the corrections used (Supplementary material 11). In all tests conducted the minor alleles at the given positions showed protective effects against migraine, indicating risk for major allele carriers.

Majority of genes behind pathway-level enrichment in migraine are regulated by RAR and RXR receptors

For further validation of an altered retinoic acid signalling raised by the GSEA results, regulatory connections between LEG69 and retinoic acid receptors *RARA*,

RARB, RARG, RXRA, RXRB, RXRG were searched for in Gene Regulatory Network database (GRNdb) [40].

Results using bulk, healthy human data from GTEx resulted in 63.76% of the LEG69 being regulated by at least one of the receptors and 28 being regulated by multiple receptor families with the default NES threshold (> =3). *CORIN* provided only one hit in pancreas tissue samples via *RXRA* (NES =3.04, Supplementary material 12).

If pathological conditions were also included, 81.16% of LEG69 were regulated by at least one of the receptors with NES over 3 (Supplementary material 12).

Migraine-associated drugs do not bind proteins of gene expression changes

To assess if migraine-associated compounds from DrugBank can act at the corresponding proteins of the identified LEG69, binding simulations with AutoDock Vina were run (since binding is a prerequisite of pharmacological effect). For LEG69, no migraine medications showed significant binding after multiple hypothesis correction (one-sided p-value < = 1.3617×10^{-5}, Supplementary material 13).

Binding predictions to proteins of gene expression changes find new drug candidates

Given the binding results for the gene expression changes and current migraine-associated compounds, we speculated that a search for potential drug candidates can help to find efficient therapies acting on the identified genes. Chemical structure-based machine learning prediction of drug-protein binding with HyperAttentionDTI showed that CYP26B1 is not bound by any investigated drug at the given significance level, while CORIN had four potential binding candidates with FDR < = 0.10 (Table 4, Supplementary material 14).

Estimating gene expression-weighted pathway-level binding showed numerous potential binding candidates (a selected list with p-values below 0.0005 can be found in Supplementary material 15). We have to note, the list of gene- and pathway-level candidates includes both (potential) migraine-promoting and -treating drugs, since the method is based on binding predictions, not direction of effects. Among the significant (p < =0.0005) substances, several drugs showed a connection to migraine, migraine symptoms, headaches and -treatments in the literature (see Table 5). In addition, several, already known migraine-associated compounds (in DrugBank, bold in Table 5) were found among pathway-level drugs, like eletriptan.

Table 4 Predicted drug candidates targeting CORIN in Migraine

Drug candidate	Predicted binding Z-score	FDR value for the Z-score	Literature reference, if any (PMIDs)
probucol	2.570	0.070	10601122
fasoracetam	2.557	0.073	
ginsenosides	2.454	0.081	1648425, 3248333, 32420095
rovafovir	2.454	0.098	

Table shows four potential drug candidates, predicted binding scores and FDR values from 11,164 drugs in DrugBank with predicted FDR significant (FDR < = 0.1) binding predictions to *CORIN*. Last column indicates PMIDs, with reference for the given substances, if any. *Abbreviations: FDR* – False Discovery Rate, *PMIDs* – PubMed IDs

Discussion

Previous transcriptomic investigations in MO yielded inconsistent findings and remained detached from existing migraine pathophysiology theories, genetic findings and the method's potential for drug development (5–10). Here, we report replicated pathway-level gene expression changes in MO at interictal state obtained after correction for underlying masking phenotypic differences, allergy and sex. The results highlight two genes, *CORIN* and *CYP26B1* and downregulated metabolic, cardiovascular, immunologic-, oxidative stress- and hormone regulation-related pathways as important transcriptomic findings. We show that genetic polymorphisms cannot explain the observed *CYP26B1* downregulation, and provide consistent proofs for its marker role of a dysregulated retinoic acid receptor signalling in patients and its potential underlying cause, reduced retinoic acid availability due to disease-associated genetic polymorphisms within the *LRP1* gene. We also show that few current migraine medications utilise the corresponding proteins of the observed transcriptomic alterations, and highlight gene- and pathway-level drug candidates, like probucol, tetrandrine and indoramin to promote future drug development.

No significantly replicated gene for MO could be found in our core analysis. This was in part due to the replication criterion, partly due to allergy-correction, as evidenced by the large number of additional pathways and the significantly replicated *CYP26B1* gene in the allergy-uncorrected analysis. Another reason may have been the relatively smaller change of expression of individual genes, corresponding to the polygenic nature of MO [41], making detection of gene-level differences harder. The latter can be compensated, at least in part, by pathway-level analysis methods.

Table 5 Potential drug candidates and associated pathways for migraine therapy

Drugs	Associated pathway(s)	Reference(s) (PMIDs)
buclizine	Positive regulation of cell death, response to lipid, transmembrane receptor protein tyrosine kinase signalling pathway	22469258
mecloxamine	Positive regulation of cell death	17373440,9265005,8366753,6342291
verapamil	Positive regulation of cell death	2668225
tetrandrine	Positive regulation of cell death	29147842
olmesartan	Transmembrane receptor protein tyrosine kinase signaling pathway	16618270,16942482,31515634
nicardipine	Cytokine mediated signaling pathway, innate immune response, phagocytosis, positive regulation of cell death, transmembrane receptor protein tyrosine kinase signaling pathway	30600979,2245455
tubocurarine	Positive regulation of cell death, transmembrane receptor protein tyrosine kinase signaling pathway	28496430,10686170
zafirlukast	Positive regulation of cell death	17691939,15648777,10759916
opc-28326	Positive regulation of cell death	17103145
indoramin	Positive regulation of cell death	49624,324566
nystatin	Positive regulation of cell death	28041915
pasireotide	Positive regulation of cell death	29925553
thioridazine	Transmembrane receptor protein tyrosine kinase signaling pathway	10667670
bicuculline	Positive regulation of cell death	36259130,12499053
bi 44370 ta	Positive regulation of cell death	32525262
berberine	Transmembrane receptor protein tyrosine kinase signaling pathway	23758551
resiniferatoxin	Positive regulation of cell death	29187670,23155193
ondansetron	Positive regulation of cell death	32433024,20661681
lomerizine	Transmembrane receptor protein tyrosine kinase signaling pathway	37194515,29221971
eletriptan	Positive regulation of cell death	15853473

