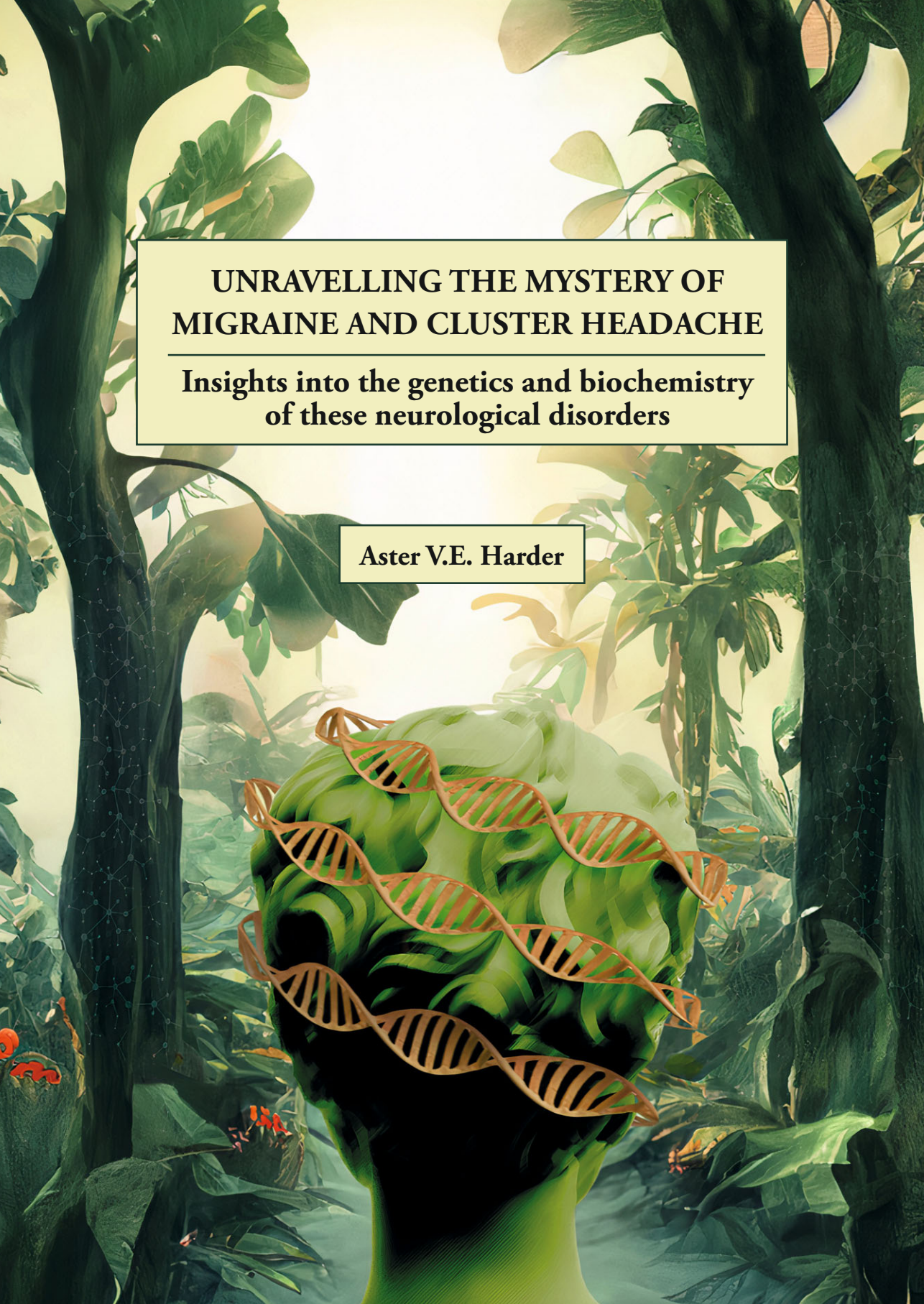


UNRAVELLING THE MYSTERY OF MIGRAINE AND CLUSTER HEADACHE

Insights into the genetics and biochemistry of these neurological disorders

Aster V.E. Harder

UNRAVELLING THE MYSTERY OF MIGRAINE AND CLUSTER HEADACHE

Insights into the genetics and biochemistry of these neurological disorders

Aster V.E. Harder

Aster Harder
Unravelling the mystery of migraine and cluster headache: Insights into the genetics and biochemistry of these neurological disorders
PhD thesis, Leiden University, Leiden The Netherlands, 2023

Cover & Lay-out: Publiss | www.publiss.nl
Print: Ridderprint | www.ridderprint.nl

Funding for publication of this thesis has been provided by the Dutch Headache Society and was gratefully accepted.

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

Proefschrift

ter verkrijging van

de graad van doctor aan de Universiteit Leiden,

op gezag van rector magnificus prof.dr.ir. H. Bijl,

volgens besluit van het college voor promoties

te verdedigen op dinsdag 21 november 2023

klokke 16:15 uur

door

Aster Victorine Elsemieke Harder

Promotiecommissie

Promotores	Prof. dr. A.M.J.M. van den Maagdenberg	
	Prof. dr. G.M. Terwindt	
	Prof. dr. M.D. Ferrari	
Leden promotiecommissie	Prof. dr. J.J.G.M. Verschuuren	
	Prof. dr. M. Wuhrer	
	Prof. dr. I. Meulenbelt	
	Prof. dr. ir. M.M. Verbeek	*Radboud Universiteit*
	Prof. dr. D.I. Boomsma	*Vrije Universiteit Amsterdam*

Table of Contents

General introduction

Adapted from:
Migraine genetics: Status and road forward
Aster V.E. Harder, Gisela M. Terwindt, Dale R. Nyholt and Arn M.J.M. van den Maagdenberg

Cephalalgia. 2023;43:3331024221145962

1

Introduction

This thesis aims to further unravel the pathophysiology of migraine and cluster headache. Both migraine and cluster headache are disabling primary headache disorders characterized by attacks of severe headache and associated symptoms.[1] Cluster headache is one of the trigeminal autonomic cephalalgias.[1] By definition, primary headache disorders are not the result of any other underlying disease or process, contrary to secondary headache disorders. Although much progress has been made with unravelling the disease mechanisms of migraine and cluster headache, their pathophysiology remains poorly understood.[2, 3] A major hurdle is that there are no diagnostic biomarkers and the diagnosis, therefore, is still made using direct interviews and/or questionnaires based on clinical consensus criteria of the International Classification of Headache disorders (ICHD-3 criteria).[1] A shortcoming of the current classification criteria is that it does not take the complexity of disease mechanisms into account. In other words, the ICHD-3 criteria do not fully capture the heterogeneity of the disease, including the underlying neurobiological and genetic factors.[4] Understanding the pathophysiology better will improve diagnosis, prognosis, and generate new treatment options.

Clinical characteristics

Migraine

Migraine is characterized by recurrent episodes of severe often unilateral pulsating headache accompanied by nausea, vomiting and/or photo- and phonophobia lasting for 4-72 hours.[1] Migraine can be subdivided in two main subtypes: migraine without aura and migraine with aura. For the latter, headaches are preceded by transient neurological symptoms, known as the aura phase, which typically lasts from 5 until 60 minutes.[1] Cortical spreading depolarization (CSD) is the presumed underlying mechanism of the aura in migraine.[5-8] A typical migraine attack consist of a preictal, ictal (aura and/or headache), and postictal (postdromal) phase.[9, 10] Clinically, a patient can be described as interictal, when there is no attack or ictal when an aura and/or migrainous headache is occurring. Migraine is three times more prevalent in women than in men with a peak prevalence of 25%.[11, 12] Migraine is associated with several neuropsychiatric disorders, among which depression.[13] Migraine is considered a multifactorial (complex) genetic disorder, with a strong familial aggregation.[14-16] Complex traits are typically brought about by a combination of multiple genetic variants, each with a small effect size, and behavioural and environmental factors. Hemiplegic migraine (HM) is a rare subtype of migraine with aura, HM is characterized by attacks that are associated with motor weakness that can lead to hemiplegia during the aura phase.[1]

Cluster headache

Cluster headache is a primary headache disorder characterized by excruciating unilateral headache or facial pain accompanied by ipsilateral facial autonomic symptoms and/or restlessness.[1,17] Attacks

may last for 15-180 minutes and can occur from once a day up to 8 times or more a day.[1] Cluster headache attacks commonly follow a circadian rhythm with attacks frequently occurring at night and according to a seasonal pattern. The majority of patients have episodic cluster headache, characterized by periods of cluster headache of weeks to months, alternating with attack-free periods of at least 3 months. A small proportion (10-15%) of patients have chronic cluster headache where the cluster periods do not remit for more than three months for at least one year. Cluster headache has a prevalence of around 0.12% and occurs more often in men than women, with a male-to-female ratio of 2:1.[18, 19] Of note, smoking and psychiatric co-morbidities are prevalent among cluster headache patients.[20] The pathophysiology of cluster headache is poorly understood with current evidence pointing at hypothalamic involvement.[3] Genetic predisposition seems to play an important role as illustrated by twin and family studies but no genetic factors have been identfied.[21]

Pathophysiology

Migraine

Different disease mechanisms are considered to be involved in migraine pathophysiology, such as neurological, cerebrovascular, and neuroinflammatory mechanisms. The aura phase is most likely caused by CSD, a wave of neuronal and glial depolarization, that is an initial hyperactivity is followed by a prolonged inactivity, resulting in a wave that propagates slowly across the cerebral cortex.[22, 23] The depolarization wave classically begins in the occipital (visual) cortex and correlates with a variety of positive aura patterns, as reported by patients.[24, 25] Mechanisms of CSD are heavily investigated in animals using various stimuli, such as topical application of KCl, injection of current, or an optogenetic stimulus, and it was shown that CSD can activate headache mechanisms.[26] However, there is only limited (neuroimaging) data that can be taken as proof of a spreading depolarization event that qualifies as an aura in humans.[27] Also whether the CSD is causally associated with the initiation of the headache phase in patients remains an enigma.[28]

It is generally accepted that the headache phase involves the activation and sensitization of the trigeminovascular system.[29] The trigeminovascular system consists of nociceptive trigeminal afferents from the trigeminal ganglion that surround cranial blood vessels and dura mater projected from the trigeminal cervical complex in the brainstem, which includes the trigeminal nucleus caudalis and the dorsal horns of cervical spinal nerves C1 and C2.[30] Following stimulation, the trigeminal afferents transfer nociceptive signals through the trigeminal ganglion to the trigeminal cervical complex. In the brainstem, the signal is modulated and further conducted to the thalamus via ascending pain pathways and reaches the cortex.[25] Upon stimulation, the trigeminal fibres release proinflammatory neuropeptides (e.g. calcitonin-gene related peptide (CGRP), pituitary adenylate cyclase-activating peptide (PACAP) substance P and neurokinin A) and other mediators that cause vasodilation of the dural and pial vessels.[25] There is ample evidence that vasodilators such as

prostaglandins may be pivotal in the development of migraine attacks.[7,31,32] Increased sensitivity of the trigeminal system is believed to be an important underlying mechanism in migraine pathology. The mechanisms underlying this hypersensitivity during a migraine attack remain unclear.

Additionally, the hypothalamus is believed to be involved in the prodromes (symptoms that precede the migraine headache).[33] Clinically this is evident by increased fatigue, food cravings, yawning and irritability in the patients. Therefore the phase before the migraine attack can also give insights into the pathophysiology of migraine.

Cluster headache

Various mechanisms/structures, e.g. the trigeminovascular system and the hypothalamus, are believed to be involved in cluster headache pathophysiology and it is thought that the interplay of these systems is responsible for the clinical presentation.[34,35] However, how these structures interact with each other and the mechanisms on the initiation of an attack remain unclear. Similar to migraine, also in cluster headache the trigeminovascular system is believed to be involved in pain processing.[34,35] Different divisions of the trigeminal nerve are primarily responsible for the innervation of cranial structures. Stimulations of the different divisions produce pain in different locations, activation of the second-order trigeminocervical neurons at the ophthalmic division is in line with the clinical presentation of pain in the peri-orbital region.[35]

The trigeminal-autonomic reflex is also associated with the physiological and anatomical landmarks of a cluster headache attack. This reflex is activated upon irritation and produces parasympathetic symptoms, such as nasal congestion and lacrimation.[36] The reflex travels from trigeminal nerve endings to second-order trigeminocervical complex, that projects to the superior salivatory nucleus located in the pons.[34,35] These projections in turn synapse in the peripheral sphenopalatine ganglion and postganglionic parasympathetic nerves and then innervate nasal, pharyngeal and lacrimal glands, inducing autonomic symptoms.[35] Activation of the trigeminovascular system and the trigeminal-autonomic reflex leads to release of neuropeptides (e.g. CGRP and PACAP).

In addition, the hypothalamus is believed to be a key player in cluster headache pathophysiology. The hypothalamus is involved in the regulation of sleep and circadian rhythms.[37] The hypothesis that the hypothalamus is involved in cluster headache is supported by the clinical feature of a circadian rhythm in cluster headache and the finding that the hypothalamus shows increased activation during glyceryl trinitrate (GTN) induced attacks of cluster headache.[38]

Rationale for biochemical studies

Identifying biochemical markers, biomarkers, can help uncover the metabolic underpinnings of human disease. Validated biomarkers can improve diagnosis, prognosis and assess the effectivity of

treatment in patients and lead to novel drug targets, and ultimately novel drugs. This has already been shown for several diseases other than migraine or cluster headache, for instance cardiac troponin helps diagnose myocardial infarction and different biomarkers have been developed for the diagnosis of ovarian cancer.[39-41] A way to investigate whether endogenous signalling molecules are involved in migraine pathology is by trying to provoke an attack in "a human model". When an attack can be provoked with a certain trigger, this suggests the involvement of a related mechanism underlying the disease. Several chemical molecules have been implicated in migraine, identified as they can trigger attacks. The triggers are mostly vasoactive substances that are present at or near the nerve fibres. It has been shown that glyceryl trinitrate (GTN), an nitric oxide (NO) donor, is able to induce an immediate headache in almost all subjects and a delayed migraine-like attack in close to 70% of migraineurs but not in controls.[42, 43] Other substances, such as calcitonin gene-related peptide (CGRP), PACAP and prostaglandin E_2 (PGE_2) and I_2 (PGI_2) are also able to trigger migraine-like attacks.[44-47] Although attacks of cluster headache have been successfully triggered with GTN and histamine, it is not common practice to investigate cluster headache using provocation studies.[48] In addition to investigating trigger mechanisms *per se*, provocation studies can also be used to study other aspects of migraine, such as consequences of attacks, as investigating spontaneous attacks is notoriously difficult as they occur unexpectedly. In contrast, in provocation studies, the set-up can be meticulously controlled.

Another way of investigating relevant substances in disease is by measuring compounds in body fluids, such as cerebrospinal fluid (CSF), blood and urine, and compare profiles in disease vs. control samples. The compounds, being proteins (proteomics) or metabolites (metabolomics), are representative of aspects of the phenotype at the molecular level. For instance, altered blood plasma levels of serotonin (5-HT) in migraine patients were found in the late eighties.[49] This finding contributed to the development of triptans, i.e. 5-HT-1D/1F receptor agonists, which are used for aborting migraine attacks. Serotonin is an amine, just like other neurotransmitters implicated in migraine pathophysiology, such as glutamate and gamma-aminobutyric acid (GABA). This led to amines to be further investigated in the pathogenesis of migraine.[50] Recently, a lot of biochemical research was done on CGRP, which is believed to play an important role in migraine and cluster headache. As mentioned earlier, infusion of CGRP is able to induce migraine-like attacks in migraine patients.[47] In addition, studies show an increase in CGRP levels in blood between cases and controls outside[51-57] or during a migraine attack,[53, 57-61] although almost as many studies have not found a difference in CGRP levels in blood outside[58, 60, 62-65] or during[59, 65] migraine attacks. Regardless, newly approved monoclonal antibodies (mAbs) that target CGRP or its receptor have a beneficial effect on the headache frequency in patients with migraine.[66, 67] In cluster headache, CGRP also seems to be involved[68, 69] with CGRP plasma levels being higher for cluster headache patients during an active period compared to those outside, after provocation with sublingual GTN.[70] In cluster headache, randomized controlled trials on CGRP antibodies in patients have been initiated but with unconvincing results sofar.[71] However, the reliability of measuring CGRP is not without controversy.[65, 72]

When identifying biochemical compounds in body fluids either using a targeted approach, which typically focuses on one or more related selected pathways of interest, or an untargeted approach, which aims to simultaneously measure a large number of metabolites, can be employed. The latter method is most commonly used in the field of metabolomics. Metabolomics is defined as the study of all low molecular weight compounds (<1500 Da) in a sample. Metabolites are the molecular endpoints of gene expression and cell activity and thereby represent, in a way, "the molecular phenotype of an organism". The various types of -omics, i.e. genomics, epigenomics, transcriptomics, proteomic and metabolomic, relate to each other (**Figure 1**). Genomics is the study of the genome at the DNA level, as does epigenomics which investigates modifications of the genome expression. Transcriptomics investigates genomic expression at the RNA level, whereas proteomics interrogates proteins. Finally, metabolomics deals with the metabolome, so the complete set of small-molecule metabolites. Logically, changes in gene expression, enzymes and environmental factors can all have an effect on the "systems biology" and metabolite concentrations.[73] The advantage of the untargeted/omics approach is that it allows for a rapid, concurrent identification and quantification of a multitude of metabolites in many samples at once. By measuring multiple metabolites, one gets a better understanding of the overall metabolomic networks involved. Important aspects and considerations of this method are the validation of the metabolites measured, validation of the used platform, as well as standardization of collection and storage methods.[74,75]

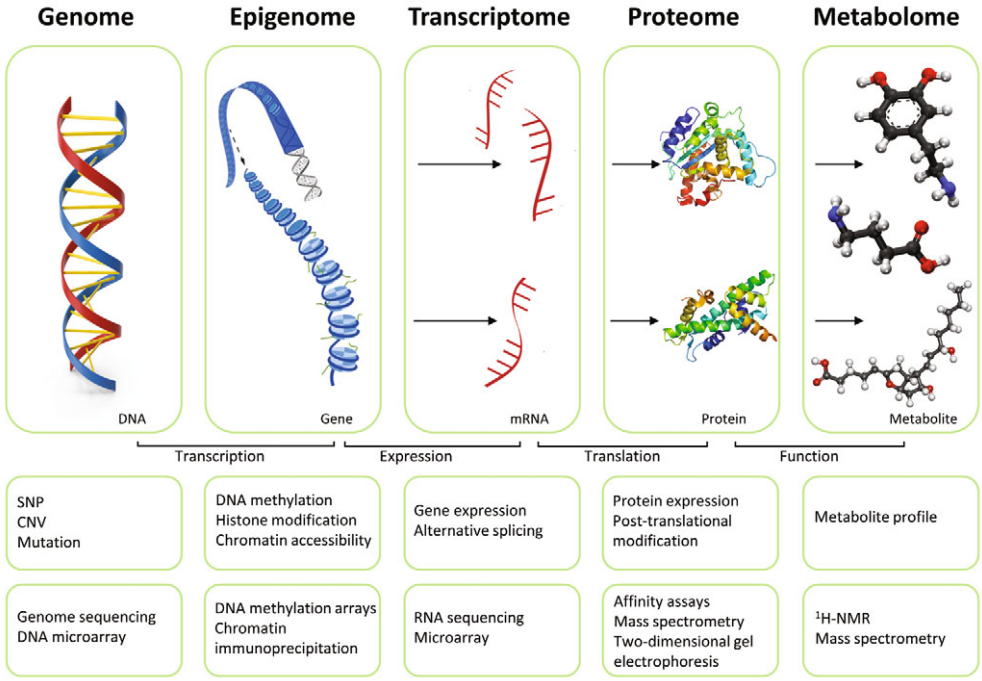
Sample collection

Although metabolomics is a proven, worthwhile approach for biomarker identification, sample collection needs to be done very meticulously as metabolically active cells in body fluids may alter the metabolomic profile *ex vivo*. It has been shown that inaccuracies in the pre-analytical steps cause low quality samples and even up to 80% of the laboratory measurement inaccuracies in daily clinical routine diagnostics.[76-78] In the field of metabolomics, the stability of many metabolites and lipids is extremely variable, therefore, systematic or pre-analytical accidents and inconsistencies can have a great effect on compounds with a low stability and lead to high variability in the analytical data. Therefore, the most critical steps regarding the quality of one's metabolomics data are related to the pre-analytical phase. Of note, each step should be well-considered, standardized and controlled to prevent degradation of sample quality and misinterpretation of findings during the analysis of data.

Important issues to consider in metabolomics research are which biochemical fluid one intends to collect, whether the materials in the collection process up to sample preparation are suitable and do not interfere with the measurement method or the low-freezing storage facility.[75] Another important step is to consider at what temperatures the body fluid will be kept during the preparation process, as lower temperatures reduce the activity of cellular metabolism. It is generally considered that it is important to centrifuge samples as soon as possible, but consistency in time until centrifugation is even more important.[75] Therefore, a standardized protocol for the process of body fluids is essential. Other issues that are more obvious to be kept consistent are,

centrifugation time and force, and temperature of sample storage. For metabolomics and lipidomics, centrifugation (between 2300 and 4000 *g* for 5–10 min) of whole blood is recommended and for CSF 2000 *g* at 4°C to separate erythrocytes, leucocytes and platelets.[75] For serum samples, the coagulation process should be standardized (brand of tubes, kind of coagulation enhancer, clotting time and ambient temperature).[75, 80] On the other hand with stable molecules changes in sample handling do not have to be of large consequences, as has been shown when comparing different aspects (temperature, centrifugation and anti-enzymatic additives) of sample handling in amines in CSF.[81] When samples are kept for long-term storage, a storage temperature of -80°C or lower is advised.[75, 82] Regarding the patient, aspects of the diet, nutritional state as well as circadian rhythm can all affect the metabolome.[75, 83-86] Hence, one should try to keep these factors consistent during sample collection. It is also highly advised to keep track of a person's medications, smoking habits, daily intake of tea/coffee, and alcohol consumption.[75] Despite that these factors are crucial in the data quality of metabolomics research they are often not described in research papers.

Figure 1 Coupling of the different -omics, i.e. the genome, epigenome, transcriptome, proteome and metabolome



Genomic data can differ at different levels due to, for example, copy number variation (CNV), single nucleotide polymorphisms (SNPs) and mutations, at the genome level; at the epigenome level DNA methylation, histone modification and chromatin accessibility; gene expression and splicing at the transcriptome level; protein expression and post-translational modification at the proteome level; at the metabolome level the metabolic profile. Each variation in each level can be assessed with different techniques either by a targeted approach or an untargeted approach. Adapted from Ritchie et al.[79]

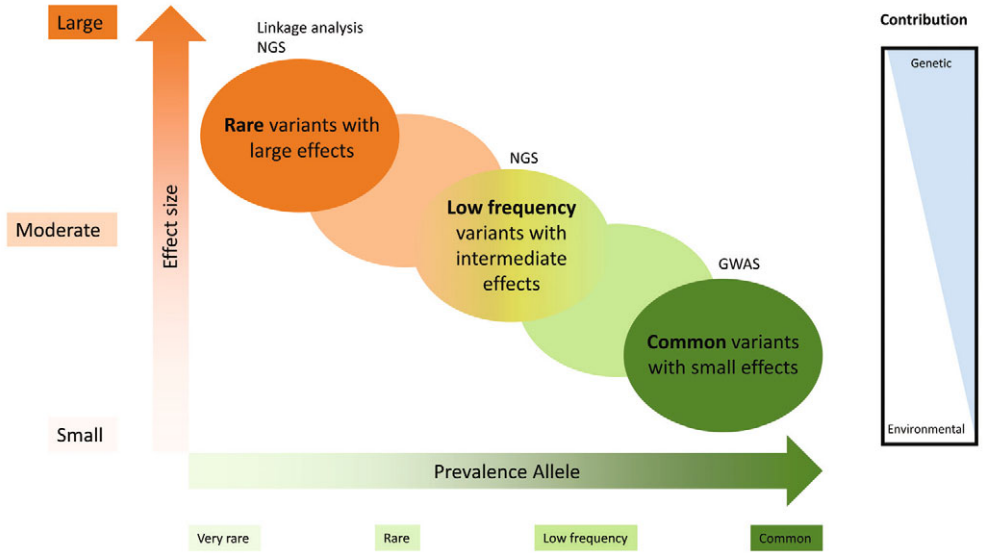
Rationale for genetic studies

When identifying genes involved in a disorder, different approaches are used depending on the disorder's architecture (i.e. monogenic, oligogenic or polygenic). The more oligogenic a disease is (i.e. the smaller the number of genes involved, with monogenic being the extreme), the larger the effect size of the associated gene variant(s) tends to be (**Figure 2**), in line with epidemiological data from disorders where rare disorders are monogenic and common disorders polygenic.

Hemiplegic migraine

Most of our knowledge of molecular mechanisms in migraine pathophysiology came from studying rare hemiplegic migraine (HM). The classical linkage method in migraine research was used to study large families with HM and this revealed a clear Mendelian (monogenic) type of inheritance. The approach led to the identification of three undisputed HM genes; *CACNA1A* (FHM1), *ATP1A2* (FHM2), and *SCN1A* (FHM3).[87-89]

Figure 2 Relationship between different types of hereditary (monogenic vs. polygenic) disorders.



Illustrating the relation to the allele frequency and the corresponding effect size as well as the contribution of genetic variants *vs.* environmental factors. Adapted from Manolio et al.[90]

In many patients with HM no pathogenic mutation has been detected in the HM genes.[91, 92] In recent years, whole-exome (next-generation) sequencing (WES) has been used to try and identify additional causal genes in patients without mutations in the known HM genes, but this has been proven difficult and no "fourth" gene has been identified thus far.[93] A study by Pelzer et al.[93] did,

however, show that patients with a more severe phenotype were more prone to have a causal mutation in one of the HM genes. Patients with a causal mutation in *CACNA1A*, *ATP1A2*, or *SCN1A* had a lower age-at-onset, more affected family members, and had attacks more frequently. Moreover, attacks were (i) brought about by mild head trauma, (ii) typically with extensive motor weakness, and (iii) with brainstem features, confusion, and brain oedema. Noteworthy, progressive ataxia and intellectual disability were only found in patients with a causal gene mutation.[93] As no mutation was found in "milder" patients, it was proposed that such HM patients may have the more extreme phenotype in the migraine with aura continuum.[91] Illustrative of this is a Finnish polygenic risk score study that showed that patients with HM, but without a high-penetrant disease-causing mutation in a known HM gene, carry an excess of genome-wide association studies (GWAS) variants associated with common migraine compared to patients suffering from the common migraine subtypes,[94] suggesting indeed a spectrum ranging from common low-risk variants to rare highly-penetrant mutations to contributing to the risk for migraine. Further support for this hypothesis are loss-of-function mutations in *PRRT2*, which do not cause HM on their own, but rather function as modifying genetic risk factors.[95] Illustrating the complex genetic architecture of HM is a recent whole-genome sequencing (WGS) where patients with HM were more likely to accumulate frameshift indels in multiple genes that have a role in synaptic signalling in the central nervous system compared to common migraine patients.[96]

Genetic studies in common migraine

Various twin and familial studies investigating the genetic and environmental susceptibility in migraine have shown that migraine is a multifactorial (complex) genetic disorder with a strong familial aggregation.[14, 15] The heritability of migraine was estimated to range from 35% to 60%.[97] Population-based studies have shown that the relative risk for a first-degree relative of a migraine patient is increased by 1.5- to 4-fold in comparison to a patient in the general population.[14] The risk was highest for those patients with a higher pain score and frequency of attacks, an early age of disease onset, and a migraine with aura phenotype.[14-16] Studies of twins identified a higher genetic load in migraine with aura compared with migraine without aura.[98] Migraine frequency, being the number of migraine days per month, appears mainly to be associated with a genetic predisposition in males.[16] A stronger family history of migraine is also associated with migraine with aura, a lower age-at-onset and more medication days.[16] For decades, identifying gene variants involved in complex disorders, such as migraine, has proven challenging.

Genetic studies in cluster headache

Twin and family studies have shown the involvement of genetic factors in cluster headache.[21] Notably, first-degree relatives have an increased relative risk between 5- and 18-fold, whereas second-degree relatives have a risk 1- to 3-fold higher than in the general population.[99] Thus far, most genetic studies have interrogated a limited number of variants in genes linked to presumed pathways in cluster headache.[100, 101] Variants in the *HCRTR2* gene were predominantly studied.

The *HCRTR2* gene encodes the G-protein coupled receptor hypocretin type 2 receptor that binds neuropeptides hypocretin-1 and -2 in the central nervous system. Such causal role of hypocretins makes sense as they have been implicated in sleep and arousal as well as pain modulation,[102] and levels were reported to be lower in CSF of patients with cluster headache.[103] However, initially positive genetic findings for *HCRTR2* associations[104-106] were not replicated in better-powered studies.[101, 107] Genes involved in circadian rhythmicity have also been investigated, but no association could be found.[108]

Genome-wide association studies

As a result of the improvement in DNA technology and the advancement of cost-effective genotyping platforms GWAS has become the method of choice to identify gene variants in complex traits in an untargeted approach in the last decade. Typically, in GWAS, several millions of single nucleotide polymorphisms (SNPs) are tested for association with a disorder by assessing differences in allele frequencies between large numbers of patients and controls. Of note, only common variants with a low to high minor allele frequency (≥0.01) are interrogated.

Since 2010, the International Headache Genetics Consortium (IHGC; www.headachegenetics.org/) has conducted several migraine GWAS, and with the increasing sample sizes, the number of associated gene variants steadily expanded. For cluster headache the first GWAS was performed in a very small, Italian study investigating patients with cluster headache.[109] They found a suggestive association with genetic variants in *ADCYAP1R1* and *MME*,[109] but the findings were not be replicated in a larger Swedish sample.[110] The hope is that larger GWAS will yield variants robustly associated with cluster headache.

Next-generation sequencing

A large part of the genetic variance and heritability in common diseases cannot be explained (usually referred to as "missing heritability") with a GWAS approach alone. One reason is that rarer variants (MAF<0.01), potentially with higher effect sizes, are not well interrogated by genotyping arrays typically utilised in the GWAS approach. Such mediate-effect-size variants can be identified using a next-generation sequencing (NGS) approach, i.e. by the simultaneous large-scale sequencing of the coding exons (whole-exome sequencing; WES) or the entire genome (whole-genome sequencing; WGS). In addition, the simultaneous sequencing of RNA transcripts ("transcriptome"; RNA-seq), either of bulk tissue or of its single nuclei can shed light on molecular mechanisms.

Only a few NGS studies have been performed in migraine thus far. Until now, WES was typically applied to cohorts of patients with HM, testing several hundred cases in an attempt to either find causal mutations in known HM genes or novel HM genes in patients that are negative for mutations in *CACNA1A*, *ATP1A2*, and *SCN1A*. Until now results have not led to additional

(undisputed) HM genes.[91-93,111,112] This may indicate that HM in mutation-negative patients may be oligogenic or polygenic, in line with the excess presence of common variants in such patients.[94] For cluster headache no gene sequencing studies have ever been performed which is logical as the gene array studies for this disease only just started.

An alternative approach to understanding molecular mechanisms involved in the pathophysiology of headache disorders is to study gene expression profiles. Contrary to genetic variation, gene expression is not fixed through life and expression is driven by both genetic and environmental factors.[113] Typically, an RNA-seq approach (i.e. simultaneous sequencing of coding (messenger) and non-coding RNAs in a sample) is for instance used to identify differences in expression between individuals with and without disease or over the course of an attack. Various RNA-seq studies have been performed in migraine, but the results are not unambiguous not in the least because of potential caveats of using peripheral blood, the main source of biomaterial for such studies in the case of migraine.[114] Gene expression studies in cluster headache are scarce. One study suggested the involvement of several brain-related mechanisms (voltage-gated channels and GABA receptor function), mitochondria, inflammation and intracellular signalling cascades.[115] Another study found an indication for inflammatory activity in the active phase of the disease.[116]

Further genetic studies

GWASs have proven successful in identifying many dozens of low-effect risk DNA variants for the more common forms of migraine with the number of associated DNA variants increasing steadily with larger sample sizes. Currently, next-generation sequencing, utilising whole-exome and -genome sequencing data, and other -omics data are being used to facilitate their functional interpretation and the discovery of additional risk factors. Various methods and analysis tools, such as genetic correlation, polygenic risk scores (PRSs) and causality analysis, are used to further characterise genetic risk factors.

Downstream bioinformatics methods

One way of making better use of the large number of small effect variants identified in migraine GWAS to have clinical benefit is the calculation of PRSs. A PRS is the combined effect of many common risk variants of genetic load for the discovery trait that can be used to estimate risk for a certain trait/phenotype in individuals in a target sample.[117] This is done by testing whether a higher PRS based on the discovery sample is associated with case status or a specific trait in the target sample via regression models. A PRS provides a promising possibility to investigate the shared genetic architecture between migraine with known and hitherto unknown co-morbidities or traits. The aggregation of migraine in families and the earlier age of onset of migraine can to some extent be contributed to common polygenic variations, where the PRS explained a larger part of the phenotype variance in familial cases, especially those with migraine with aura and hemiplegic migraine compared to population cases.[94]

Another way in which GWAS data can be used is by investigating the genetic relationships between traits, one of these analysis is Mendelian randomisation (MR). MR is able to entangle the pleiotropy that exists across many traits. In an MR analysis, genetic variants associated with an exposure are identified and regressed upon an outcome measurement to infer causality (*i.e.*, direction) of the association. Given the random assortment of alleles at gametogenesis in early life, this method is less likely to suffer from issues of confounding and reverse causation than methods used in conventional observational epidemiological studies.[118] For a successful MR analysis, three assumptions need to be fulfilled.[119] (I) Variants used as instrumental variables (IVs) need to be associated with the exposure. (II) The IVs only affect the outcome through the exposure, not through any other causal pathway. Factors that may lead to violation of this assumption include population stratification, LD and horizontal pleiotropy, the latter means that there is an (in)direct independent association of the IV (or another SNP in LD with the IV) with another trait that is not in the causal pathway of the investigated relation. (III) The IVs must not be associated with confounders. In one-directional MR the possible causal relation between trait X on trait Y is investigated, in bidirectional MR studies the directional effect from trait Y on trait X is also investigated.

Outline of this thesis

The research conducted for this thesis is divided in two parts. **Part 1** of the thesis focuses on biochemical studies in migraine. Here the biochemistry of migraine is investigated in: (i) a targeted approach, focusing on one or more related selected pathways of interest or, (ii) an untargeted approach aiming to simultaneously measure as many metabolites as possible from a biological sample. **Part 2** of the thesis focusses on genetical studies in both migraine and cluster headache using next-generation sequencing data and array genotyping data.

Part I Biochemistry of migraine

In **Chapter 2** we investigated whether the overall metabolic profile in blood of patients with migraine differed from those without migraine. Close to 100 metabolites were measured with ^{1}H-NMR spectroscopy in blood serum of 289 individuals with migraine and 1,360 individuals without migraine, all derived from a genetic isolate in the South-West of the Netherlands. **Chapter 3** describes whether CSF levels of amines, measured using an untargeted approach correlate with blood plasma levels in healthy volunteers. The study was then extrapolated to migraine patients. This chapter illustrates to what extent amine levels of CSF and blood relate to each other and seems to emphasize the role of blood-brain-barrier transport. **Chapter 4** investigates whether the endocannabinoid system is disrupted in interictal patients with migraine. To this end, the levels of three endocannabinoids in CSF were investigated in interictal (e.g. outside an attack) individuals with migraine with aura (n = 97) and without aura (n = 97) compared to healthy volunteers (n =

94). Endocannabinoids were measured using a previously validated micro-liquid chromatography-tandem mass spectrometry (micro-LC-MS/MS) technique. In **Chapter 5** the role of PGE_2 in the (early) phase of an induced migraine attack was investigated. To this end, PGE_2 plasma levels were measured towards in the (pre)ictal state of a glyceryl trinitrate (GTN) provoked attack in women with and without migraine.

Part II Genetics of different headache forms

In **Chapter 6** a GWAS in cluster headache is described. This study aims to demonstrate whether there are robust genetic associations for cluster headache. The study investigated 840 Dutch patients and a replication was performed in 144 Norwegian patients. In **Chapter 7** a meta-analysis of multiple GWAS studies of cluster headache with patients from Norway, Sweden, UK Germany, Denmark, Greece, Spain, Italy and the Netherlands (in total 4,043 patients) is conducted to not only confirm previous risk loci but also identify new disease risk loci. Using a Mendelian randomization approach it is investigated whether the intensity of cigarette smoking has a causal effect on cluster headache. **Chapter 8** describes a meta-analysis of multiple GWAS studies in migraine (102,084 migraine cases and 771,257 controls). Specific risk loci for migraine subtypes are investigated in a clinical sample. The aim of **Chapter 9** is to investigate whether there is an increased burden in hemiplegic migraine of missense variants in *CACNA1X* genes in patients without a high-penetrant disease-causing mutation in one of the well-known hemiplegic migraine genes (*CACNA1A*, *ATP1A2*, and *SCN1A*). The study illustrates the genetic complexity of hemiplegic migraine and the possibility of a spectrum ranging from high-risk rare mutations to low-risk common variants contributing to the risk for all forms of migraine.

Finally, **Chapter 10** provides a general discussion of the thesis together with suggestions for future research.

References

1. Headache Classification Committee of the International Headache Society (IHS) The International Classification of Headache Disorders, 3rd edition. *Cephalalgia*. 2018;38(1):1-211.
2. Ferrari MD, Klever RR, Terwindt GM, Ayata C, van den Maagdenberg AM. Migraine pathophysiology: lessons from mouse models and human genetics. *Lancet Neurol*. 2015;14(1):65-80.
3. May A, Schwedt TJ, Magis D, et al. Cluster headache. *Nat Rev Dis Primers*. 2018;4:18006.
4. Ashina M, Terwindt GM, Al-Karagholi MA, et al. Migraine: disease characterisation, biomarkers, and precision medicine. *Lancet*. 2021;397(10283):1496-1504.
5. Ayata C. Cortical spreading depression triggers migraine attack: pro. *Headache*. 2010;50(4):725-730.
6. Charles A. Does cortical spreading depression initiate a migraine attack? Maybe not. *Headache*. 2010;50(4):731-733.
7. Eikermann-Haerter K, Ayata C. Cortical spreading depression and migraine. *Curr Neurol Neurosci Rep*. 2010;10(3):167-173.
8. Pietrobon D, Striessnig J. Neurobiology of migraine. *Nat Rev Neurosci*. 2003;4(5):386-398.
9. Headache Classification Committee of the International Headache Society (IHS). The International Classification of Headache Disorders, 3rd edition. *Cephalalgia*. 2018;38(1):1-211.
10. Dodick DW. A Phase-by-Phase Review of Migraine Pathophysiology. *Headache*. 2018;58 Suppl 1:4-16.
11. Launer LJ, Terwindt GM, Ferrari MD. The prevalence and characteristics of migraine in a population-based cohort: the GEM study. *Neurology*. 1999;53(3):537-542.
12. Louter MA, Pelzer N, de Boer I, et al. Prevalence of lifetime depression in a large hemiplegic migraine cohort. *Neurology*. 2016;87(22):2370-2374.
13. Minen MT, Begasse De Dhaem O, Kroon Van Diest A, et al. Migraine and its psychiatric comorbidities. *J Neurol Neurosurg Psychiatry*. 2016;87(7):741-749.
14. Russell MB, Olesen J. Increased familial risk and evidence of genetic factor in migraine. *BMJ*. 1995;311(7004):541-544.
15. Stewart WF, Bigal ME, Kolodner K, et al. Familial risk of migraine: variation by proband age at onset and headache severity. *Neurology*. 2006;66(3):344-348.
16. Pelzer N, Louter MA, van Zwet EW, et al. Linking migraine frequency with family history of migraine. *Cephalalgia*. 2018:333102418783295.
17. Nesbitt AD, Goadsby PJ. Cluster headache. *BMJ*. 2012;344:e2407.
18. Frederiksen HH, Lund NL, Barloese MC, Petersen AS, Jensen RH. Diagnostic delay of cluster headache: A cohort study from the Danish Cluster Headache Survey. *Cephalalgia*. 2020;40(1):49-56.
19. Fischera M, Marziniak M, Gralow I, Evers S. The incidence and prevalence of cluster headache: a meta-analysis of population-based studies. *Cephalalgia*. 2008;28(6):614-618.
20. Lund N, Petersen A, Snoer A, Jensen RH, Barloese M. Cluster headache is associated with unhealthy lifestyle and lifestyle-related comorbid diseases: Results from the Danish Cluster Headache Survey. *Cephalalgia*. 2019;39(2):254-263.
21. Sjaastad O, Shen JM, Stovner LJ, Elsas T. Cluster headache in identical twins. *Headache*. 1993;33(4):214-217.
22. Lauritzen M. Pathophysiology of the migraine aura. The spreading depression theory. *Brain*. 1994;117 (Pt 1):199-210.
23. Leao aap. spreading depression of activity in the cerebral cortex. *J Neurophysiol*. 1944;7(6):359-390.
24. Ferrari MDP, Klever RRM, Terwindt GMMD, Ayata CMD, van den Maagdenberg AMJMP. Migraine pathophysiology: lessons from mouse models and human genetics. *Lancet Neurology, The*. 2015;14(1):65-80.

25. Goadsby PJ, Holland PR, Martins-Oliveira M, et al. Pathophysiology of Migraine: A Disorder of Sensory Processing. *Physiol Rev*. 2017;97(2):553-622.
26. Zhang X, Levy D, Kainz V, et al. Activation of central trigeminovascular neurons by cortical spreading depression. *Ann Neurol*. 2011;69(5):855-865.
27. Hadjikhani N, Sanchez Del Rio M, Wu O, et al. Mechanisms of migraine aura revealed by functional MRI in human visual cortex. *Proc Natl Acad Sci U S A*. 2001;98(8):4687-4692.
28. Charles AC, Baca SM. Cortical spreading depression and migraine. *Nat Rev Neurol*. 2013;9(11):637-644.
29. Bernstein C, Burstein R. Sensitization of the trigeminovascular pathway: perspective and implications to migraine pathophysiology. *J Clin Neurol*. 2012;8(2):89-99.
30. Pietrobon D. Ion channels in migraine disorders. *Current Opinion in Physiology*. 2018;2:98-108.
31. Davis RJ, Murdoch CE, Ali M, et al. EP4 prostanoid receptor-mediated vasodilatation of human middle cerebral arteries. *Br J Pharmacol*. 2004;141(4):580-585.
32. Yokota C, Inoue H, Kuge Y, et al. Cyclooxygenase-2 expression associated with spreading depression in a primate model. *J Cereb Blood Flow Metab*. 2003;23(4):395-398.
33. Maniyar FH, Sprenger T, Monteith T, Schankin C, Goadsby PJ. Brain activations in the premonitory phase of nitroglycerin-triggered migraine attacks. *Brain*. 2014;137(Pt 1):232-241.
34. Hoffmann J, May A. Diagnosis, pathophysiology, and management of cluster headache. *The Lancet Neurology*. 2018;17(1):75-83.
35. Wei DY, Goadsby PJ. Cluster headache pathophysiology - insights from current and emerging treatments. *Nat Rev Neurol*. 2021;17(5):308-324.
36. May A, Goadsby PJ. The trigeminovascular system in humans: pathophysiologic implications for primary headache syndromes of the neural influences on the cerebral circulation. *J Cereb Blood Flow Metab*. 1999;19(2):115-127.
37. Saper CB, Scammell TE, Lu J. Hypothalamic regulation of sleep and circadian rhythms. *Nature*. 2005;437(7063):1257-1263.
38. May A, Bahra A, Büchel C, Frackowiak RS, Goadsby PJ. Hypothalamic activation in cluster headache attacks. *Lancet*. 1998;352(9124):275-278.
39. Chan D, Ng LL. Biomarkers in acute myocardial infarction. *BMC Med*. 2010;8:34.
40. Molina R, Escudero JM, Augé JM, et al. HE4 a novel tumour marker for ovarian cancer: comparison with CA 125 and ROMA algorithm in patients with gynaecological diseases. *Tumour Biol*. 2011;32(6):1087-1095.
41. Høgdall C, Fung ET, Christensen IJ, et al. A novel proteomic biomarker panel as a diagnostic tool for patients with ovarian cancer. *Gynecol Oncol*. 2011;123(2):308-313.
42. Olesen J, Iversen HK, Thomsen LL. Nitric oxide supersensitivity: a possible molecular mechanism of migraine pain. *Neuroreport*. 1993;4(8):1027-1030.
43. Thomsen LL, Kruuse C, Iversen HK, Olesen J. A nitric oxide donor (nitroglycerin) triggers genuine migraine attacks. *Eur J Neurol*. 1994;1(1):73-80.
44. Antonova M, Wienecke T, Olesen J, Ashina M. Prostaglandin E(2) induces immediate migraine-like attack in migraine patients without aura. *Cephalalgia*. 2012;32(11):822-833.
45. Wienecke T, Olesen J, Ashina M. Prostaglandin I2 (epoprostenol) triggers migraine-like attacks in migraineurs. *Cephalalgia*. 2010;30(2):179-190.
46. Ashina M, Hansen JM, Olesen J. Pearls and pitfalls in human pharmacological models of migraine: 30 years' experience. *Cephalalgia*. 2013;33(8):540-553.
47. Hansen JM, Hauge AW, Olesen J, Ashina M. Calcitonin gene-related peptide triggers migraine-like attacks in patients with migraine with aura. *Cephalalgia*. 2010;30(10):1179-1186.
48. Bogucki A. Studies on nitroglycerin and histamine provoked cluster headache attacks. *Cephalalgia*. 1990;10(2):71-75.

49. Ferrari MD, Odink J, Tapparelli C, et al. Serotonin metabolism in migraine. *Neurology*. 1989;39(9):1239-1242.
50. Onderwater GLJ, van Dongen RM, Harms AC, et al. Cerebrospinal fluid and plasma amine profiles in interictal migraine. *Ann Neurol*. 2022; 93(4):715-728
51. Ashina M, Bendtsen L, Jensen R, Schifter S, Olesen J. Evidence for increased plasma levels of calcitonin gene-related peptide in migraine outside of attacks. *Pain*. 2000;86(1-2):133-138.
52. Cernuda-Morollon E, Larrosa D, Ramon C, et al. Interictal increase of CGRP levels in peripheral blood as a biomarker for chronic migraine. *Neurology*. 2013;81(14):1191-1196.
53. Fan PC, Kuo PH, Chang SH, et al. Plasma calcitonin gene-related peptide in diagnosing and predicting paediatric migraine. *Cephalalgia*. 2009;29(8):883-890.
54. Fekrazad R, Sardarian A, Azma K, et al. Interictal levels of calcitonin gene related peptide in gingival crevicular fluid of chronic migraine patients. *Neurol Sci*. 2018;39(7):1217-1223.
55. Fusayasu E, Kowa H, Takeshima T, Nakaso K, Nakashima K. Increased plasma substance P and CGRP levels, and high ACE activity in migraineurs during headache-free periods. *Pain*. 2007;128(3):209-214.
56. Jang MU, Park JW, Kho HS, Chung SC, Chung JW. Plasma and saliva levels of nerve growth factor and neuropeptides in chronic migraine patients. *Oral Dis*. 2011;17(2):187-193.
57. Rodriguez-Osorio X, Sobrino T, Brea D, et al. Endothelial progenitor cells: a new key for endothelial dysfunction in migraine. *Neurology*. 2012;79(5):474-479.
58. Gallai V, Sarchielli P, Floridi A, et al. Vasoactive peptide levels in the plasma of young migraine patients with and without aura assessed both interictally and ictally. *Cephalalgia*. 1995;15(5):384-390.
59. Goadsby PJ, Edvinsson L, Ekman R. Vasoactive peptide release in the extracerebral circulation of humans during migraine headache. *Ann Neurol*. 1990;28(2):183-187.
60. Juhasz G, Zsombok T, Modos EA, et al. NO-induced migraine attack: strong increase in plasma calcitonin gene-related peptide (CGRP) concentration and negative correlation with platelet serotonin release. *Pain*. 2003;106(3):461-470.
61. Sarchielli P, Alberti A, Codini M, Floridi A, Gallai V. Nitric oxide metabolites, prostaglandins and trigeminal vasoactive peptides in internal jugular vein blood during spontaneous migraine attacks. *Cephalalgia*. 2000;20(10):907-918.
62. Gupta R, Ahmed T, Banerjee B, Bhatia M. Plasma calcitonin gene-related peptide concentration is comparable to control group among migraineurs and tension type headache subjects during inter-ictal period. *J Headache Pain*. 2009;10(3):161-166.
63. Lee MJ, Lee SY, Cho S, Kang ES, Chung CS. Feasibility of serum CGRP measurement as a biomarker of chronic migraine: a critical reappraisal. *J Headache Pain*. 2018;19(1):53.
64. Oterino A, Toriello M, Palacio E, et al. Analysis of endothelial precursor cells in chronic migraine: a case-control study. *Cephalalgia*. 2013;33(4):236-244.
65. Tvedskov JF, Lipka K, Ashina M, et al. No increase of calcitonin gene-related peptide in jugular blood during migraine. *Ann Neurol*. 2005;58(4):561-568.
66. Diener H-C. CGRP antibodies for migraine prevention — new kids on the block. *Nature Reviews Neurology*. 2019;15(3):129-130.
67. 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;14(6):338-350.
68. Goadsby PJ, Edvinsson L. Human in vivo evidence for trigeminovascular activation in cluster headache. Neuropeptide changes and effects of acute attacks therapies. *Brain*. 1994;117 (Pt 3):427-434.
69. Snoer A, Vollesen ALH, Beske RP, et al. Calcitonin-gene related peptide and disease activity in cluster headache. *Cephalalgia*. 2019;39(5):575-584.

70. Fanciullacci M, Alessandri M, Figini M, Geppetti P, Michelacci S. Increase in plasma calcitonin gene-related peptide from the extracerebral circulation during nitroglycerin-induced cluster headache attack. *Pain*. 1995;60(2):119-123.

71. Ashina H, Schytz HW, Ashina M. CGRP in human models of primary headaches. *Cephalalgia*. 2016;38(2):353-360.

72. Messlinger K, Vogler B, Kuhn A, et al. CGRP measurements in human plasma - a methodological study. *Cephalalgia*. 2021;41(13):1359-1373.

73. Kuehnbaum NL, Britz-McKibbin P. New advances in separation science for metabolomics: resolving chemical diversity in a post-genomic era. *Chem Rev*. 2013;113(4):2437-2468.

74. Fiehn O, Robertson D, Griffin J, et al. The metabolomics standards initiative (MSI). *Metabolomics*. 2007;3(3):175-178.

75. Lehmann R. From bedside to bench-practical considerations to avoid pre-analytical pitfalls and assess sample quality for high-resolution metabolomics and lipidomics analyses of body fluids. *Anal Bioanal Chem*. 2021;413(22):5567-5585.

76. Carraro P, Zago T, Plebani M. Exploring the initial steps of the testing process: frequency and nature of pre-preanalytic errors. *Clin Chem*. 2012;58(3):638-642.

77. Lippi G, Guidi GC, Mattiuzzi C, Plebani M. Preanalytical variability: the dark side of the moon in laboratory testing. *Clin Chem Lab Med*. 2006;44(4):358-365.

78. Szecsi PB, Ødum L. Error tracking in a clinical biochemistry laboratory. *Clin Chem Lab Med*. 2009;47(10):1253-1257.

79. Ritchie MD, Holzinger ER, Li R, Pendergrass SA, Kim D. Methods of integrating data to uncover genotype–phenotype interactions. *Nature Reviews Genetics*. 2015;16(2):85-97.

80. Liu X, Hoene M, Wang X, et al. Serum or plasma, what is the difference? Investigations to facilitate the sample material selection decision making process for metabolomics studies and beyond. *Anal Chim Acta*. 2018;1037:293-300.

81. Noga MJ, Zielman R, van Dongen RM, et al. Strategies to assess and optimize stability of endogenous amines during cerebrospinal fluid sampling. *Metabolomics*. 2018;14(4):44.

82. Teunissen CE, Petzold A, Bennett JL, et al. A consensus protocol for the standardization of cerebrospinal fluid collection and biobanking. *Neurology*. 2009;73(22):1914-1922.

83. Bar N, Korem T, Weissbrod O, et al. A reference map of potential determinants for the human serum metabolome. *Nature*. 2020;588(7836):135-140.

84. Dallmann R, Viola AU, Tarokh L, Cajochen C, Brown SA. The human circadian metabolome. *Proc Natl Acad Sci U S A*. 2012;109(7):2625-2629.

85. Kasukawa T, Sugimoto M, Hida A, et al. Human blood metabolite timetable indicates internal body time. *Proc Natl Acad Sci U S A*. 2012;109(37):15036-15041.

86. Chua EC, Shui G, Lee IT, et al. Extensive diversity in circadian regulation of plasma lipids and evidence for different circadian metabolic phenotypes in humans. *Proc Natl Acad Sci U S A*. 2013;110(35):14468-14473.

87. Ophoff RA, Terwindt GM, Vergouwe MN, et al. Familial hemiplegic migraine and episodic ataxia type-2 are caused by mutations in the Ca2+ channel gene CACNL1A4. *Cell*. 1996;87(3):543-552.

88. Dichgans M, Freilinger T, Eckstein G, et al. Mutation in the neuronal voltage-gated sodium channel SCN1A in familial hemiplegic migraine. *Lancet*. 2005;366(9483):371-377.

89. De Fusco M, Marconi R, Silvestri L, et al. Haploinsufficiency of ATP1A2 encoding the Na+/K+ pump alpha2 subunit associated with familial hemiplegic migraine type 2. *Nat Genet*. 2003;33(2):192-196.

90. Manolio TA, Collins FS, Cox NJ, et al. Finding the missing heritability of complex diseases. *Nature*. 2009;461(7265):747-753.

91. Hiekkala ME, Vuola P, Artto V, et al. The contribution of CACNA1A, ATP1A2 and SCN1A mutations in hemiplegic migraine: A clinical and genetic study in Finnish migraine families. *Cephalalgia*. 2018;38(12):1849-1863.

92. Sutherland HG, Maksemous N, Albury CL, et al. Comprehensive Exonic Sequencing of Hemiplegic Migraine-Related Genes in a Cohort of Suspected Probands Identifies Known and Potential Pathogenic Variants. *Cells*. 2020;9(11) :2368.

93. Pelzer N, Haan J, Stam AH, et al. Clinical spectrum of hemiplegic migraine and chances of finding a pathogenic mutation. *Neurology*. 2018;90(7):e575-e582.

94. Gormley P, Kurki MI, Hiekkala ME, et al. Common Variant Burden Contributes to the Familial Aggregation of Migraine in 1,589 Families. *Neuron*. 2018;99(5):1098.

95. Pelzer N, de Vries B, Kamphorst JT, et al. PRRT2 and hemiplegic migraine: a complex association. *Neurology*. 2014;83(3):288-290.

96. Rasmussen AH, Olofsson I, Chalmer MA, Olesen J, Hansen TF. Higher burden of rare frameshift indels in genes related to synaptic transmission separate familial hemiplegic migraine from common types of migraine. *J Med Genet*. 2020;57(9):610-616.

97. Polderman TJ, Benyamin B, de Leeuw CA, et al. Meta-analysis of the heritability of human traits based on fifty years of twin studies. *Nat Genet*. 2015;47(7):702-709.

98. Ulrich V, Gervil M, Kyvik KO, Olesen J, Russell MB. Evidence of a genetic factor in migraine with aura: a population-based Danish twin study. *Ann Neurol*. 1999;45(2):242-246.

99. Russell MB. Epidemiology and genetics of cluster headache. *Lancet Neurol*. 2004;3(5):279-283.

100. May A, Bahra A, Buchel C, Frackowiak RS, Goadsby PJ. Hypothalamic activation in cluster headache attacks. *Lancet*. 1998;352(9124):275-278.

101. Gibson KF, Santos AD, Lund N, Jensen R, Stylianou IM. Genetics of cluster headache. *Cephalalgia*. 2019;39(10):1298-1312.

102. Razavi BM, Hosseinzadeh H. A review of the role of orexin system in pain modulation. *Biomed Pharmacother*. 2017;90:187-193.

103. Barloese M, Jennum P, Lund N, et al. Reduced CSF hypocretin-1 levels are associated with cluster headache. *Cephalalgia*. 2015;35(10):869-876.

104. Rainero I, Gallone S, Valfrè W, et al. A polymorphism of the hypocretin receptor 2 gene is associated with cluster headache. *Neurology*. 2004;63(7):1286-1288.

105. Fourier C, Ran C, Steinberg A, et al. Analysis of HCRTR2 Gene Variants and Cluster Headache in Sweden. *Headache*. 2019;59(3):410-417.

106. Schürks M, Kurth T, Geissler I, et al. Cluster headache is associated with the G1246A polymorphism in the hypocretin receptor 2 gene. *Neurology*. 2006;66(12):1917-1919.

107. Weller CM, Wilbrink LA, Houwing-Duistermaat JJ, et al. Cluster headache and the hypocretin receptor 2 reconsidered: a genetic association study and meta-analysis. *Cephalalgia*. 2015;35(9):741-747.

108. Ofte HK, Tronvik E, Alstadhaug KB. Lack of association between cluster headache and PER3 clock gene polymorphism. *J Headache Pain*. 2015;17:18.

109. Bacchelli E, Cainazzo MM, Cameli C, et al. A genome-wide analysis in cluster headache points to neprilysin and PACAP receptor gene variants. *J Headache Pain*. 2016;17(1):114.

110. Ran C, Fourier C, Michalska JM, et al. Screening of genetic variants in ADCYAP1R1, MME and 14q21 in a Swedish cluster headache cohort. *J Headache Pain*. 2017;18(1):88.

111. Ibrahim O, Sutherland HG, Maksemous N, et al. Exploring Neuronal Vulnerability to Head Trauma Using a Whole Exome Approach. *J Neurotrauma*. 2020;37(17):1870-1879.

112. Gil-Perotín S, Jaijo T, Verdú AG, et al. Epilepsy, status epilepticus, and hemiplegic migraine coexisting with a novel SLC4A4 mutation. *Neurol Sci*. 2021;42(9):3647-3654.

113. Bar-Joseph Z, Gitter A, Simon I. Studying and modelling dynamic biological processes using time-series gene expression data. *Nature reviews Genetics*. 2012;13(8):552-564.

114. Gerring Z, Rodriguez-Acevedo AJ, Powell JE, et al. Blood gene expression studies in migraine: Potential and caveats. *Cephalalgia*. 2016;36(7):669-678.

115. Eising E, Pelzer N, Vijfhuizen LS, et al. Identifying a gene expression signature of cluster headache in blood. *Sci Rep*. 2017;7:40218.

116. Sjöstrand C, Duvefelt K, Steinberg A, et al. Gene expression profiling in cluster headache: a pilot microarray study. *Headache*. 2006;46(10):1518-1534.

117. Purcell SM, Wray NR, Stone JL, et al. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature*. 2009;460(7256):748-752.

118. Zheng J, Baird D, Borges MC, et al. Recent Developments in Mendelian Randomization Studies. *Current epidemiology reports*. 2017;4(4):330-345.

119. Haycock PC, Burgess S, Wade KH, et al. Best (but oft-forgotten) practices: the design, analysis, and interpretation of Mendelian randomization studies. *Am J Clin Nutr*. 2016;103(4):965-978.

PART I

Biochemistry of migraine

2

Metabolic profile changes in serum of migraine patients detected using ^{1}H-NMR spectroscopy

Aster V.E. Harder, Lisanne S. Vijfhuizen, Peter Henneman, Ko Willems van Dijk, Cornelia M. van Duijn, Gisela M. Terwindt and Arn M.J.M. van den Maagdenberg

J Headache Pain 2021;22:142

Abstract

Background: Migraine is a common brain disorder but reliable diagnostic biomarkers in blood are still lacking. Our aim was to identify, using proton nuclear magnetic resonance (^{1}H-NMR) spectroscopy, metabolites in serum that are associated with lifetime and active migraine by comparing metabolic profiles of patients and controls.

Methods: Fasting serum samples from 313 migraine patients and 1,512 controls from the Erasmus Rucphen Family (ERF) study were available for ^{1}H-NMR spectroscopy. Data was analysed using elastic net regression analysis.

Results: A total of 100 signals representing 49 different metabolites were detected in 289 cases (of which 150 active migraine patients) and 1,360 controls. We were able to identify profiles consisting of 6 metabolites predictive for lifetime migraine status and 22 metabolites predictive for active migraine status. We estimated with subsequent regression models that after correction for age, sex, BMI and smoking, the association with the metabolite profile in active migraine remained. Several of the metabolites in this profile are involved in lipid, glucose and amino acid metabolism.

Conclusion: This study indicates that metabolic profiles, based on serum concentrations of several metabolites, including lipids, amino acids and metabolites of glucose metabolism, can distinguish active migraine patients from controls.

Introduction

Migraine is a common multifactorial brain disorder with a lifetime prevalence of 15-20%, causing disability worldwide and a three times higher prevalence in woman compared to men.[1,2] Migraine is characterized by recurrent episodes of severe often unilateral pulsating headache accompanied by nausea, vomiting and/or photo- and phonophobia lasting for 4-72 hours.[3] Although much progress has been made with unravelling its (non)genetic disease mechanisms,[4] a diagnosis of migraine is still made by interview and physical examination or questionnaire, as no diagnostic biomarker is available. The lack of biomarkers, for instance in a biofluid such as blood, has also hampered the development of novel treatments.

Metabolomics is an established valuable approach for biomarker identification and has been successful in revealing the metabolic underpinnings of various human diseases.[5-10] Validated biomarkers can greatly improve diagnosis, prognosis and assessing effectivity of treatment of patients, as was already shown for several diseases other than migraine.[11,12] Various attempts have been made to identify reliable biomarkers (either clinical, genetic, radiological or biochemical) in migraine.[13-16] without much success, also not for biochemical studies in blood[17] or cerebrospinal fluid.[14] Especially the identification of metabolites in an easily accessible body fluid such as peripheral blood is urgently needed.[18] When using metabolomics either a targeted approach that typically focuses on one or more related selected pathways of interest or an untargeted approach that aims to simultaneously measure as many metabolites as possible from a biological sample, can be employed. Several biochemical studies in migraine in the past two decades explored the targeted approach by examining a limited number of compounds, such as amino acids,[19,20] inflammatory markers,[21-23] vasoactive neuropeptides,[24-26] and (cardio)vascular risk factors,[27-29] because of their presumed role in migraine pathophysiology. More recently, mainly because of the advent of novel treatment antagonizing calcitonin gene-related peptide (CGRP) or its receptor,[30,31] the field of biomarker research in peripheral blood regained interest,[32] with reports of promising possible peripheral biomarkers in migraine.[33,34]

To search for migraine metabolite profiles in serum we used an untargeted, hypothesis-free, approach and performed high-throughput proton nuclear magnetic resonance (^{1}H-NMR) spectroscopy. This method allows for a rapid, robust, simultaneous identification and quantification of a variety of metabolites in large numbers of samples.[35] Here we analysed metabolite profiles in serum samples of migraine patients and controls from the Erasmus Rucphen Family population, a large Dutch population-based family study from the Southwest of the Netherlands in which we previously had identified migraine cases.[36] We set out to investigate whether metabolites identified by ^{1}H-NMR spectroscopy are associated with migraine by comparing metabolic profiles of migraine patients and controls in a "real-life variation" cohort.

Methods

Study population

The study included participants from the Erasmus Rucphen Family (ERF) study [37, 38]. This study population is based on a genetically isolated community in the Southwest of the Netherlands. In brief, the ERF study population includes 3,465 living descendants of 22 couples that had at least six children baptized in the community church between 1850 and 1900. Hence, study participants were all members of a large extended pedigree and all of European ancestry. All individuals 18 years and older were invited to participate.

Migraine diagnoses

Migraine was diagnosed using a validated three-stage screening procedure,[2] based on International Classification of Headache Disorder formerly ICHD-II, now ICHD-III criteria.[3, 39] Details on the migraine case-finding procedure have been published previously.[36] In short, first, participants filled out a five-item screening questionnaire on headache and aura symptoms. Next, screen-positives completed an additional detailed questionnaire on headache and aura symptoms. Finally, the diagnosis was validated with a telephone interview by a physician trained in headache disorders. Probable migraine patients were excluded. ERF participants who were negative for severe headache and/or migraine based on the aforementioned three-stage screening procedure were included as controls.[2, 36] Samples from participants were collected after overnight fasting.

^{1}H-NMR spectroscopy metabolite profiling: data processing and quality control

Venous blood samples had been drawn by venipuncture from the median cubital vein from participants of the ERF study after at least 8 h fasting period. Samples were centrifuged at 1,000-2,000 x g for 10 minutes at 4°C and serum was aliquoted in cryovials and stored at -80°C until further use. The ^{1}H-NMR data were generated as part of a larger project and described by Vaarhorst et al.[40] All ^{1}H-NMR spectroscopy experiments had been acquired on a 600 MHz Bruker Avance II spectrometer (Bruker) equipped with a 5-mm triple resonance inverse (TCI) cryogenic probe head with Z-gradient system and automatic tuning and matching. All experiments were recorded at 310K. Temperature calibration was done prior to each batch of measurements using the method of Findeisen et al.[41] The duration of the π/2 pulses were automatically calibrated for each individual sample using a homonuclear-gated nutation experiment on the locked and shimmed samples after automatic tuning and matching of the probe head.[42]

Then, stored samples were thawed at 4°C and mixed by inverting the containers ten times. Samples (300 µL) were mixed with 300 µL 75 mM disodium phosphate buffer in H_2O/D_2O (80/20) (pH 7.4), containing 6.15 mM NaN_3 and 4.64 mM sodium 3-[trimethylsilyl] d4-propionate (TSP), using a Gilson 215 liquid handler in combination with a Bruker SampleTrack system (Bruker, Karlsruhe, Germany). Samples were transferred into 5-mm SampleJet NMR tubes (Bruker) in 96-tube racks using a modified Gilson 215 tube filling station (Gilson, Middleton, WI, USA) and kept at 6°C on a SampleJet sample changer (Bruker) while queued for acquisition.

For water suppression pre-saturation of the water resonance with an effective field of γB_1 = 25 Hz was applied during the relaxation delay.[43] J-resolved spectra (JRES)[44] were recorded with a relaxation delay of 2 s and a total of one scan for each increment in the indirect dimension. A data matrix of 40 x 12,288 data points was collected covering a sweep width of 78 x 10,000 Hz. A sine-shaped window function was applied and the data was zero-filled to 256 x 16,384 complex data points prior to Fourier transformation. The resulting data matrix was tilted along the rows by shifting each row (k) by 0.4992*(128-k) points and symmetrised about the central horizontal lines to compensate for the skew of the multiplets in the F1 dimension. For T2-filtered ^{1}H-NMR spectra, a standard 1D Carr-Purcell-Meiboom-Gill (CPMG) pulse sequence[45,46] was used with a relaxation delay of 4 s. A pulse train of 130 refocusing pulses with individual spin echo delays of 0.6 ms were applied resulting in a total T2 filtering delay of 78 ms. A total of 73,728 data points covering a spectral width of 12,019 Hz were collected using 16 scans. The Free Induction Delay (FID) was zero-filled to 131,072 complex data points and an exponential window function was applied with a line broadening factor of 1.0 Hz prior to Fourier transformation. The spectra were automatically phase and baseline corrected.

Quality control, scaling and calibration of the NMR spectra

Further data processing was performed in Matlab® (R2009a; The Mathworks Inc., Natick, MA, USA) and described in Vaarhorst et al.[40] In brief, the spectra and associated data were converted into Matlab files using in-house code. First, the spectra were combined into one file while removing superfluous information. For CPMG this included dropping the imaginary part of the spectrum, while for the JRES spectra the sum projection along the indirect dimension was taken. Quality control (QC) on the set of ^{1}H-NMR spectra was carried out by examining a set of spectroscopic parameters such as shim values and intensity of the water signal, and subsequently visually inspecting the spectra. Spectra that failed the quality control were not included for further analysis. The remaining spectra were scaled with respect to the sensitivity of the receiver coil. This sensitivity is inversely proportional to the pulse length, which is dependent on the tuning of the RF coil. After subtracting a constant value as a simple baseline correction, the spectra were calibrated with respect to the anomeric resonance of α-D-glucose (δ = 5.23 ppm).[47] Since there are small deviations of the signal position in the different ^{1}H-NMR spectra, alignment was performed using the correlation optimized warping algorithm by Tomasi et al.[48] This was performed actively for the CPMG spectra, after which the same warping was applied to the JRES projection. The peaks in the JRES projection were automatically deconvoluted by fitting the spectra with mixed Gauss-Lorentz line shapes using the Simplex method. As the fitting algorithm incidentally converges to a local minimum, values further from the median than three times the interquartile range were discarded. Using partial least square regression, the remaining signal intensities were used to build a linear model that predicts the intensities directly from the non-warped spectrum, yielding also reasonable values for the cases where the deconvolution or warping algorithms failed.

Finally, metabolites were assigned using information from the Human Metabolome Database (HMDB) and the Pearson correlation coefficients between the peak intensities.[49]

Statistical analyses and data processing

Student's *t*-test and Chi-square tests were used to compare demographic characteristics between cases and controls. Raw ^{1}H-NMR signal data were processed as follows. Values below [mean - 4 * SD] and above [mean + 4 * SD] were filtered out. Then normality was assessed and data were $\log_{10}$-transformed when necessary, using SPSS software version 20.0 (SPSS Inc., IBM, Armonk, NY, USA). Signal data was adjusted for kinship by linear regression in GenABEL version 1.7-0, using R version 2.14.2 (R Foundation for Statistical Computing, Vienna, Austria).[50] Finally, the residuals from this linear regression model were transformed into Z-scores to approximate normality using SPSS software version 25.0 (SPSS Inc., IBM, Armonk, NY, USA). To reduce the dimensionality of the data and due to possible correlations between the parameters, elastic net regression was used to select a subset of the most informative signals for: (1) lifetime migraine diagnosis, and (2) a diagnosis of active migraine (defined as having at least one severe migraine in the last 12 months). Of note, patients likely had many attacks in the last year as is typical in migraine patients when they still have migraines, but data are lacking to assess how many attacks they had and when the last attack was before blood withdrawal nor do we know whether they were on medication. Hence we consider our migraine cases a sample with "real-life variation" with respect to attack frequency and severity. The R package glmnet was used with alpha set to 0.5 and 50-fold cross-validation using R software version 3.6.1.[51] In this cross-validation step we validated the selection of the signals by performing our regression analysis on 50 randomly chosen samples of our study population. Elastic net regression reduces variance and error and increases bias and the predictive power, which leads to better long-term prediction. However, the inferential capability decreases, which makes interpretation difficult as there are no uncertainties in terms of confidence intervals or hypothesis testing.

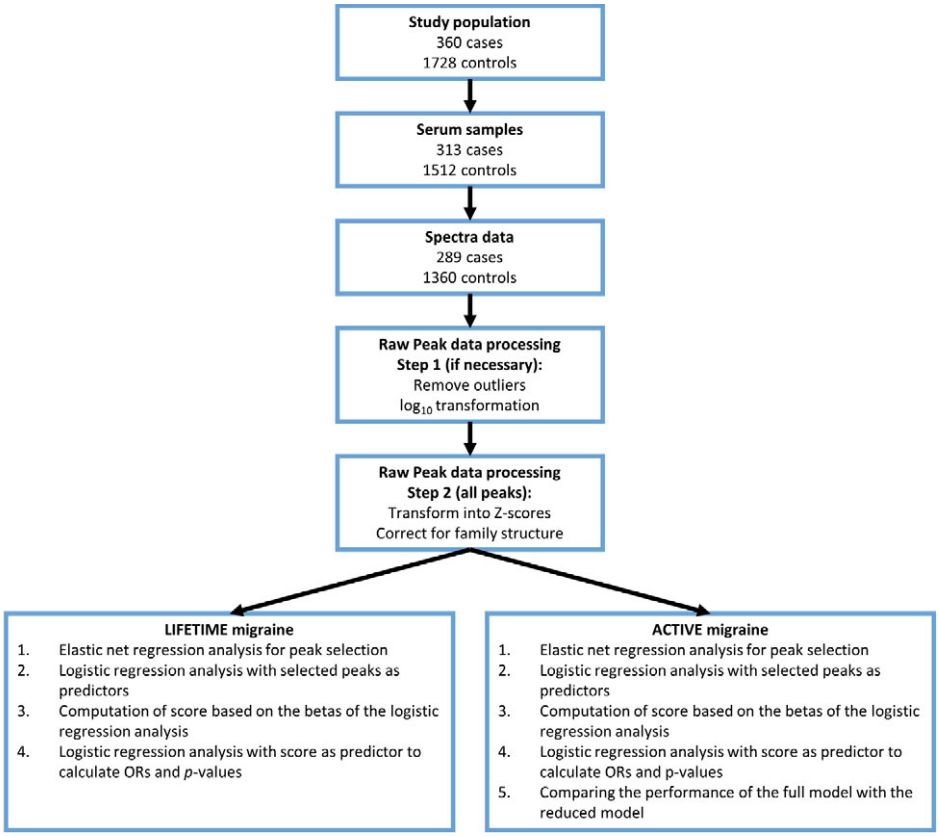
In an attempt to interpret our findings, we performed subsequent regression models. Because we had to perform the regression models within the unique cohort the exact *p*-values of these models are no longer valid, although the results may provide at least some information whether metabolites may be involved. For the regression models we entered the metabolites of the metabolic profiles in a logistic regression model to determine the weights for each signal for this population. The linear predictor of the logistic regression model was used as a "weighted metabolite score" (sum of regression coefficients multiplied by the corresponding covariate values). This score was used in a second logistic regression analysis to calculate odds ratios (ORs), *p*-values and the proportion of explained variance. To determine whether we had to correct our logistic regression model we independently assessed the influence of sex, age, body mass index (BMI) and smoking status on the "weighted metabolite score", by visually inspecting stratification plots and performing a linear model, where the "weighted metabolite score" was modelled as a function of migraine status. We included age, sex, BMI and current smoking status as covariates in the logistic regression model. To validate the findings from the previous analysis we performed analysis of variance (ANOVA) in which we compared the performance of the full model with the identified scores for migraine with the performance of a model containing only information on age, sex, BMI and smoking.

Results

Study population

We conducted a case-control study with in the ERF population cohort and included 2,088 participants in the study of which 360 were lifetime migraine patients and 1,728 without severe headache served as controls. Eight-hour fasting serum samples were available from 313 migraine patients and 1,512 controls that were used for ^{1}H-NMR spectroscopy profiling (see **Figure 1**).

Figure 1 Flow chart of the patient flow and analysis steps



Signal detection, assignment and processing

A total of 100 metabolite signals were detected in the JRES projection and quantified in the ^{1}H-NMR spectra. For 82 signals, metabolites could be assigned. These 82 signals represented 49 different metabolites (See Additional file 1 for signal assignment). The other 18 signals could not be annotated. In total, good-quality ^{1}H-NMR spectra were obtained from 289 migraine patients and 1,360 controls. For 19 signals (out of 100) outliers were removed and nine signals were $\log_{10}$-transformed (See Additional file 1). The remaining data points of the 100 signals were used for the association analyses with migraine.

Demographic characteristics

The demographic characteristics of the study population of whom good ^{1}H-NMR data were obtained are shown in **Table 1**. Migraine patients tended to be younger (p = 0.013) and more often were female than controls (p < 0.001). In addition, lifetime migraine patients more often than controls were smokers (44.6% cases vs. 35.1% controls p = 0.006). No difference in BMI was observed between cases and controls (p = 0.934). Of the 289 lifetime migraine patients, 150 (52%) reported at least one severe migraine attack in the 12 months preceding the interview and were assigned to the group of "active migraine patients". The active migraine patients consisted of 124 women (83%), had a mean age of 44 (SD ± 11.4), 71 (47.3%) were currently smoking and had a mean BMI of 26 (SD ± 4.9). Next, we assessed the influence of age, sex, BMI and smoking (see Additional file 2) on the weighted metabolite score. All covariates showed to be of influence on the weighted metabolite score and were added to the logistic regression model.

Association of metabolites with lifetime migraine diagnosis

Elastic net regression analysis of all 289 migraine patients and 1,360 controls for all 100 signals identified six ^{1}H-NMR signals as the best prediction subset. These signals were representative of four different metabolites (isoleucine, methionine, 1,5-anhydrosorbitol and creatine) and one unknown signal (**Table 2**). Subsequent logistic regression analysis showed support for association (odds ratio (OR) = 2.72; 95% confidence interval (CI) 1.97 - 3.75; $p = 1.28 \times 10^{-9}$) explaining 3.9% of the variance in migraine status (Nagelkerke R^2). After correction for age, sex, BMI and smoking the association no longer showed support for association (OR = 1.49; 95% CI 0.99 - 2.23; p = 0.051).