The table shows 1) selected migraine-associated drugs (bold) and potential drug candidates with significant (p-value < = 0.0005) weighted binding to significantly replicated pathways between migraine and controls and 2) supporting literature references. See also Supplementary material 15. *Abbreviations*: *PMIDs* – PubMed IDs

Indeed, numerous replicable results emerged with pathway-level analyses and these aligned to existing migraine pathophysiology theories: downregulated and altered amino acid-, lipid- and carbohydrate metabolism [42–45], immunologic processes [8, 46], cardiovascular comorbidities [38, 47] and hormonal influences [48] have been implicated in MO previously. It remains uncertain, if in contrast to our interictal study, during a migraine attack different findings would have emerged (as suggested by ref. [6]). Nonetheless, agreement of the observed changes with numerous proposed attack mechanisms in the literature [8, 38, 42–48] suggests that our findings may not only represent interictal characteristics, but are more general. Future studies need to decipher if the present results correspond only to interictal state and if during migraine attack new gene expression changes (e.g., other pathways) emerge or expression level of the found pathways change.

Sex- and allergy-dependent analyses indicated a promising direction for future studies investigating MO-relevant pathways. The more than 3-times larger number of significantly replicated pathways and the additional genes in allergy-uncorrected analysis clearly indicated that allergy may have common mechanisms with MO, while analyses directly comparing the two sexes suggested that male-specific pathways that were also significant in females were found by general migraine versus control comparison confirming the stability of our findings, but leaving room for, especially, male-specific studies for the future.

Among sex- and allergy-independent migraine-associated genes, *CYP26B1* and *CORIN* deserve particular attention. Both 1) the downregulated *CYP26B1*-levels, which are reliant on the substrate, all-trans retinoic acid (atRA) in the heart and vasculature [49–51] and 2) the lack of *CYP26B1* SNPs associated with migraine even on a nominal level, indicated a marker role for *CYP26B1* for reduced atRA levels in MO. Analyses of vitamin A pathway SNPs implicated *LRP1* polymorphisms as potential underlying cause for reduction in *CYP26B1* levels. LRP1 helps retinyl-ester containing chylomicron remnant uptake [28] in the liver, from where vitamin A derivatives are transported to the periphery [28] LRP1 is also involved in the uptake of serum amyloid A-retinol complexes in immune cells of the intestine [52]. Thus, LRP1 is capable of influencing both direct retinol uptake into immune cells and peripheral atRA abundance available for signaling (Fig. 4). Impairment of the latter processes