Table 1 Demographic characteristics

Variable	Lifetime migraine patients (N = 289)	Controls (N = 1360)	p-value	Active migraine patients[b] (N = 150)	p-value
Age (years)	46.5 ± 12.1	48.7 ± 14.5	0.013[c*]	44.0 ± 11.4	< 0.001[c*]
Female sex (%)	220 (76.1)	673 (49.5)	< 0.001[d*]	124 (82.7)	< 0.001[d*]
BMI	26.9 ± 5.0	26.8 ± 4.6	0.934[c]	26.3 ± 4.9	0.219[c]
Smoking[a] (yes) (%)	129 (44.9)	481 (36.0)	0.006[d*]	71 (47.3)	0.008[d*]
MO patients	163 (56.4)	-	-	77 (51.3)	-

Values are expressed as absolute values and percentage or mean ± SD. Numbers and proportions may not add up to total of 100 due to rounding or missing values. [a]Defined as currently cigarette smoking. [b]Defined as having at least one severe migraine attack in the last 12 months. [c]Student's t-test. [d]Chi-square Test. *Significant p-values (p < 0.05). Missing values in lifetime migraine patients for BMI (n = 2), smoking status (n = 2) and in controls for BMI (n = 24) smoking status (n = 27). MO = migraine without aura, BMI = body mass index

Association of metabolites with active migraine diagnosis

Next, we performed an elastic net regression on all 150 active migraine patients and 1,360 controls for all 100 signals. This analysis identified 22 predictive signals. The subsequent logistic regression analysis was performed on 146 cases and 1,343 controls, as not all subjects had sufficient signal data

for all 22 signals. The regression showed support for association between the signal data and active migraine status (OR = 2.72; 95% CI 2.09 - 3.54; $p = 1.35 \times 10^{-13}$) explaining 8.5% of the variance, this association remained after correction for sex, age, BMI and current smoking status (OR = 1.84; 95% CI 1.34 - 2.53; $p = 1.64 \times 10^{-4}$) with a total explained variance of 12.3% (Nagelkerke R^2). Hosmer and Lemeshow shows a good fit of the final model ($p = 0.688$). The outcome of our ANOVA analysis ($p = 7.1 \times 10^{-5}$) added to the evidence for involvement of these metabolites in active migraine patients. The majority of the 22 signals have been annotated to metabolites, but four remained unknown (**Table 2**). The known metabolites that were relevant to distinguish metabolic profile of migraine patients from controls were cholesterol, isoleucine, leucine, lipids (CH2 and CH*2CH=CH), acetate, pyruvate, methionine, dimethylglycine, 1,5-anhydrosorbitol, valine, myoinositol, glucose, serine, creatinine, and proline. Our data suggests that there is a metabolic profile for active migraine that distinguishes them from controls even after correcting for age, sex, BMI and smoking status. Remarkable is that five of the six signals predictive for lifetime migraine status are also predictive for active migraine status (**Table 2**).

Table 2 ^{1}H-NMR signals associated with lifetime migraine patients and active migraine patients

Life time migraine patients		*Active migraine patients*	
Metabolite	**Chemical shift (ppm)**	**Metabolite**	**Chemical shift (ppm)**
Isoleucine	0.92847	Cholesterol	0.89006
Isoleucine	0.99919	Isoleucine	0.92847
Unknown	1.40660	Unknown	0.95118
Methionine	2.63742	Leucine	0.95702
1,5-Anhydrosorbitol	3.58832	Isoleucine	0.99919
Creatine	3.92001	Lipids (CH2)†	1.26482
		Unknown	1.40660
		Acetate	1.90859
		Lipids (CH*2CH=CH)†	2.22215
		Pyruvic acid	2.36196
		Methionine	2.63742
		Dimethylglycine	2.91618
		Unknown	3.35396
		1,5-Anhydrosorbitol	3.58832
		Valine	3.59782
		Myoinositol	3.62232
		Glucose	3.72103
		Serine	3.95567
		Creatinine	4.04386
		Proline	4.12106
		Unknown	4.50117
		Glucose	5.22921

Ppm: parts per million. †The term in parenthesis indicates the structural feature of the lipids measured by ^{1}H-NMR spectroscopy

Discussion

Here we investigated metabolites identified by ^{1}H-NMR spectroscopy in serum of migraine patients and controls to assess whether metabolic profiles can distinguish the two groups. We identified 22 metabolites that were predictive for active migraine and estimated that they would remain predictive after correction for age, sex, BMI and smoking status. Active migraine status was linked with metabolic profiles with more (22) metabolites, when compared with lifetime migraine (6), suggesting that active migraine patients may have a more disturbed metabolic profile compared to lifetime migraine patients, at least among the 100 measured metabolites.

Although, based on our study, it is not possible to directly interpret the *p*-values nor to make association on an individual metabolite level among the total 22 compounds associated with active migraine, it is remarkable that the majority of these 22 metabolites have been (indirectly) implicated in migraine before. In our study, we found metabolites involved in lipid metabolism namely; cholesterol, and two types of lipids (CH2 and CH*2CH=CH). A number of studies have previously implicated lipid metabolism in migraine for instance, epidemiologic studies have shown that obesity is a risk factor for migraine and that there is a comorbidity of cerebrovascular and cardiovascular disease and migraine.[52, 53] Some studies found an elevated total cholesterol, LDL-cholesterol, or triglycerides, and decreased levels of HDL-cholesterol in migraine,[54, 55] whereas several other studies found no significant differences in lipid profiles.[54, 56] A recent meta-analysis encompassing 2,800 migraine patients and 7,353 controls from eight Dutch cohorts, using a different ^{1}H-NMR metabolomics platform in a systematic approach, also showed alterations in HDL metabolism, in that study defined by a decreased level of lipoprotein A1 and a decreased free cholesterol to total lipid ratio in small HDL subspecies.[55] Neurovascular and endothelial dysfunction are believed to be an underlying cause for the increased risk in cerebrovascular and cardiovascular diseases in migraine patients.[57, 58] At the basis of this involvement lies a possible higher prevalence of risk factors, such as hypertension and hyperlipidemia, in migraine patients.[57] Also the involvement of lipids in migraine pathophysiology has been shown in various studies.[54, 55] Regardless, the exact role lipids play is complex and needs further investigation.

Glucose is another metabolite we found that has previously been associated with migraine either directly or via metabolically associated pathways. Glucose levels and insulin metabolism, as well as mitochondrial dysfunction have been known to play a role in migraine pathology.[59, 60] Still, no association was found between migraine and diabetes type 2.[59, 61, 62] It has been suggested that outside attacks, migraine patients have an impaired insulin sensitivity and higher fasting plasma insulin levels compared to controls.[63, 64] Recently it was shown that glucose levels were higher during a spontaneous migraine attack compared to outside of an attack.[65] Both 1,5-anhydrosorbitol and myoinositol, which were part of our prediction model, are involved in glucose metabolism. 1,5-Anhydrosorbitol is a naturally occurring monosaccharide found in nearly all foods and myoinositol, which is a carbocyclic sugar that is abundant in brain and other mammalian tissues,

is synthesized from glucose 6-phosphate. Pyruvate is the conjugate base of pyruvic acid and is a key intermediate in several metabolic pathways throughout the cell. Pyruvic acid can be produced from glucose through glycolysis and it can supply energy to the cell via the Krebs cycle in the mitochondria. One study has investigated the lactic and pyruvic acid levels in the plasma of the migraine patients it was shown that both were significantly higher in migraine patients than in normal controls.[66]

In addition, multiple amino acids were part of our prediction model for active migraine status, namely leucine, isoleucine, methionine, valine, proline and serine. Over the last decades, multiple amino acids have been hypothesized to play a role in migraine pathophysiology.[19] Leucine, isoleucine and valine are branched-chain amino acids (BCAAs), BCAAs have emerged as potential biomarkers of disease as they are associated with risk of cardiovascular disease, end-stage renal failure, and ischemic stroke.[67] In a small study of 37 migraine patients and 40 controls elevated levels of isoleucine in blood serum were found.[68] A recent study investigated amine pathways in 20 patients with migraine without aura, and 20 healthy subjects in serum with liquid chromatography coupled to mass spectrometry (LC-MS).[69] This LC-MS study found decreased levels of leucine, isoleucine and methionine in migraine patients compared to controls. The valine, proline and serine concentration was not assessed directly in this study.[69] Although glutamate/glutamine has been repeatedly linked to migraine,[70, 71] in our study the levels of glutamine/glutamate were not part of the predictive profile for migraine status.

As far as we know the other metabolites we found to be associated with active migraine status (acetate, dimethylglycine and creatinine) have thus far not been associated with migraine. Acetate is a monocarboxylic acid anion, which is metabolized mostly in peripheral tissues. Dimethylglycine, which is a derivative of the amino acid glycine, but it can also be a by-product of the metabolism of choline. Dimethylglycine has been suggested as a treatment for mitochondrial diseases[72] and in that sense might be associated with the migraine-glucose dysregulation. Creatine is involved in the conversion from adenosine diphosphate (ADP) back to ATP in muscle and is synthesized mainly in the liver from amino acids glycine and arginine.

We here identified a metabolite profile predictive for active migraine, a finding supported by the observation that several of its metabolites have already been reported in literature to be individually (in)directly associated with migraine. We would like to emphasize again that the focus of this study was to explore whether metabolite profiles can be linked to migraine status and less to show direct clinical relevance of individual metabolites.

A limitation of our study is that, the set of metabolites we studied using our metabolic profiling method covers only a small part of the human metabolome. Future, complementary, studies using different, more advanced, platforms may identify additional metabolites associated with migraine status. Additionally, in our study population we know to what extent patients are related and have the opportunity to correct for this. Future studies have to show to what extent our findings

are also applicable to well-selected groups of migraine patients, for instance, with respect to frequency of attacks, time of last attack to blood withdrawal, possible comorbidities, etc. Another possible limitation is that in our study, model selection by elastic net regression was used for predictor selection to eliminate high correlations among predictors. This might lead to reduced transferability of prediction models, because correlation structures of predictors can vary between studies. Although we corrected for age, sex, BMI and smoking in our analysis, we cannot exclude a residual confounding effect of this variable nor of any other variable that we have not tested. Another limitation is that we used the same population for discovery of the associated signals and for assessing the magnitude of the association. Ideally, a replication study, to validate our findings, should be performed.

Conclusion

In conclusion, using hypothesis-free metabolic profiling, by measuring a large set of metabolites using ^{1}H-NMR spectroscopy, we identified a metabolomic profile consisting of 22 metabolite signals (lipids, amino acids and metabolites of glucose metabolism) that was predictive for active migraine status.

Supplementary information

Additional file 1.
DOC. file with an overview of signals identified in the 2-dimensional J-resolved ^{1}H-NMR spectrum and performed transformations

https://static-content.springer.com/esm/art%3A10.1186%2Fs10194-021-01357-w/MediaObjects/10194_2021_1357_MOESM1_ESM.docx

Additional file 2.
PDF. file with ratio stratification plots for sex, age BMI and smoking status in lifetime migraine patients and active migraine patients
https://static-content.springer.com/esm/art%3A10.1186%2Fs10194-021-01357-w/MediaObjects/10194_2021_1357_MOESM2_ESM.png

Abbreviations: ^{1}H-NMR: Proton nuclear magnetic resonance spectroscopy; ERF: Erasmus Rucphen Family; JRES: J-resolved spectra; CPMG: Purcell-Meiboom-Gill; OR: Odds ratio; CI: Confidence interval; CGRP: Calcitonin gene-related peptide

Ethics approval and consent to participate: The study was approved by the medical ethics committee of Erasmus University Medical Centre. All subjects provided written informed consent prior to the study.

Availability of data and materials: The data used and/or analysed during the current study are available from the corresponding author on reasonable request.

Competing interests: The authors declare that they have no competing interests

Funding: This work was supported by grants from the Netherlands Organization for Scientific Research (NWO) (907-00-217, G.M.T; VIDI 91711319, G.M.T.); the Centre for Medical Systems Biology (CMSB) and Netherlands Consortium for Systems Biology (NCSB), both within the framework of the Netherlands Genomics Initiative (NGI)/NWO (to A.M.J.M.v.d.M.); and the FP7 EU project EUROHEADPAIN (no. 602633) (to A.M.J.M.v.d.M. and G.M.T.). They had

no role in the design or conduct of the study. The ERF study is part of EUROSPAN (European Special Populations Research Network) FP6 STRP (no.018947 LSHG-CT-2006-01947) and also received funding from the European Community's Seventh Framework Programme (FP7/2007-2013)/grant agreement HEALTH-F4-2007-201413 by the European Commission under the programme "Quality of Life and Management of the Living Resources" of 5th Framework Programme (no. QLG2-CT-2002-01254) as well as FP7 project EUROHEADPAIN (nr 602633).

Author's contributions: AMJMvdM, GMT designed the study design. PH and AVEH were involved in the ^{1}H-NMR measurements and/or annotation. PH, AVEH and LSV performed the statistical analyses. All authors contributed to the interpretation of results. AVEH made the figures and wrote the initial draft of the manuscript. All authors critically revised and approved the final version for submission.

Acknowledgements: Thanks to all study participants from the Erasmus Rucphen Family and their relatives, general practitioners, and neurologists for their contributions. We thank Dr. Oleg A. Mayboroda, from the Center for Proteomics and Metabolomics (LUMC), in which the NMR analyses were done, for critical reading of the manuscript and Dr. Boukje de Vries for technical help and support.

References

1. GBD 2019 Diseases and Injuries Collaborators. Global burden of 369 diseases and injuries in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. *Lancet*. 2020;396(10258):1204-1222.
2. Launer LJ, Terwindt GM, Ferrari MD. The prevalence and characteristics of migraine in a population-based cohort: the GEM study. *Neurology*. 1999;53(3):537-542.
3. Headache Classification Committee of the International Headache Society (IHS) The International Classification of Headache Disorders, 3rd edition. *Cephalalgia*. 2018;38(1):1-211.
4. Ferrari MD, Klever RR, Terwindt GM, Ayata C, van den Maagdenberg AM. Migraine pathophysiology: lessons from mouse models and human genetics. *Lancet Neurol*. 2015;14(1):65-80.
5. Wang TJ, Larson MG, Vasan RS, et al. Metabolite profiles and the risk of developing diabetes. *Nat Med*. 2011;17(4):448-453.
6. Kaddurah-Daouk R, Krishnan KR. Metabolomics: a global biochemical approach to the study of central nervous system diseases. *Neuropsychopharmacology*. 2009;34(1):173-186.
7. Orešič M, Hyötyläinen T, Herukka SK, et al. Metabolome in progression to Alzheimer's disease. *Translational psychiatry*. 2011;1(12):e57.
8. Patel S, Ahmed S. Emerging field of metabolomics: big promise for cancer biomarker identification and drug discovery. *J Pharm Biomed Anal*. 2015;107:63-74.
9. Shah SH, Sun JL, Stevens RD, et al. Baseline metabolomic profiles predict cardiovascular events in patients at risk for coronary artery disease. *Am Heart J*. 2012;163(5):844-850.e1.
10. Deng Y, Huang C, Su J, Pan CW, Ke C. Identification of biomarkers for essential hypertension based on metabolomics. *Nutr Metab Cardiovasc Dis*. 2021;31(2):382-395.
11. Chan D, Ng LL. Biomarkers in acute myocardial infarction. *BMC Med*. 2010;8:34.
12. Riely GJ, Marks J, Pao W. KRAS mutations in non-small cell lung cancer. *Proc Am Thorac Soc*. 2009;6(2):201-205.
13. De Vries B, Haan J, Frants RR, Van den Maagdenberg AM, Ferrari MD. Genetic biomarkers for migraine. *Headache*. 2006;46(7):1059-1068.
14. van Dongen RM, Zielman R, Noga M, et al. Migraine biomarkers in cerebrospinal fluid: A systematic review and meta-analysis. *Cephalalgia*. 2017;37(1):49-63.
15. Gormley P, Anttila V, Winsvold BS, et al. Meta-analysis of 375,000 individuals identifies 38 susceptibility loci for migraine. *Nat Genet*. 2016;48(8):856-866.
16. Schwedt TJ, Chiang CC, Chong CD, Dodick DW. Functional MRI of migraine. *Lancet Neurol*. 2015;14(1):81-91.
17. Frederiksen SD, Bekker-Nielsen Dunbar M, Snoer AH, Deen M, Edvinsson L. Serotonin and Neuropeptides in Blood From Episodic and Chronic Migraine and Cluster Headache Patients in Case-Control and Case-Crossover Settings: A Systematic Review and Meta-Analysis. *Headache*. 2020;60(6):1132-1164.
18. Ferroni P, Barbanti P, Spila A, et al. Circulating Biomarkers in Migraine: New Opportunities for Precision Medicine. *Curr Med Chem*. 2019;26(34):6191-6206.
19. Ferrari MD, Odink J, Bos KD, Malessy MJ, Bruyn GW. Neuroexcitatory plasma amino acids are elevated in migraine. *Neurology*. 1990;40(10):1582-1586.
20. Martinez F, Castillo J, Rodriguez JR, Leira R, Noya M. Neuroexcitatory amino acid levels in plasma and cerebrospinal fluid during migraine attacks. *Cephalalgia*. 1993;13(2):89-93.
21. Guldiken S, Guldiken B, Demir M, et al. Soluble CD40 ligand and prolactin levels in migraine patients during interictal period. *J Headache Pain*. 2011;12(3):355-360.
22. Vanmolkot FH, de Hoon JN. Increased C-reactive protein in young adult patients with migraine. *Cephalalgia*. 2007;27(7):843-846.

23. Pavelek Z, Soucek O, Krejsek J, et al. The role of the immune system and the biomarker CD3 + CD4 + CD45RA-CD62L- in the pathophysiology of migraine. *Sci Rep*. 2020;10(1):12277.
24. Fusayasu E, Kowa H, Takeshima T, Nakaso K, Nakashima K. Increased plasma substance P and CGRP levels, and high ACE activity in migraineurs during headache-free periods. *Pain*. 2007;128(3):209-214.
25. Cernuda-Morollon E, Larrosa D, Ramon C, et al. Interictal increase of CGRP levels in peripheral blood as a biomarker for chronic migraine. *Neurology*. 2013;81(14):1191-1196.
26. Cernuda-Morollón E, Martínez-Camblor P, Alvarez R, et al. Increased VIP levels in peripheral blood outside migraine attacks as a potential biomarker of cranial parasympathetic activation in chronic migraine. *Cephalalgia*. 2015;35(4):310-316.
27. Hamed SA, Hamed EA, Ezz Eldin AM, Mahmoud NM. Vascular risk factors, endothelial function, and carotid thickness in patients with migraine: relationship to atherosclerosis. *J Stroke Cerebrovasc Dis*. 2010;19(2):92-103.
28. Kurth T, Ridker PM, Buring JE. Migraine and biomarkers of cardiovascular disease in women. *Cephalalgia*. 2008;28(1):49-56.
29. Tietjen GE, Khubchandani J, Herial N, et al. Migraine and vascular disease biomarkers: A population-based case-control study. *Cephalalgia*. 2018;38(3):511-518.
30. 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;14(6):338-350.
31. Edvinsson L, Warfvinge K. Recognizing the role of CGRP and CGRP receptors in migraine and its treatment. *Cephalalgia*. 2019;39(3):366-373.
32. Ramon C, Cernuda-Morollon E, Pascual J. Calcitonin gene-related peptide in peripheral blood as a biomarker for migraine. *Curr Opin Neurol*. 2017;30(3):281-286.
33. Greco R, De Icco R, Demartini C, et al. Plasma levels of CGRP and expression of specific microRNAs in blood cells of episodic and chronic migraine subjects: towards the identification of a panel of peripheral biomarkers of migraine? *J Headache Pain*. 2020;21(1):122.
34. Hagen K, Stovner LJ, Zwart JA. High sensitivity C-reactive protein and risk of migraine in a 11-year follow-up with data from the Nord-Trondelag health surveys 2006-2008 and 2017-2019. *J Headache Pain*. 2020;21(1):67.
35. Dunn WB, Broadhurst D, Begley P, et al. Procedures for large-scale metabolic profiling of serum and plasma using gas chromatography and liquid chromatography coupled to mass spectrometry. *Nat Protoc*. 2011;6(7):1060-1083.
36. Stam AH, de Vries B, Janssens AC, et al. Shared genetic factors in migraine and depression: evidence from a genetic isolate. *Neurology*. 2010;74(4):288-294.
37. Aulchenko YS, Heutink P, Mackay I, et al. Linkage disequilibrium in young genetically isolated Dutch population. *Eur J Hum Genet*. 2004;12(7):527-534.
38. Sleegers K, de Koning I, Aulchenko YS, et al. Cerebrovascular risk factors do not contribute to genetic variance of cognitive function: the ERF study. *Neurobiol Aging*. 2007;28(5):735-741.
39. Headache Classification Subcommittee of the International Headache S. The International Classification of Headache Disorders: 2nd edition. *Cephalalgia : an international journal of headache*. 2004;24 Suppl 1:9-160.
40. Vaarhorst AA, Verhoeven A, Weller CM, et al. A metabolomic profile is associated with the risk of incident coronary heart disease. *Am Heart J*. 2014;168(1):45-52 e7.
41. Findeisen M, Brand T, Berger S. A H-1-NMR thermometer suitable for cryoprobes. *Magn Reson Chem*. 2007;45(2):175-178.
42. Wu PSC, Otting G. Rapid pulse length determination in high-resolution NMR. *J Magn Reson*. 2005;176(1):115-119.
43. Price WS. Water Signal Suppression in NMR Spectroscopy. *Annu Rep NMR Spectrosc*. 1999;38:289-354.
44. Aue WP, Karhan J, Ernst RR. Homonuclear Broad-Band Decoupling and 2-Dimensional J-Resolved Nmr-Spectroscopy. *J Chem Phys*. 1976;64(10):4226-4227.
45. Meiboom S, Gill D. Modified Spin-Echo Method for Measuring Nuclear Relaxation Times. *Rev Sci Instrum*. 1958;29(8):688-691.

46. Nicholson JK, Foxall PJD, Spraul M, Farrant RD, Lindon JC. 750-Mhz H-1 and H-1-C-13 Nmr-Spectroscopy of Human Blood-Plasma. *Anal Chem*. 1995;67(5):793-811.

47. MacIntyre DA, Jimenez B, Lewintre EJ, et al. Serum metabolome analysis by H-1-NMR reveals differences between chronic lymphocytic leukaemia molecular subgroups. *Leukemia*. 2010;24(4):788-797.

48. Tomasi G, van den Berg F, Andersson C. Correlation optimized warping and dynamic time warping as preprocessing methods for chromatographic data. *J Chemometr*. 2004;18(5):231-241.

49. Wishart DS, Jewison T, Guo AC, et al. HMDB 3.0-The Human Metabolome Database in 2013. *Nucleic Acids Res*. 2013;41(D1):D801-D807.

50. Aulchenko YS, Ripke S, Isaacs A, Van Duijn CM. GenABEL: an R library for genome-wide association analysis. *Bioinformatics*. 2007;23(10):1294-1296.

51. Friedman J, Hastie T, Tibshirani R. Regularization Paths for Generalized Linear Models via Coordinate Descent. *J Stat Softw*. 2010;33(1):1-22.

52. Bigal ME, Lipton RB. Obesity is a risk factor for transformed migraine but not chronic tension-type headache. *Neurology*. 2006;67(2):252-257.

53. Linstra KM, Ibrahimi K, Terwindt GM, Wermer MJ, MaassenVanDenBrink A. Migraine and cardiovascular disease in women. *Maturitas*. 2017;97:28-31.

54. Tietjen GE, Khubchandani J. Vascular biomarkers in migraine. *Cephalalgia*. 2015;35(2):95-117.

55. Onderwater GLJ, Ligthart L, Bot M, et al. Large-scale plasma metabolome analysis reveals alterations in HDL metabolism in migraine. *Neurology*. 2019;92(16):e1899-e1911.

56. Rubino E, Vacca A, Govone F, et al. Investigating the role of adipokines in chronic migraine. *Cephalalgia*. 2017;37(11):1067-1073.

57. Bigal ME, Kurth T, Hu H, Santanello N, Lipton RB. Migraine and cardiovascular disease: possible mechanisms of interaction. *Neurology*. 2009;72(21):1864-1871.

58. Paolucci M, Altamura C, Vernieri F. The Role of Endothelial Dysfunction in the Pathophysiology and Cerebrovascular Effects of Migraine: A Narrative Review. *J Clin Neurol*. 2021;17(2):164-175.

59. Rainero I, Govone F, Gai A, Vacca A, Rubino E. Is Migraine Primarily a Metaboloendocrine Disorder? *Curr Pain Headache Rep*. 2018;22(5):36.

60. Yorns WR, Jr., Hardison HH. Mitochondrial dysfunction in migraine. *Semin Pediatr Neurol*. 2013;20(3):188-193.

61. Burch RC, Rist PM, Winter AC, et al. Migraine and risk of incident diabetes in women: a prospective study. *Cephalalgia*. 2012;32(13):991-997.

62. López-de-Andrés A, Luis Del Barrio J, Hernández-Barrera V, et al. Migraine in adults with diabetes; is there an association? Results of a population-based study. *Diabetes Metab Syndr Obes*. 2018;11:367-374.

63. Cavestro C, Rosatello A, Micca G, et al. Insulin metabolism is altered in migraineurs: a new pathogenic mechanism for migraine? *Headache*. 2007;47(10):1436-1442.

64. Rainero I, Limone P, Ferrero M, et al. Insulin sensitivity is impaired in patients with migraine. *Cephalalgia*. 2005;25(8):593-597.

65. Zhang DG, Amin FM, Guo S, et al. Plasma Glucose Levels Increase During Spontaneous Attacks of Migraine With and Without Aura. *Headache*. 2020;60(4):655-664.

66. Okada H, Araga S, Takeshima T, Nakashima K. Plasma lactic acid and pyruvic acid levels in migraine and tension-type headache. *Headache*. 1998;38(1):39-42.

67. Batch BC, Hyland K, Svetkey LP. Branch chain amino acids: biomarkers of health and disease. *Curr Opin Clin Nutr Metab Care*. 2014;17(1):86-89.

68. Domitrz I, Koter MD, Cholojczyk M, et al. Changes in Serum Amino Acids in Migraine Patients without and with Aura and their Possible Usefulness in the Study of Migraine Pathogenesis. *CNS Neurol Disord Drug Targets*. 2015;14(3):345-349.

69. Ren C, Liu J, Zhou J, et al. Low levels of serum serotonin and amino acids identified in migraine patients. *Biochem Biophys Res Commun*. 2018;496(2):267-273.

70. Cananzi AP, Dandrea G, Perini F, Zamberlan F, Welch KMA. Platelet and Plasma-Levels of Glutamate and Glutamine in Migraine with and without Aura. *Cephalalgia*. 1995;15(2):132-135.

71. Zielman R, Wijnen JP, Webb A, et al. Cortical glutamate in migraine. *Brain*. 2017;140(7):1859-1871.

72. Pfeffer G, Majamaa K, Turnbull DM, Thorburn D, Chinnery PF. Treatment for mitochondrial disorders. *Cochrane Database Syst Rev*. 2012;2012(4):Cd004426.

3

Correlating ratios of amines in plasma and cerebrospinal fluid

Aster V.E. Harder*, Jan B. van Klinken*, Robin M. van Dongen, Gerrit L.J. Onderwater, Michel M.D. Ferrari, Amy C. Harms, Thomas Hankemeier, Gisela M. Terwindt* and Arn M.J.M. van den Maagdenberg*

Ann Neurol 2023 Jul; 10.1002/ana.26743

Abstract

The cerebrospinal fluid (CSF) is regarded a suitable body fluid for metabolomics studies on central nervous system (CNS) disorders as it, due to its close proximity, best reflects changes in the brain. However, as repeated CSF sampling is not patient-centered, blood sampling should be considered the better alternative for, for instance, longitudinal studies. A relevant question then is how well metabolites in blood and CSF correlate. Therefore, we here studied concentrations of amines, including amino acids and biogenic amines in human blood plasma and CSF to investigate in control individuals and patients with migraine how these fluids are associated. Amines were measured using a validated ultra-performance liquid chromatography mass spectrometry (UPLC-MS) platform. Using single-metabolite correlation, a minority of 39 amines (4/39, 10.3%) had a correlation coefficient ≥ .70. Ratio correlations were significantly better than single-metabolite correlations for 308 of 741 (41.5%) of amine combinations after multiple testing correction (FDR < .05). Primarily, amino acids showed high correlation when studied as ratios (i.e. L-leucine/L-methionine, L-leucine/L-asparagine, L-valine/L-phenylalanine), likely because these ratios are tightly regulated by blood-brain barrier transport systems. We extended our analyses to migraine. Our analyses revealed that some amine ratios seem affected in participants with migraine when compared to healthy controls.

KEYWORDS: Omics, comparison, neurotransmitters, migraine, ultra-performance liquid chromatography mass spectrometry (UPLC-MS)

Introduction

The cerebrospinal fluid (CSF) is the body fluid closest to the brain and, therefore, of major interest for central nervous system (CNS) disorders. Hence, it is believed that the CSF should more faithfully reflect changes in brain metabolites than blood, which is typically sampled to study metabolites related to the brain's neurochemistry, also in the context of a brain disease. Novel metabolomics techniques can quantify an increasing number of metabolites, making it possible to better study the metabolomic composition of CSF in both health and disease.[1,2] However, CSF collected via lumbar puncture putt subjects at risk of post-dural puncture headache,[3] although small non-traumatic needles lead to an important reduction of this risk.[4,5] Hence, ideally, blood measurements are used to reflect disease-relevant metabolomic changes in the brain over time. Moreover, as repeated CSF collection is not a realistic and patient-centered, alternative blood measurements are especially needed when repeated measures are necessary to study short-interval longitudinal disease pathophysiology, such as in the case of paroxysmal brain disorders where repeated measurement towards attacks may be particularly insightful. With that approach, metabolomic profiles may be obtained in the interictal, pre-ictal, ictal, and postictal phases of subjects, who suffer from a paroxysmal brain disorder. Also for studying pharmacodynamic effects of drugs targeting the CNS, it may be helpful to better understand the relationship between blood and CSF. Although some CNS disorder, such as cerebral amyloid angiopathy display important biomarkers, such as amyloid that can be used for follow-up studies[6-8] this is less helpful for paroxysmal CNS disorder where the time-frame for repeated measurements typically is within hours (up to days), instead of years.

One relevant class of metabolites for paroxysmal CNS disorders are amines such as amino acids and biogenic amines. This class of metabolites includes important neurotransmitters and their precursors, such as glutamate, γ-aminobutyric acid (GABA), glutamine, serotonin, and tryptophan.[9] Studies in the distant past showed that the levels of only a few amines seem to correlate between blood plasma and CSF.[10, 11] Of note, it was shown that there was a strong correlation of the "CSF/plasma ratio" between different amines, which was explained by the fact that amines are transported across the blood-brain barrier by the same transporter.[11] Correlation was particularly prominent with five neutral amino acids, namely isoleucine, leucine, phenylalanine, valine and tyrosine.[11] This finding may imply that any correlation between blood plasma and CSF is not shown by the concentrations themselves, but mere ratios of amines.

As current measurement techniques have greatly advanced - whereas older studies were primarily focused on proteinogenic amino acids - and are able to measure a much broader range of amines with higher precision, we aimed to employ our validated amine metabolomics platform for CSF and plasma,[12] to investigate to which extent amines correlate between both body fluids. We extrapolated our approach from healthy participants to those with migraine, a common paroxysmal brain disease, to assess whether concentrations and/or ratios of amines differed between those with and without migraine.

Methods

Study Design

One aim of this study is to determine to what extent amine levels in blood plasma correlate to those in CSF from healthy controls. To this end, we examined: (1) single-metabolite correlations of amine levels between CSF and plasma, and (2) plasma/CSF ratio-metabolite correlations. In addition, we examined both types of correlation in a paroxysmal CNS disease, namely migraine, as to investigate if correlations can provide insight into disease pathophysiology. The study was conducted according to the criteria of the Declaration of Helsinki and was approved by the Leiden University Medical Center institutional ethics committee. All participants provided written informed consent prior to the study.

Healthy controls

Participants were recruited as healthy controls in a larger biochemical study on migraine.[13] CSF and plasma were collected purely for research purposes. Exclusion criteria for healthy controls were severe neurological (including migraine or other high-frequent headache disorder) or psychiatric disorders, any oncological history, or contraindication for lumbar puncture (including signs and symptoms of increased intracranial pressure, local skin infection, or coagulopathy, including the use of anti-coagulant drugs or platelet inhibitors).

Migraine participants

Participants with episodic migraine were recruited as part of the same study as that of the healthy controls. Migraine was diagnosed according to the International Classification of Headache Disorders.[14] CSF and plasma were collected outside migraine attacks (interictal) [4,5,13]. Apart from migraine diagnosis the same exclusion criteria were used for the healthy controls.

Sample collection

CSF sampling was performed via lumbar puncture before 13.00 p.m. and after overnight fasting (participants were only allowed to drink water during 8 h prior to sampling). Lumbar puncture was performed between the L3/L4, L4/L5 or L5/S1 interspace using a .9 (20 G) x 90 mm Quincke traumatic needle (MediPlast, Malmö, Sweden) or .9 (20 G), .7 (22 G), or .5 (24 G) x 90 mm Sprotte atraumatic needle (Pajunk, Geisingen, Germany). Intracranial pressure was measured, and first 3 mL CSF was sampled for routine diagnostics (cell count, glucose and total protein levels), prior CSF collection for metabolic measurements. For metabolomics measurements 3.8 mL of CSF was sampled directly in a 15-mL polypropylene falcon tube (tube P1a; Cat. No. 188 271, (Greiner Bio-One, Alphen aan de Rijn, The Netherlands) and centrifuged at 4°C for 5 min (2000 rpm, 747 g).[15] The supernatant was transferred to a new 15-mL polypropylene falcon tube, inverted several times, and divided in .5 mL aliquots into 1.8-mL Nunc™ cryotubes (Cat. No.

368632, Sigma-Aldrich, Saint Louis, MO, USA) that already contained 1.0 mL of cold ethanol (Prod. No.:8098, ethanol absolute, J.T.Baker, Phillipsburg, NJ, USA). The Nunc™ cryotubes were inverted several times to thoroughly mix CSF and ethanol and were placed on dry ice within 30 min from sampling. Thereafter, aliquots were transferred to -80˚C for storage within 60 min from sampling. Plasma samples were collected straight after lumbar puncture, while participants rested in supine position samples were drawn from the median cubital vein. Venous blood was collected in a EDTA plasma tube (Cat. No. 366643, BD Medical, Franklin Lakes, NJ, USA) and centrifuged at 4˚C for 20 min (2,000 rpm, 747 g). The supernatant (plasma) was transferred to a new 15-mL polypropylene falcon tube, inverted several times, and divided in .5 mL aliquots into 1.0-mL Nunc™ cryotubes (Cat. No. 366656, Sigma-Aldrich). Thereafter, all aliquots were transferred to -80˚C for storage within 60 min from sampling. All CSF and plasma samples remained at -80˚C until sample preparation.

Amine measurements

Both CSF and plasma samples were measured with an ultra-performance liquid chromatography mass spectrometry (UPLC-MS) method that was shown to reliably quantify 74 biogenic amines for mouse CSF samples [12]. For details on sample preparation and amine measurements see Onderwater et al.[13] In brief, quantitation of amino acids and biogenic amines utilized an AccQ-Tag derivatization strategy adapted from the protocol supplied by Waters. 5.0 µL of each sample was spiked with an internal standard solution. Proteins were precipitated by the addition of MeOH after which the samples were dried in a speedvac. The residue was reconstituted in a borate buffer (pH 8.8) with AQC reagent. LC-MS measurements were performed with an Acquity UPLC System (Waters, Milford, MA, USA) coupled to a QTRAP 6500 Triple-Quadruple MS System (Sciex, Framingham, MA, USA). For the chromatographic separation 1 µL of the sample was injected on an AccQ·Tag Ultra 100 x 2.1 mm column with a particle size of 1.7 µm (Waters, Milford, MA, USA) and chromatographic separation was achieved with the flow rate of .7 mL/min. over an 11-min LC gradient program.

The target analytes were detected in electrospray ionization (ESI) positive ion mode. The derivatized target metabolites and their internal standards were identified by their retention times and using their specific Multiple Reaction Monitoring (MRM) at nominal mass resolution. Data were pre-processed with MultiQuant Software for Quantitative Analysis v3.0.2 (Sciex, Framingham, MA, USA). Peak areas of target analytes relative to their corresponding internal standards were calculated as area ratios.

All samples were measured in ten separate batches including calibration (injected at the start and the end), QC, blank, and randomized study samples. Plasma samples were measured in the first five batches, whereas CSF samples were measured later. QC samples were analyzed after every ten injections and used to monitor the intra-batch data quality and to correct for inter-batch analytical variance. To subtract the inter-batch analytical variance, all area ratios (study samples, calibration

and blank samples) were corrected with a regression model generated through the data acquired from regular QC injections.[16]

Concentration calculations were performed by using the Pearson's linear correlation within the calibration range for both CSF and plasma samples. The R^2 for each calibration line was selected ≥ .95. Based on corrected area ratios, the analytes were defined as not detected if 30% of the mean value of all study samples is equal or less than the mean value of the lowest calibration level. The final concentration values were reported in μmol/L (μM) unit.

Statistical analysis

Quality control

Outlier detection was performed using principal component analysis (PCA) with a 99% confidence interval. Prior to the statistical analyses data were ^{10}log-transformed and values below [mean - 4 * SD] and above [mean + 4 * SD] were replaced with these cut-offs.

Single-metabolite plasma/CSF correlations

After $\log_{10}$-transformation, metabolite levels were adjusted for age, sex and age*sex and were subsequently inverse normal transformed to obtain robust estimates of the plasma/CSF correlation coefficients. Plasma/CSF correlation coefficients *r* were compared between groups by normalizing *r* with Fisher's transformation

$$z = 0.5\frac{\ln(1+r)}{\ln(1-r)}$$

and subsequently testing the difference in *z*-values between groups against the centered normal distribution $N\left(0, \frac{1}{n_1-3}+\frac{1}{n_2-3}\right)$, with *n1* and *n2* the sample sizes of groups 1 and 2, respectively.