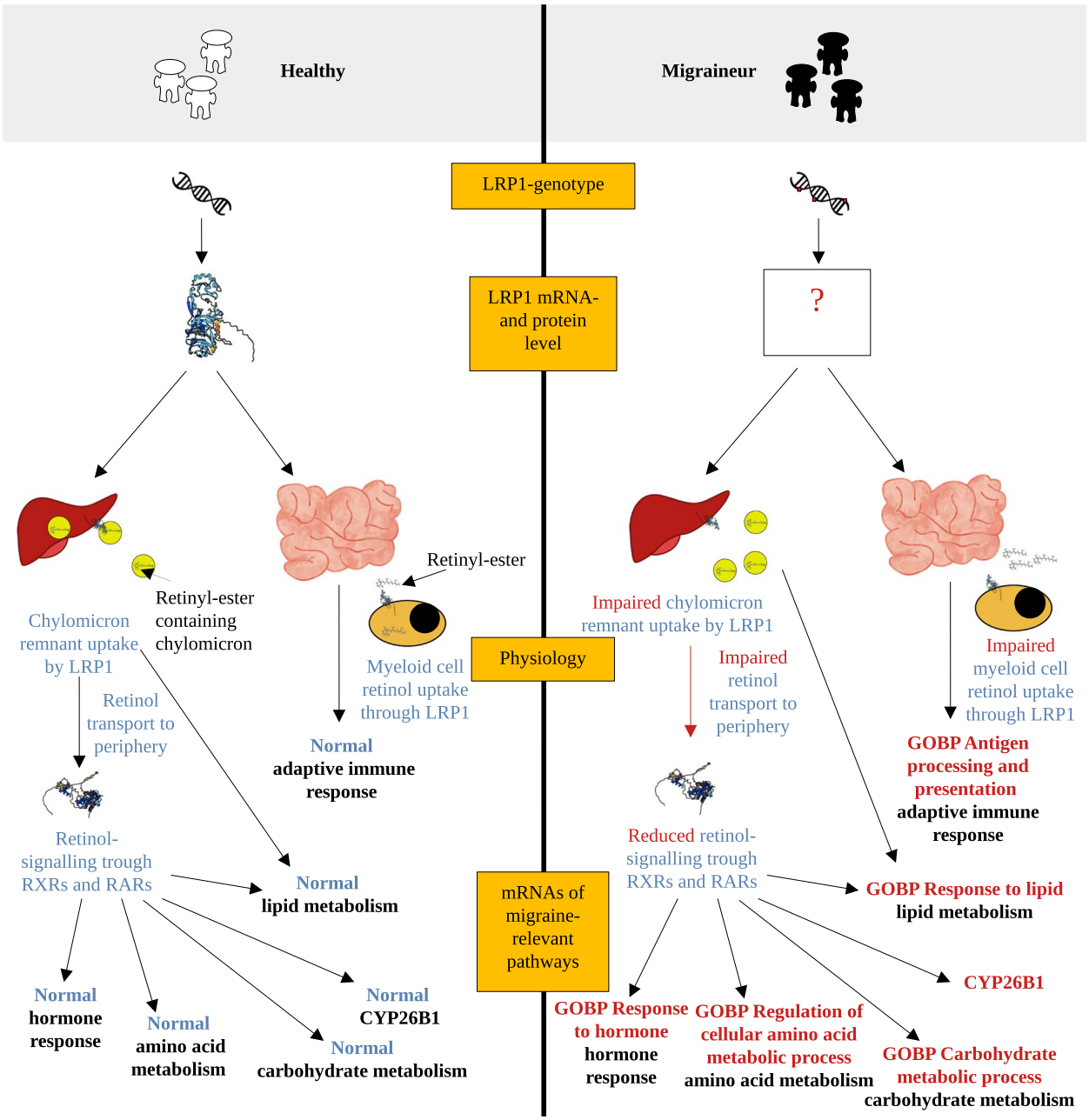

Fig. 4 Proposed Mechanistic Model: Role of Vitamin A Signaling in Migraine Pathogenesis. Figure depicts how LRP1 polymorphisms, impacting LRP1 gene expression (not measured due to low blood-expression) and protein levels, may lead to migraine-related transcriptomic changes. Key processes involve retinyl-ester and retinol uptake in hepatocytes and myeloid cells. The changes ultimately result in the downregulation of CYP26B1 in a replicated allergy-uncorrected analysis, indicative of underlying physiological alterations, and other downregulated migraine-related pathways, where one example pathway each has been mentioned. Measured changes are highlighted in bold black fonts

naturally can lead to decreased levels of hepatic- and peripheral retinoic acid derivatives and result in diminished atRA-dependent signalling and *CYP26B1* expression.

RXR/RAR receptors are key players in atRA-dependent signaling and known to influence gene expression. The found regulatory connections between 63.76% of LEGs and RXR/RAR receptors show how a reduced retinoic acid level can be connected to the LEGs and pathway-level changes. In agreement, pathway-level downregulations including amino acid- and carbohydrate metabolism and oxidative stress can also be related

to reduced atRA levels, as vitamin A deficiency could induce growth restriction and reduced gluconeogenesis, glycolysis, and likely, protein catabolism in the liver [53] and a reduced antioxidant response [54]. Furthermore, LRP1-dependent reduced retinol uptake may be the cause for downregulation in immune system-related pathways by restricting retinol availability for immune cells [52], while impaired chylomicron remnant uptake [28] may be responsible for the observed downregulated lipid metabolism and elevated lipid availability in the periphery, an often comorbid condition with migraine.

All in all, *CYP26B1* downregulation together with the genetic polymorphisms of *LRP1*, the regulatory connections between LEGs and RXR/RAR receptors and several downregulated pathway categories paint a consistent picture about a reduced retinoic acid signaling in MO. It extends previous meta-analyses [3, 4], which found consistent associations between *LRP1* polymorphisms and migraine, by providing a genetically founded mechanism for the transcriptomic and pathophysiological changes, which were independent of the different vitamin A intake between migraine patients and controls. The proposed mechanism also establishes a connection between migraine subtypes, since 1) meta-analysis showed relevance for *LRP1* lead variant for both MO and migraine with aura (4) and 2) reduced retinol binding protein 4, a protein with high sensitivity and specificity to measure vitamin A deficiency, was found in both subtypes [55]. In sum, the found changes 1) integrate alterations in various pathways corresponding to migraine pathophysiology theories via an altered retinoic acid signaling, 2) and raise the possibility of therapeutic approaches targeting this pathway in MO in clinical settings.

CORIN downregulation seemed to be an independent alteration in MO, as regulatory connections between retinoic acid receptors and *CORIN* were scarce. CORIN is a key enzyme in atrial and brain-derived natriuretic peptide (ANP, BNP, respectively) synthesis [56]. Reduced level of CORIN may explain for 1) a report about decreased blood ANP levels in migraine [57] and 2) increased BNP precursors in patients [58]. Furthermore, reduced BNP level, a likely result of downregulated *CORIN* in our study, has been shown to mimic effects of a key mutation in the familial form of migraine on P2X3 receptors in trigeminal neurons and to facilitate trigeminal sensitization, the primary cause for migraine pain [59, 60].

Additional LEGs of interest also emerged. The overall picture of these candidates (Fig. 2) shows that the relationship between genetic background and observed transcriptomic changes in MO is complex. As Bonferroni-correction is known to filter positive findings in favor to control type 1 errors [61–63], some genes, such as *EFNB2*, *CBR3*, *TEK*, *KCNMA1*, *HTR7* with nominally significant SNPs and connection to migraine or its characteristics [64–68], may (also) have genetically determined changes in their gene expression. Another gene of interest may be *SGCD*, which is involved in cardiomyopathy and age-related macular degeneration [69], and provided a large majority of SNP-level nominally significant hits, but remained unassociated previously with MO. These findings propose that different underlying mechanisms are reflected in the current findings: 1) with some directly linked to underlying SNPs and 2) some connected indirectly, via probable, functional, secondary links.

Despite replicable, consistent findings on gene- and pathway-levels, existing migraine medications showed limited binding to corresponding proteins in accordance with their significant, but modest absolute therapeutic benefit [70]. This is not due to non-existing chemical agents to manipulate these targets. Multiple drug candidates from Drugbank were shown to bind LEGs or significant pathways. In fact, some of these have been tested for migraine or its symptoms already, like probucol, ginsenosides, nicardipin, tetrandrine (Tables 4–5). These may serve as templates for future drug developments. At the same time, results contain several hundred additional small molecule compounds predicted to act at the identified LEGs and pathways. These, especially, if combined with expert knowledge and well-formulated hypotheses, can deliver more efficient MO medications in the future.

Our approach comes with limitations. First, the use of whole blood samples was limited in identifying differentially expressed genes with low or no absolute expression in blood (like *LRP1*). It has to be noted, however, that 1) the reported *CYP26B1* and *CORIN* downregulation and metabolic, immunologic, cardiovascular and hormonal pathways's effects have mostly tissue-independent implications, 2) expression from blood samples correlated well with e.g., that of brain samples especially, when relevant genes were examined [71], 3) genetic analyses by nature report tissue-independent associations. Second, single-cell gene expression may have provided detailed results, but our intention in the present study was not a detailed temporal and spatial resolution of cell-type specific changes. Third, we did not measure protein levels, thus, it cannot be excluded that some gene expression changes do not manifest in protein level changes. However, this is unlikely to substantially influence our conclusions, since mRNA and protein levels show reasonable correlation with divergence often attributed to post-transcriptional regulation [72], a phenomenon less relevant in case of mRNA downregulations. Fourth, we could not directly measure retinoic acid/vitamin A levels in our small sample due to its likely degradation at the time we realised

its importance. Few subjects also reported taking vitamin A containing supplements in the control group (Supplementary material 1) with variable amounts and uncertain regimen. Despite these uncertainties, we conducted confirmatory analyses with corrections for vitamin A containing supplement intake to test the independence of our results. Results have shown marginal changes in log2 FC and significance values of key findings (Supplementary material 16), indicating that our core results are independent from supplement intake. Furthermore, vitamin A deficiency is a rare condition in the examined population [73] and the chance that such individuals were accidentally assigned disproportionately to the MO group is improbable. Fifth, we did not correct for contraceptive use due to the various active ingredients used by the subjects (Supplementary material 1). Sixth, one participant took an anti-allergic medication at the second time point. The latter two limitations may have contributed to heterogeneity of findings, but for this very reason were unlikely to influence those that remained significant at both time points. Seventh, additional factors, like diet, sleep and exercise may have influenced our results, for which, however, we had less reliable data. However, for the only factor showing significant differences between MO and controls, a measure of sleep quality (Supplementary material 1), we conducted confirmatory analyses. Similarly to vitamin A containing supplement intake, results showed that inclusion of sleep quality as co-variate only slightly changed log2 FC and significance values (Supplementary material 16). Eighth, genomic analyses use Bonferroni correction as standard despite the potential for false negatives, therefore, nominally significant results in genetic analyses may be potential candidates for future studies to test. Ninth, CYP26B1 protein levels have not been validated by further experiments and the connection between LRP1 and CYP26B1 remains hypothetical, albeit with strong empirical evidence for both being involved in retinoic acid signaling. The above are unlikely to severely impact the interpretability or generalizability of the present findings due to the study's rigorous selection criteria, methodological detail and replication criterion., Future studies compensating these, nonetheless, could definitely add further insights.

Conclusion

All in all, the presented results provide solid support for existing pathophysiology theories through the stable, interictal transcriptomic map of MO, which provided both gene- and pathway level candidates for the underlying pathophysiology. In addition, our study also delivered drug candidates and interventions to manipulate these MO-associated gene- and pathway level targets. Furthermore, the present study provided consistent evidences for a factor behind a share of previous pathological observations in metabolic-, immunologic- and cardiovascular processes in migraine in the form of an altered retinoic acid signaling and implicitly suggests that targeting elements of this pathway in patients may alleviate some of the previously found pathological alterations in the disorder.

Abbreviations

atRA	All-trans Retinoic Acid
BP	Biological Processes
CNN	Convolutional Neural Network
FDR	False Discovery Rate
GO	Gene Ontology
GRNdb	Gene Regulatory Network database
GSEA	Gene Set Enrichment Analysis
LEGs	Leading Edge Genes
LEG5	Leading Edge Genes with an absolute mean logarithm twofold change value > = 2
LEG69	Leading Edge Genes with an absolute mean logarithm twofold change value > = 0.5
log2 FC	Logarithm twofold change
MF	Molecular Function
MIDAS	Migraine Disability Assessment
MINI	Mini International Neuropsychiatric Interview
MO	Migraine without Aura
MSigDB	Molecular Signature Database
NES	Normalized Enrichment Score
RNA-Seq	MRNA Sequencing
S1	Time Point 1
S2	Time Point 2
SGCD	Sarcoglycan Delta
SNP	Single Nucleotide Polymorphism
UKB	UK Biobank
QC	Quality control

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s10194-025-02033-z.

Supplementary Material 1
Supplementary Material 2
Supplementary Material 3
Supplementary Material 4
Supplementary Material 5
Supplementary Material 6
Supplementary Material 7
Supplementary Material 8
Supplementary Material 9
Supplementary Material 10
Supplementary Material 11
Supplementary Material 12
Supplementary Material 13
Supplementary Material 14
Supplementary Material 15
Supplementary Material 16

Authors' contributions
Conceptualization: PP, SK, GJ Data curation: KG, DB, GK Formal analysis: PP, SK, DAN, DT, ZG Funding acquisition: GJ, HM Methodology: PP, SK, DAN, DT, HM, GJ Project administration: PP, GJ Writing – original draft: PP, SK Writing – review & editing: PP, SK, DAN, DT, ZG, DB, KG, GK, HM, GJ.