Plasma/CSF metabolite ratio correlations

In addition to determining plasma/CSF correlation coefficients for all single metabolites, also correlation coefficients were calculated for the ratios between metabolites. Prior to the correlation analysis the ratios were $\log_{10}$-transformed, adjusted for age, sex, age*sex and inverse normal transformed. Since some ratio correlation coefficients seemed higher than the correlation coefficients of the related single metabolites, we tested for this using the Steiger's approach for comparing correlations coefficients in overlapping data.[17] Specifically for a ratio of metabolites M1 and M2, we tested the difference between the Fisher transformed correlation coefficient of the ratio $z_{M1/M2}$ and the average *z* of both metabolites $\frac{1}{2}z_{M1}+\frac{1}{2}z_{M2}$ one-sided against the centered normal distribution $N\left(0, \frac{1-(0.5s_1+0.5s_2)}{n-3}\right)$, with s_i the correction factor for overlapping data of M1/M2 relative to respectively M1 and M2, as defined in the Steiger's approach.[17] Finally, we

transformed the difference in *z*-value back with the inverse Fisher transformation to get a measure of the *gain* in the ratio correlation coefficient with respect to single metabolites.

$$r \text{ ratio gain}(M1, M2) = \frac{\exp\left(2z_{M1/M2} - z_{M1} - z_{M2}\right) - 1}{\exp\left(2z_{M1/M2} - z_{M1} - z_{M2}\right) + 1}$$

Multiple test correction was performed by using Benjamini Hochberg's procedure. A p value < .05 was considered significant after false discovery rate (FDR) correction. Coefficients of determination (R^2) and R^2 ratio gains (R^2 gain) for all metabolite ratios were plotted in a heatmap using R software (version 3.6.3 with package *gplots* and heatmap.2 function).

Results

Amines in blood plasma and CSF of healthy controls

Concentrations of amines

In blood plasma, 40 amines were reliably measured in healthy controls. In CSF, o-phosphoethanolamine showed abnormally high concentrations (> 3 SDs above median) in multiple samples from one batch and was, therefore, excluded leaving 39 amines for further analysis. No other metabolite in plasma nor CSF showed such batch effect. At the individual level, PCA identified one clear outlier in both plasma and CSF that, therefore, was excluded, leaving amine profiles of 95 healthy controls for further analysis (**Table 1**). Later, it turned out that the excluded participant suffered from cardiac disease.

Blood plasma concentrations were higher than CSF concentrations for most amines in healthy controls (**Table 2**). Only concentrations of ethanolamine and putrescine were higher in CSF than plasma. L-Glutamine was the most abundant amine in CSF (median: 591 mM, interquartile range: 539-638 mM), with similar concentrations in blood (median: 665 mM, interquartile range: 600-715 mM).

Single-metabolite correlations between plasma and CSF

Both correlation coefficients (*r*), and coefficients of determination (R^2) were calculated to show the percentage of variation explained by a linear model. Most amines showed a low single/direct correlation between plasma and CSF concentrations ($r = \frac{\text{Amine x}_{plasma}}{\text{Amine x}_{CSF}}$) (**Table 3**). Four of 39 amines (10.3%) had a correlation coefficient ≥ .70, namely, homocitrulline, S-methylcysteine, methionine sulfone, and L-alpha-aminobutyric acid (Table 3). **Figure 1** shows the correlation plots of the $\log_{10}$-transformed uncorrected data of the four best correlating amines, the remaining plots can be found in the Supplementary Data (**Figure S1**). The four amines did not show

remarkable sex differences (**Table S1**). L-Threonine additionally showed higher correlation in females and SDMA showed a lower concentration in females compared to males.

Plasma/CSF ratios metabolite correlations

Next, we tested whether metabolite ratios in plasma correlated with the same ratio in CSF. Correlations between all possible ratios were calculated based on their R^2 (**Figure 2**, **Table S2**). As expected, metabolites that already had strong single-correlations (i.e. homocitrulline, S-methylcysteine, methionine sulfone, L-alpha-aminobutyric acid, and L-threonine), also showed strong correlation for most ratios with other amines. Additionally, several other metabolite pairs showed high correlation. For example, for the single L-Valine correlation the R^2 = .11 and for the single L-Phenylalanine correlation the R^2 was .01, but the plasma ratio of L-Valine/L-Phenylalanine had a higher correlation with the CSF ratio of L-Valine/L-Phenylalanine ($\frac{\text{L-Valine}_{plasma}}{\text{L-Phenylalanine}_{plasma}} / \frac{\text{L-Valine}_{CSF}}{\text{L-Phenylalanine}_{CSF}} = R^2 = .67$). This means that 67% of the variance in y (the ratio between L-Valine/L-Phenylalanine in CSF) is explained by x (the ratio between L-Valine/L-Phenylalanine in plasma) (**Figure 3A-C**). Sex differences in the correlations were minor for any the ratio pairs (**Table S2**, **Figure S2A**, **B**). Primarily, the amino acids showed a high correlation in their ratios.

The improvement in R^2 ratio correlation (R^2 gain) compared to single-correlation is illustrated in **Figure 4**. The ratio correlations were significantly higher than the related single-metabolite correlations for 308 of 741 (41.5%) amine combinations, i.e. after multiple testing correction (FDR < .05) (**Figure 4**, **Table S3**). As a control, the correlations between the product of two amines were tested, but did not show clear significance (data not shown).

Table 1 Clinical characteristics of healthy controls

	Healthy controls
Number of participants	95
Subject characteristics	
Females (n, (%))	56 (58.9)
Years of age (± SD‡)	38.8 (± 14.5)
BMI† (± SD‡)	23.7 (± 2.8)
Smoking (n, (%))	20 (21.1)
Overnight fasting	
Fasting time in h (± SD‡)	11.6 (± 2.4)
Sampling characteristics	
Opening pressure in mmH2O (± SD‡)	19.1 (± 4.4)
CSF characteristics	
Erythrocytes count/3µL (± SD‡)	154 (± 934)
Leukocytes count/3µL (± SD‡)	6 (± 6)
Protein concentration in g/L (± SD‡)	.35 (± .13)
Glucose in mmol/L (± SD‡)	3.2 (± .3)

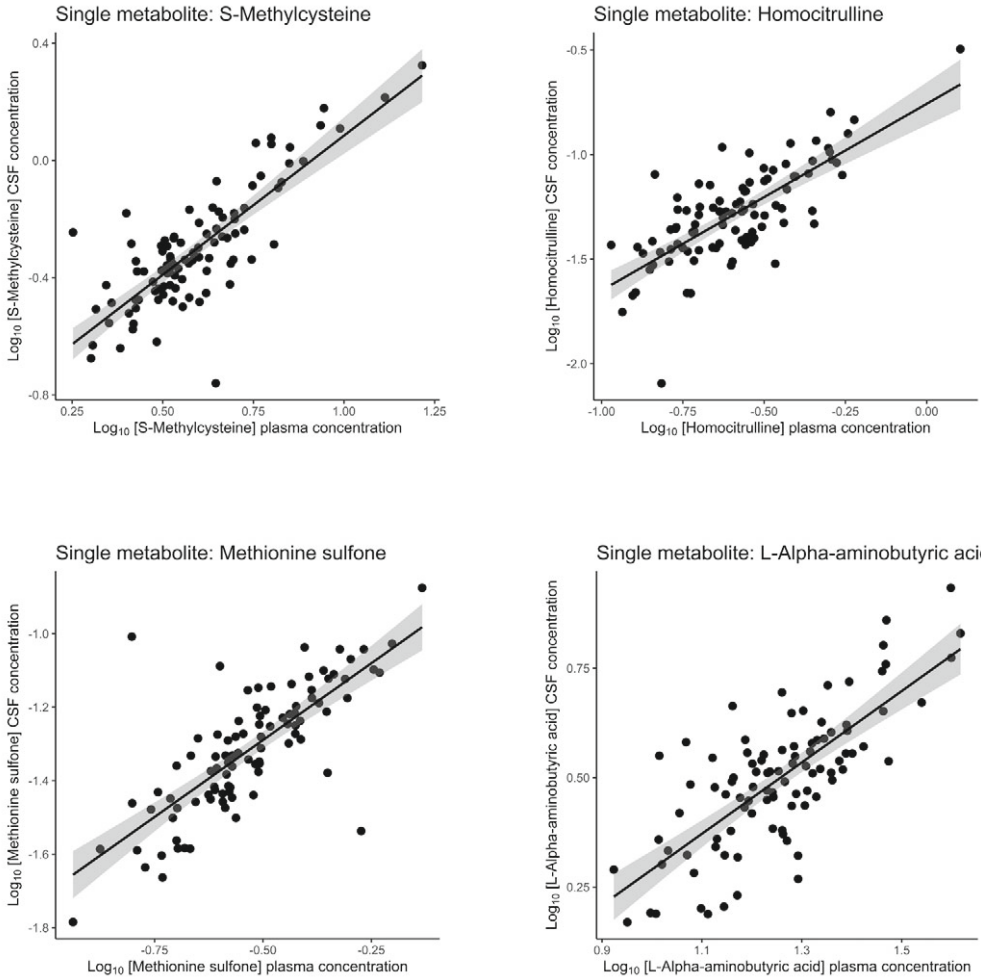
† BMI = body mass index;‡ SD = standard deviation

Table 2 Plasma and CSF concentrations in healthy controls

Metabolite	Plasma (µM)			CSF† (µM)			Plasma/ CSF Ratio
	Median	Q1	Q3	Median	Q1	Q3	Median
Ethanolamine	12.088	11.004	13.338	22.679	20.239	24.404	.55
Putrescine	.105	.083	.13	.161	.123	.198	.67
L-Glutamine	665.43	600.158	714.841	591.26	538.944	637.634	1.12
N6-N6-N6-Trimethyl-L-lysine	.468	.39	.568	.404	.365	.447	1.17
Gamma-aminobutyric acid	.146	.123	.169	.099	.071	.121	1.51
L-Homoserine	.134	.12	.15	.06	.053	.066	2.25
Gamma-Glutamylglutamine	4.693	4.109	5.715	1.98	1.778	2.226	2.36
L-Arginine	78.796	66.678	94.937	24.216	21.507	26.295	3.4
L-Threonine	135.285	117.762	152.372	36.591	31.929	42.588	3.63
SDMA	.61	.549	.68	.163	.139	.199	3.65
L-Serine	108.619	92.113	121.252	28.161	25.711	30.606	3.78
L-Histidine	60.536	55.485	65.125	15.007	13.649	16.842	4.11
Homocitrulline	.258	.186	.331	.051	.037	.073	4.81
L-Phenylalanine	55.817	51.336	62.03	10.273	9.267	11.781	5.41
L-Alpha-aminobutyric acid	18.218	14.484	21.797	3.242	2.518	3.757	5.72
L-Methionine	23.506	21.018	26.282	3.904	3.544	4.455	5.8
L-Tyrosine	50.099	42.364	55.329	7.791	6.986	9.175	5.97
3-Methoxytyrosine	.089	.077	.106	.015	.012	.017	6.11
Methionine sulfone	.285	.242	.376	.047	.037	.063	6.17
L-Lysine	186.795	167.22	204.311	29.905	26.963	32.829	6.22
L-Asparagine	52.278	47.707	58.596	8.114	7.484	9.141	6.47
L-Methionine sulfoxide	.736	.648	.828	.099	.082	.114	7.49
S-Methylcysteine	3.64	3.106	4.86	.461	.369	.622	7.92
L-2-aminoadipic acid	.699	.583	.906	.082	.072	.1	8.22
L-Leucine	277.057	244.39	318.336	33.685	29.966	39.018	8.45
L-Alanine	359.072	317.329	421.116	34.465	29.743	43.473	10.06
L-Isoleucine	116.684	103.751	132.595	11.702	10.009	13.24	10.46
Ornithine	65.305	55.212	74.791	5.359	4.799	6.309	11.6
L-Valine	242.729	212.533	277.108	20.728	17.63	23.222	11.88
L-4-hydroxy-L-proline	7.569	5.951	10.286	.664	.474	.838	11.97
Gamma-L-glutamyl-L-alanine	.507	.411	.61	.041	.036	.046	12.01
Citrulline	29.169	25.484	33.029	2.168	1.836	2.638	12.93
Taurine	97.824	71.535	127.601	7.397	6.461	8.244	13.03
Glutathione	5.577	4.595	6.355	.337	.275	.428	15.83
L-Tryptophan	56.353	49.209	63.568	2.385	2.146	2.703	23.28
Glycine	296.343	237.466	333.362	8.686	7.506	10.336	33.39
L-Pipecolic acid	11.26	9.687	13.694	.34	.258	.41	34.28
L-Glutamic acid	44.036	34.697	63.195	.399	.35	.447	110.32
L-Proline	178.184	140.29	218.478	.922	.751	1.233	183.31

Metabolites are ranked by plasma/CSF ratio from smallest to largest ratio. The table is based on pre-processed data.
†CSF = cerebrospinal fluid; Q1 = First quartile; Q3 = Third quartile.

Figure 1 Correlation plots of amines with strongest correlation between plasma and CSF in healthy controls



Log_{10}-transformed concentrations are plotted with plasma concentrations on the X-axis and CSF concentrations on the Y-axis. Values below [mean - 4 * SD] and above [mean + 4 * SD] were replaced with these cut-offs. Plots are based on uncorrected data. The black line indicates the linear regression line and the grey area the 95% confidence interval. CSF = cerebrospinal fluid.

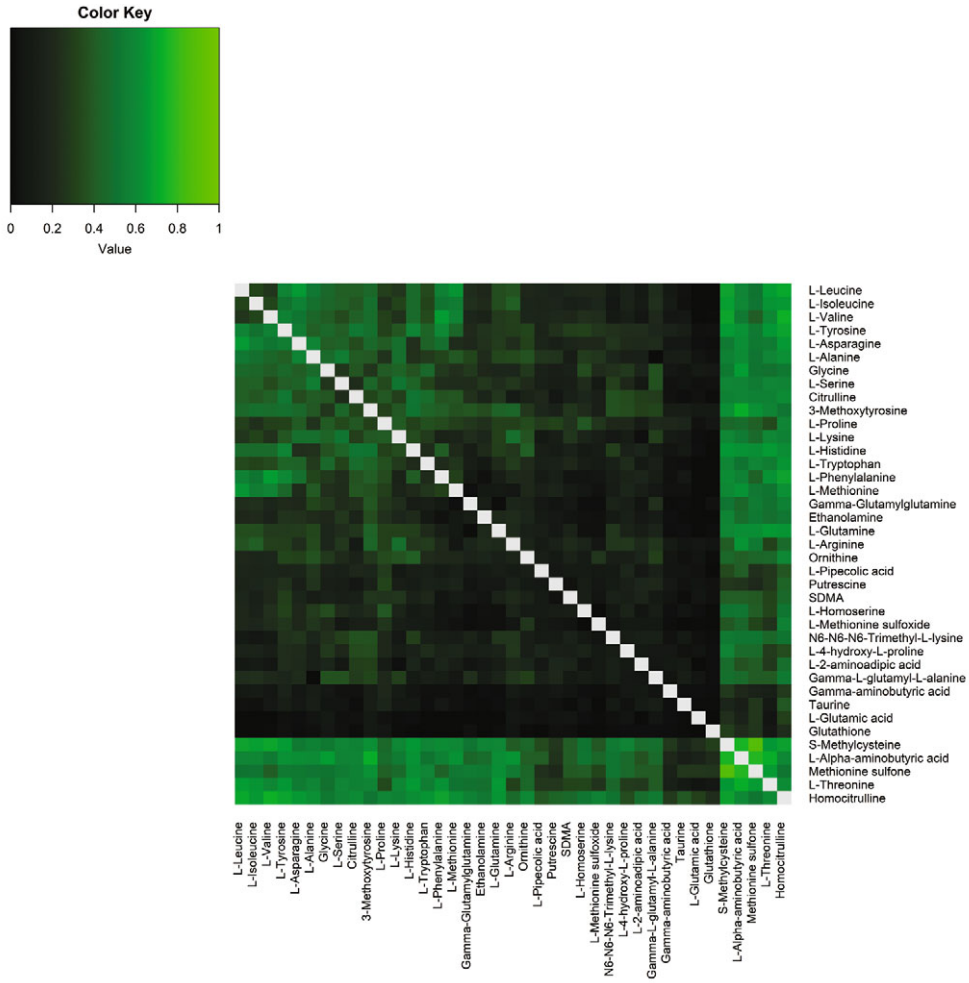
Amines in blood plasma and CSF of participants with migraine

Concentrations of amines

Amine concentrations in plasma and CSF were reliably measured for 197 migraine participants (n = 98 for migraine without aura MO, n = 99 for migraine with aura MA) (for clinical characteristics see **Table 4**). Plasma concentrations were higher than CSF concentrations for most amines in MO participants (**Table 5**) and MA participants (**Table 6**), similar to observations in healthy controls. Also in migraine samples ethanolamine and putrescine had higher concentrations in CSF than

in plasma; L-Glutamine was the most abundant amine in CSF (MO: median: 589 mM and interquartile range: 516-637 mM, MA: median: 568 mM and interquartile range: 516-616 mM) with similar concentrations in blood (MO: median: 674 mM and interquartile range: 595-752 mM, MA: median: 653 mM and interquartile range 587-712 mM).

Figure 2 Correlation matrix of all possible metabolite ratios



Coefficients of determination (R^2) of all possible metabolite ratios are plotted in healthy controls. The higher the R^2 the brighter the square. Grey squares: ratio not applicable.

Single plasma/CSF metabolite correlations

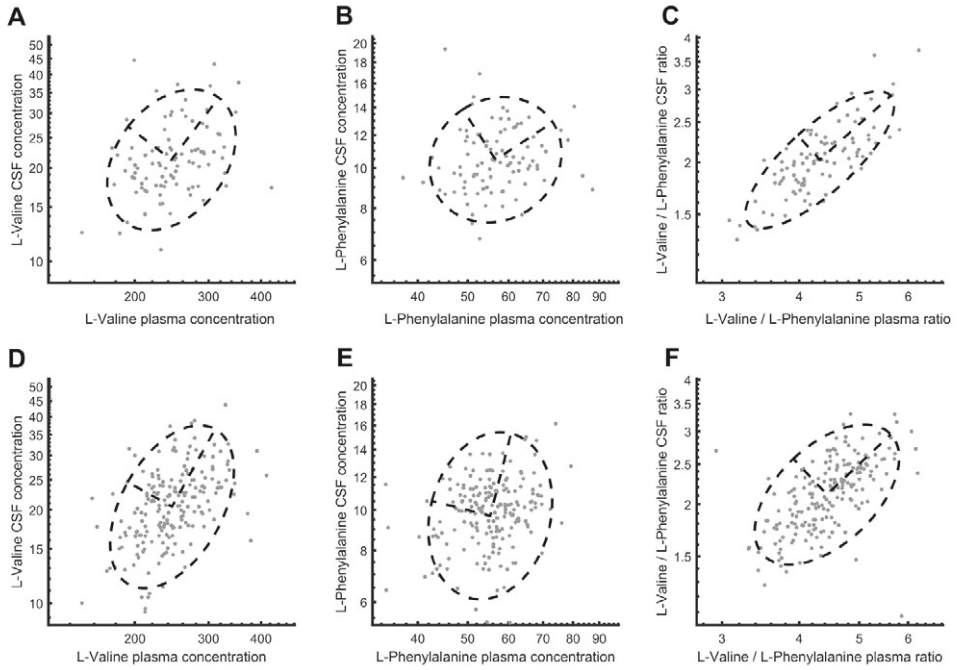
Additionally, we assessed single correlations in samples from participants with migraine (**Table 7**, **Table S4**). Similar as in healthy controls, most amines had low correlation between plasma and

CSF concentrations. Four (homocitrulline, S-methylcysteine, methionine sulfone and L-alpha-aminobutyric acid) out of 39 amines (10.3%) had a correlation coefficient ≥ .70. These four amines were the same as in healthy controls, and were the same for both MO and MA participants. Overall, there were no clear differences between correlations between the migraine subgroups.

Plasma/CSF ratios metabolite correlations

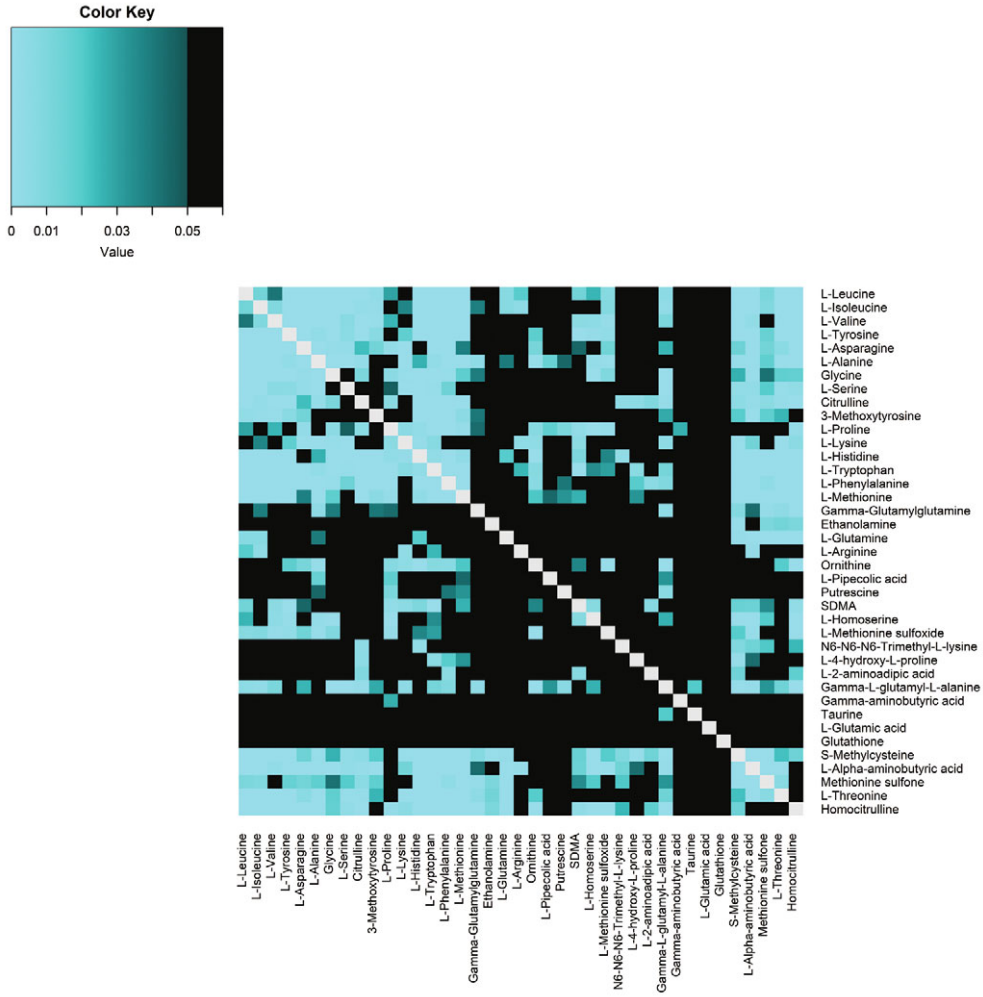
Next, we assessed the same metabolite ratios in plasma and CSF, as illustrated in **Figure 5** (**Table S5**). On visual inspection the results largely look the same as obtained for the healthy controls, that is, the correlation of all amines with S-methylcysteine, L-alpha-aminobutyric acid, methionine sulfone, L-threonine and homocitrulline remain high. However, also a clear difference was observed when comparing the correlation matrix plot of healthy controls (**Figure 2**) and that of participants with migraine (**Figure 5**).

Figure 3 Example of single-metabolite *vs.* metabolite ratio correlation



Concentrations are plotted plasma concentrations on the $\log_{10}$ transformed X-axis and CSF concentrations on the $\log_{10}$ transformed Y-axis. Values below [mean - 4 * SD] and above [mean + 4 * SD] were replaced with these cut-offs. Plots are based on uncorrected data. (a) Correlation between plasma and CSF concentrations of L-Valine in healthy volunteers. (b) Correlation between plasma and CSF concentrations of L-Phenylalanine in healthy controls. (c) Correlation between plasma ratio of L-Valine/L-Phenylalanine and CSF ratio of L-Valine/L-Phenylalanine in healthy controls. (d) Correlation between plasma and CSF concentrations of L-Valine in participants with migraine. (e) Correlation between plasma and CSF concentrations of L-Phenylalanine in participants with migraine. (f) Correlation between plasma ratio of L-Valine/L-Phenylalanine and CSF ratio of L-Valine/L-Phenylalanine in participants with migraine. The dashed circle represents the 95% confidence interval of the Gaussian distribution. CSF = cerebrospinal fluid.

Figure 4 *P*-values of R^2 gain



R^2 is significantly higher than the single R^2 in healthy controls (FDR < .05). The lower the p-value value the brighter blue the square, black squares are not significantly different (FDR < .05). Grey squares: *p*-value not applicable. R^2 gain = the improvement of Plasma/CSF ratio correlations compared to single-metabolite ratio correlations.

Comparison between healthy controls and participants with migraine

To determine whether the metabolite ratio model gives insight in disease pathology, we statistically compared the R^2 correlations from healthy controls and participants with migraine. For single-metabolite correlations, there were no significant differences between healthy controls and participants with migraine (**Table S6**). However, for the plasma/CSF ratios metabolite correlations, differences in R^2 ratio were observed when comparing data of healthy controls and participants with migraine (**Table S7**, **Figure S3**). When comparing these correlations, significant differences after multiple testing correction (FDR < .05) were observed for the following combinations; L-Valine/L-Phenylalanine (FDR = .003), L-Arginine/S-Methylcysteine (FDR = .003), L-Alanine/L-4-hydroxy-L-proline (FDR = .015), L-Valine/L-4-hydroxy-L-proline (FDR = .025), L-Leucine/L-Phenylalanine (FDR = .028), and Gamma-L-glutamyl-L-alanine/L-4-hydroxy-L-proline (FDR = .041) $((\frac{\text{Amine x}_{plasma}}{\text{Amine y}_{plasma}} / \frac{\text{Amine x}_{CSF}}{\text{Amine y}_{CSF}})control \neq (\frac{\text{Amine x}_{plasma}}{\text{Amine y}_{plasma}} / \frac{\text{Amine x}_{CSF}}{\text{Amine y}_{CSF}})migraineur)$. Most relevant are probably L-Valine/L-Phenylalanine and L-Leucine/L-Phenylalanine, because the single-metabolite plasma/CSF ratios between healthy controls and participants with migraine did not largely differ, where those of 4-hydroxy-proline and S-Methylcysteine seem to do, albeit not significant (**Table S6**). This means that the difference between healthy volunteers and patients with migraine in plasma/CSF R^2 ratio is not merely driven by one metabolite but rather the mechanism behind it. For L-Valine/L-Phenylalanine ratio and the L-Leucine/L-Phenylalanine the ratio correlation was lower in migraine participants compared to that in healthy controls (**Figure 3E-F**).

Table 3 Single CSF-plasma correlations for single metabolites in healthy controls

Metabolite	*r*	R^2	95% c.i.	FDR
Homocitrulline	**.75**	.56	.41 - .68	6.16E-17
S-Methylcysteine	**.75**	.56	.41 - .68	6.16E-17
Methionine sulfone	**.75**	.56	.41 - .68	6.16E-17
L-Alpha-aminobutyric acid	**.73**	.54	.39 - .67	2.71E-16
L-Threonine	.69	.48	.32 - .62	5.03E-14
3-Methoxytyrosine	.64	.40	.24 - .55	3.11E-11
L-Arginine	.57	.32	.17 - .48	1.33E-08
L-Glutamine	.47	.22	.09 - .38	5.97E-06
L-Lysine	.46	.21	.08 - .37	1.11E-05
L-Serine	.46	.21	.08 - .37	1.20E-05
L-Tyrosine	.45	.21	.08 - .36	1.37E-05
L-Proline	.44	.20	.07 - .35	2.27E-05
Citrulline	.44	.19	.07 - .35	2.58E-05
Glycine	.43	.18	.06 - .34	4.56E-05
L-Asparagine	.42	.17	.05 - .33	6.85E-05
Ethanolamine	.41	.17	.05 - .32	9.12E-05
L-Alanine	.41	.17	.05 - .32	9.96E-05
N6-N6-N6-Trimethyl-L-lysine	.40	.16	.05 - .32	1.04E-04
L-4-hydroxy-L-proline	.36	.13	.03 - .28	6.05E-04
L-Isoleucine	.35	.12	.03 - .27	8.95E-04
L-Leucine	.35	.12	.02 - .26	1.12E-03
Ornithine	.34	.12	.02 - .26	1.29E-03
L-Valine	.33	.11	.02 - .25	1.70E-03
Gamma-Glutamylglutamine	.33	.11	.02 - .25	2.09E-03
Putrescine	.32	.10	.02 - .25	2.13E-03
L-2-aminoadipic acid	.30	.09	.01 - .23	4.59E-03
L-Pipecolic acid	.29	.08	.01 - .21	7.34E-03
L-Histidine	.25	.06	.00 - .18	2.31E-02
Taurine	.18	.03	.00 - .14	1.01E-01
L-Tryptophan	.17	.03	.00 - .13	1.38E-01
L-Homoserine	.16	.03	.00 - .13	1.40E-01
L-Methionine	.16	.02	.00 - .12	1.59E-01
Gamma-L-glutamyl-L-alanine	.15	.02	.00 - .12	1.64E-01
SDMA	.12	.01	.00 - .10	2.97E-01
L-Glutamic acid	.09	.01	.00 - .08	4.56E-01
L-Phenylalanine	.08	.01	.00 - .08	4.56E-01
Gamma-aminobutyric acid	.08	.01	.00 - .08	4.80E-01
L-Methionine sulfoxide	.06	.00	.00 - .07	5.58E-01
Glutathione	.02	.00	.00 - .05	8.63E-01

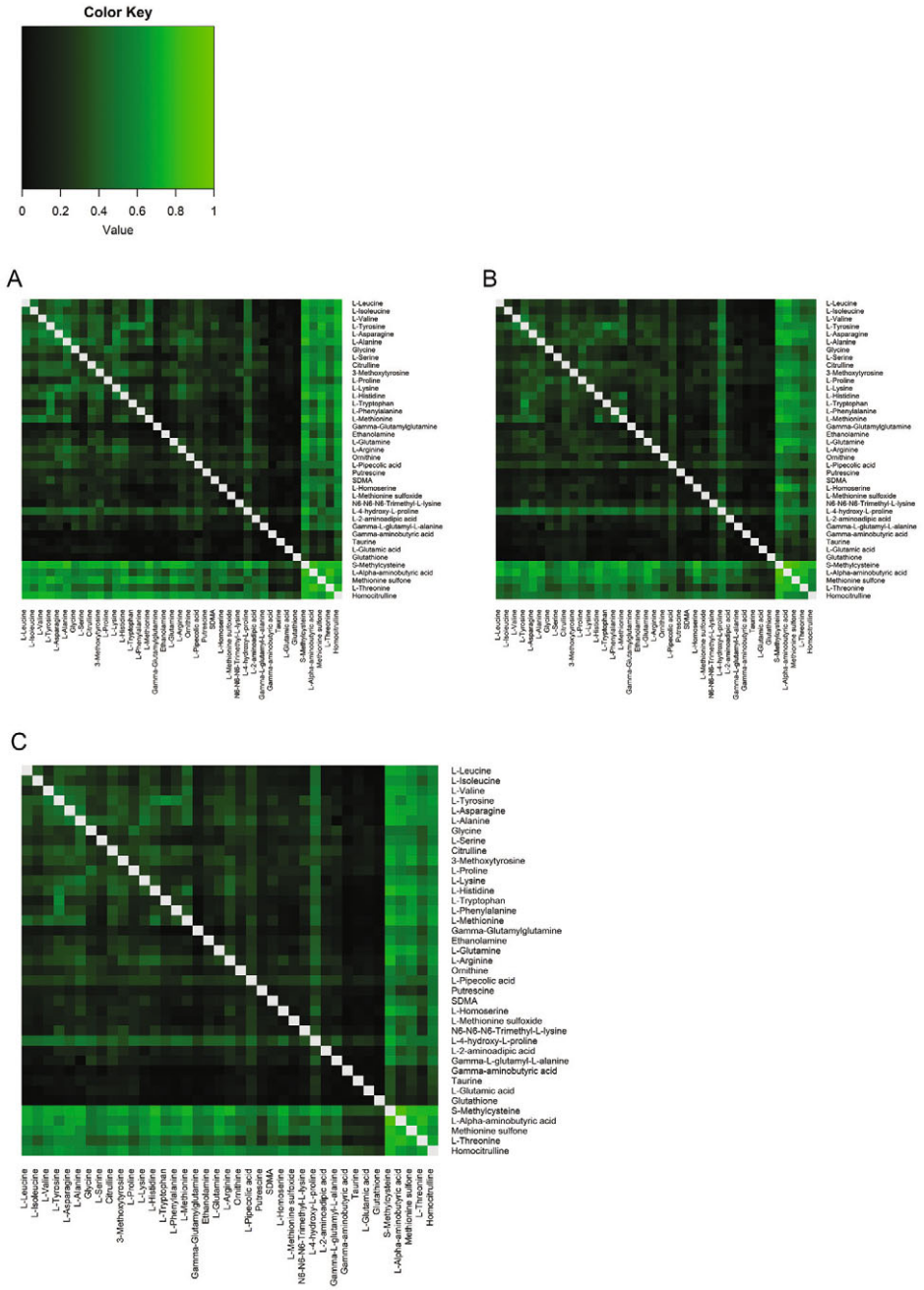
Metabolites are ranked by value of this coefficient and values ≥ .7 are printed in bold. Table is based on corrected data. *r* = Pearson correlation coefficient; R^2 = coefficient of determination; 95% c.i. = 95% confidence interval of R^2; FDR = false discovery rate.

Table 4 Clinical characteristics of migraine participants.

	MO[§] participants	MA[¶] participants
Number of participants	98	99
Subject characteristics		
Females (n, (%))	60 (61.2)	66 (66.7)
Years of age (± SD‡)	42.01 (± 12.9)	41.7 (± 13.5)
BMI† (± SD‡)	23.6 (± 2.5)	24.0 (± 2.8)
Smoking (n, (%))	13 (13.3)	13 (13.1)
Overnight fasting		
Fasting time in h (± SD‡)	11.9 (± 1.6)	11.7 (± 1.6)
Sampling characteristics		
Opening pressure in mmH2O (± SD‡)	18.0 (± 4.7)	18.8 (± 4.1)
CSF characteristics		
Erythrocytes count/3µL (± SD‡)	130 (± 505)	2276 (± 20577)
Leukocytes count/3µL (± SD‡)	5 (± 7)	21 (± 89)
Protein concentration in g/L (± SD‡)	.35 (± .10)	.36 (± .25)
Glucose in mmol/L (± SD‡)	3.1 (± .2)	3.2 (± .3)

†BMI = body mass index; ‡SD = standard deviation; §MO = migraine without aura; ¶MA = migraine without aura.

Figure 5 Correlation matrix of all possible metabolite ratios in migraine patients



Coefficients of determination (R^2) of all possible metabolite ratios are plotted in (a) migraine without aura participants (b) migraine with aura participants. (c) all migraine participants together. The higher the correlation the brighter the square.

Discussion

By investigating 40 amino acids and biogenic amines, targeted by our validated mass spectrometry platform, in plasma and CSF in healthy controls, we show that amine concentrations in plasma do not correlate well with those in CSF. Hence, plasma concentrations *per se* are poor predictors of CSF concentrations for the majority of these metabolites. However, when studied as ratios, several amines correlated much better, indicating that ratios have an increased predictive ability. A possible explanation for this finding is that ratios are tightly regulated by blood-brain-barrier (BBB) transport systems and most likely indicating the cotransport of amines, but not so much the distribution of individual amines over the various body fluid compartments.

There have been a limited number of studies on the correlation between plasma and CSF concentrations of amines. The studies typically focused on proteinogenic amino acids and similar to our results did not yield overt correlations.[10, 11] However, in our study, we measured additional amines and four of them showed strong correlation, namely homocitrulline, S-methylcysteine, methionine sulfone, and L-alpha-aminobutyric acid. Homocitruline has been associated with energy metabolism and cerebellar disfunction in the animal studies [18, 19] and S-methylcysteine can act neuroprotective after certain types of damage,[20] however, their exact biological roles in the brain are not fully known.

Previous studies did not study whether *ratios of amines* correlate between plasma and CSF. In our study, ratios showed better correlation, especially for pairs of amino acids. The regulation of nutrient and metabolite movement between blood and the brain is co-regulated by the action of the blood-brain-barrier. As one of its consequences, there is a concentration gradient between the brain and the rest of the body with the concentration of amino acids in the brain typically being lower than in plasma,[21, 22] this was in line with our study where we saw lower concentration in the CSF for most amines compared to plasma. Both this gradient and several transporters regulate the transport of amino acids across the BBB.[23-26] The amino acids that showed high correlation in their ratios (L-leucine/L-methionine, L-leucine/L-asparagine, L-valine/L-phenylalanine) are all transported by the L amino acid transport system (LAT/L1), this is a sodium-independent transmembrane antiporter, consisting of two protein subunits, either a catalytic permease SLC7A5/LAT1 or SLC7A8/LAT2 and a regulatory glycoprotein (SLC3A2).[24, 27, 28] Substrates carried by L1 are asparagine, glutamine, leucine, valine, methionine, histidine, isoleucine, tyrosine, tryptophan, phenylalanine, and threonine.[24] The L1 system is the most important source by which essential neutral amino acids (NAAs) gain access to the brain.[24] L1 imports large and neutral amino acids in exchange for intracellular amino acids. [29] The high correlations that were observed in our study hence may reflect tight regulation by this transporter. It is likely that the relative concentration, i.e. the abundance of one amino acid compared to a second amino acid, is important for transport.

Classic neurotransmitters glutamic acid and GABA did not show strong correlations between plasma and CSF concentrations, neither on their own nor as ratio. Concentrations of amines

related to the synthesis of neurotransmitters, such as glutamine and glutathione for glutamic acid and GABA and tryptophan for serotonin,[9] also did not show strong correlation between plasma and CSF. Thus, plasma levels of these metabolites seem not be useful to infer or predict CSF levels.