Funding
Open access funding provided by Semmelweis University. The present research has been supported by Hungarian Academy of Sciences, MTA-SE Neuropsychopharmacology and Neurochemistry Research Group; KTIA_NAP_13 - 1–2013- 0001, KTIA_NAP_13 - 2- 2015–0001, 2017–1.2.1-NKP- 2017–00002; Hungarian Brain Research Program 3.0, NAP2022-I- 4/2022; Thematic Excellence Program of the Ministry for Innovation and Technology in Hungary, 2020–4.1.1.-TKP2020; Ministry of Innovation and Technology of Hungary from the National Research, Development and Innovation Fund, TKP2021-EGA-25; National Research, Development and Innovation Office, Hungary, K143391; ERA-NET COFUND Program (ERAPERMED2019 - 108), 2019–2.1.7-ERA-NET-2020–00005; UNKP- 23–4-I-SE- 31 (KG), UNKP- 23–4-II-SE- 2 (DB); MEXT KAKENHI 19H04169, 20F20809, 21H05027, 22H03645 and 25H01144 (HM); AIPSE program by the Academy of Finland (HM); Japan Society for the Promotion of Science Postdoctoral Fellowships for Research in Japan, standard program, P20809 (PP); International Short-term Exchange Program for Young Researchers by the ICR-iJURC, Kyoto University (DT and KG); University Research Scholarship Programme - EKÖP-2024-68 (DT) and EKÖP-2024-164 (KG); the Hungarian National Research, Development and Innovation Office (K143764 and ADVANCED 150815). This research has been conducted using the UK Biobank Resource under Application Number 71718, Copyright © (2023), NHS England. Re-used with the permission of the NHS England [and/or UK Biobank]. All rights reserved.
This research used data assets made available by National Safe Haven as part of the Data and Connectivity National Core Study, led by Health Data Research UK in partnership with the Office for National Statistics and funded by UK Research and Innovation (research which commenced between 1 st October 2020–31 st March 2021 grant ref MC_PC_20029; 1 st April 2021 - 30 th September 2022 grant ref MC_PC_20058).

Data availability
The RNA-Seq datasets generated and analysed during the current study are available in the ArrayExpress repository, https://www.ebi.ac.uk/biostudies/arrayexpress/studies/E-MTAB- 13397.

Declarations

Ethics approval and consent to participate
The study adhered to the Helsinki Declaration for Research and was approved by the Hungarian ethical committee of the Medical Research Council (number: 23609–1/2011-EKU, 23421–1/2015-EKU). All participants provided written informed consent before participation.

Consent for publication
The authors approved the publication of this paper.

Competing interests
The authors declare no competing interests.

Received: 18 February 2025 Accepted: 11 April 2025
Published online: 12 May 2025

References

1. Steiner TJ, Stovner LJ, Jensen R, Uluduz D, Katsarava Z, The L, Burden: the Global Campaign against Headache, (2020) Migraine remains second among the world's causes of disability, and first among young women: findings from GBD2019. J Headache Pain 21(1):137. https://doi.org/10.1186/s10194-020-01208-0
2. Charles A (2018) The pathophysiology of migraine: implications for clinical management. Lancet Neurol 17(2):174–182. https://doi.org/10.1016/s1474-4422(17)30435-0
3. Gormley P, Anttila V, Winsvold BS, Palta P, Esko T, Pers TH, Farh K-H, Cuenca-Leon E, Muona M, Furlotte NA et al (2016) Meta-analysis of 375,000 individuals identifies 38 susceptibility loci for migraine. Nat Genet 48(8):856–866. https://doi.org/10.1038/ng.3598
4. Hautakangas H, Winsvold BS, Ruotsalainen SE, Bjornsdottir G, Harder AVE, Kogelman LJA, Thomas LF, Noordam R, Benner C, Gormley P et al (2022) Genome-wide analysis of 102,084 migraine cases identifies 123 risk loci and subtype-specific risk alleles. Nat Genet 54(2):152–160. https://doi.org/10.1038/s41588-021-00990-0
5. Kogelman LJA, Falkenberg K, Halldorsson GH, Poulsen LU, Worm J, Ingason A, Stefansson H, Stefansson K, Hansen TF, Olesen J (2019) Comparing migraine with and without aura to healthy controls using RNA sequencing. Cephalalgia 39(11):1435–1444. https://doi.org/10.1177/0333102419851812
6. Kogelman LJA, Falkenberg K, Buil A, Erola P, Courraud J, Laursen SS, Michoel T, Olesen J, Hansen TF (2021) Changes in the gene expression profile during spontaneous migraine attacks. Sci Rep 11(1):8294. https://doi.org/10.1038/s41598-021-87503-5
7. Aczél T, Körtési T, Kun J, Urbán P, Bauer W, Herczeg R, Farkas R, Kovács K, Vásárhelyi B, Karvaly GB et al (2021) Identification of disease- and headache-specific mediators and pathways in migraine using blood transcriptomic and metabolomic analysis. J Headache Pain 22(1):117. https://doi.org/10.1186/s10194-021-01285-9
8. Gerring ZF, Powell JE, Montgomery GW, Nyholt DR (2017) Genome-wide analysis of blood gene expression in migraine implicates immune-inflammatory pathways. Cephalalgia 38(2):292–303. https://doi.org/10.1177/0333102416686769
9. Perry CJ, Blake P, Buettner C, Papavassiliou E, Schain AJ, Bhasin MK, Burstein R (2016) Upregulation of inflammatory gene transcripts in periosteum of chronic migraineurs: Implications for extracranial origin of headache. Ann Neurol 79(6):1000–1013. https://doi.org/10.1002/ana.24665
10. Nagata E, Hattori H, Kato M, Ogasawara S, Suzuki S, Shibata M, Shimizu T, Hamada J, Osada T, Takaoka R et al (2009) Identification of biomarkers associated with migraine with aura. Neurosci Res 64(1):104–110. https://doi.org/10.1016/j.neures.2009.02.001
11. Stewart WF, Lipton RB, Dowson AJ, Sawyer J. Development and testing of the Migraine Disability Assessment (MIDAS) Questionnaire to assess headache-related disability. Neurology. 2001;56(6 Suppl 1):S20–S28. https://https://doi.org/10.1212/wnl.56.suppl_1.s20.
12. The International Classification of Headache Disorders (2013) 3rd edition (beta version). Cephalalgia 33:629–808. https://doi.org/10.1177/0333102413485658
13. Sheehan DV, Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E, Hergueta T, Baker R, Dunbar GC (1998) The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. J Clin Psychiatry, 59(Suppl 20), 22–33
14. Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: A flexible trimmer for Illumina sequence data. Bioinformatics 30(15):2114–2120. https://doi.org/10.1093/bioinformatics/btu170
15. Andrews S (2010) FastQC A Quality Control tool for High Throughput Sequence Data. https://www.bioinformatics.babraham.ac.uk/projects/fastqc/. Accessed 15 Jan 2025
16. Kim D, Langmead B, Salzberg SL (2015) HISAT: A fast spliced aligner with low memory requirements. Nat Methods 12(4):357–360. https://doi.org/10.1038/nmeth.3317
17. Pertea M, Kim D, Pertea GM, Leek JT, Salzberg SL (2016) Transcript-level expression analysis of RNA-seq experiments with HISAT. StringTie and Ballgown, Nat Protoc 11:1650–1667. https://doi.org/10.1038/nprot.2016.095
18. R Core Team (2018) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/. Accessed 15 Jan 2025
19. Huber W, Carey V, Gentleman R, Anders S, Carlson M, Carvalho BS, Bravo HC, Davis S, Gatto L, Girke T et al (2015) Orchestrating high-throughput genomic analysis with Bioconductor. Nat Methods 12:115–121. https://doi.org/10.1038/nmeth.3252
20. Robinson MD, McCarthy DJ, Smyth GK (2010) edgeR: A Bioconductor package for differential expression analysis of digital gene expression

data. Bioinformatics 26(1):139–140. https://doi.org/10.1093/bioinformatics/btp616
21. Benjamini Y, Hochberg Y (1995) Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J R Stat Soc Series B Stat Methodol 57(1):289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x
22. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES et al (2005) Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A 102(43):15545–15550. https://doi.org/10.1073/pnas.0506580102
23. Human MSigDB Collections. https://www.gsea-msigdb.org/gsea/msigdb/human/collections.jsp. Accessed 25 March 2025.
24. Korotkevich G, Sukhov V, Budin N, Shpak B, Artyomov MN, Sergushichev A (2021) Fast gene set enrichment analysis. BioRxiv, 060012. https://doi.org/10.1101/060012 [PREPRINT]
25. Rapaport F, Khanin R, Liang Y, Pirun M, Krek A, Zumbo P, Mason C, Socci ND, Betel D. Comprehensive evaluation of differential gene expression analysis methods for RNA-seq data. Genome Biol. 2013;14(9):R95. https://https://doi.org/10.1186/gb-2013-14-9-r95.
26. Torok D, Petschner P, Baksa D, Juhasz G. Improved polygenic risk prediction in migraine-first patients. J Headache Pain. 2024;25:161. https://https://doi.org/10.1186/s10194-024-01870-8.
27. Wang K, Li M, Hakonarson H (2010) ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res 38(16):e164. https://doi.org/10.1093/nar/gkq603
28. D'Ambrosio DN, Clugston RD, Blaner WS (2011) Vitamin A metabolism: an update. Nutrients 3(1):63–103. https://doi.org/10.3390/nu3010063
29. Carazo A, Macáková K, Matoušová K, Krčmová LK, Protti M, Mladěnka P (2021) Vitamin A Update: Forms, Sources, Kinetics, Detection, Function, Deficiency. Therapeutic Use and Toxicity Nutrients 13:1703. https://doi.org/10.3390/nu13051703
30. Stelzer G, Rosen N, Plaschkes I, Zimmerman S, Twik M, Fishilevich S, Stein TI, Nudel R, Lieder I, Mazor Y et al. (2016) The GeneCards suite: from gene data mining to disease genome sequence analyses. Curr Protoc Bioinformatics, 54:1.30.1- 1.30.33. https://doi.org/10.1002/cpbi.5
31. Wishart DS, Feunang YD, Guo AC, Lo EJ, Marcu A, Grant JR, Sajed T, Johnson D, Li C, Sayeeda Z et al. (2017) DrugBank 5.0: a major update to the DrugBank database for 2018. Nucleic Acids Res, 46(D1):D1074-D1082. https://doi.org/10.1093/nar/gkx1037
32. Kim S, Chen J, Cheng T, Gindulyte A, He J, He S, Li Q, Shoemaker BA, Thiessen PA, Yu B et al (2023) PubChem 2023 update. Nucleic Acids Res 51(D1):D1373–D1380. https://doi.org/10.1093/nar/gkac956
33. The UniProt Consortium (2021) UniProt: the universal protein knowledgebase in 2021. Nucleic Acids Res 49(D1):D480–D489. https://doi.org/10.1093/nar/gkaa1100
34. Eberhardt J, Santos-Martins D, Tillack AF, Forli S. (2021) AutoDock Vina 1.2.0: New Docking Methods, Expanded Force Field, and Python Bindings. J Chem Inf Model, 61(8):3891–3898. https://doi.org/10.1021/acs.jcim.1c00203
35. Luo Q, Zhao L, Hu J, Jin H, Liu Z, Zhang L (2017) The scoring bias in reverse docking and the score normalization strategy to improve success rate of target fishing. PLoS ONE 12(2):e0171433. https://doi.org/10.1371/journal.pone.0171433
36. Zhao Q, Zhao H, Zheng K, Wang J (2022) HyperAttentionDTI: improving drug-protein interaction prediction by sequence-based deep learning with attention mechanism. Bioinformatics 38(3):655–662. https://doi.org/10.1093/bioinformatics/btab715
37. González-Hernández A, Condés-Lara M (2014) The multitarget drug approach in migraine treatment: the new challenge to conquer. Headache 54(1):197–199. https://doi.org/10.1111/head.12237
38. Buse DC, Reed ML, Fanning KM, Bostic R, Dodick DW, Schwedt TJ, Munjal S, Singh P, Lipton RB (2020) Comorbid and co-occurring conditions in migraine and associated risk of increasing headache pain intensity and headache frequency: results of the migraine in America symptoms and treatment (MAST) study. J Headache Pain 21(1):23. https://doi.org/10.1186/s10194-020-1084-y
39. Vetvik KG, MacGregor EA (2017) Sex differences in the epidemiology, clinical features, and pathophysiology of migraine. Lancet Neurol 16(1):76–87. https://doi.org/10.1016/s1474-4422(16)30293-9
40. Fang L, Li Y, Ma L, Xu Q, Tan F, Chen G (2021) GRNdb: decoding the gene regulatory networks in diverse human and mouse conditions. Nucleic Acids Res 49(D1):D97–D103. https://doi.org/10.1093/nar/gkaa995