In our interictal migraine group, significant changes of six amine (which were mainly amino acids) related ratios were observed when comparing the plasma/CSF R^2 ratios of participants with migraine to those of healthy controls. The most likely explanation for this finding would be an underlying malfunctioning of the respective amine cotransporter in migraine, resulting in a difference between R^2 ratios of patients and healthy volunteers. Most of the changed ratios are related to the L1 transporter system. It has previously been demonstrated that the ratio between amino acids is important for the entry to the brain and that there is competition between these NAAs for the entry into the brain.[30]

In addition, the L-Arginine/S-Methylcysteine ratio was increased in migraine participants, while these amines are not transported by the L1 system. Arginine is transported by the y^+ or the cationic amino acid (CAA) transport (CAT) system.[24, 26] The CAT system is primarily a CAA transporter, but also exhibits weak interactions with NAAs (phenylalanine, threonine, histidine, valine, methionine, serine, glutamine, alanine, and glycine).[24, 31] Arginine levels in CSF were previously found to be associated with migraine by our group [13]. Of note, arginine forms a caveolar complex with endothelial nitric oxide synthase (eNOS) to form nitric oxide (NO) [32]. Interestingly, NO has previously been implicated in migraine pathophysiology [33-35].

Although the current study does not allow for a detailed analysis of the mentioned transport mechanisms, the ratios can still be useful for assessing CNS metabolism of amines. For example, if a person has a high leucine/methionine ratio in plasma, this is likely also the case in CSF. If this plasma ratio changes over time within this person, due to a certain disease process, one can envisage that the CSF ratio may also change. Of course, this must be proven in a longitudinal design with additional evidence that there is indeed a relationship between the specific ratio and the disease process.

The strengths of this study are the large number of participants and the wide coverage of amine molecules measured. There are also some limitations. We did not have repeated measurements for individual participants and could not validate our findings for dynamic changes. Hence, future (dynamic) studies should replicate the identified correlations.

In conclusion, for individual amine metabolites there is generally poor correlation between the concentrations in plasma and CSF. However, ratios of certain amines show a good correlation of plasma with CSF. Plasma amine measurements thus have the potential to be used for predicting CSF metabolite levels, which can be very relevant to assess fluctuations of amino acids and biogenic amines metabolism in paroxysmal CNS disorders.

Table 5 Plasma and CSF concentrations in migraine without aura participants

Metabolite	Plasma (µM)			CSF† (µM)			Plasma/ CSF Ratio
	Median	Q1	Q3	Median	Q1	Q3	Median
Ethanolamine	12.297	11.175	13.553	21.821	18.822	23.871	.58
Putrescine	.100	.077	.125	.144	.115	.206	.62
L.Glutamine	674.580	595.338	752.180	589.130	515.600	637.313	1.19
O.Phosphoethanolamine	10.010	6.639	13.745	7.752	6.437	8.995	1.22
N6.N6.N6.Trimethyl.L.lysine	.513	.418	.610	.384	.343	.454	1.28
Gamma.aminobutyric.acid	.141	.127	.163	.100	.074	.129	1.41
L.Homoserine	.132	.119	.144	.058	.053	.064	2.31
Gamma.Glutamylglutamine	5.014	4.231	5.832	2.091	1.833	2.386	2.49
L.Arginine	80.671	60.536	94.243	23.307	20.515	26.424	3.50
L.Threonine	134.514	116.573	156.030	37.440	32.806	44.239	3.62
SDMA	.600	.540	.680	.152	.131	.190	3.79
L.Serine	104.086	95.945	119.437	27.274	24.831	31.276	3.95
L.Histidine	60.834	57.208	66.340	14.658	13.246	16.115	4.33
Homocitrulline	.242	.170	.332	.049	.037	.074	4.83
L.Phenylalanine	56.070	52.176	60.344	10.093	8.901	11.138	5.77
L.Alpha.aminobutyric.acid	18.149	15.071	22.688	3.119	2.453	3.855	6.00
L.Tyrosine	49.079	43.117	55.010	8.053	6.849	9.219	6.20
L.Methionine	23.253	21.219	25.536	3.905	3.331	4.257	6.22
3.Methoxytyrosine	.093	.083	.108	.015	.013	.017	6.31
Methionine.sulfone	.304	.220	.381	.045	.035	.060	6.37
L.Asparagine	51.826	46.566	57.312	7.849	6.839	9.082	6.66
L.Lysine	200.450	175.718	217.740	29.707	24.925	34.051	6.68
L.Methionine.sulfoxide	.723	.619	.806	.097	.081	.117	7.22
S.Methylcysteine	3.744	3.037	4.854	.464	.371	.647	8.14
L.Leucine	280.280	255.741	320.453	32.286	28.673	39.077	8.59
L.2.aminoadipic.acid	.775	.648	.997	.092	.080	.104	8.61
L.Alanine	374.163	326.391	429.543	35.577	30.439	41.706	10.84
L.Isoleucine	118.753	104.690	133.144	11.132	9.200	13.814	11.03
Taurine	92.333	69.557	116.352	7.763	6.767	8.755	12.16
L.Valine	245.782	223.515	277.096	20.551	17.680	24.981	12.36
Gamma.L.glutamyl.L.alanine	.532	.441	.672	.044	.040	.051	12.60
L.4.hydroxy.L.proline	8.812	6.417	12.451	.685	.554	.938	12.73
Ornithine	65.407	51.951	79.256	5.250	4.162	6.138	12.83
Citrulline	28.087	23.808	35.088	2.151	1.697	2.619	13.21
Glutathione	5.661	4.636	6.485	.400	.312	.493	14.56
L.Tryptophan	57.474	51.837	61.938	2.284	2.000	2.700	24.75
Glycine	289.064	235.273	334.399	8.261	7.287	9.859	32.34
L.Pipecolic.acid	10.778	8.670	14.545	.323	.257	.403	35.21
L.Glutamic.acid	40.028	32.485	54.316	.387	.355	.478	107.70
L.Proline	171.004	140.167	209.404	.984	.805	1.235	181.27

Metabolites are ranked by plasma/CSF ratio from smallest to largest ratio. The table is based on uncorrected data.
†CSF = cerebrospinal fluid. Q1 = First quartile; Q3 = Third quartile.

Table 6 Plasma and CSF concentrations in migraine with aura participants.

Metabolite	Plasma (µM)			CSF† (µM)			Plasma /CSF Ratio
	Median	Q1	Q3	Median	Q1	Q3	Median
Ethanolamine	11.578	10.470	12.732	20.911	18.792	22.715	.55
Putrescine	.096	.079	.119	.157	.122	.198	.63
L.Glutamine	653.450	586.971	711.809	567.820	516.139	615.613	1.15
N6.N6.N6.Trimethyl.L.lysine	.469	.395	.564	.396	.364	.433	1.18
O.Phosphoethanolamine	9.092	6.046	13.232	6.789	6.028	7.943	1.24
Gamma.aminobutyric.acid	.141	.126	.162	.098	.073	.123	1.46
L.Homoserine	.132	.117	.147	.060	.053	.065	2.28
Gamma.Glutamylglutamine	5.079	4.119	5.593	1.968	1.682	2.271	2.52
L.Arginine	72.078	60.850	81.213	22.499	19.179	24.445	3.22
SDMA	.575	.511	.640	.155	.135	.182	3.58
L.Threonine	131.388	114.403	146.662	37.090	30.923	41.640	3.61
L.Serine	109.030	93.263	124.880	28.032	24.813	32.959	3.86
L.Histidine	60.959	55.442	65.949	14.569	13.181	16.302	4.12
Homocitrulline	.221	.183	.321	.049	.039	.066	4.78
L.Phenylalanine	55.540	50.081	60.634	10.001	8.832	11.155	5.47
L.Alpha.aminobutyric.acid	18.415	15.313	21.751	3.141	2.386	4.094	5.78
L.Tyrosine	49.443	42.673	56.030	8.247	6.846	9.526	5.81
L.Methionine	22.550	20.440	24.734	3.704	3.175	4.290	5.87
3.Methoxytyrosine	.088	.077	.105	.014	.013	.017	6.05
Methionine.sulfone	.290	.230	.363	.048	.036	.061	6.21
L.Asparagine	49.761	46.959	54.779	7.571	6.818	8.638	6.57
L.Lysine	188.289	171.233	214.701	29.191	25.893	32.633	6.64
L.Methionine.sulfoxide	.730	.624	.828	.098	.079	.120	7.51
S.Methylcysteine	3.529	2.837	4.965	.491	.364	.640	7.56
L.2.aminoadipic.acid	.716	.593	.899	.088	.074	.102	8.26
L.Leucine	264.337	240.759	307.927	32.152	27.056	37.816	8.44
L.Isoleucine	112.672	99.368	130.394	10.764	9.395	12.632	10.49
L.Alanine	379.229	326.339	442.442	34.173	30.171	40.508	10.74
Taurine	83.719	64.708	113.254	7.117	6.093	8.013	11.14
L.Valine	235.246	213.624	269.877	21.665	16.751	24.563	11.86
L.4.hydroxy.L.proline	7.770	6.020	10.363	.644	.485	.832	12.10
Gamma.L.glutamyl.L.alanine	.540	.462	.651	.042	.037	.048	12.55
Citrulline	27.375	22.837	31.007	2.021	1.758	2.428	12.69
Ornithine	62.759	51.411	78.293	5.143	4.161	6.015	13.09
Glutathione	5.404	4.490	6.696	.327	.267	.445	15.62
L.Tryptophan	54.546	47.679	59.223	2.338	2.098	2.608	23.17
Glycine	286.608	235.756	352.980	8.037	7.021	10.223	34.24
L.Pipecolic.acid	10.626	8.529	12.854	.313	.258	.397	34.61
L.Glutamic.acid	39.946	28.893	58.014	.391	.359	.458	99.13
L.Proline	168.383	133.030	213.782	.910	.749	1.177	174.50

Metabolites are ranked by plasma/CSF ratio from smallest to largest ratio. The table is based on uncorrected data.
†CSF = cerebrospinal fluid. Q1 = First quartile; Q3 = Third quartile.

Table 7 Plasma/CSF correlations for single metabolites in migraine participants

	MO				MA			
Metabolite	***r***	**R^2**	**95% c.i.**	**FDR**	***r***	**R^2**	**95% c.i.**	**FDR**
S-Methylcysteine	**.84**	.71	.60 - .80	4.44E-26	**.86**	.74	.64 - .82	1.04E-28
Methionine sulfone	**.79**	.63	.50 - .74	2.94E-21	**.79**	.62	.49 - .73	3.25E-21
Homocitrulline	**.78**	.61	.47 - .72	2.86E-20	**.72**	.51	.36 - .64	8.16E-16
L-Alpha-aminobutyric acid	**.73**	.54	.39 - .66	9.66E-17	**.80**	.63	.50 - .74	1.43E-21
L-Threonine	.68	.47	.31 - .61	6.85E-14	.62	.39	.23 - .53	4.40E-11
L-4-hydroxy-L-proline	.63	.40	.24 - .55	1.86E-11	.64	.41	.26 - .56	6.02E-12
L-Arginine	.56	.32	.17 - .47	9.56E-09	.45	.21	.08 - .36	7.27E-06
3-Methoxytyrosine	.55	.30	.15 - .45	2.93E-08	.60	.36	.20 - .51	3.77E-10
Citrulline	.54	.29	.15 - .45	3.71E-08	.48	.23	.09 - .38	2.25E-06
L-Alanine	.54	.29	.14 - .45	3.71E-08	.44	.19	.07 - .34	1.84E-05
L-Lysine	.53	.28	.14 - .44	5.99E-08	.35	.13	.03 - .27	5.73E-04
L-Pipecolic acid	.52	.27	.12 - .42	1.57E-07	.56	.32	.17 - .47	6.38E-09
L-Tyrosine	.48	.23	.10 - .39	1.55E-06	.51	.26	.12 - .41	3.72E-07
Ornithine	.46	.21	.08 - .37	4.85E-06	.33	.11	.02 - .24	1.66E-03
L-Asparagine	.45	.21	.08 - .36	7.05E-06	.37	.14	.03 - .28	2.96E-04
Glycine	.41	.17	.05 - .32	7.11E-05	.36	.13	.03 - .27	4.58E-04
L-Proline	.38	.15	.04 - .29	2.27E-04	.47	.22	.09 - .38	2.78E-06
L-Valine	.37	.14	.03 - .28	3.74E-04	.38	.14	.04 - .29	2.23E-04
N6-N6-N6-Trimethyl-L-lysine	.36	.13	.03 - .28	5.13E-04	.43	.18	.06 - .34	2.53E-05
L-Glutamine	.35	.12	.02 - .26	9.34E-04	.40	.16	.05 - .31	8.91E-05
L-Histidine	.32	.11	.02 - .24	2.03E-03	.25	.06	.00 - .19	1.54E-02
L-Isoleucine	.32	.10	.02 - .24	2.04E-03	.22	.05	.00 - .16	4.06E-02
Gamma-L-glutamyl-L-alanine	.31	.09	.01 - .23	3.78E-03	.21	.04	.00 - .15	4.76E-02
L-2-aminoadipic acid	.30	.09	.01 - .22	4.49E-03	.41	.16	.05 - .31	8.31E-05
Putrescine	.29	.08	.01 - .21	6.09E-03	.19	.04	.00 - .14	7.30E-02
L-Serine	.28	.08	.01 - .21	7.59E-03	.40	.16	.05 - .31	8.41E-05
L-Leucine	.25	.06	.00 - .18	1.94E-02	.29	.08	.01 - .21	5.45E-03
SDMA	.23	.05	.00 - .17	2.88E-02	.10	.01	.00 - .09	3.79E-01
L-Methionine	.21	.04	.00 - .15	5.37E-02	.30	.09	.01 - .22	4.12E-03
L-Methionine sulfoxide	.20	.04	.00 - .15	6.20E-02	.19	.04	.00 - .14	6.49E-02
Gamma-Glutamylglutamine	.13	.02	.00 - .11	2.42E-01	.23	.05	.00 - .17	2.79E-02
Ethanolamine	.12	.01	.00 - .10	3.11E-01	.40	.16	.05 - .31	8.79E-05
Gamma-aminobutyric acid	-.11	.01	.00 - .09	3.29E-01	.04	.00	.00 - .06	7.21E-01
Taurine	-.11	.01	.00 - .09	3.35E-01	.20	.04	.00 - .15	6.29E-02
L-Homoserine	.11	.01	.00 - .09	3.37E-01	.32	.10	.02 - .24	2.19E-03
L-Phenylalanine	-.05	.00	.00 - .06	6.95E-01	.33	.11	.02 - .25	1.59E-03
L-Glutamic acid	.04	.00	.00 - .06	7.15E-01	.05	.00	.00 - .06	6.31E-01
Glutathione	.03	.00	.00 - .05	7.67E-01	.09	.01	.00 - .08	3.91E-01
L-Tryptophan	-.03	.00	.00 - .05	7.67E-01	.08	.01	.00 - .08	4.36E-01

Metabolites are ranked by value of this coefficient and values ≥ .7 are printed in bold. Table is based on corrected data. *r* = Pearson correlation coefficient; R^2 = coefficient of determination; 95% c.i. = 95% confidence interval of R^2; FDR = false discovery rate MO = migraine with aura, MA = migraine without aura.

Conflict of interest statement: The authors declare no conflict of interest.

Author's contributions: *Conceptualization*: A.V.E.H., J.B.v.K., R.M.v.D., G.L.J.O., T.H., G.M.T. and A.M.J.M.v.d.M.; *Methodology*: A.V.E.H., J.B.v.K., R.M.v.D., and A.C.H.; *Investigation*: R.M.v.D., G.L.J.O., and A.C.H.; *Formal analysis*: A.V.E.H., J.B.v.K. and R.M.v.D.; *Resources*: T.H., M.D.F., G.M.T. and A.M.J.M.v.d.M.; *Data curation*: A.V.E.H., J.B.v.K., R.M.v.D., G.L.J.O. A.C.H.; *Writing – original draft*: A.V.E.H., J.B.v.K. and R.M.v.D.; *Writing – review & editing*: G.L.J.O., A.C.H., T.H., G.M.T. and A.M.J.M.v.d.M.; *Visualisation*: A.V.E.H., J.B.v.K. and R.M.v.D.; *Supervision*: T.H., G.M.T. and A.M.J.M.v.d.M.; *Project administration*: A.V.E.H., R.M.v.D., and G.L.J.O.; *Funding acquisition*: T.H., G.M.T. and A.M.J.M.v.d.M.

Funding information: This work was supported by the Netherlands Organization for Scientific Research (VICI grant 918.56.601 and Spinoza 2009 to M.D.F.), the Netherlands Organization for Health Research and Development (Clinical Fellowship grant 90700217 and Vidi grant 917-11-31 to G.M.T.), and European Community-funded FP7-EUROHEADPAIN (grant 602633 to M.D.F. and A.M.J.M.v.d.M.).

Data accessibility statement: Data not published within the article is available from the corresponding author upon reasonable request.

Supplementary information

Figure S1 Remaining correlation plots of amines between plasma and CSF in healthy controls.
Figure S2 Correlation matrix of all possible metabolite ratios in healthy controls split by sex.
Figure S3 The difference in coefficients of determination (R2) of all possible metabolite ratios in migraine participants compared to healthy controls.
Table S1 Single-plasma/CSF metabolite correlations in heathy controls.
Table S2 Plasma/CSF ratio metabolite correlations in heathy controls.
Table S3 The improvement in R2 ratio correlation (R2 gain) compared to r in healthy controls.
Table S4 Single-plasma/CSF metabolite correlations in migraine participants.
Table S5 Plasma/CSF ratio metabolite correlations in migraine participants.
Table S6 Comparison of single-metabolite correlations of plasma and CSF between heathy controls and migraine participants.
Table S7 Comparison of plasma/CSF ratio metabolite correlations between heathy controls and migraine participants.

References

1. Patti GJ, Yanes O, Siuzdak G. Innovation: Metabolomics: the apogee of the omics trilogy. *Nat Rev Mol Cell Biol*. 2012;13(4):263-269.
2. Wishart DS. Metabolomics for Investigating Physiological and Pathophysiological Processes. *Physiol Rev*. 2019;99(4):1819-1875.
3. Monserrate AE, Ryman DC, Ma S, et al. Factors associated with the onset and persistence of post-lumbar puncture headache. *JAMA Neurol*. 2015;72(3):325-332.
4. van Oosterhout WP, van der Plas AA, van Zwet EW, et al. Postdural puncture headache in migraineurs and nonheadache subjects: a prospective study. *Neurology*. 2013;80(10):941-948.
5. van Dongen RM, Onderwater GLJ, Pelzer N, et al. The effect of needle size on cerebrospinal fluid collection time and post-dural puncture headache: A retrospective cohort study. *Headache*. 2021;61(2):329-334.
6. De Kort AM, Kuiperij HB, Marques TM, et al. Decreased Cerebrospinal Fluid Amyloid β 38, 40, 42, and 43 Levels in Sporadic and Hereditary Cerebral Amyloid Angiopathy. *Ann Neurol*. 2023; 93(6):1173-118
7. van Etten ES, Verbeek MM, van der Grond J, et al. β-Amyloid in CSF: Biomarker for preclinical cerebral amyloid angiopathy. *Neurology*. 2017;88(2):169-176.
8. Vervuurt M, de Kort AM, Jäkel L, et al. Decreased ratios of matrix metalloproteinases to tissue-type inhibitors in cerebrospinal fluid in sporadic and hereditary cerebral amyloid angiopathy. *Alzheimers Res Ther*. 2023;15(1):26.
9. Lajtha A, Galoyan A, Besedovsky H. *Handbook of Neurochemistry and Molecular Neurobiology: Neuroimmunology*. Springer; 2007.
10. Hagenfeldt L, Bjerkenstedt L, Edman G, Sedvall G, Wiesel FA. Amino acids in plasma and CSF and monoamine metabolites in CSF: interrelationship in healthy subjects. *J Neurochem*. 1984;42(3):833-837.
11. Kruse T, Reiber H, Neuhoff V. Amino acid transport across the human blood-CSF barrier. An evaluation graph for amino acid concentrations in cerebrospinal fluid. *J Neurol Sci*. 1985;70(2):129-138.
12. Noga MJ, Dane A, Shi S, et al. Metabolomics of cerebrospinal fluid reveals changes in the central nervous system metabolism in a rat model of multiple sclerosis. *Metabolomics*. 2012;8(2):253-263.
13. Onderwater GLJ, van Dongen RM, Harms AC, et al. Cerebrospinal fluid and plasma amine profiles in interictal migraine. *Ann Neurol*. 2022; 93(4):715-728
14. Headache Classification Committee of the International Headache Society (IHS). The International Classification of Headache Disorders, 3rd edition. *Cephalalgia*. 2018;38(1):1-211.
15. Noga MJ, Zielman R, van Dongen RM, et al. Strategies to assess and optimize stability of endogenous amines during cerebrospinal fluid sampling. *Metabolomics*. 2018;14(4):44.
16. van der Kloet FM, Bobeldijk I, Verheij ER, Jellema RH. Analytical error reduction using single point calibration for accurate and precise metabolomic phenotyping. *J Proteome Res*. 2009;8(11):5132-5141.
17. Steiger JH. Tests for comparing elements of a correlation matrix. *Psychol Bull*. 1980;87(2):245-251.
18. Zanatta A, Viegas CM, Tonin AM, et al. Disturbance of redox homeostasis by ornithine and homocitrulline in rat cerebellum: a possible mechanism of cerebellar dysfunction in HHH syndrome. *Life Sci*. 2013;93(4):161-168.
19. Viegas CM, Zanatta A, Knebel LA, et al. Experimental evidence that ornithine and homocitrulline disrupt energy metabolism in brain of young rats. *Brain Res*. 2009;1291:102-12.
20. Khovarnagh N, Seyedalipour B. Antioxidant, histopathological and biochemical outcomes of short-term exposure to acetamiprid in liver and brain of rat: The protective role of N-acetylcysteine and S-methylcysteine. *Saudi Pharm J*. 2021;29(3):280-289.

21. Westergren I, Nyström B, Hamberger A, Nordborg C, Johansson BB. Concentrations of amino acids in extracellular fluid after opening of the blood-brain barrier by intracarotid infusion of protamine sulfate. *J Neurochem*. 1994;62(1):159-165.
22. O'Kane RL, Hawkins RA. Na+-dependent transport of large neutral amino acids occurs at the abluminal membrane of the blood-brain barrier. *Am J Physiol Endocrinol Metab*. 2003;285(6):E1167-1173.
23. del Amo EM, Urtti A, Yliperttula M. Pharmacokinetic role of L-type amino acid transporters LAT1 and LAT2. *Eur J Pharm Sci*. 2008;35(3):161-174.
24. Hawkins RA, O'Kane RL, Simpson IA, Vina JR. Structure of the blood-brain barrier and its role in the transport of amino acids. *J Nutr*. 2006;136(1 Suppl):218S-226S.
25. Schweikhard ES, Ziegler CM. Amino acid secondary transporters: toward a common transport mechanism. *Curr Top Membr*. 2012;70:1-28.
26. Nałęcz KA. Solute Carriers in the Blood-Brain Barier: Safety in Abundance. *Neurochem Res*. 2017;42(3):795-809.
27. Boado RJ, Li JY, Nagaya M, Zhang C, Pardridge WM. Selective expression of the large neutral amino acid transporter at the blood-brain barrier. *Proc Natl Acad Sci U S A*. 1999;96(21):12079-12084.
28. Poncet N, Mitchell FE, Ibrahim AF, et al. The catalytic subunit of the system L1 amino acid transporter (slc7a5) facilitates nutrient signalling in mouse skeletal muscle. *PLoS One*. 2014;9(2):e89547.
29. Zaragozá R. Transport of Amino Acids Across the Blood-Brain Barrier. *Front Physiol*. 2020;11:973-973.
30. Fernstrom JD, Wurtman RJ. Brain serotonin content: physiological regulation by plasma neutral amino acids. *Science*. 1972;178(4059):414-416.
31. O'Kane RL, Vina JR, Simpson I, et al. Cationic amino acid transport across the blood-brain barrier is mediated exclusively by system y+. *Am J Physiol Endocrinol Metab*. 2006;291(2):E412-419.
32. McDonald KK, Zharikov S, Block ER, Kilberg MS. A caveolar complex between the cationic amino acid transporter 1 and endothelial nitric-oxide synthase may explain the "arginine paradox". *J Biol Chem*. 1997;272(50):31213-3121 6.
33. Olesen J. The role of nitric oxide (NO) in migraine, tension-type headache and cluster headache. *Pharmacol Ther*. 2008;120(2):157-171.
34. Bagdy G, Riba P, Kecskemeti V, Chase D, Juhasz G. Headache-type adverse effects of NO donors: vasodilation and beyond. *Br J Pharmacol*. 2010;160(1):20-35.
35. Neeb L, Reuter U. Nitric oxide in migraine. *CNS Neurol Disord Drug Targets*. 2007;6(4):258-264.

4

Endocannabinoid cerebrospinal fluid levels in migraine

Aster V.E. Harder[*], Xinyu Di[*], Wendy M. Winter, Bingshu He, Robin M. van Dongen, Gerrit L.J. Onderwater, Erik W. van Zwet, Elke H.J. Krekels, Isabelle Kohler, Michel D. Ferrari, Amy C. Harms, Arn M.J.M. van den Maagdenberg[*], Thomas Hankemeier[*] and Gisela M. Terwindt[1*]

Submitted

Abstract

Disruption of the endocannabinoid system may play a role in both migraine and depression, two conditions that are often comorbid. However, the evidence for this is limited. Here we measured the levels of endocannabinoids (eCBs) anandamide (AEA) and 2-arachidonoylglycerol (2-AG), and the endocannabinoid analogue docosahexaenoylethanolamine (DHEA) in the cerebrospinal fluid (CSF) of n = 194 individuals with episodic migraine between attacks (n = 97 migraine with aura and n = 97 migraine without aura) and n = 94 healthy volunteers, using micro-liquid chromatography – tandem mass spectrometry (micro-LC-MS/MS). Groups were age- and sex-matched. Timing and processing of sampling was protocolized. Multivariate linear regression analyses for $\log_{10}$-transformed eCB concentrations were performed to compare groups. Covariates that were included were sex, age, BMI, alcohol and cigarette consumption and lifetime depression. No differences were detected between individuals with migraine with or without aura and healthy controls in the CSF for AEA (P = 0.46, P = 0.07), 2-AG (P = 0.40, P = 1.00) and DHEA (P = 0.65, P = 0.79), respectively. The covariates sex, smoking, age, and BMI had an effect on various eCBs. For 2-AG, levels were positively correlated with lifetime depression (β = 0.060; 95% CI: 0.020 to 0.10, P = 0.003). CSF endocannabinoid 2-AG is associated with depression, but CSF endocannabinoids seem not to be useful biomarkers to differentiate interictal migraine individuals from controls. This study emphasizes the complexity of the endocannabinoid system in relation to migraine and depression.

KEYWORDS: Migraine, depression, cerebrospinal fluid, endocannabinoids, biomarker

Introduction

Migraine is a highly prevalent, disabling, primary brain disorder clinically characterized by attacks of severe headache and associated symptoms.[1] There are no diagnostic biomarkers and migraine is diagnosed according to clinical criteria[1] Depression and anxiety are common comorbid disorders.[2, 3] Although much progress has been made in unravelling the pathophysiology of attacks, little is known about why migraine attacks develop. Research into neurochemistry in people with migraine between attacks may help elucidate its etiology and provide useful biomarkers.

The endocannabinoid system has a strong influence on neurotransmission, the neuroendocrine and neuroimmune systems,[4-6] which are also involved in migraine pathophysiology.[7, 8] Previous studies have indicated that endocannabinoids (eCBs) play a role in the trigeminovascular system, a key component in migraine pathophysiology.[9, 10]

The endocannabinoid system consists of two cannabinoid receptors (CB_1 en CB_2) and their lipid signaling ligands, the eCBs.[11] CB_1 receptors are expressed throughout the peripheral and central nervous system, whereas CB_2 receptors are mainly located on immune cells.[12] The two most widely studied eCBs are *N*-arachidonoylethanolamine, also called anandamide (AEA) and 2-arachidonoylglycerol (2-AG). AEA is a partial agonist for both cannabinoid receptors, with a high affinity for the CB_1 and a low affinity for the CB_2.[11] Additionally, AEA is a full agonist at the transient receptor potential vanilloid 1 (TRPV1) where it modulates the release of calcitonin gene-related peptide (CGRP). CGRP is a key neurotransmitter of the trigeminovascular system and instrumental in causing migraine attacks.[13-15] 2-AG is a full agonist at CB_1 and CB_2 receptors and probably the major synaptic eCB ligand in nervous tissue.[11, 16]

Endocannabinoid signaling is a key regulator of synaptic communication through retrograde inhibition of the release of inhibitory and excitatory neurotransmitters.[17, 18] The CB_1 receptor particularly is strongly involved in inhibiting the release of gamma-aminobutyric acid (GABA), serotonin, glutamate, and CGRP, all believed to play an important role in migraine.[6, 9, 15, 19]

Several studies suggest that the endocannabinoid system modulates the trigeminovascular system in migraine.[9, 20-22] In a small study, AEA was suggested to be lower in the cerebrospinal fluid (CSF) and palmitoylethanolamide (PEA) to be higher in chronic migraine.[20] Fatty acid amide hydrolase (FAAH), which degrades AEA, and endocannabinoid membrane transporter (EMT) activity in platelets were increased in women with migraine but not in men, suggesting an altered endocannabinoid system that seems more profound in female patients.[23] Gene expressions and peripheral protein levels of endocannabinoid system components were elevated in the blood of people with migraine.[24] At the same time, evidence has emerged for a role of the endocannabinoid system in depression,[25] which has a bidirectional relationship with migraine.[26] Women with depression have reduced endocannabinoid levels in blood serum,[27] and chronic treatment with antidepressants can alter the endocannabinoid system.[28]

Here, we compared CSF levels of eCBs in a large cohort of people with migraine with aura and without aura and healthy controls. We sought to detect differences in the endocannabinoid system and identify diagnostic biomarkers that distinguish these clinical entities taking comorbid depressive symptomatology into account.

Methods

Study design and participants

The study population has been described in Onderwater et al.[29] In brief, we enrolled n = 198 individuals with migraine (n = 100 migraine with aura and n = 98 migraine without aura) and n = 96 healthy controls and group-matched them for age (using 5-year strata) and sex. Migraine was diagnosed according to the International Classification of Headache Disorders (ICHD) by experienced physicians (supervised by G.M.T. and M.D.F.).[1] All participants were between 18 and 65 years and had a body mass index (BMI) between 18 and 28. Individuals suffering from migraine did not use acute migraine drugs on more than 8 days per month. Healthy volunteers had no obvious signs or symptoms of a disease and had no history of headache (except for infrequent tension-type headaches) or other pain-related syndromes. They also did not have first-degree relatives with migraine or trigeminal autonomic cephalalgia. Other exclusion criteria were (1) severe psychiatric disorder, (2) history of oncological disease, or (3) contra-indication for lumbar puncture (LP) (signs and symptoms of increased intracranial pressure, local skin infection, or a coagulopathy including use of anti-coagulant drugs or platelet-inhibitors).

Participants were asked to fill out a questionnaire on the day of the LP that included questions related to their migraine characteristics and possible history of depression. To compute lifetime depression, additional questionnaires were preferentially completed on the day of LP, but otherwise extracted from electronic medical records. The presence of lifetime depression was determined based on previously described criteria.[30] In short, symptoms of depression were determined using validated cut-off scores for the Hospital Anxiety and Depression Scale (HADS) and the Center for Epidemiologic Studies Depression Scale (CES-D). The HADS questionnaire has been used in clinical studies as it intrinsically corrects for the overlap between symptoms of somatic diseases and depression (e.g., lack of sleep, changes in appetite). Lifetime depression was defined as HADS-D ≥ 8 or CES-D ≥ 16 or (past) depression diagnosed by a physician or (past) use of antidepressants for depression.

The study was conducted according to the criteria of the Declaration of Helsinki and was approved by the Leiden University Medical Center institutional ethics committee. All participants provided written informed consent prior to participation in the study.

Cerebrospinal fluid collection

Sampling and handling of CSF samples was described in Onderwater et al.[29] In brief, the time of the LP was between 8:30 a.m. and 1:00 p.m. and in random order to minimize diurnal and seasonal

variation. Participants were asked to refrain from eating or drinking, apart from water, for at least 8 h prior to sampling. Sampling occurred interictally with migraine individuals being attack-free for a minimum of 3 days. Participants were urged to refrain from the use of prophylactic or any other chronic medication (other than oral contraceptives) in the four weeks preceding the LP. The LP and subsequent CSF collection were performed in a standardized matter to limit inter-individual variation.

Two mL CSF was transferred to a 12-mL polypropylene Falcon tube (Cat. No. 160172; Greiner Bio-One B.V., Alphen aan den Rijn, the Netherlands) for routine diagnostics (cell count, glucose and total protein levels). Next, for endocannabinoid measurements, 3 mL CSF was sampled directly in a 15-mL polypropylene Falcon tube (Cat. No. 188271; Greiner Bio-One B.V.) that already contained 6 mL of ice-cold ethanol (to instantly stop any metabolic activity). The tube was immediately put back on ice after CSF collection. After gently shaking the tube, the CSF was divided in aliquots of 1.5 mL in 1.8-mL cryotubes (Art. No. 368632; NUNC Brand, Sigma-Aldrich, Zwijndrecht, the Netherlands). The cryotubes were placed on dry ice within 30 minutes of sampling and transferred to a -80˚C freezer as soon as possible. All CSF samples remained at -80˚C until sample preparation for analysis.

Analysis of eCBs using micro-liquid chromatography – tandem mass spectrometry quantification

The analysis of AEA, 2-AG and DHEA levels in CSF was performed using a validated method.[31] Sample preparation, except the evaporation step, was carried out under ice-cold conditions. Analytes were extracted using liquid-liquid extraction. In brief, 750 µL of CSF samples (CSF:ethanol = 1:2, v/v) was evaporated to approximately 100 µL using a SpeedVac (Thermo Fisher Scientific, Waltham, MA, USA). Then, 10 µL of isotopically labelled internal standard solution and 50 µL of a buffer solution (0.1 M ammonium acetate solution, adjusted to pH 4 with formic acid) were added to each sample. One milliliter of methyl-tert-butylether (MTBE) was then added prior to agitation for 10 min using a bullet blender and subsequent centrifugation for 5 min at 16,000 g. The organic supernatant (900 µL) was collected and evaporated to dryness using the SpeedVac. The dry residues were then reconstituted in 20 µL of an ice-cold solution of water and acetonitrile (ACN) 50:50 (v/v), prior to agitation (20 min) and centrifugation at 16,000 g for 50 min. Finally, 15 µL of the supernatant was transferred into a glass vial and 3 µL was injected into a nanoAcquity micro-LC system (Waters Corporation, Mildford, MA, USA) hyphenated with a Shimadzu 8060 triple quadrupole mass spectrometer (Shimadzu Corporation, Kyoto, Japan). Separation was performed using a micro C18 column (0.3 × 150 mm, 2.6 µm) from Phenomenex (Torrance, CA, USA) maintained at 45˚C. The mobile phase was composed of 2 mM ammonium formate and 10 mM formic acid in water (A), and acetonitrile (B). The flow rate was 4 µL/min. Ionization of the compounds was performed using a customized micro-electrospray ionization (micro-ESI) source in positive mode. Selected Reaction Mode (SRM) was used for MS/MS acquisition. SRM

transitions were individually optimized for targeted analytes and respective isotopically labelled internal standards. Data acquisition and pre-processing was performed using LabSolutions LCMS Version 5.97 SP1 (Shimadzu Corporation). Data quality was monitored using regular injection of quality control (QC) samples prepared from pooled CSF. Samples were injected in a randomized order. QC samples were used to correct for inter-batch variations using the in-house developed mzQuality workflow (available at http://www.mzQuality.nl). Relative standard deviations (RSDs) of peak area ratios were calculated for each targeted analyte detected in the QC samples. The absolute concentration of the eCBs in CSF samples was calculated using calibration samples. Quality control was set on RSDqc < 15% and blank effect < 15%. The data as relative response ratios (target area/ISTD area; unit free) of these metabolites after QC correction were calculated and converted to concentrations.

Statistical analysis

Principal component analysis (PCA) was performed on the concentration measurements of all measured eCBs together. Based on the first two principal components outliers were excluded by visually inspecting the plots. Log_{10}-transformed eCB concentrations (pM) were used for further statistical analyses. Multivariate linear regression was used to detect metabolite differences between migraineurs and healthy controls for each eCB separately. Based on literature, BMI, sex, age, weekly alcohol consumption, cigarettes per day, and lifetime depression were identified as potential covariates and therefore included as independent variables. Prior to data analysis missing data for relevant covariates was imputed with Multiple Imputation by Chained Equations (MICE).[32] For MICE, we used 10 imputations and 30 iterations as settings and included the following variables: Log_{10} eCB concentrations, age, sex, BMI, migraine without aura/migraine with aura/control-stratus, lifetime depression score, and the amount of alcohol and smoking per week. Regression analyses were performed on the imputed data and results were pooled. All data analyses were performed using R version 4.0.1. and RStudio version 1.4.1717 (RStudio: Integrated Development for R. RStudio, PBC).[33] *P*-values of ≤ 0.05 were considered statistically significant.

Data Availability

The data that support the findings of this study are available from the corresponding author, upon reasonable request.

Results

Clinical characteristics

We reliably detected three endocannabinoids, AEA, 2-AG, and DHEA in the CSF of 195 individuals with migraine (n = 98 migraine with aura; n = 97 migraine without aura) and 94 healthy controls. Other endocannabinoids did not comply to the quality control criteria and for 5

individuals there was not enough CSF sample available (< 1.5 mL). Next, one population outlier with migraine with aura was excluded based on the PCA plot, leaving 194 people with migraine and 94 controls for the statistical analysis. After inclusion, n = 5 healthy controls were found to have a first-degree relative with migraine. However, their results were kept in the study, because CSF from healthy volunteers is very difficult to obtain. The clinical characteristics are shown in **Table 1**. There were no differences between the migraine groups, except for a higher monthly attack frequency in migraine with aura (1.5 IQR 1.0 - 3.0) versus migraine without aura (2.0 IQR 1.0 - 3.5). However, there was a difference between the three groups in, stress during LP (P = 0.024), all HADS related scores (p < 0.001), total CES-D score (p < 0.001) and a lower weekly alcohol consumption (P = 0.009), the latter three being in line with prior results on migraine comorbidities and lifestyle characteristics (**Table 1**).[2, 3, 34] Data was assumed to be missing at random for weekly alcohol consumption (n = 6), lifetime depression (n = 45), HADS scores (n = 9), total CES-D (n = 48) and drug use (n = 1). The missing values used for the regression analysis were imputed. No group differences were observed for fasting time (in hours) and time of CSF collection (data not shown).