41. Sutherland HG, Albury CL, Griffiths LR (2019) Advances in genetics of migraine. J Headache Pain 20(1):72. https://doi.org/10.1186/s10194-019-1017-9
42. Alam Z, Coombes N, Waring RH, Williams AC, Steventon GB (1998) Plasma levels of neuroexcitatory amino acids in patients with migraine or tension headache. J Neurol Sci 156(1):102–106. https://doi.org/10.1016/s0022-510x(98)00023-9
43. Biringer RG (2022) Migraine signaling pathways: amino acid metabolites that regulate migraine and predispose migraineurs to headache. Mol Cell Biochem 477(9):2269–2296. https://doi.org/10.1007/s11010-022-04438-9
44. Harder AVE, Vijfhuizen LS, Henneman P, van Dijk KW, van Duijn CM, Terwindt GM, van den Maagdenberg AMJM (2021) Metabolic profile changes in serum of migraine patients detected using 1H-NMR spectroscopy. J Headache Pain 22(1):142. https://doi.org/10.1186/s10194-021-01357-w
45. Gross EC, Lisicki M, Fischer DSándor PS, Schoenen J, (2019) The metabolic face of migraine - from pathophysiology to treatment. Nat Rev Neurol 15(11):627–643. https://doi.org/10.1038/s41582-019-0255-4
46. Biscetti L, De Vanna G, Cresta E, Bellotti A, Corbelli I, Letizia Cupini M, Calabresi P, Sarchielli P (2022) Immunological findings in patients with migraine and other primary headaches: a narrative review. Clin Exp Immunol 207(1):11–26. https://doi.org/10.1093/cei/uxab025
47. Elgendy IY, Nadeau SE, Bairey Merz CN, Pepine CJ (2019) American College of Cardiology Cardiovascular Disease in Women Committee. Migraine Headache: An Under-Appreciated Risk Factor for Cardiovascular Disease in Women. J Am Heart Assoc, 8(22):e014546. https://doi.org/10.1161/jaha.119.014546
48. Krause DN, Warfvinge K, Haanes KA, Edvinsson L (2021) Hormonal influences in migraine - interactions of oestrogen, oxytocin and CGRP. Nat Rev Neurol 17(10):621–633. https://doi.org/10.1038/s41582-021-00544-2
49. Isoherranen N, Zhong G (2019) Biochemical and physiological importance of the CYP26 retinoic acid hydroxylases. Pharmacol Ther 204:107400. https://doi.org/10.1016/j.pharmthera.2019.107400
50. Reijntjes S, Gale E, Maden M (2004) Generating gradients of retinoic acid in the chick embryo: Cyp26C1 expression and a comparative analysis of the Cyp26 enzymes. Dev Dyn 230(3):509–517. https://doi.org/10.1002/dvdy.20025
51. Kashimada K, Svingen T, Feng CW, Pelosi E, Bagheri-Fam S, Harley VR, Schlessinger D, Bowles J, Koopman P(2011) Antagonistic regulation of Cyp26b1 by transcription factors SOX9/SF1 and FOXL2 during gonadal development in mice. FASEB J, 25(10):3561–3569. https://doi.org/10.1096/fj.11-184333
52. Bang YJ, Hu Z, Li Y, Gattu S, Ruhn KA, Raj P, Herz J, Hooper LV (2021) Serum amyloid A delivers retinol to intestinal myeloid cells to promote adaptive immunity. Science, 373(6561):eabf9232. https://doi.org/10.1126/science.abf9232
53. Chen W, Chen G (2014) The roles of vitamin A in the regulation of carbohydrate, lipid, and protein metabolism. J Clin Med 3(2):453–479. https://doi.org/10.3390/jcm3020453
54. Blaner WS, Shmarakov IO, Traber MG (2021) Vitamin A and vitamin E: will the real antioxidant please stand up? Annu Rev Nutr 41:105–131. https://doi.org/10.1146/annurev-nutr-082018-124228
55. Tanik N, Celikbilek A, Metin A, Gocmen AY, Inan LE (2015) Retinol-binding protein-4 and hs-CRP levels in patients with migraine. Neurol Sci 36:1823–1827. https://doi.org/10.1007/s10072-015-2262-6
56. Shen X, Dong N, Xu Y, Han L, Yang R, Liao J, Zhang X, Xie T, Wang Y, Chen C et al (2022) Analyzing Corin–BNP–NEP Protein Pathway Revealing Differential Mechanisms in AF-Related Ischemic Stroke and No AF-Related Ischemic Stroke. Front Aging Neurosci 14:863489. https://doi.org/10.3389/fnagi.2022.863489
57. Araki N (1995) Autonomic nervous activity in migraine. Rinsho Shinkeigaku 35(12):1336–1338
58. Uzar E, Evliyaoglu O, Yucel Y, Cevik MU, Acar A, Guzel I, Islamoglu Y, Colpan L, Tasdemir N (2011) Serum cytokine and pro-brain natriuretic peptide (BNP) levels in patients with migraine. Eur Rev Med Pharmacol Sci 15(10):1111–1116
59. Marchenkova A, Vilotti S, Fabbretti E, Nistri A (2015) Brain natriuretic peptide constitutively downregulates P2X3 receptors by controlling

their phosphorylation state and membrane localization. Mol Pain 11(71):s12990–s13015. https://doi.org/10.1186/s12990-015-0074-6
60. Marchenkova A, Vilotti S, Ntamati N, van den Maagdenberg AMJM, Nistri A (2016) Inefficient constitutive inhibition of P2X3 receptors by brain natriuretic peptide system contributes to sensitization of trigeminal sensory neurons in a genetic mouse model of familial hemiplegic migraine. Mol Pain 12:1744806916646110. https://doi.org/10.1177/1744806916646110
61. Conneely KN, Boehnke M (2007) So many correlated tests, so little time! Rapid adjustment of *P* values for multiple correlated tests. Am J Hum Genet 81(6):1158–1168. https://doi.org/10.1086/522036
62. van den Oord EJ (2008) Controlling false discoveries in genetic studies. Am J Med Genet B Neuropsychiatr Genet 147B(5):637–644. https://doi.org/10.1002/ajmg.b.30650
63. Petschner P, Baksa D, Hullam G, Torok D, Millinghoffer A, Deakin JFW, Bagdy G, Juhasz G (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
64. Thomas JM, Surendran S, Abraham M, Rajavelu A, Kartha CC (2016) Genetic and epigenetic mechanisms in the development of arteriovenous malformations in the brain. Clin Epigenetics 8:78. https://doi.org/10.1186/s13148-016-0248-8
65. Anttila V, Nyholt DR, Kallela M, Artto V, Vepsäläinen S, Jakkula E, Wennerström A, Tikka-Kleemola P, Kaunisto MA, Hämäläinen E et al (2008) Consistently replicating locus linked to migraine on 10q22-q23. Am J Hum Genet 82(5):1051–1063. https://doi.org/10.1016/j.ajhg.2008.03.003
66. Al-Karagholi MA, Ghanizada H, Nielsen CAW, Skandarioon C, Snellman J, Lopez CL, Hansen JM, Ashina M (2020) Opening of BKCa channels alters cerebral hemodynamic and causes headache in healthy volunteers. Cephalalgia 40(11):1145–1154. https://doi.org/10.1177/0333102420940681
67. Miller JP, Moldenhauer HJ, Keros S, Meredith AL (2021)An emerging spectrum of variants and clinical features in KCNMA1-linked channelopathy. Channels (Austin), 15(1):447–464. https://doi.org/10.1080/19336950.2021.1938852
68. Cox HC, Lea RA, Bellis C et al (2012) A genome-wide analysis of "Bounty" descendants implicates several novel variants in migraine susceptibility. Neurogenetics 13(3):261–266. https://doi.org/10.1007/s10048-012-0325-x
69. Perez-Ortiz AC, Luna-Angulo A, Zenteno JCPerez-Ortiz AC, Luna-Angulo A, Zenteno JC, Rendon A, Cortes-Ballinas LG, Jimenez-Collado D, Antonio-Aguirre B, Peralta-Ildefonso MJ, Ramírez I, Jacob-Kuttothara S, et al (2008) Significant Association Between Variant in SGCD and Age-Related Macular Degeneration. Genes 9(10):467. https://doi.org/10.3390/genes9100467
70. Tsou AY, Rouse B, Bloschichak A, Treadwell JR (2021) Drugs and Devices for Migraine Prevention: Interactive Evidence Maps., Patient-Centered Outcomes Research Institute (PCORI). https://doi.org/10.25302/emv1.2021.2
71. Tylee DS, Kawaguchi DM, Glatt SJ (2013) On the outside, looking in: a review and evaluation of the comparability of blood and brain "-omes". Am J Med Genet B Neuropsychiatr Genet. 162 (7):595–603. https://doi.org/10.1002/ajmg.b.32150
72. Buccitelli C, Selbach M (2020) mRNAs, proteins and the emerging principles of gene expression control. Nat Rev Genet 21(10):630–644. https://doi.org/10.1038/s41576-020-0258-4
73. Han X, Ding S, Lu J, Li Y (2022) Global, regional, and national burdens of common micronutrient deficiencies from 1990 to 2019: A secondary trend analysis based on the Global Burden of Disease 2019 study. EClinicalMedicine 44:101299. https://doi.org/10.1016/j.eclinm.2022.101299

Publisher's Note