CSF levels of AEA, 2-AG and DHEA in migraine patients and controls

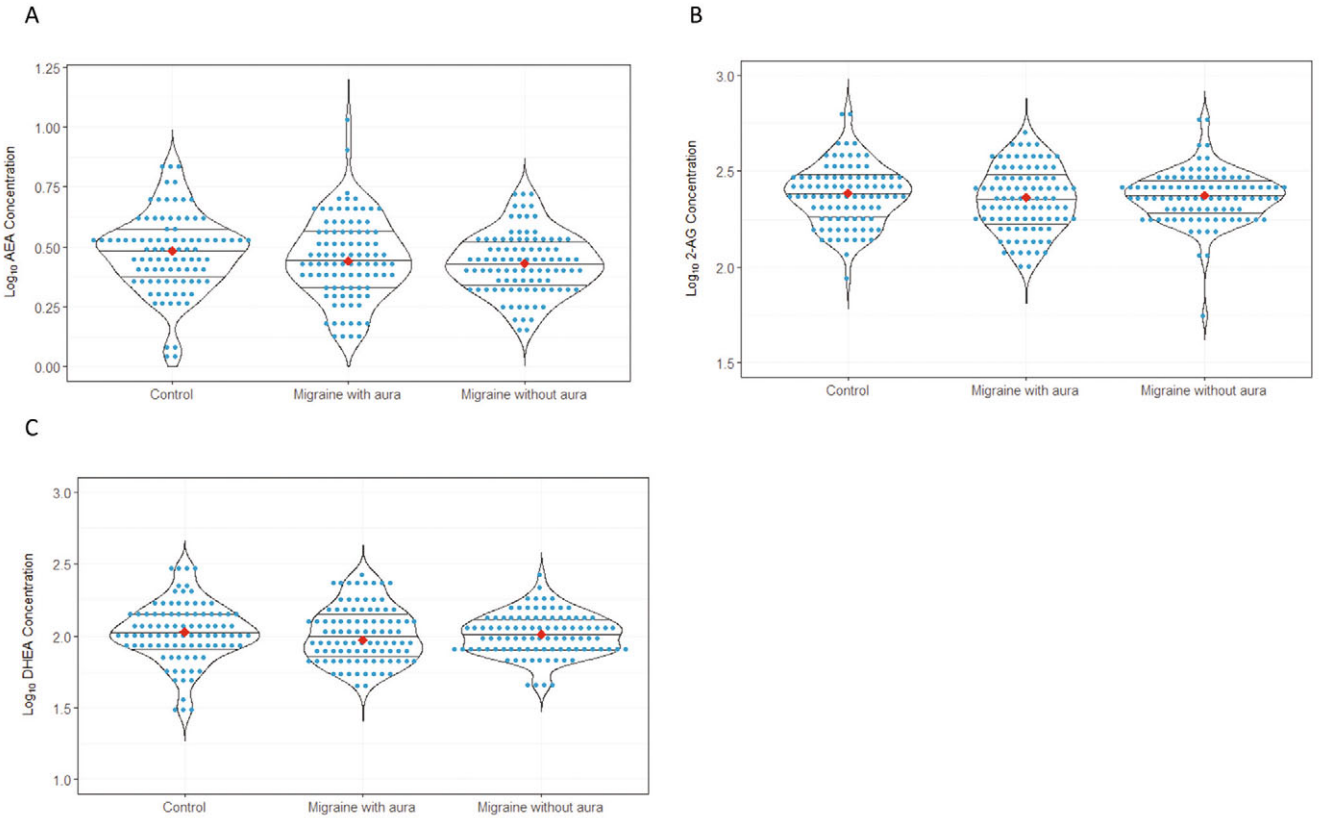
Multivariate linear regression analyses showed that there were no differences in concentrations in migraine with aura and migraine without aura versus control for each of the three eCBs (**Figure 1A-C**) P = 0.46 and P = 0.07 for AEA, P = 0.40 and P = 1.00 for 2-AG and P = 0.65 and P = 0.79 for DHEA, respectively. Some of the covariates, however, did have a significant effect on eCB concentration in the regression model. AEA concentration was associated with sex (β = -0.077; 95% CI: -0.12 to -0.04, P < 0.001) and number of cigarettes per day (β = 0.007; 95% CI: 0.00 to 0.01, P = 0.008). Female sex was associated with a lower AEA concentration (**Figure 2A**) and the number of cigarettes per day was positively correlated with the AEA concentration (**Figure 2B**). 2-AG concentration was negatively correlated with female sex (β = -0.083; 95% CI: -0.12 to -0.05, P < 0.001) (**Figure 3A**) and positively correlated with lifetime depression (β = 0.060; 95% CI: 0.020 to 0.10, P = 0.003) (**Figure 3B and 3C**). DHEA concentration was positively correlated with age (β = 0.003; 95% CI: 0.00 to 0.00, P < 0.001) and BMI (β = 0.010; 95% CI: 0.00 to 0.02, P = 0.010) (**Figure 4A-B**).

Table 1 Characteristics of control and migraine patient cohort

Participants Characteristics	MA (n = 97)	IQR	MO (n = 97)	IQR	Control (n = 94)	IQR	*P*-value MaMoCo
General characteristics							
Sex, female (n, %)	63 (64.9)		59 (60.8)		54 (57.4)		0.567[a]
BMI	23.9	22.0-25.8	23.8	21.9-25.4	23.5	21.5 - 25.3	0.433[b]
Age	40	29-52	42	32-53	39.0	26.0 - 51.0	0.262[c]
Alcohol use (n, %)	78 (80.4)		67 (69.1)		77 (81.9)		0.068[a]
Units alcohol per week[#]	4	1.7-7	2.8	1.0-5.1	5.0	1.4 - 8.0	**0.009**[c]
Smoking (n, %)	13 (13.4)		13 (13.4)		20 (21.3)		0.231[a]
Cigarettes per day[#]	4	1.5-15	6.5	2.5-10	6.0	2.8 - 10.5	0.693[c]
Migraine characteristics							
Migraine days (days/month)	1.5	1.0-3.0	2.0	1.0-3.5	NA	NA	**0.017**[d]
Depression/Anxiety characteristics							
Lifetime depression (n, %)	32.0 (33.0)		26.0 (26.8)		17.0 (27.4)		0.424[a]
Total HADS score	8.0	5.0-13.0	8.0	4.0-11.3	4.0	2.0 - 7.0	**<0.001**[c]
HADS Anxiety	5.0	3.0-8.0	4.0	2.0-7.25	2.0	1.0-4.5	**<0.001**[c]
HADS Depression	3.0	1.0-5.0	3.0	1.0-3.0	1.0	0.0-3.0	**<0.001**[c]
CES-D score	10.5	5.0-17.3	8.0	4.0-12.0	3.5	1.0-7.8	**<0.001**[c]
Drugs related characteristics							
Drug use, (n, %)	4 (4)		4 (4)		9 (10)		0.188[a]
Soft drugs	2 (2)		4 (4)		8 (9)		0.125[e]
Hard drugs	2 (2)		0		1 (1)		0.535[e]
Drug use per month[#]	1.5	0.6-3.4	31.5	5.5-73.0	3	0.5 - 8.0	0.299[c]
LP-related characteristics							
Pain during the LP (VAS)	3.0	2.0-6.0	3	1.0-5.0	2.0	1.0 - 4.8	0.056[c]
Stress during the LP (VAS)	3.0	1.0-5.0	2.0	0.0-4.0	2.0	0.0 - 3.8	**0.024**[c]

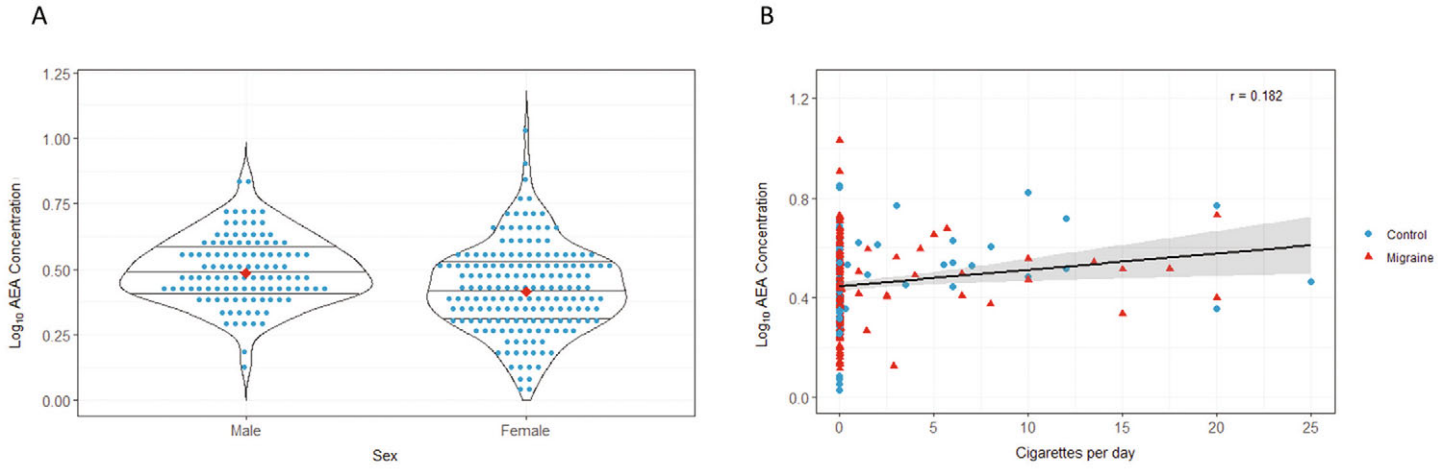
IQR = inter quartile range; MA = migraine with aura, MO = migraine without aura, LP = lumbar puncture; BMI = body mass index; VAS = visual analogue scale. Values are expressed as absolute values and percentage or median and IQR. *P*-values ≤ 0.05 are marked in bold. [a] Chi-squared Test, [b] One way Anova, [c] Kruskal-Wallis test, [d] Mann-Whitney U test, [e] Fisher's Exact Test, [#]Only users (alcohol/smokers/drugs) are used for number per day/week/month. Missing values in: units alcohol per week (cases n = 6 (3 MA and 3 MO), controls n = 0), migraine days (cases n = 7 (5 MA and 2 MO), lifetime depression (cases n = 13 (8 MA and 5 MO), controls n = 32), HADS scores (cases n = 2 (1 MA and 1 MO), controls n = 7) CES-D score (cases = 16 (9 MA and 7 MO), controls = 32) , and Drug use (cases n = 1 (1 MA), controls n = 0).

Figure 1 Endocannabinoid concentration per subgroup



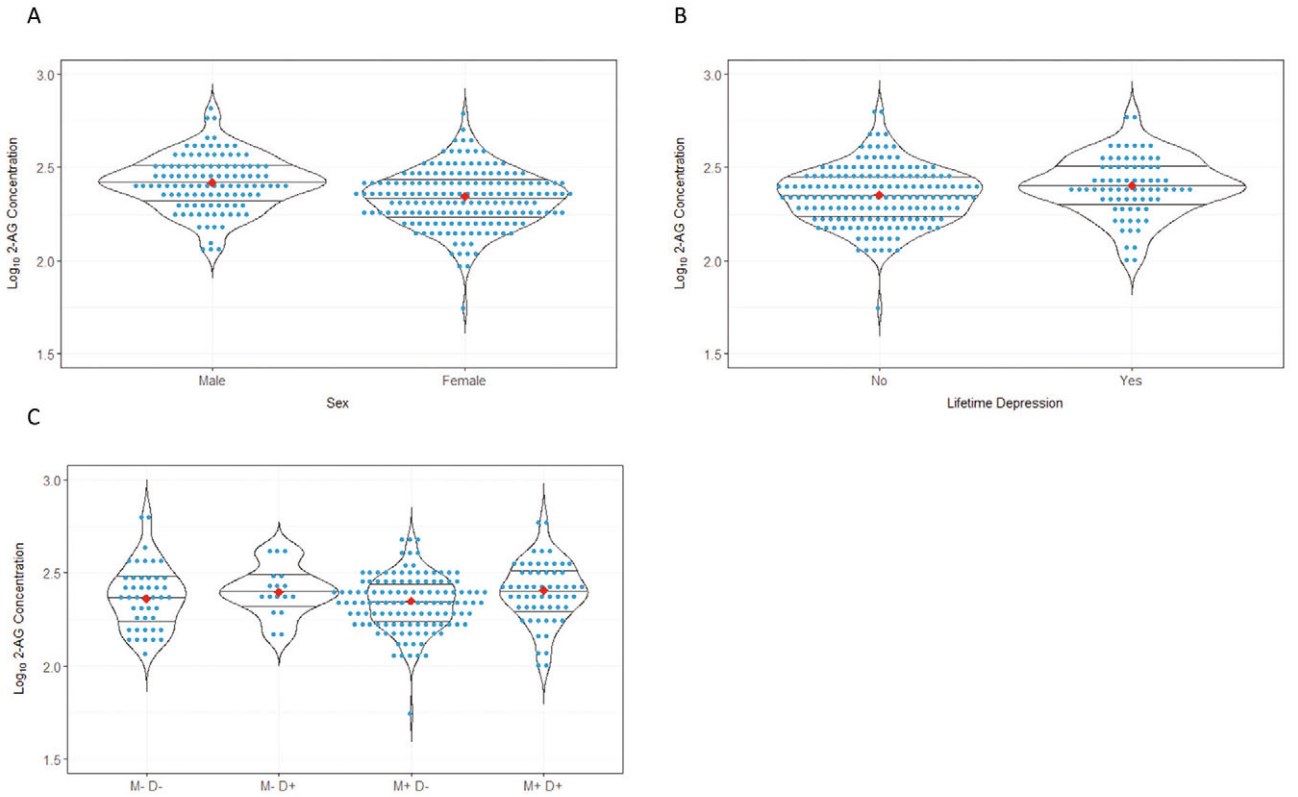
A) Violin plot of log10 transformed AEA concentrations in each subgroup. **B)** Violin plot of log10 transformed 2-AG concentrations in each subgroup. **C)** Violin plot of log10 transformed DHEA concentrations in each subgroup. Each dot represents an individual, the median is marked by a red diamond. The horizontal lines represent the 0.25, 0.5 and 0.75 quantiles. AEA = N-Arachidonoylethanolamine, 2-AG = Arachidonoylglycerol, DHEA = Docosahexaenoylethanolamine. Plots are based on pre-imputed data and uncorrected.

Figure 2 Relation of AEA concentration with covariates



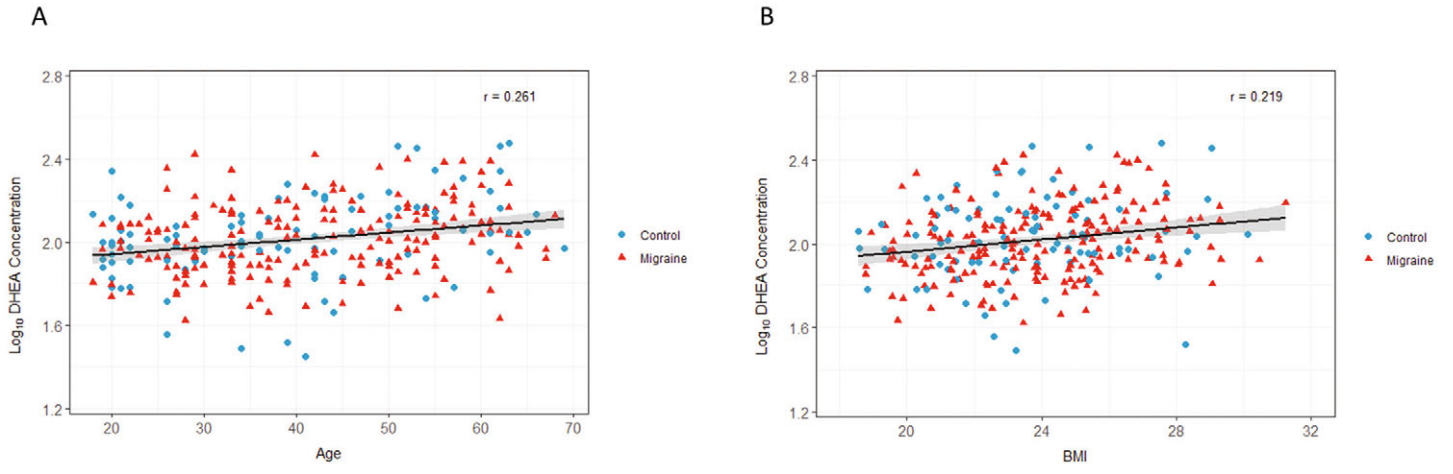
A) Violin plot of log10 transformed AEA concentrations and sex. Each dot represents an individual, the median is marked by a red diamond. The horizontal lines represent the 0.25, 0.5 and 0.75 quantiles. **B)** Correlation plot of log10 transformed AEA concentration and cigarettes per day. The black line indicates the linear regression line and grey area the 95% confidence interval. Cases are depicted in red triangles and controls in blue circles. r = Pearson's correlation coefficient. AEA = N-Arachidonoylethanolamine. Plots are based on pre-imputed data and uncorrected.

Figure 3 Relation of 2-AG concentration with covariates



A) Violin plot of log10 transformed 2-AG concentrations and sex. **B)** Violin plot of log10 transformed 2-AG concentrations and lifetime depression. **C)** Violin plot of log10 transformed 2-AG concentrations and lifetime depression subdivided by diagnosis with and without depression (control = M-, case = M+, positive for lifetime depression = D+ and negative for lifetime depression = D-). Each dot represents an individual, the median is marked by a red diamond. The horizontal lines represent the 0.25, 0.5 and 0.75 quantiles. 2-AG = Arachidonoylglycerol. Plots are based on pre-imputed data and uncorrected.

Figure 4 Relation of DHEA concentration with covariates



A) Correlation plot of log10 transformed DHEA concentration and age. **B)** Correlation plot of log10 transformed DHEA concentration and BMI. The black line indicates the linear regression line and grey area the 95% confidence interval. Cases are depicted in red triangles and controls in blue circles. r = Pearson's correlation coefficient. DHEA = Docosahexaenoylethanolamine, BMI = Body mass index. Plots are based on pre-imputed data and uncorrected.

Discussion

Here we measured and compared concentrations of the eCBs AEA, 2-AG, and DHEA in CSF of 194 individuals with migraine with (n = 97) or without aura (n = 97) in the interictal state and 94 healthy controls. The three study populations were matched for age and sex. The CSF sampling and processing were strictly protocoled, and the analysis was performed with our validated micro-liquid chromatography – tandem mass spectrometry quantification (micro-LC-MS/MS).[31] We found no differences in eCBs related to migraine status. However, covariates age, sex, daily number of cigarettes, and BMI had an effect on the different eCB concentrations. The effect on the covariates emphasizes the complexity of the endocannabinoid system regarding confounding factors. Interestingly, 2-AG levels were positively correlated with lifetime depression, which has implications for further research as comorbidity with depression in migraine patients makes them more prone for chronification and medication overuse.[35]

So far, only two studies have assessed the levels of eCBs in body fluids of individuals with migraine, but the results were contradictory.[20, 36] In one study, concentrations of AEA were lower and those of the eCB analogue PEA were higher in the CSF of 15 individuals with chronic migraine versus 20 age-matched controls.[20] However, in that study, the analysis of eCBs was based on a different extraction method (i.e., the Bligh and Dyer liquid-liquid extraction method), which may generate AEA during the analytical step (see **e-Figure 1**, supplementary material). Most importantly, there was no correction for confounding factors as we did in our study. Therefore, the results of that study should be interpreted with caution. In the other study, 16 different eCBs were measured by LC-MS/MS in blood plasma from 38 females with episodic migraine between attacks and 26 healthy controls collected on days 1 and 14 of the menstrual cycle.[36] There were no differences for AEA and N-acylethanolamines. In addition, in a subset of study participants eCBs were measured in plasma collected during attacks and compared with the interictal measurement; again no differences were found.[36] In both studies, the number of participants was limited, the body fluids in which the measurements were done were different, the sampling conditions were not protocoled, and no correction for all possible confounding factors was done.

Contrary to the previous suggestion of the influence of endocannabinoids in migraine pathophysiology, no association between the endocannabinoid system and migraine was detected in our CSF study.[9, 20-22] A dysregulation of the endocannabinoid system might have remained undetected in case it involves only a localized (e.g., near the synapse) or migraine-attack related disturbance of the endocannabinoid system and not a generalized phenomenon that shows in CSF. Basal levels of eCBs in the brain are low, with 2-AG levels much higher than those of AEA in the brain.[37] This may be an explanation for not finding differences in CSF in migraine patients in the interictal phase. Further limitations of our study are that due to the size of the study and the great care taken to match our study groups, including for age, sex and diurnal and seasonal time of sampling, the samples were collected over a fairly long time. This could have potentially

affected the stability of the compounds studied, even though they were stored at the recommended -80°C.[38] However, only minor changes of oxylipins were observed after prolonged storage at -80°C of blood samples and the metabolome and lipidome of CSF were found to be stable for longer periods.[38, 39]

Our study has several strengths. First, we studied and compared CSF from a uniquely large group of people with migraine and controls, who were carefully matched for major confounding factors such as sex, age, and timing of sampling. Moreover, unlike other CSF studies, control samples were collected from people who were truly healthy and had not undergone a lumbar punction because of neurological symptoms and almost all had no first-degree relatives with migraine.[40]

Second, when investigating a putative role of eCBs in disease, it is crucial to use standardized protocols and to correct the statistical model for possible confounders such as BMI, sex, age, weekly alcohol consumption, daily cigarettes use,[41-46] and most importantly, lifetime depression,[25, 27, 47-49] which is an established comorbid condition in migraine. Although people with major depression were excluded in our study, we found a positive correlation between 2-AG and lifetime depression as screened with validated tools (HADS-D and CESD) for depressive symptomatology. Previous research on the role of AEA and 2-AG in depression has been conducted primarily in individuals with major depressive disorder (MDD).[27, 47, 48] The hypothalamic–pituitary–adrenal (HPA) -axis, neurogenesis and neuroinflammation have all been implicated in the etiology of MDD,[50] and all are dependent on proper functioning of the endocannabinoid system,[25] suggesting a crucial role for the endocannabinoid system in MDD. Given the increased risk of depression in migraine, it is important to realize that depressive symptoms may have a significant effect on eCB levels, even when MDD has been ruled out. The association between migraine and eCBs may therefore depend on depressive symptomatology rather than on migraine itself. Future studies should therefore take depressive symptomology into account when studying the role of the endocannabinoid system in migraine.

In conclusion, CSF endocannabinoid 2-AG levels are associated with depression, but CSF endocannabinoids seem not to be useful biomarkers to differentiate interictal migraine individuals from controls. This study emphasizes the complexity of the endocannabinoid system in relation to migraine and depression.

Supplementary Information

Supplementary e-Figure 1

Conflicts of interest statement: G.M. Terwindt reports consultancy support from Novartis, Allergan/Abbvie, Lilly, and Teva, Lundbeck and independent support from Dutch Research Council, the Dutch Heart & Brain Foundations, Dioraphte, and the European Community. A.M.J.M. van den Maagdenberg reports independent support from Schedule 1 Therapeutics, Praxis Precision Medicine, and AbbVie as well as the European Community. The other authors report no disclosures.

Acknowledgements: We greatly acknowledge all study participants, the medical students who assisted with recruitment of participants, physicians F. Bakels, N. Pelzer, L.A. Wilbrink and A.A. van der Plas for sample collection.

This study was supported by the Netherlands Organization for Scientific Research (VICI grant 918.56.601 and Spinoza 2009 to M.D. Ferrari), the Netherlands Organization for Health Research and Development (Clinical Fellowship 90700217 and Vidi 917-11-31 to G.M. Terwindt), and European Community (EC) FP7-EUROHEADPAIN – no. 602633 (to M.D. Ferrari and A.M.J.M. van den Maagdenberg).

A.V.E. Harder, X. Di, R.M. van Dongen, G.L.J. Onderwater, M.D. Ferrari , A.M.J.M. van den Maagdenberg, T. Hankemeier and G.M. Terwindt were responsible for the study concept and design. A.V.E. Harder, W.M. Winter, E.W. van Zwet, A.M.J.M. van den Maagdenberg and G.M. Terwindt contributed to data collection and interpretation. X. Di, E.H.J. Krekels, I. Kohler and A.C. Harms were responsible for the endocannabinoid measurements. A.V.E. Harder and W.M. Winter carried out the statistical analyses under supervision of E.W. van Zwet. A.V.E. Harder wrote the initial draft of the manuscript. All authors critically revised the article and approved the final version for submission.

References

1. Headache Classification Committee of the International Headache Society (IHS). The International Classification of Headache Disorders, 3rd edition. *Cephalalgia*. 2018;38(1):1-211.
2. Louter MA, Pijpers JA, Wardenaar KJ, et al. Symptom dimensions of affective disorders in migraine patients. *J Psychosom Res*. 2015;79(5):458-463.
3. Breslau N, Lipton RB, Stewart WF, Schultz LR, Welch KM. Comorbidity of migraine and depression: investigating potential etiology and prognosis. *Neurology*. 2003;60(8):1308-1312.
4. Hillard CJ, Beatka M, Sarvaideo J. Endocannabinoid Signaling and the Hypothalamic-Pituitary-Adrenal Axis. *Compr Physiol*. 2016;7(1):1-15.
5. Pandey R, Mousawy K, Nagarkatti M, Nagarkatti P. Endocannabinoids and immune regulation. *Pharmacol Res*. 2009;60(2):85-92.
6. Szabo B, Schlicker E. Effects of Cannabinoids on Neurotransmission. In: Pertwee RG, ed. *Cannabinoids*. Springer Berlin Heidelberg; 2005:327-365.
7. Goadsby PJ, Holland PR, Martins-Oliveira M, et al. Pathophysiology of Migraine: A Disorder of Sensory Processing. *Physiol Rev*. 2017;97(2):553-622.
8. Bruno PP, Carpino F, Carpino G, Zicari A. An overview on immune system and migraine. *Eur Rev Med Pharmacol Sci*. 2007;11(4):245-248.
9. Akerman S, Kaube H, Goadsby PJ. Anandamide is able to inhibit trigeminal neurons using an in vivo model of trigeminovascular-mediated nociception. *J Pharmacol Exp Ther*. 2004;309(1):56-63.
10. Akerman S, Holland PR, Goadsby PJ. Cannabinoid (CB1) receptor activation inhibits trigeminovascular neurons. *J Pharmacol Exp Ther*. 2007;320(1):64-71.
11. Zou S, Kumar U. Cannabinoid Receptors and the Endocannabinoid System: Signaling and Function in the Central Nervous System. *Int J Mol Sci*. 2018;19(3):833.
12. Howlett AC. The cannabinoid receptors. *Prostaglandins Other Lipid Mediat*. 2002;68-69:619-631.
13. Ross RA. Anandamide and vanilloid TRPV1 receptors. *Br J Pharmacol*. 2003;140(5):790-801.
14. Zygmunt PM, Petersson J, Andersson DA, et al. Vanilloid receptors on sensory nerves mediate the vasodilator action of anandamide. *Nature*. 1999;400(6743):452-457.
15. Engel MA, Izydorczyk I, Mueller-Tribbensee SM, et al. Inhibitory CB1 and activating/desensitizing TRPV1-mediated cannabinoid actions on CGRP release in rodent skin. *Neuropeptides*. 2011;45(3):229-237.
16. Sugiura T, Kondo S, Kishimoto S, et al. Evidence that 2-arachidonoylglycerol but not N-palmitoylethanolamine or anandamide is the physiological ligand for the cannabinoid CB2 receptor. Comparison of the agonistic activities of various cannabinoid receptor ligands in HL-60 cells. *J Biol Chem*. 2000;275(1):605-612.
17. Katona I, Freund TF. Endocannabinoid signaling as a synaptic circuit breaker in neurological disease. *Nat Med*. 2008;14(9):923-930.
18. Ohno-Shosaku T, Maejima T, Kano M. Endogenous cannabinoids mediate retrograde signals from depolarized postsynaptic neurons to presynaptic terminals. *Neuron*. 2001;29(3):729-738.
19. Rea K, Roche M, Finn DP. Supraspinal modulation of pain by cannabinoids: the role of GABA and glutamate. *Br J Pharmacol*. 2007;152(5):633-648.
20. Sarchielli P, Pini LA, Coppola F, et al. Endocannabinoids in chronic migraine: CSF findings suggest a system failure. *Neuropsychopharmacology*. 2007;32(6):1384-1390.
21. Greco R, Gasperi V, Sandrini G, et al. Alterations of the endocannabinoid system in an animal model of migraine: evaluation in cerebral areas of rat. *Cephalalgia*. 2010;30(3):296-302.

22. Greco R, Mangione AS, Sandrini G, et al. Effects of anandamide in migraine: data from an animal model. *J Headache Pain*. 2011;12(2):177-183.
23. Cupini LM, Bari M, Battista N, et al. Biochemical changes in endocannabinoid system are expressed in platelets of female but not male migraineurs. *Cephalalgia*. 2006;26(3):277-281.
24. Greco R, Demartini C, Zanaboni AM, et al. Peripheral changes of endocannabinoid system components in episodic and chronic migraine patients: A pilot study. *Cephalalgia*. 2021;41(2):185-196.
25. Gallego-Landin I, García-Baos A, Castro-Zavala A, Valverde O. Reviewing the Role of the Endocannabinoid System in the Pathophysiology of Depression. Review. *Front Pharmacol*. 2021;12 :762738
26. Peres MFP, Mercante JPP, Tobo PR, Kamei H, Bigal ME. Anxiety and depression symptoms and migraine: a symptom-based approach research. *J Headache Pain*. 2017;18(1):37.
27. Hill MN, Miller GE, Ho WS, Gorzalka BB, Hillard CJ. Serum endocannabinoid content is altered in females with depressive disorders: a preliminary report. *Pharmacopsychiatry*. 2008;41(2):48-53.
28. Hill MN, Ho WS, Hillard CJ, Gorzalka BB. Differential effects of the antidepressants tranylcypromine and fluoxetine on limbic cannabinoid receptor binding and endocannabinoid contents. *J Neural Transm (Vienna)*. 2008;115(12):1673-1679.
29. Onderwater GLJ, van Dongen RM, Harms AC, et al. Cerebrospinal fluid and plasma amine profiles in interictal migraine. *Ann Neurol*. 2022; 93(4):715-728
30. Louter MA, Pelzer N, de Boer I, et al. Prevalence of lifetime depression in a large hemiplegic migraine cohort. *Neurology*. 2016;87(22):2370-2374.
31. He B, Di X, Guled F, et al. Quantification of endocannabinoids in human cerebrospinal fluid using a novel micro-flow liquid chromatography-mass spectrometry method. *Anal Chim Acta*. 2022;1210:339888.
32. van Buuren S, Groothuis-Oudshoorn K. mice: Multivariate Imputation by Chained Equations in R.2011;45(3):67-%
33. *RStudio: integrated development for R*. RStudio, Inc.; 2015. http://www.rstudio.com
34. Onderwater GLJ, van Oosterhout WPJ, Schoonman GG, Ferrari MD, Terwindt GM. Alcoholic beverages as trigger factor and the effect on alcohol consumption behavior in patients with migraine. *Eur J Neurol*. 2019;26(4):588-595.
35. Ashina S, Terwindt GM, Steiner TJ, et al. Medication overuse headache. *Nat Rev Dis Primers*. 2023;9(1):5.
36. Gouveia-Figueira S, Goldin K, Hashemian SA, et al. Plasma levels of the endocannabinoid anandamide, related N-acylethanolamines and linoleic acid-derived oxylipins in patients with migraine. *Prostaglandins Leukot Essent Fatty Acids*. 2017;120:15-24.
37. Buczynski MW, Parsons LH. Quantification of brain endocannabinoid levels: methods, interpretations and pitfalls. *Br J Pharmacol*. 2010;160(3):423-442.
38. Lehmann R. From bedside to bench-practical considerations to avoid pre-analytical pitfalls and assess sample quality for high-resolution metabolomics and lipidomics analyses of body fluids. *Anal Bioanal Chem*. 2021;413(22):5567-5585.
39. Jonasdottir HS, Brouwers H, Toes REM, Ioan-Facsinay A, Giera M. Effects of anticoagulants and storage conditions on clinical oxylipid levels in human plasma. *Biochimica et biophysica acta Molecular and cell biology of lipids*. 2018;1863(12):1511-1522.
40. van Dongen RM, Zielman R, Noga M, et al. Migraine biomarkers in cerebrospinal fluid: A systematic review and meta-analysis. *Cephalalgia*. 2017;37(1):49-63.
41. Hillard CJ. Circulating Endocannabinoids: From Whence Do They Come and Where are They Going? *Neuropsychopharmacology*. 2018;43(1):155-172.
42. Engeli S, Böhnke J, Feldpausch M, et al. Activation of the peripheral endocannabinoid system in human obesity. *Diabetes*. 2005;54(10):2838-2843.

43. Piyanova A, Lomazzo E, Bindila L, et al. Age-related changes in the endocannabinoid system in the mouse hippocampus. *Mech Ageing Dev*. 2015;150:55-64.

44. Wagner EJ. Sex differences in cannabinoid-regulated biology: A focus on energy homeostasis. *Front Neuroendocrinol*. 2016;40:101-109.

45. Craft RM, Marusich JA, Wiley JL. Sex differences in cannabinoid pharmacology: a reflection of differences in the endocannabinoid system? *Life Sci*. 2013;92(8-9):476-481.

46. Hirvonen J, Zanotti-Fregonara P, Gorelick DA, et al. Decreased Cannabinoid CB1 Receptors in Male Tobacco Smokers Examined With Positron Emission Tomography. *Biol Psychiatry*. 2018;84(10):715-721.

47. Kranaster L, Hoyer C, Aksay SS, et al. Electroconvulsive therapy enhances endocannabinoids in the cerebrospinal fluid of patients with major depression: a preliminary prospective study. *Eur Arch Psychiatry Clin Neurosci*. 2017;267(8):781-786.

48. Hill MN, Miller GE, Carrier EJ, Gorzalka BB, Hillard CJ. Circulating endocannabinoids and N-acyl ethanolamines are differentially regulated in major depression and following exposure to social stress. *Psychoneuroendocrinology*. 2009;34(8):1257-1262.

49. Patel S, Hillard CJ. Role of endocannabinoid signaling in anxiety and depression. *Curr Top Behav Neurosci*. 2009;1:347-371.

50. Jesulola E, Micalos P, Baguley IJ. Understanding the pathophysiology of depression: From monoamines to the neurogenesis hypothesis model - are we there yet? *Behav Brain Res*. 2018;341:79-90.

5

Prostaglandin-E_2 levels over the course of glyceryl trinitrate provoked migraine attacks

Aster V.E. Harder*, Gerrit L.J. Onderwater*, Robin M. van Dongen, Marieke Heijink, Erik W. van Zwet, Martin Giera, Arn M.J.M. van den Maagdenberg* and Gisela M. Terwindt*

Neurobiol. Pain. 2023;13:100112

Abstract

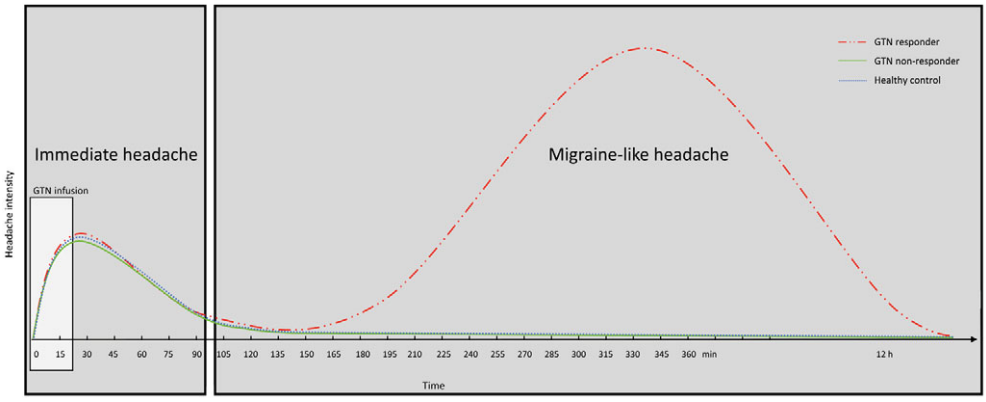
Administration of glyceryl trinitrate (GTN), a donor of nitric oxide, can induce migraine-like attacks in subjects with migraine. Provocation with GTN typically follows a biphasic pattern; it induces immediate headache in subjects with migraine, as well as in healthy controls, whereafter only subjects with migraine may develop a migraine-like headache several hours later. Interestingly, intravenous infusion with prostaglandin-E_2 (PGE_2) can also provoke a migraine-like headache, but seems to have a more rapid onset compared to GTN. The aim of the study was to shed light on the mechanistic aspect PGE_2 has in migraine attack development. Therefore, PGE_2 plasma levels were measured towards the (pre)ictal state of an attack, which we provoked with GTN. Blood samples from women with migraine (n = 37) and age-matched female controls (n = 25) were obtained before and ~140 min and ~320 min after GTN infusion. PGE_2 levels were measured using liquid chromatography tandem mass spectrometry (LC-MS/MS) analysis. Data was analyzed using a generalized linear mixed-effect model. Immediate headache after GTN infusion occurred in 85% of migraine participants and in 75% of controls. A delayed onset migraine-like attack was observed in 82% of migraine subjects and in none of the controls. PGE_2 levels were not different between the interictal and preictal state ($P = 0.527$) nor between interictal and ictal state (defined as having migraine-like headache) ($P = 0.141$). Hence, no evidence was found that a rise in PGE_2 is an essential step in the initiation of GTN-induced migraine-like attacks.

KEYWORDS: Migraine, Glyceryl trinitrate, Prostaglandin E_2, Preictal, Plasma

Introduction

Migraine is a common multifactorial paroxysmal brain disorder with a life-time prevalence of 15-20%, causing disability worldwide.[1, 2] A typical migraine attack consists of a preictal, an ictal (aura and/or headache), and a postictal (postdromal) phase.[3] The pathophysiological mechanisms underlying migraine attacks, however, remain to be fully elucidated. Notably, migraine-like attacks can be induced in subjects with migraine, but not in healthy controls, by the administration of glyceryl trinitrate (GTN), a donor of nitric oxide (NO). Two types of NO-induced headaches have been reported (**Figure 1**).[4] First, in both migraine subjects and healthy controls an immediate headache develops within the first hour of GTN infusion. This headache is of mild to medium severity and typically resolves within an hour after GTN administration. Second, only in subjects with migraine, a delayed onset migraine-like headache (moderate to severe, accompanied by associated symptoms such as nausea, vomiting, photo- and/or phonophobia) may develop within 12 hours after GTN infusion.[5, 6] This different response to GTN in cases compared to controls may provide clues for mechanisms underlying migraine attacks. Whereas the immediate headache seems related to a direct action of the NO-cGMP pathway via vasodilation by smooth muscle relaxation,[7], independent of neuropeptide calcitonin gene-related peptide (CGRP) release,[8] the delayed migraine-like attack is thought to be the result of trigeminovascular activation mediated via CGRP release.[5, 7, 9]

Figure 1 Schematic headache pattern after the start of the GTN infusion consisting of the immediate headache and the migraine-like attack



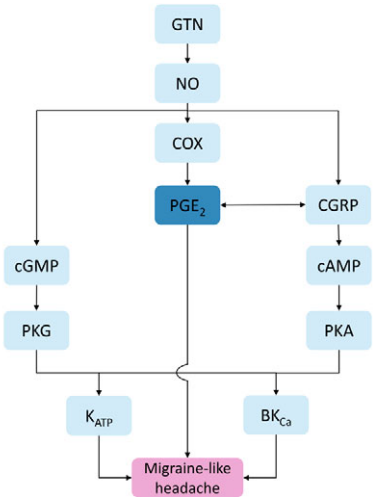
Three different response groups can be distinguished. The red two-dot chain line represents a typical headache pattern for a subject with migraine who responded to GTN (GTN responder),this is combined with typical patterns for a subject with migraine who did not respond to GTN (GTN non-responder) represented by the contineous green line, and a healthy control represented by the dotted blue line. GTN, glyceryl trinitrate. Adapted from Onderwater et al.[21]

Besides CGRP there is ample evidence that prostaglandins may be pivotal in the development of GTN-induced migraine-like attacks, and possibly spontaneous migraine attacks.[10] NO stimulates cyclooxygenase (COX-1 and COX-2) synthesis, which are enzymes that produce prostaglandins. [11] Non-steroidal anti-inflammatory drugs (NSAIDs), which inhibit prostaglandin synthesis, are a first

line treatment for migraine headaches. Cortical spreading depolarization, the underlying mechanism for the migraine aura, causes COX-2 upregulation potentially leading to increased prostaglandin levels.[12,13] The role of prostaglandins has also been investigated in provocation experiments in migraine subjects, in most cases in those without aura, demonstrating that intravenous infusion of prostaglandin I_2 (PGI_2) and E_2 (PGE_2) induces migraine like-attacks in 75% of participants with migraine.[14,15] Remarkably, subjects with migraine typically developed rapid onset migraine-like attacks, with a median onset of 20 minutes, in 25% (PGI_2) and 58% (PGE_2) of cases, which is in contrast to provocation with GTN, pituitary adenylate cyclase–activating peptide (PACAP) and CGRP for which the majority of cases develops a delayed onset migraine-like attack after at least a few hours.[14,16]

It has been shown that PGE_2 is mediated via CGRP release, and *vice versa*,[10] as evidenced by observations that PGE_2 stimulates the release of CGRP in rat trigeminal neurons,[17] trigeminal nucleus caudalis [18], and trigeminal ganglia,[19] while CGRP induces secondary release of PGE_2.[20] All the above suggests that PGE_2 may be closely upstream of GTN-induced migraine attacks (**Figure 2**).

Figure 2 Pathway relevant to nitroglycerin (GTN)-induced migraine-like headache



Nitroglycerine (GTN) liberates nitric oxide (NO) in peripheral and cerebral structures. NO subsequently, by binding to soluble guanylyl cyclase (sGC), increases cyclic guanosine monophosphate (cGMP).[22] Furthermore, NO can interact with superoxide to form peroxynitrite. Peroxynirite ($ONOO^-$) is a proinflammatory compound and has been implicated in the pathophysiology of not only stroke, but also pain and is gaining interest in the migraine field.[23,24] Additionally, NO on the one hand stimulates COX synthesis and prostaglandin E_2 (PGE_2) production,[11] and on the other hand stimulates CGRP, independent of the cGMP signaling pathway.[8] Subsequently, CGRP has been shown to induce PGE_2,[20] and vice versa.[17-19] In turn, it has been shown that $ONOO^-$ when inducing inflammation-derived hyperalgesia acts via the COX-to-PGE_2 pathway,[25] and $ONOO^-$ is also implicated along the trigeminovascular migraine pathway associated with CGRP.[26] PKG-mediated phosphorylation opens ATP-sensitive potassium channels (K_{ATP}) channels and large (big)-conductance calcium-activated K^+ (BK_{Ca}) via the NO/cGMP/PKG pathway.[27,28] CGRP activates vascular smooth muscle K_{ATP} channels and BK_{Ca} channels via cyclic adenosine monophosphate (cAMP) and protein kinase A (PKA) phosphorylation.[29,30] PGE_2 can also either increase or decrease the amount of cAMP depending on to which receptor it binds.[31] Opening of K_{ATP} and BK_{Ca} channels generates outward K^+ currents and causes vasodilation,[32] and can eventually lead to a migraine-like attack.[33,34] Provocation with PGE_2 in subjects with migraine leads to a rapid-onset migraine attack,[14] which suggests that PGE_2 is closely upstream of a migraine-like attack.

We here aimed to shed light on the mechanistic aspect PGE_2 has in migraine attack development, as it might serve as a possible drug target. We measured PGE_2 plasma levels in female subjects with migraine and age-matched female healthy controls in the (pre)ictal phases of GTN provoked migraine-like attacks to assess whether PGE_2 levels change as part of GTN-induced migraine attacks.

Methods

Participants

This study was conducted as part of an extensive migraine provocation study, described in Onderwater et al.[21] In total, 37 female subjects with migraine (without aura) and 25 age-matched female healthy controls were included. Due to the predominance of migraine in females only female subjects were included in the study. Migraine was diagnosed in accordance with the International Classification of Headache Disorders (ICHD-3).[3] Participants with migraine experienced one or more migraine attacks per month during the past six months. Subjects with chronic migraine or medication-overuse headache were excluded. Healthy controls were free of (severe) headaches, neurological or psychiatric disorders and had no family history of severe primary headaches, but were allowed to occasionally have tension-type headaches. None of the participants used chronic medication other than oral contraceptives. The study was approved by the ethics committee of the Leiden University Medical Center and in accordance with the World Medical Association Declaration of Helsinki. All participants provided written informed consent prior to the study.

Study design

During the study day, each participant was subjected to detailed interviews over the course of the day and underwent three blood withdrawals. Samples were drawn by venipuncture from the medial cubital vein. Participants were attack-free at least three days prior to the investigation and had been instructed to refrain from using prophylactic medication for at least four weeks. Apart from abstaining from alcoholic beverages, caffeinated beverages, and smoking for at least 8 hours prior to and during the study day there were no dietary restrictions. Before GTN infusion, all participants underwent a baseline assessment consisting of a neurological examination, headache assessment, and a blood withdrawal in ethylenediaminetetraacetic acid (EDTA)-containing tubes was performed for baseline measurement [T0]. Following the baseline measurement, participants received an intravenous infusion of GTN (0.5 μg/kg/min over 20 minutes) between 9:45 and 10:45 AM, in supine position. After GTN infusion, blood was again drawn from participants at two time points, namely ~140 minutes [T1] and ~320 minutes after the start of GTN infusion [T2]. To avoid biochemical interference in the processes related to the initiation and onset of a migraine-like headache, participants were requested to abstain from using acute migraine attack medication until after the 3rd and final blood measurement [T2]. Blood was centrifuged at room

temperature for 20 minutes (2,000 rpm, 622 *g*). The supernatant was transferred to a 15-mL polypropylene tube (Greiner Bio-One CELLSTAR®), inverted several times, and divided in 0.5-mL aliquots (1.0 mL Nunc™ cryotubes). Plasma samples were stored at -80˚C until further use; no extra freeze-thaw cycles were allowed.

Migraine-like headache and criteria

Participants were notified that GTN could potentially induce a headache, without any information regarding the expected onset or course. Questionnaires were performed, as described in Onderwater et al. [21] In short, during the 20-minute GTN infusion, headache characteristics and associated symptoms were documented every 5 minutes. After the infusion period, the occurrence of premonitory symptoms, headache, and associated symptoms was documented every 15 minutes until 5 hours after GTN infusion. After the study day (6 hours after GTN infusion), to determine GTN responder status, participants filled in a headache diary and were asked for headache fitting migraine-like attack onset in a telephone follow-up ~3 days after participation. Headache intensity was scored with a verbal rating scale (VRS) from 0 to 10 (0 indicating no headache, 1 indicating a very mild headache and 10 indicating the worst possible headache pain imaginable). In addition, the response form included the type of pain, localization, associated symptoms, premonitory symptoms, and adverse events. Furthermore, subjects with migraine were asked whether the reported headache resembled their usual migraine attacks. Despite the resemblance with spontaneous attacks, induced attacks are referred to as 'migraine-like headaches', as they cannot fulfil all criteria of a migraine without aura attack; for this the attack needs to be spontaneous and last (untreated) at least 4 hours.[3] Therefore, in accordance with earlier provocation studies,[16] migraine-like attack onset (ictal) was determined as either (1) a moderate to severe headache (VRS ≥ 4) fulfilling ICHD-3 criteria C and D for migraine without aura or (2) a headache described as mimicking the subject's usual migraine attack and treated with acute migraine medication.

PGE_2 quantification

PGE_2 was quantified in EDTA plasma using a method analogue for the quantification of 8-iso-PGF2α, previously described.[35] In short, 250 mL EDTA plasma was diluted with 2.0 mL sodium acetate buffer (0.1 M, pH 3.5) and 3 mL PGE2-d4 (50 ng/mL) in methanol (MeOH) was added. The samples were loaded onto C18 SPE cartridges (200 mg, 3cc; Waters, Sep-Pak, Milford, MA) that had been conditioned and equilibrated with MeOH and water. After a wash with water and n-hexane samples were eluted using methyl formate. Eluates were then dried under a gentle stream of nitrogen at 40˚C and reconstituted in 150 mL 40% MeOH.

Samples were measured by Liquid Chromatography (Shimadzu SIL-30AC autosampler, two Shimadzu LC-30AD pumps and a Shimadzu CTO-20AC column oven) coupled to a Sciex Qtrap 6500 mass spectrometer. Forty-mL samples were injected and separated on a C18 column

(Phenomenex, 50 × 2.1 mm, 1.7 μm). A gradient of 0.01% acetic acid in water (A) and 0.01% acetic acid in MeOH (B) was used to elute the components of interest from the column. The total flow rate was 400 mL/min. The column oven was set to 50˚C. The mass spectrometer (MS) was equipped with an ESI source and operated in negative scheduled MRM mode. The needle voltage was set to -4,500 V, the drying temperature to 450˚C, ion source gas 1/nebulizer gas (air) at 40 psi, ion source gas 2/drying gas (air) at 30 psi and the nebulizer gas (nitrogen) at 30 psi. For PGE_2 the transition used was 351/271, for PGE_2-d4 355/193. PGE_2 was identified based on its tandem MS transition and relative retention time and, quantified using external calibration.

Statistical analysis

We aimed to investigate the role of PGE_2 over the course of a provoked migraine attack, healthy controls were included to ensure that direct pharmacological effects of the provocation substance itself is not incorrectly labelled as a marker for provoked attacks. As we were primarily interested in the effect of different phases on PGE_2 levels in blood, we distinguished three phases: interictal (outside a migraine-like attack), preictal (before a migraine-like headache of which the onset is ≤ 12 hours after GTN infusion), and ictal (migraine-like headache). To account for repeated measurements within each subject, we used a linear mixed model with a random effect per person and unstructured correlation, the same model was used previously.[36] The outcome (dependent variable) was the measured PGE_2 concentration. Predictors (independent variables) were age, diagnosis (migraine or control), time point (T0, T1, T2) and migraine phase (interictal, preictal, ictal). Controls were coded as "interictal" at all time points. Furthermore, we added the interaction between time point and diagnosis to account for subjects with migraine possibly reacting differently to GTN than controls, irrespective of migraine phase. Statistical analyses were performed using SPSS (version 25.0, IBM SPSS Statistics for Windows, IBM Corp, Armonk, NY). In this study, data was collected as part of an extensive larger study and, therefore, no *a priori* power calculations were performed for this sub-study.

Results

Clinical characteristics

We initially included n = 37 participants with migraine and n = 25 healthy controls, of which five participants were excluded for further analyses. Two cases were removed as GTN infusion was not performed, both participants withdrew from participation after the baseline measurement. Two cases were excluded, because we were unable to classify the provoked headache attack (one not fully fulfilling a migraine-like headache nor classifying as a non-responder and the other developed a migraine-like attack, but already proceeded to a postdrome state during the study day). One healthy control was excluded due to a (first) provoked migraine-like headache. In total, data from n = 33 participants with migraine and n = 24 healthy controls were included in the

analyses. The demographic and clinical characteristics of cases and controls are shown in **Table 1**. There were no adverse events reported.

Table 1 Demographic and clinical characteristics of the study population

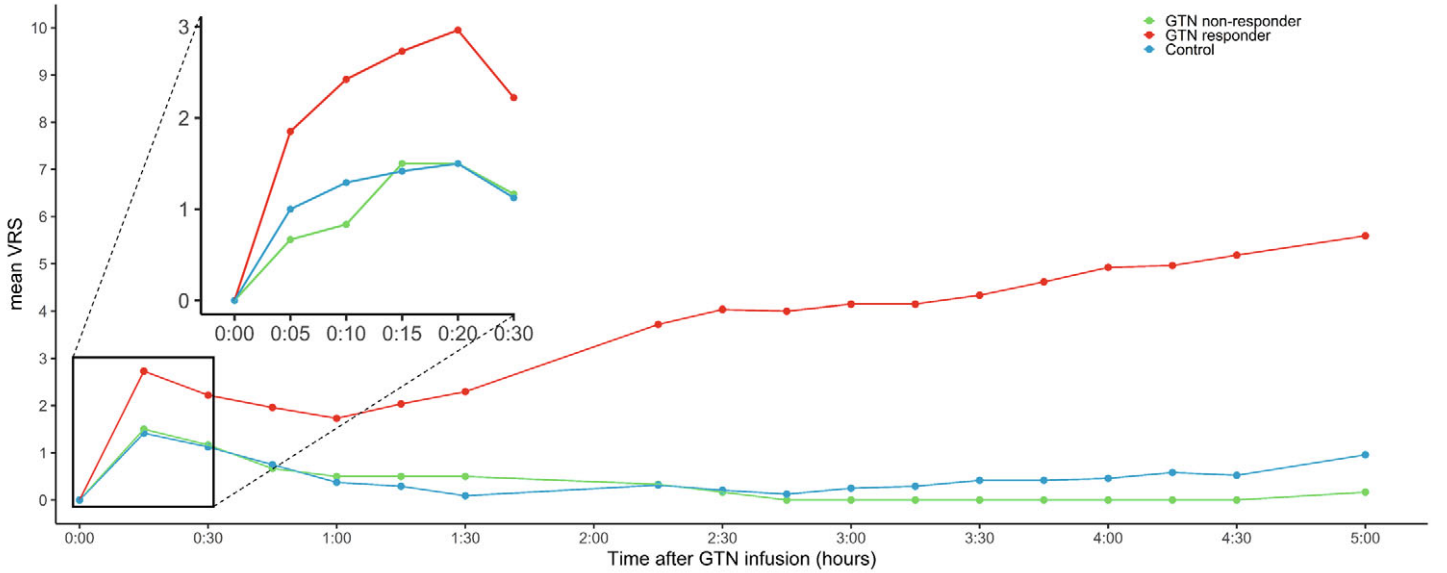
Participants Characteristics	Migraine cases (n = 33)	Healthy controls (n = 24)	P value	GTN responders (n = 27)	GTN non-responders (n = 6)
General characteristics					
Age	34.3 ± 8.2	35.2 ± 9.1	0.709†	35.2 ± 8.4	30.3 ± 6.4
BMI	22.9 ± 2.6	23.2 ± 2.7	0.714†	23.3 ± 2.7	21.6 ± 1.4
Smoking (n, %)	5 (15.1%)	3 (12.5%)	1.000‡	5 (18.5%)	0 (0%)
Migraine characteristics					
Age of onset	16.3 ± 5.6		-	17.4 ± 4.7	11.2 ± 6.7
Migraine days (attack/month)	4.7 ± 2.7		-	5.1 ± 2.8	2.7 ± 0.8

Values are expressed as absolute values and percentage or mean ± SD., *P* values are calculated with † Student's *t*-test, ‡ Fisher's Exact Test. GTN, glyceryl trinitrate, BMI, body mass index.

GTN response

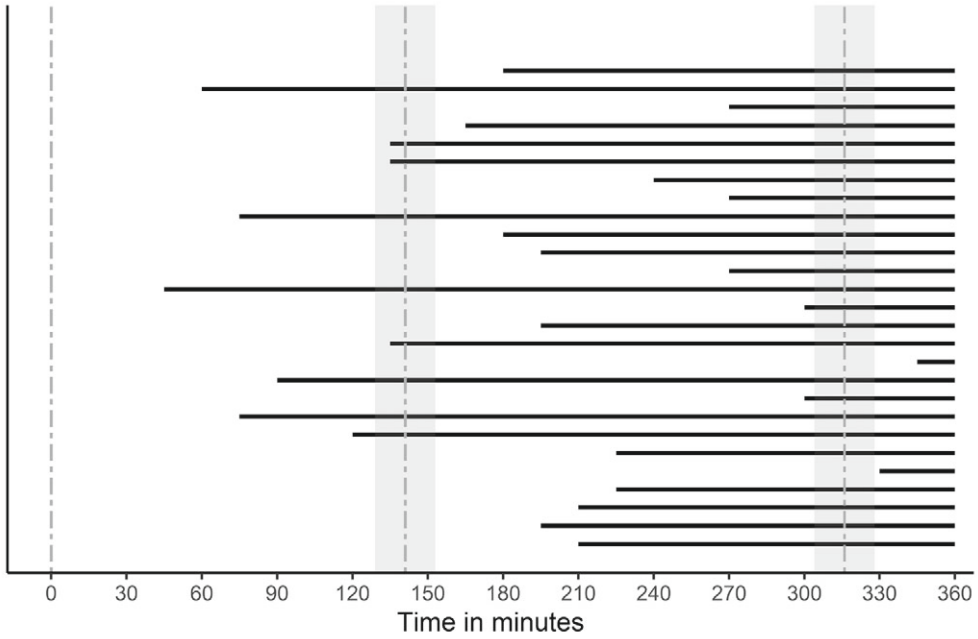
In total, n = 28 subjects with migraine (85%) and n = 18 healthy controls (75%) developed an immediate headache (VRS ≥ 1) during the GTN infusion. At 5 minutes after the start of GTN infusion, the mean VRS value was 1.6 for those with migraine (1.8 for responders and 0.7 for non-responders) and 1 for controls. In total, n = 20 subjects with migraine (61%) and n = 10 healthy controls (42%) had an immediate headache. The mean VRS value increased until the end of the GTN infusion to 2.6 (3 for responders and 1.5 for non-responders) and 1.5, for subjects with migraine and controls, respectively. At 20 minutes, 26 subjects with migraine (79%) and 13 controls (54%) experienced a headache. Overall, the immediate headache was mild to moderate in severity and generally resolved rapidly after termination of the infusion (**Figure 3, Figure S1**). In some subjects with migraine a "headache-free" interval was absent (**Figure S1**), in those subjects the headache continued after infusion and eventually became more severe with characteristics of a migraine-like attack. The mean VRS for those who responded to GTN (responders) continued to increase, as the headache became more severe although only at a later stage met the criteria of migraine and in those who classified as non-responders the headache severity decreased. Generally, the immediate phase is considered to be 0-90 minutes post infusion. Four subjects developed migraine within this timeframe. One subject with migraine developed a headache fulfilling the migraine-like criteria within one hour after the start of GTN infusion, one at 60 minutes, and two at 75 minutes. Eventually, 27 (82%) subjects with migraine receiving GTN experienced a migraine-like attack (**Figure 4**) during the study day and 6 (18%) did not experience such an attack, hence they were labelled as GTN responders and GTN non-responders, respectively (**Table 1**). Migraine-like attack onset ranged between 45 and 345 minutes (mean 192 ± 84 minutes) (**Figure 4**).

Figure 3 Mean verbal rating scale (VRS) for headache severity per responder group



The X-axis represents time after GTN infusion and the Y-axis the mean VRS. In red, participants with migraine who responded to GTN (GTN responders), participants with migraine who did not respond to GTN (GTN non-responders) in green and healthy controls in blue. GTN, glyceryl trinitrate; VRS, verbal rating scale. All subjects (including those without headache) were used in the calculation of the average. For some time points there were many missing values, this resulted in exclusion of these time points from the figure. Whiskers represent the standard deviation from the mean.

Figure 4 Timing migraine onset in GTN responders



The onset of migraine is plotted for each glyceryl trinitrate (GTN) responder with respect to time after GTN infusion. The start of the black continuous line represents the timing of onset of migraine attack per individual. The dotted line represents the blood draw timepoints T0, T1 and T2 at 0, ~140 and ~320 minutes, respectively, after the start of the GTN infusion.

Table 2 Median PGE_2 concentrations over time independent of migraine phase

Group	[T0]	[T1]	[T2]
GTN responders	0.044 (0.02-0.10)	0.053 (0.03-0.10)	0.049 (0.03-0.08)
GTN non-responders	0.052 (0.01-0.09)	0.031 (0.01-0.07)	0.040 (0.02-0.07)
Controls	0.044 (0.02-0.08)	0.043 (0.03-0.09)	0.060 (0.03-0.09)

Values are the uncorrected medians of absolute concentrations in ng/mL with their interquartile range.

[T0] = baseline, [T1] = ~140 minutes after the start of GTN infusion, [T2] = ~320 minutes after GTN infusion. GTN responder, migraine patients who responded to GTN; GTN non-responder, migraine patients who did not respond to GTN. GTN, glyceryl trinitrate.

PGE_2 in relation to migraine-like attack onset

The level of PGE_2 per individual varied per time point (**Table 2**). To determine whether PGE_2 levels were linked to the various phases (baseline, preictal and ictal) of a migraine attack, a generalized linear mixed model was used. The transition from an interictal state towards a migraine-like attack had no influence on PGE_2 concentration ($F(2, 69.70) = 1.235$, $P = 0.297$). Both the transition from "interictal to preictal" ($P = 0.527$) and "interictal to ictal" ($P = 0.141$) phase of GTN-induced migraine-like attacks had no influence on PGE_2 concentration (Table S1).

Discussion

We performed a GTN provocation study in subjects with migraine and healthy controls and found that 82% of migraine participants developed a delayed onset migraine-like attack. We prospectively assessed PGE_2 levels at three time points selected over the course of provoked migraine-like attacks and compared these to those without provoked attacks and controls. We found no evidence that GTN-induced migraine-like headaches are characterized by changes in plasma PGE_2 levels towards the (pre)ictal state. This suggests that a rise in PGE_2 is not an essential step in the initiation of GTN-induced migraine-like attacks.

PGE_2 is able to induce rapid-onset migraine-like attacks in subjects with migraine within 90 minutes,[14] in contrast to provocation with substances such as PACAP, CGRP and GTN that result in a delayed (after a few hours) onset of a migraine-like attack.[14, 16] Thus, we hypothesized that PGE_2 could be one of the molecules involved in a(n experimentally induced) migraine attack. Given that administration of PGE_2 can cause a rapid-onset migraine-like attack, in contrast to the other provocative substances, PGE_2 may perhaps serve as a marker for upcoming migraine attacks, albeit that the timing of blood sampling is important. In our study, we used the GTN provocation model to assess the role of PGE_2. It has been hypothesized that the time it takes to develop delayed migraine-like attack is due to various processes that include the regulation of gene expression and proteins ultimately resulting in migraine-like attacks in subjects with migraine with a median attack onset of 3 to 6 hours, after infusion of the provocation substance. Afterall, in animal models of migraine, GTN activates the COX-2-PGE_2 pathway in the brainstem not before 4 hours after GTN administration.[37] However, based on our proposed mechanism and the PGE_2 human provocation studies with rapid onset of provoked migraine-like headaches, we expected a rise in PGE_2 to be close to the start of a migraine attack as an early marker of migraine, which would fit our time points of blood withdrawal. The alternative explanation that we did not find a rise in PGE_2 levels might indicate that the pathway activated by GTN towards a migraine-like attack does not primarily act via PGE_2. One can envisage that pathways, independent of PGE_2 via for instance cGMP or cAMP, are more strongly activated than the PGE_2-pathway when GTN is administered. Another explanation might be that a rise in PGE_2 is very locally and hence not measurable in blood.

To our knowledge no other study measured PGE_2 levels over the course of GTN-induced migraine-like attack in subjects with migraine. Still, few studies reporting measurements of PGE_2 levels during spontaneous migraine attacks suggested those to be elevated in blood,[38, 39] and saliva.[40] More specifically, in contrast to our study, a much smaller study of only five subjects with migraine reported an increase in PGE_2 levels in jugular venous blood peaking between 2 and 6 hours after the start of a spontaneous migraine attack and normalizing towards the end of the attack.[38] In our study the mean attack onset was ~192 minutes, hence many cases were over 2 hours into their delayed migraine-like attack at the ~320-minute time point, which suggests that our timing was

not different from the spontaneous migraine attack study and thus could have picked up a similar rise in PGE_2 levels. In addition, two studies found that PGE_2 levels in plasma,[39] (18 cases and 12 controls) and saliva [40] (6 cases and 9 controls) in subjects with migraine were lower compared to controls outside attacks and increased during a spontaneous attack surpassing the levels found in controls. Although this was not our primary question, we tested this and did not find a difference in baseline PGE_2 levels between cases and controls. Giving the small number of participants in previous studies, our larger study should have been able to reveal differences in PGE_2 levels during GTN-induced migraine-like attacks. Another reason for the discrepancy with earlier studies might be in the measuring techniques used and/or the matching and correction of data. For our study we used a highly reliable, standardized technique for measuring PGE_2 levels and additionally have minimized external effects on PGE_2 levels, by careful matching and correcting for multiple factors to single out the effect of PGE_2 on a migraine attack. Whereas such external effects do not seem to have affected our results, they might have played a role in earlier studies. Another possibility is that spontaneous attacks are not always the same as provoked attacks (e.g. GTN provocation in migraine patients with aura leads to a migraine-like attack, but not an aura). This may indicate that in spontaneous attacks different pathways may be initiated depending on headache (sub)type, none the less these pathways ultimately lead to the same migraine headache.

We envisage several possible explanations why we found no evidence for a change in PGE_2 levels over the course of a GTN-induced attack. PGE_2 acts via four distinct G protein–coupled receptors EP1, EP2, EP3 and EP4. Ligand binding to the different EP receptors leads to the activation of distinct downstream signaling pathways, resulting in distinct biological outcomes,[31, 41] one of these second messengers being cyclic adenosine monophosphate (cAMP).[31] Via its receptors, PGE_2 is known to play a role in nociceptive pain processing and inflammation,[42, 43] exerting both damaging pro-inflammatory and protective anti-inflammatory effects in the brain.[44-46] Thus, the PGE_2 response is dependent on the array of receptors cells express as well as on intracellular pathways to which they are coupled [46, 47]. Hence, any involvement of PGE_2 in the pathogenesis of migraine may be very complex.

As mentioned previously, the immediate headache is thought to be the result of vasodilation via the NO-cGMP pathway,[7] independent of CGRP release [8], whereas the delayed migraine-like attack is thought to be the result of trigeminovascular activation mediated via CGRP.[5, 7, 9] However, there likely is extensive cross talk between both pathways (for details see **Figure 2**). For instance, on a cellular level multiple components in the migraine pathway are known to be vasodilators, but can also lead to migraine attacks. As exemplified by ATP-sensitive potassium (K_{ATP}) channel openers (levcromakalim) and (big)-conductance calcium-activated K^+ (BK_{Ca}) channel opener (MaxiPost), both activated via the NO-cGMP pathway, which is known to play a role in the immediate headache, but activation of these channels can also induce migraine-like attacks.[33, 34] However, the rather long delay of several hours between infusion of levcromakalim/MaxiProst and the occurrence of a migraine-like attack (with a median time of 3 hours) indicates

that various mediators must be involved in slower cascades of events leading to a migraine-like attack. Evidence for cross talk is that both the administration of CGRP (pathway via cAMP) and sildenafil (pathway via cGMP) can lead to a migraine-like attack, suggesting that convergence to a common cellular determinator seem to exist ultimately triggering similar attacks.[48] Given that the median time until an attack for CGRP is ~165 minutes, so much shorter than the ~285 minutes for sildenafil, one can envisage that CGRP acts more downstream in the generation of a migraine attack. Such sequential actions, given the cross talk between CGRP and PGE_2, especially in GTN-induced attacks, seems in line with a chain-of-events-pathway (**Figure 2**).

Our study has several limitations. We collected a large number of blood samples of migraine participants and healthy controls. Each participant was sampled at three fixed times during the study day in an attempt to measure PGE_2 concentrations during attack development and during the attack itself. We find that the 95% confidence interval of the change in PGE2 levels from interictal to another phase extends from -0.02 to 0.05 ng/mL. Of course, the onset of the attack varies between subjects and did not align perfectly with the measurement times. Moreover, we must account for a possible temporal effect of the GTN infusion on PGE_2 concentrations. Combined with within and between subject measurement variation, we must acknowledge that not finding a statistically significant difference in PGE_2 levels over the course of an induced migraine-like attack does not prove the absence of such an effect. There may yet exist subtle, short-duration, variations in PGE_2 levels that we could not detect. Furthermore, we have used LC-MS/MS which is distinct from the more often used ELISA kits to measure PGE_2, this might make it difficult to compare absolute concentrates between studies. However, by using this method we were able to detect very low levels of PGE_2 with good accuracy, despite the short half-life of PGE_2. Furthermore, whereas we did not observe changes in PGE_2 levels in blood it is conceivable that levels may be different in cerebrospinal fluid, as increased PGE_2 levels have been reported indicative for probable Alzheimer's disease.[49] However, we deem it too unethical and unlikely that subjects with migraine (and controls) are willing to participate in a provocation study with, logically, repeated lumbar punctures to get information on PGE_2 levels over time. Finally, our study only consists of females to prevent any sex effects, which may limit the generalizability of our findings to male migraine patients. Additionally, although we have performed our study in a female only population to account for the most notable sex hormone differences, small differences in cycle and use of contraceptives might be of influence in the downstream provocation pathways.

Supplementary information

Table S1 https://ars.els-cdn.com/content/image/1-s2.0-S2452073X22000290-mmc2.docx

Figure S1 https://ars.els-cdn.com/content/image/1-s2.0-S2452073X22000290-mmc1.pdf

Author's contributions: Aster V.E. Harder: Conceptualization, Data curation, Formal analysis, Investigation, Writing original draft, Methodology, Project Administration. Gerrit L.J. Onderwater: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project Administration, Writing – Review & Editing. Robin M. van Dongen: Conceptualization, Investigation, Writing – Review & Editing Marieke Heijink: Methodology, Validation, Writing – Review & Editing. Erik W. van Zwet: Formal analysis, Methodology, Writing – Review & Editing. Martin Giera: Conceptualization, Resources, Validation, Writing – Review & Editing. Arn M.J.M. van den Maagdenberg: Conceptualization, Funding Acquisition, Resources, Supervision, Writing – Review & Editing. Gisela M. Terwindt: Conceptualization, Funding Acquisition, Resources, Supervision, Writing – Review & Editing.

Conflict of interests: G.M.T. reports consultancy support from Novartis, Allergan/Abbvie, Lilly, Teva, Lundbeck and independent support from Dutch Research Council, the Dutch Heart & Brain Foundations, IRRF and Dioraphte. A.M.J.M.v.d.M reports research support from Schedule 1 Therapeutics, AbbVie and Praxis Precision Medicine. The remaining authors have no conflicts of interest to declare.

Acknowledgements: We greatly acknowledge the subjects who participated in this study.

Funding information: This study was supported by the Netherlands Organization for Health Research and Development (ZonMw) (Clinical Fellowship 90700217 and VIDI grant 917.11.31 to G.M.T.), and European Community (EC) FP7-EUROHEADPAIN (602633 to A.M.J.M.v.d.M.).

Data availability: Data not published within the article is available from the corresponding author upon reasonable request.

References

1. GBD 2019 Diseases and Injuries Collaborators. Global burden of 369 diseases and injuries in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. *Lancet*. 2020;396(10258):1204-1222.
2. Launer LJ, Terwindt GM, Ferrari MD. The prevalence and characteristics of migraine in a population-based cohort: the GEM study. *Neurology*. 1999;53(3):537-542.
3. Headache Classification Committee of the International Headache Society (IHS). The International Classification of Headache Disorders, 3rd edition. *Cephalalgia*. 2018;38(1):1-211.
4. Olesen J. The role of nitric oxide (NO) in migraine, tension-type headache and cluster headache. *Pharmacol Ther*. 2008;120(2):157-171.
5. Bagdy G, Riba P, Kecskemeti V, Chase D, Juhasz G. Headache-type adverse effects of NO donors: vasodilation and beyond. *Br J Pharmacol*. 2010;160(1):20-35.
6. Iversen HK, Olesen J, Tfelt-Hansen P. Intravenous nitroglycerin as an experimental model of vascular headache. Basic characteristics. *Pain*. 1989;38(1):17-24.
7. Akerman S, Williamson DJ, Kaube H, Goadsby PJ. Nitric oxide synthase inhibitors can antagonize neurogenic and calcitonin gene-related peptide induced dilation of dural meningeal vessels. *Br J Pharmacol*. 2002;137(1):62-68.
8. Bellamy J, Bowen EJ, Russo AF, Durham PL. Nitric oxide regulation of calcitonin gene-related peptide gene expression in rat trigeminal ganglia neurons. *Eur J Neurosci*. 2006;23(8):2057-2066.
9. Juhasz G, Zsombok T, Modos EA, et al. NO-induced migraine attack: strong increase in plasma calcitonin gene-related peptide (CGRP) concentration and negative correlation with platelet serotonin release. *Pain*. 2003;106(3):461-470.
10. Davis RJ, Murdoch CE, Ali M, et al. EP4 prostanoid receptor-mediated vasodilatation of human middle cerebral arteries. *Br J Pharmacol*. 2004;141(4):580-585.
11. Mollace V, Muscoli C, Masini E, Cuzzocrea S, Salvemini D. Modulation of prostaglandin biosynthesis by nitric oxide and nitric oxide donors. *Pharmacol Rev*. 2005;57(2):217-252.
12. Eikermann-Haerter K, Ayata C. Cortical spreading depression and migraine. *Curr Neurol Neurosci Rep*. 2010;10(3):167-173.
13. Yokota C, Inoue H, Kuge Y, et al. Cyclooxygenase-2 expression associated with spreading depression in a primate model. *J Cereb Blood Flow Metab*. 2003;23(4):395-398.
14. Antonova M, Wienecke T, Olesen J, Ashina M. Prostaglandin E(2) induces immediate migraine-like attack in migraine patients without aura. *Cephalalgia*. 2012;32(11):822-833.
15. Wienecke T, Olesen J, Ashina M. Prostaglandin I2 (epoprostenol) triggers migraine-like attacks in migraineurs. *Cephalalgia*. 2010;30(2):179-190.
16. Ashina M, Hansen JM, Olesen J. Pearls and pitfalls in human pharmacological models of migraine: 30 years' experience. *Cephalalgia*. 2013;33(8):540-553.
17. Jenkins DW, Feniuk W, Humphrey PP. Characterization of the prostanoid receptor types involved in mediating calcitonin gene-related peptide release from cultured rat trigeminal neurones. *Br J Pharmacol*. 2001;134(6):1296-1302.
18. Jenkins DW, Langmead CJ, Parsons AA, Strijbos PJ. Regulation of calcitonin gene-related peptide release from rat trigeminal nucleus caudalis slices in vitro. *Neurosci Lett*. 2004;366(3):241-424.
19. Neeb L, Hellen P, Boehnke C, et al. IL-1beta stimulates COX-2 dependent PGE(2) synthesis and CGRP release in rat trigeminal ganglia cells. *PLoS One*. 2011;6(3):e17360.

20. Kress M, Guthmann C, Averbeck B, Reeh PW. Calcitonin gene-related peptide and prostaglandin E2 but not substance P release induced by antidromic nerve stimulation from rat skin in vitro. *Neuroscience*. 1999;89(1):303-310.
21. Onderwater GLJ, Dool J, Ferrari MD, Terwindt GM. Premonitory symptoms in glyceryl trinitrate triggered migraine attacks: a case-control study. *Pain*. 2020;161(9):2058-2067.
22. Arnold WP, Mittal CK, Katsuki S, Murad F. Nitric oxide activates guanylate cyclase and increases guanosine 3':5'-cyclic monophosphate levels in various tissue preparations. *Proc Natl Acad Sci U S A*. 1977;74(8):3203-3207.
23. Taffi R, Vignini A, Lanciotti C, et al. Platelet membrane fluidity and peroxynitrite levels in migraine patients during headache-free periods. *Cephalalgia*. 2005;25(5):353-358.
24. Bredt DS. Endogenous nitric oxide synthesis: biological functions and pathophysiology. *Free Radic Res*. 1999;31(6):577-596.
25. Ndengele MM, Cuzzocrea S, Esposito E, et al. Cyclooxygenases 1 and 2 contribute to peroxynitrite-mediated inflammatory pain hypersensitivity. *FASEB J*. 2008;22(9):3154-3164.
26. Akerman S, Salvemini D, Romero-Reyes M. Targeting reactive nitroxidative species in preclinical models of migraine. *Cephalalgia*. 2021;41(11-12):1187-1200.
27. Schubert R, Nelson MT. Protein kinases: tuners of the BKCa channel in smooth muscle. *Trends Pharmacol Sci*. 2001;22(10):505-512.
28. Murphy ME, Brayden JE. Nitric oxide hyperpolarizes rabbit mesenteric arteries via ATP-sensitive potassium channels. *J Physiol*. 1995;486 (Pt 1)(Pt 1):47-58.
29. Miyoshi H, Nakaya Y. Calcitonin gene-related peptide activates the K+ channels of vascular smooth muscle cells via adenylate cyclase. *Basic Res Cardiol*. 1995;90(4):332-336.
30. Hosokawa S, Endoh T, Shibukawa Y, et al. Calcitonin gene-related peptide- and adrenomedullin-induced facilitation of calcium current by different signal pathways in nucleus tractus solitarius. *Brain Res*. 2010;1327:47-55.
31. Markovič T, Jakopin Ž, Dolenc MS, Mlinarič-Raščan I. Structural features of subtype-selective EP receptor modulators. *Drug discovery today*. 2017;22(1):57-71.
32. Chrissobolis S, Sobey CG. Inwardly rectifying potassium channels in the regulation of vascular tone. *Curr Drug Targets*. 2003;4(4):281-289.
33. Al-Karagholi MA, Hansen JM, Guo S, Olesen J, Ashina M. Opening of ATP-sensitive potassium channels causes migraine attacks: a new target for the treatment of migraine. *Brain*. 2019;142(9):2644-2654.
34. Al-Karagholi MA, Ghanizada H, Waldorff Nielsen CA, et al. Opening of BKCa channels causes migraine attacks: a new downstream target for the treatment of migraine. *Pain*. 2021;162(10):2512-2520.
35. Dekker J, Martherus T, Lopriore E, et al. The Effect of Initial High vs. Low FiO2 on Breathing Effort in Preterm Infants at Birth: A Randomized Controlled Trial. Clinical Trial. *Front Pediatr*. 2019;7504.:
36. Onderwater GLJ, Wijnen JP, Najac C, et al. Cortical glutamate and gamma-aminobutyric acid over the course of a provoked migraine attack, a 7 Tesla magnetic resonance spectroscopy study. *Neuroimage Clin*. 2021;32:102889.
37. Tassorelli C, Greco R, Armentero MT, et al. A role for brain cyclooxygenase-2 and prostaglandin-E2 in migraine: effects of nitroglycerin. *Int Rev Neurobiol*. 2007;82:373-382.
38. Sarchielli P, Alberti A, Codini M, Floridi A, Gallai V. Nitric oxide metabolites, prostaglandins and trigeminal vasoactive peptides in internal jugular vein blood during spontaneous migraine attacks. *Cephalalgia*. 2000;20(10):907-918.
39. Nattero G, Allais G, De Lorenzo C, et al. Relevance of prostaglandins in true menstrual migraine. *Headache*. 1989;29(4):233-238.
40. Vardi J, Flechter S, Alguati A, Regev I, Ayalon D. Prostaglandin--E2 levels in the saliva of common migrainous women. *Headache*. 1983;23(2):59-61.

41. Negishi M, Sugimoto Y, Ichikawa A. Molecular mechanisms of diverse actions of prostanoid receptors. *Biochim Biophys Acta*. 1995;1259(1):109-119.
42. Kawabata A. Prostaglandin E2 and pain--an update. *Biol Pharm Bull*. 2011;34(8):1170-3117.
43. Uda R, Horiguchi S, Ito S, Hyodo M, Hayaishi O. Nociceptive effects induced by intrathecal administration of prostaglandin D2, E2, or F2 alpha to conscious mice. *Brain Res*. 1990;510(1):26-32.
44. Pradhan SS, Salinas K, Garduno AC, et al. Anti-Inflammatory and Neuroprotective Effects of PGE(2) EP4 Signaling in Models of Parkinson's Disease. *J Neuroimmune Pharmacol*. 2017;12(2):292-304.
45. Liang X, Wang Q, Shi J, et al. The prostaglandin E2 EP2 receptor accelerates disease progression and inflammation in a model of amyotrophic lateral sclerosis. *Ann Neurol*. 2008;64(3):304-314.
46. Andreasson K. Emerging roles of PGE2 receptors in models of neurological disease. *Prostaglandins Other Lipid Mediat*. 2010;91(3-4):104-112.
47. Tilley SL, Coffman TM, Koller BH. Mixed messages: modulation of inflammation and immune responses by prostaglandins and thromboxanes. *J Clin Invest*. 2001;108(1):15-23.
48. Younis S, Christensen CE, Toft NM, et al. Investigation of distinct molecular pathways in migraine induction using calcitonin gene-related peptide and sildenafil. *Cephalalgia*. 2019;39(14):1776-1788.
49. Montine TJ, Sidell KR, Crews BC, et al. Elevated CSF prostaglandin E2 levels in patients with probable AD. *Neurology*. 1999;53(7):1495-1498.

PART II

Genetics of different headache forms

Genetic susceptibility loci in genome-wide association study of cluster headache

Aster V.E. Harder*, Bendik S. Winsvold*, Raymond Noordam*, Lisanne S. Vijfhuizen, Sigrid Børte, Lisette J.A. Kogelman, Irene de Boer, Erling Tronvik, Frits R. Rosendaal, Ko Willems van Dijk, Emer O'Connor, Carmen Fourier, Laurent F. Thomas, Espen S. Kristoffersen, Cluster Headache Genetics Working Group, Rolf Fronczek, Patricia Pozo-Rosich, Rigmor H. Jensen, Michel D. Ferrari, Thomas F. Hansen, John-Anker Zwart*, Gisela M. Terwindt* and Arn M.J.M van den Maagdenberg*

*shared authorship

Ann Neurol 2021;90:203-216

Abstract

Objective: Identifying common genetic variants that confer genetic risk for cluster headache.

Methods: We conducted a case-control study in the Dutch Leiden University Cluster headache neuro-Analysis program (LUCA) study population (n = 840) and controls from the Netherlands Epidemiology of Obesity Study (NEO) (n = 1,457), representing the general population. Replication was performed in a Norwegian sample of 144 cases from the Trondheim Cluster headache sample and 1,800 controls from the Nord-Trøndelag Health Survey (HUNT). Gene set and tissue enrichment analyses, blood cell-derived RNA-sequencing of genes around the risk loci and linkage disequilibrium score regression were part of the downstream analyses.

Results: An association was found with cluster headache for four independent loci ($r^2 < 0.1$) with genome-wide significance ($p < 5 \times 10^{-8}$), rs11579212 (odds ratio (OR) = 1.51, 95% CI 1.33-1.72 near *RP11-815M8.1*), rs6541998 (OR = 1.53, 95% CI 1.37-1.74 near *MERTK*), rs10184573 (OR = 1.43, 95% CI 1.26-1.61 near *AC093590.1*), and rs2499799 (OR = 0.62, 95% CI 0.54-0.73 near *UFL1/FHL5*), collectively explaining 7.2% of the variance of the phenotype. SNPs rs11579212, rs10184573 and rs976357, as proxy SNP for rs2499799 ($r^2 = 1.0$), replicated in the Norwegian sample ($p < 0.05$). Gene-based mapping yielded *ASZ1* as possible fifth locus. RNA-sequencing indicated differential expression of *POLR1B* and *TMEM87B* in cluster headache.

Interpretation: This GWAS identified and replicated genetic risk loci for cluster headache with effect sizes larger than those typically seen in complex genetic disorders.

Introduction

Cluster headache (CH) is a primary headache disorder characterized by attacks of intense unilateral orbital, supraorbital and/or temporal pain that lasts for 15-180 minutes and is associated with ipsilateral facial autonomic symptoms and/or restlessness. The majority of patients have episodic CH, with periods of cluster headache of weeks to months, alternating with attack-free periods of at least 3 months. In 10-15% of patients cluster periods never remit for longer than three months for at least one year, classifying them as chronic CH. The male-to-female ratio is 2:1.[1] Smoking and psychiatric co-morbidities are prevalent.[2] Current treatment strategies include aborting acute attacks and aim to reduce attack frequency with preventive treatment.[3] CH shows some phenotypic overlap with other trigeminal neuralgias, but also with migraine, e.g. in that some patients with migraine may also report autonomic features. Certain similar pathophysiological pathways are hypothesized to be involved in both CH and migraine.[4] Although these disorders share prominent features, they are clinically well distinguishable.[5]

The pathophysiology of CH is poorly understood, although vasomotor changes, inflammation, hypothalamic dysfunction, and dysregulation of the autonomic nervous system have been implicated as potential disease mechanisms.[6] Twin and family studies have highlighted the involvement of genetic factors in CH.[7] Thus far, most genetic studies used a hypothesis-driven approach and have examined a limited number of variants in genes linked to presumed pathways in CH. Most studied are variants in *HCRTR2*, which encodes the hypocretin (orexin) type 2 receptor that binds neuropeptides hypocretin-1 and -2 in the central nervous system. Still, initially positive genetic findings for *HCRTR2* associations[8-10] were not replicated in better-powered studies.[11, 12] Finally, the first, though very small hypothesis-free, Italian genome-wide association study (GWAS) investigating 99 patients with CH reported suggestive associations with genetic variants in *ADCYAP1R1* and *MME*,[13] but these findings were not replicated in a larger Swedish sample.[14]

To detect genetic variants for CH, we conducted a GWAS in a Dutch sample of 840 patients with CH and 1,457 controls from the same geographical region. Results were replicated in a Norwegian sample. Downstream analyses further assessed genes and mechanisms contributing to the pathogenesis of CH.

Methods

Patient recruitment and sample collection

The Dutch cluster headache study included 862 Dutch CH patients from the clinic-based Leiden University Cluster headache neuro-Analysis program (LUCA) that were recruited between 2010 and 2015 via the project's website. CH patients aged 18 years or older were included. Participants

fulfilling the screening criteria were asked to complete an extended questionnaire that focused on signs and symptoms of CH as outlined in the International Classification of Headache Disorders III (ICHD-III) criteria for CH.[5] Individual diagnoses were made upon visiting the outpatient clinic or using an validated algorithm (positive predictive value: ~92%) based on ICHD criteria.[15] CH cases were diagnosed in specialized headache centers to minimize misclassification. Controls (n = 1,671) were obtained from the Netherlands Epidemiology of Obesity Study (NEO) study,[16] a population-based sample that includes individuals aged 45-65 years living in a nearby municipality (Leiderdorp, The Netherlands) recruited between 2008 and 2012. Most cases and all controls originated from the same geographical region, in the Western part of the Netherlands. All participants were unrelated and of European ancestry. The local ethics committees approved the study. Written informed consent was obtained from all participants.

Genotyping, quality control, and imputation in the discovery stage

Genomic DNA was extracted from peripheral blood leukocytes according to standard protocols and genotyping of both cases and controls was performed using the Illumina Infinium CoreExome-24 v1.1 array according to the protocol from the manufacturer. Cases were genotyped at the Genomics-Core Facility at the Norwegian University of Science and Technology (Trondheim, Norway) and controls at the Centre National de Génotypage (Paris, France). For the cases, variant calling was performed with Genome Studio 2.0 following a standard quality protocol,[17] and the CHARGE best practice calling of the HumanExome Bead chip.[18] For the controls, calling was performed using the GenCall algorithm using standard settings as provided by Illumina. Quality control (QC) was performed according to standard procedures.[19] Markers with high missingness rates (≥ 2%), monomorphic variants and those failing the Hardy-Weinberg equilibrium were excluded. Individuals were excluded if they had a high proportion of missing genotype data (≥ 2%), inconsistent sex information, were related (PI-HAT ≥ 0.2), or were heterozygosity outliers. Principal component analysis (PCA) was performed on the pruned data set (with a 50-kb sliding window, $r^2 > 0.2$) using PLINK and population outliers were excluded. No overt population substructure between cases and controls was observed (data not shown). After combining the genotyped SNP information from LUCA and NEO, imputation was performed on the Michigan Imputation Server (https://imputationserver.sph.umich.edu/) using Haplotype Reference Consortium (HRC v1.1 2016) as a reference panel after phasing by Eagle (v2.3),[20] using the default parameters.

In total 345,064 SNPs from 2,297 individuals (840 cases and 1,457 controls) were available for imputation. Prior to analyses, variants with MAF ≤ 0.01 or imputation INFO score ≤ 0.6 were excluded, resulting in 7,578,399 SNPs available for analysis.

Statistical analysis in the discovery stage

Case-control SNP association analysis was performed using a logistic regression model implemented in SNPTEST (v2.5.2) for autosomal variants, with case-control status as outcome and assuming

additive allelic effects. The model was adjusted for sex. In addition, the model was adjusted for the first four principal components to minimize effects of confounding and population stratification. A Manhattan and a quantile-quantile (QQ) plot for test statistics were generated using R v3.6.1 (R Foundation for Statistical Computing, Vienna, Austria). We determined lead SNPs that were independent from each other at $r^2 < 0.1$ and further apart than 500 kb, with association $p < 5 \times 10^{-8}$. Positional gene mapping and fine mapping of significant loci was performed using Functional Mapping and Annotation v1.3.6 (FUMA), Probabilistic Identification of Causal SNPs (PICS), and Locuszoom.[21-23] The proportion of variance explained by a given SNP was calculated using Nagelkerke pseudo R^2.

Patient recruitment and data generation in the replication stage

Cases were recruited at the Norwegian Advisory Unit on Headaches, St Olav's Hospital, Trondheim (Norway) between 2005 and 2016, with the inclusion criterion being the definite diagnosis of CH according to ICHD-2 or ICHD-3,[24] made by a neurologist with special competence in headache disorders to minimize misclassification. As controls we used a random subset of 1,800 adult participants from the Nord-Trøndelag Health Study (HUNT) who did not have CH defined by the ICD-10 diagnosis G44.0 (Cluster headache syndrome) or the ICD-9 diagnosis 346.2 (Migraine variants, including cluster headache).[25]

A sample of 159 CH cases were genotyped with the Illumina Infinium CoreExome-24 v1.1. Calling was performed with Genome Studio 2.0, using the cluster file from the largest batch of 58,996 HUNT All-in controls (see below). The analysis followed the Genome Studio quality protocol,[17] and the CHARGE best practice calling of the HumanExome Bead chip.[18] The HUNT control samples were genotyped on three different Illumina HumanCoreExome arrays (HumanCoreExome12 v1.0, HumanCoreExome12 v1.1 and UM HUNT Biobank v1.0), and called as described elsewhere.[26] Markers with high missingness rates (≥ 2%), monomorphic variants and those failing the Hardy-Weinberg equilibrium were excluded. Individuals with high missingness rates (≥ 2%) or whose inferred sex contradicted with reported sex were excluded. A second round of quality control was performed after merging cases and all HUNT controls, excluding variants that were monomorphic, deviated from Hardy-Weinberg equilibrium or had different genotype rate between cases and controls. Individuals were excluded if they had missingness ≥ 2%, outlying heterozygosity rate or were duplicates. Population outliers and non-European samples were excluded. No overt population substructure between cases and controls was observed (data not shown). A total of 69,440 individuals passed quality control, including 144 cases. A dataset including the 144 cases and 1,800 randomly selected controls was imputed using Minimac3 (v2.0.1) and the Hapmap r22 CEU panel. Variants with minor allele count < 3 or with imputation quality $r^2 < 0.3$ were excluded, resulting in 2,363,678 well-imputed variants for 144 cases (38 women and 106 men) and 1,800 controls (952 women and 848 men). The study was approved by the local ethics committees. Written informed consent was obtained from all participants.

Replication Analysis

Association analysis was performed using a mixed logistic regression model implemented in SAIGE (v0.35.8.3), where CH was modeled as the dependent variable, and the genetic variants as the independent variable. Sex and the first eight principal components were included as covariates. From each independent significant locus ($p < 5 \times 10^{-8}$) in the discovery sample, the lead SNP, or a proxy SNP, was selected for replication. To correct for multiple testing, Bonferroni correction was applied for the number of loci tested (n = 4).

Sex stratified analysis

Analyses stratified for men and women were performed in SNPTEST to examine possible sex-specific genetic effects. Both models were adjusted for the first four principal components.

Previously reported cluster headache loci

The 9 different SNPs which have previously been reported for a significant association with CH were tested for association in our discovery analysis, see details in **Table 4**.[8-10, 13, 27-31] *P*-values were adjusted for multiple-testing using Bonferroni correction.

Univariate LD-score regression

Linkage Disequilibrium Score Regression (LDSC v1.0.1) was used to estimate the proportion of a true polygenic signal versus confounding factors such as population stratification, and to calculate SNP-based heritability.[32] Variants present in the HapMap 3 reference set were used, after excluding variants (1) with large-effect, explaining > 1% of phenotype variation, or variants in LD with such; (2) with MAF ≤ 0.01 or imputation INFO score ≤ 0.9; and (3) in the HLA region. Heritability estimates were converted to the liability scale assuming a population prevalence of CH of 0.1%.[6]

Colocalization analysis

To test whether the association signals for CH and migraine, on chromosome 6 near *UFL1/FHL5*, are consistent with a shared causal variant, we used a Bayesian colocalization procedure using the R package 'coloc' with default settings.[33] This test generates posterior probabilities for each locus weighting the evidence for five competing hypotheses regarding the sharing of causal variants, namely H0 (no causal variant for either trait); H1 or H2 (a causal variant only for trait one or two); H3 (distinct causal variants, for each trait); and H4 (a single causal variant common to both traits). The analysis assumes a single causal SNP for each trait. For CH we used the summary statistics from the discovery cohort and for migraine we used the summary statistics from Gormley *et al.*[34] without 23andMe (30,465 migraine cases and 143,147 controls); both populations are of European ancestry. Colocalization was tested for the region between the two nearest recombination hotspots.

Genetic correlation

LDSC was also used to calculate genetic correlation between CH and migraine.[32] For migraine we used summary statistics from Gormley *et al*.[34] without 23andMe (30,465 migraine cases and 143,147 controls), excluding variants with MAF ≤ 0.01,INFO score ≤ 0.6, large-effect variants or variants in an HLA region. In addition, the 38 genome-wide significant migraine loci were tested for association with CH.[34] Using the *cor.test* function in R, the correlation of the effect size (beta) between migraine and CH (current study) was calculated.

Gene-based analysis

We performed the MAGMA gene-based association analysis implemented in FUMA, using default settings to identify genes associated with CH.[22] This calculates a gene test-statistic (*p*-value) based on all SNPs located within genes. SNPs were assigned to the genes obtained from Ensembl build 85 (only protein-coding genes).

Tissue specificity analyses

To further test the relationship between tissue-specific expression and genetic associations to CH, we examined all SNPs and their respective effect on the expression of genes up to 1 Mb away (cis-eQTL), using FUMA quantitative trait locus (eQTL) mapping (https://fuma.gtlab.nl/tutorial#eQTLs); all SNPs were mapped based on each of the tissues in the Genotype Tissue Expression (GTEx) v8 dataset using default setting.[22] Additionally, we performed tissue expression analysis based on the MAGMA gene property in FUMA.[22] This analysis tests for positive relationships between tissue-specific gene expression in 30 general tissue types and 54 specific tissue types in the GTEx v8 RNA-seq data and gene-based *p*-values from the gene-based analysis described above.

RNA-sequencing of CH patients and controls

The genes identified by eQTL mapping with FUMA (see above) were further interrogated using existing RNA sequencing (RNA-seq) data generated from peripheral venous blood samples from 39 CH patients and 20 controls matched for age, sex and smoking habits. Data generation and quality control is described in detail elsewhere.[35] In short, RNA was extracted, using the PAXgene Blood miRNA kit, and sequenced using Illumina Hiseq4000. RNA-seq reads were aligned and processed using the in house transcriptome analysis pipeline Gentrap (version 0.3.1). Within this pipeline, sequencing reads were aligned to the human genome reference GRCh38 using TopHat (version 2.0.13) and counted per gene using Htseq (version 0.6.1p1). The data was normalized for between-sample variation and for within-sample variation, using Limma voom transformation. Differential expression analysis was performed in Limma, fitting a linear model correcting for age, gender, current smoking status and leukocyte counts. *P*-values were adjusted for multiple-testing using Bonferroni correction.

To determine the specificity of differential expression results obtained for CH, we examined the (nominally) significant genes from the CH RNA-seq analysis in RNA-seq data obtained from 26 migraine patients and 20 age- and sex-matched controls. Data generation and quality control have been previously described.[36] In short, peripheral venous blood samples were drawn when the migraine patients were migraine-free for at least five days and headache-free for 24 hours. RNA was extracted using PAXgene Blood RNA kit, sequenced (using Illumina Novaseq) and aligned. RNA-seq reads were, after quality control, aligned to the human reference transcriptome, using kallisto (version 0.42.5). Resulting count matrices were corrected for library size and gene length, and normalized using the R package DESeq2. Differential expression was performed using the R package DESeq2 by fitting a generalized linear model, correcting for age.

Results

Study participants

The clinical characteristics of cases and controls of the discovery sample are summarized in **Table 1**. There was a higher proportion of men (69% *vs* 44%) and smokers (52% *vs* 14%) among the cases compared to controls. Most patients had episodic CH (69%). A total of 13% of cases had migraine.

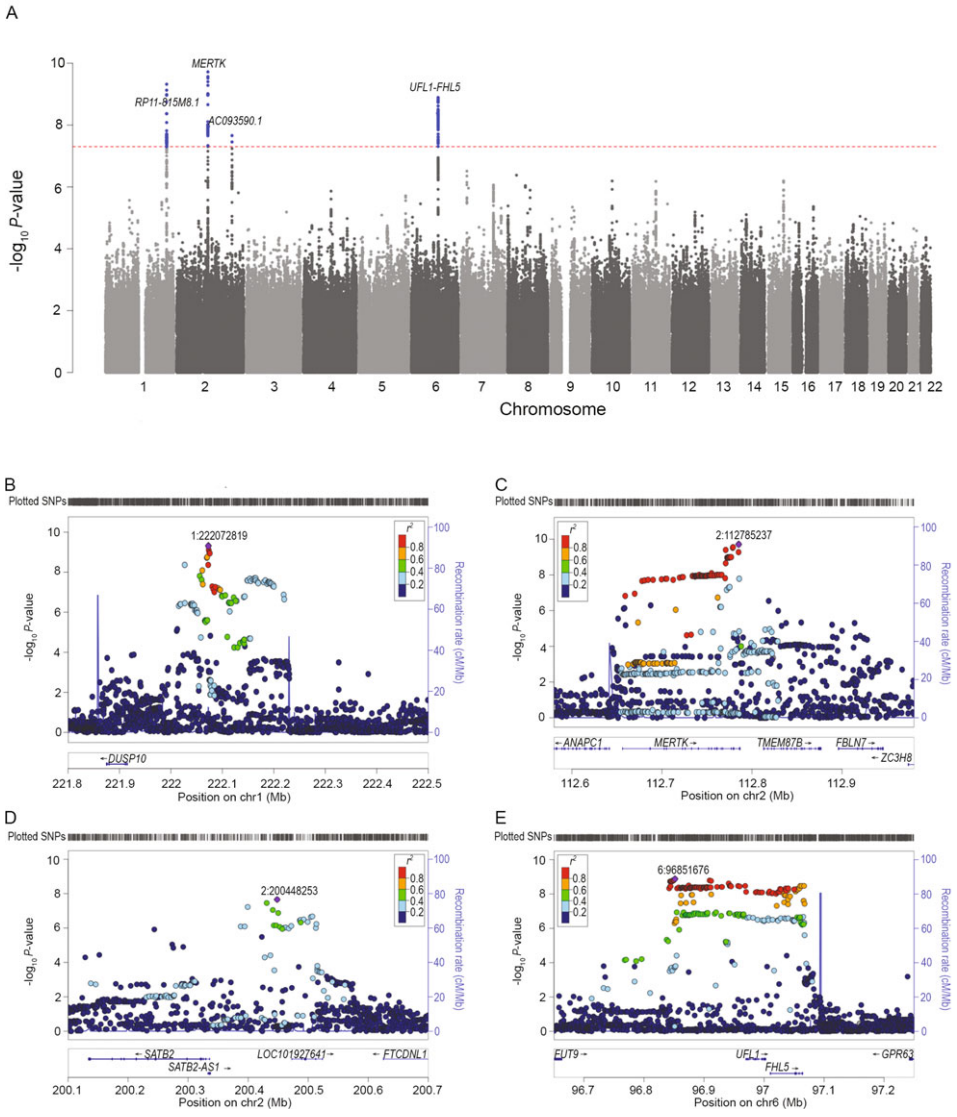
Fine mapping with PICS identified two variants with causal probability larger than 0.2, at rs11579212 (PICS probability = 0.40) and rs10184573 (PICS probability = 1.0), respectively.

Association analysis

Overall association results are shown in the Manhattan plot (**Figure 1A**) and the QQ plot (**Figure 2**). In total, four independent loci showed genome-wide significant ($p < 5 \times 10^{-8}$) associations with CH (**Figure 1B-E**) with a combined explained variance of 7.2%. More specifically, we identified rs11579212 (odds ratio (OR) = 1.51, 95% CI 1.33-1.72 near *RP11-815M8.1*), rs6541998 (OR = 1.53, 95% CI 1.37-1.74 near *MERTK*), rs10184573 (OR = 1.43, 95% CI 1.26-1.61 near *AC093590.1*), and rs2499799 (OR = 0.62, 95% CI 0.54-0.73 near *UFL1/FHL5*) (**Table 2**). These lead SNPs had either a call rate or imputation metric close to 100%. Three of the four lead SNPs were present in the replication sample (rs11579212, rs6541998 and rs10184573), while for the SNP on chromosome 6 (rs2499799) we selected a proxy SNP (rs976357, $r^2 = 1.0$, D' = 1). Lead SNPs of loci rs11579212, (OR = 1.58, 95% CI 1.16-2.15) rs10184573 (OR = 1.74, 95% CI 1.29-2.34) and rs976357 (OR = 0.44, 95% CI 0.30-0.64) replicated after Bonferroni correction (**Table 3**).

The genomic inflation factor (λ) was 1.069 in the discovery analysis, while the linkage disequilibrium score regression (LDSR) intercept was 1.044 (SE 0.0077), indicating moderate inflation due to factors other than polygenic architecture. We estimated the SNP-based heritability (h^2) of CH at 30.3% (SE 19.4%) on the observed scale. Assuming a population prevalence of 0.1% for CH this corresponds to a h^2 of 11.5% (SE 7.4%) on the liability scale.

Figure 1 Manhattan plot and reginal plots for the discovery analysis



A) Manhattan plot showing the -log10 for individual SNP. Each marker was tested for association using an additive genetic model by logistic regression. The horizontal axis shows the chromosomal position and the vertical axis shows the significance of tested markers from logistic regression. The threshold for genome wide significance ($p < 5 \times 10^{-8}$) is indicated by a red dotted line. Markers that reach genome-wide significance are shown in blue. **B-E**) Regional Manhattan plots of the four genome-wide significant cluster headache loci, with +/- 600 kb windows. Each dot represents a SNP, the horizontal axis gives the genomic coordinate and the vertical axis the significance level (-log10 p- value). The index SNP for each locus is marked with a purple diamond and annotated with its corresponding location number (CRCh37/hg19). SNPs are colored based on their correlation (r^2) with the labelled lead SNP according to the legend. The solid blue line shows the recombination rate from 1000 Genomes (EUR) data (right vertical axis). Gencode genes are shown. Figures were obtained from LocusZoom.[23] (**B**) Locus: rs11579212; 1:222072819. (**C**) Locus: rs6541998; 2:112785237. (**D**) Locus: rs10184573; 2:200448253. (**E**) Locus: rs2499799; 6:96851676. SNP = single nucleotide polymorphisms

Genetic correlation of cluster headache with migraine

The observed h^2 for migraine was 17.1% (SE 1.56%). The genetic correlation between CH and migraine was 0.33 (SE 0.021, $p = 0.12$). Next, we examined the 38 migraine-associated loci reported in Gormley *et al.*[34] Of the 37 migraine loci that were represented in our data set directly or by variants in high LD, one, located on chromosome 6, was associated with CH rs2971606, a proxy ($r^2 = 1.0$) for the migraine index variant rs67338227, in *FHL5*, $p = 1.39 \times 10^{-8}$ (Bonferroni corrected $pcorr = 0.5 \times 10^{-6}$). The association had the same effect direction for migraine and CH. There was also moderate LD between the lead SNP for migraine and CH ($r^2 = 0.64$ in data from 1000 Genomes Project Phase 3 CEU). Still, colocalization analysis revealed a 67.5% posterior probability for the hypothesis (H3) that the causal variants for CH and migraine are distinct, higher than the 32.5% posterior probability for hypothesis (H4) that CH and migraine share a causal variant in this region.

The other 36 migraine loci were not associated with CH (data not shown), with the second strongest association seen for rs10786156 in *PLCE1* ($p = 2.82 \times 10^{-3}$, $pcorr = 0.10$). The migraine locus near *MED14* on chromosome X (rs12845494) was not represented in our dataset. The effect sizes for the 37 loci combined correlated with those of CH, Pearson's $r(35) = 0.59$ ($p = 1.36 \times 10^{-4}$), even disregarding the overlapping *FHL5* locus (Pearson's $r(34) = 0.58$, $p = 2.18 \times 10^{-4}$).

Sex-stratified analyses

The low number of female cases gave limited power for the women-only analysis. Rs6541998was genome-wide significant in men using sex-stratified analyses; all other loci were nominally significant ($p < 1 \times 10^{-3}$) for both men and women with effects in the same direction. Using the method suggested by Clogg et al.,[37] we found no significant differences for the regression coefficients between men and women at the four lead SNPs (p-value 0.54 for rs11579212; p-value 0.62 for rs6541998; p-value 0.57 for rs10184573; p-value 0.59 for rs2499799).

Previously reported cluster headache loci

Of the 9 different SNPs previously associated with CH, one replicated, rs1800759 in *ADH4* ($p = 0.00039$, Bonferroni corrected $pcorr = 0.0035$) (**Table 4**). In contrast, none of the previously reported associations in *HCRTR2, ADCYAP1R1, CLOCK, CHRNA3-CHRNA5*, and *MME* were replicated in our sample (**Table 4**).

Table 1 Clinical characteristics of the discovery sample[a]

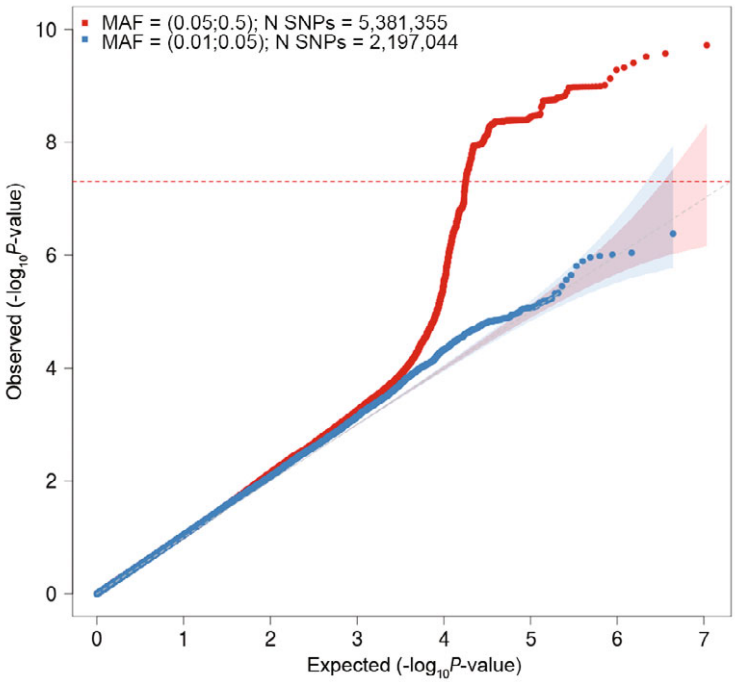
Characteristics	CH patients (n = 840)	Controls (n = 1,457)	*P*-value[b]
Men	579 (68.9)	636 (43.7)	< 0.001
Current daily smoking	440 (52.4)	202 (13.9)	< 0.001
Episodic cluster headache	577 (68.7)	-	-
Chronic cluster headache	233 (27.7)	-	-
Migraine co-morbidity	106 (12.6)	-	-

Data are expressed as numbers (percentages) unless otherwise stated. [a] Numbers and proportions may not add up to total of 100 due to rounding or missing values. [b] *P*-values of chi-square test for categorical variables. CH = cluster headache

Downstream bioinformatic analysis

Using FUMA gene-based eQTL mapping, 16 genes were mapped to the four loci (**Table 2**). Additionally, gene-mapping with MAGMA identified five genes whose expression was significantly influenced by the CH loci, *TMEM87B*, *MERTK*, *FHL5*, *UFL1* and *ASZ1* (**Figure 3**). Finally, we performed a MAGMA tissue expression analysis which did not render any significant results (data not shown).

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

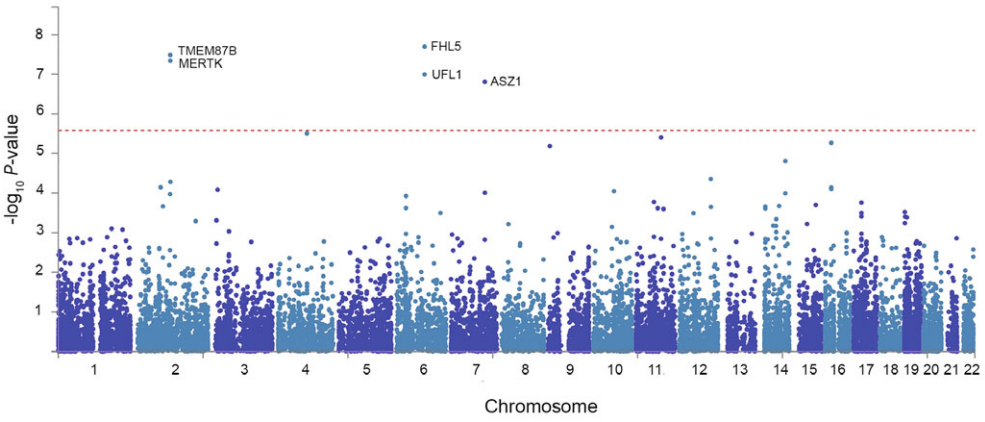


The horizontal axis shows -log10 *p*-values expected under the null distribution; The vertical axis shows observed -log10 p-values. Red = common SNPs (MAF ≥ 0.05), blue = low frequency SNPs (MAF = 0.005-0.05). Genomic inflation factor (λ) = 1.069. MAF = minor allele frequency

RNA-seq analyses

Using RNA-seq data from white blood cells of 39 CH patients and 20 controls we assessed the 16 eQTL-mapped genes derived from FUMA. Eleven genes were expressed in the samples, of which one was differentially lower expressed (*POLR1B*, $p = 7.50 \times 10^{-5}$, $pcorr = 8.3 \times 10^{-4}$) and one was nominally differentially lower expressed (*TMEM87B*, $p = 0.014$, $pcorr = 0.15$) in CH cases than in controls, both genes representing the rs6541998 locus (**Table 5**). The two genes were not differentially expressed in RNA-seq data when comparing 26 migraine with 20 controls ($p = 0.50$ and 0.45, respectively) (**Table 5**).

Figure 3 Gene-based Manhattan plot



Input SNPs were mapped to 18,795 protein coding genes. The horizontal axis shows the chromosomal position and the vertical axis shows the significance of tested markers. The threshold for genome wide significance (p = 0.05/18,795 = 2.66 x 10^{-6}) is indicated by a dotted line.

Table 2 SNPs at the four loci associated with cluster headache discovery sample

SNP	Chr:Pos EA:NEA[a]	EAF	OR [95% CI]	*P*-value[b]	Nearest Gene[c]	eQTL mapped genes[d]
rs11579212	1:222072819 C:A	0.34	1.51 [1.33-1.72]	4.78 x 10^{-10}	*RP11-815M8.1*	*DUSP10*
rs6541998	2:112785237 C:T	0.63	1.53 [1.37-1.74]	1.91 x 10^{-10}	*MERTK*	*TTL* *POLR1B* *FBLN7* *ZC3H8* *MERTK* *TMEM87* *RGPD8* *ZC3H6*
rs10184573	2:200448253 T:G	0.44	1.43 [1.26-1.61]	2.20 x 10^{-8}	*AC093590.1*	*SATB2* *FTCDNL1*
rs2499799	6:96851676 C:T	0.81	0.62 [0.54-0.73]	1.29 x 10^{-9}	*UFL1/FHL5*	*UFL1* *FHL5* *GPR63* *MMS22L* *FUT9*

[a] Chromosomal positions in GRCh37/hg19 coordinates. [b] Significant result (p < 5 x 10^{-8}). [c] The nearest gene is based on ANNOVAR annotations with Ensembl build 85. [d] eQTL mapping was done in FUMA based on GTEx v8. SNP = single nucleotide polymorphism; Chr = chromosome; Pos = position; EA = effect allele; NEA = non effect allele; EAF = effect allele frequency; OR = odds ratio; CI = confidence interval

Table 3 Replication of the significant loci in an independent sample

L	Chr	SNP discovery sample	SNP replication sample	Pos[a]	OR [95% CI]	*P*-value[b]	Direction[c]
1	1	rs11579212	rs11579212	222072819	1.58 [1.16-2.15]	3.50×10^{-3}	+
2	2	rs6541998	rs6541998	112785237	1.04 [0.78-1.40]	0.78	+
3	2	rs10184573	rs10184573	200448253	1.74 [1.29-2.34]	2.78×10^{-4}	+
4	6	rs2499799	rs976357 (r^2=1.0)	96849679	0.44 [0.30-0.64]	2.76×10^{-5}	+

[a] Chromosomal positions in GRCh37/hg19 coordinates for the replication SNP. [b] Significant result ($p < 0.05/4$). [c] Direction; Same (+) or opposite (-) direction of association for discovery and replication analyses. L = locus number; Chr = chromosome; SNP = single nucleotide polymorphism; Pos = chromosomal position; OR = odds ratio; CI = confidence interval.

Table 4 The association of previously reported cluster headache loci in our discovery cluster headache sample

Previously reported CH loci						Association with CH discovery sample	
Nearest coding gene	Index SNP	EA	OR [95%CI]	*P*-value	Ref	OR [95%CI]	*Pcorr*-value[a]
ADCYAP1R1	rs12668955	G	0.48 [0.34-0.07]	9.1×10^{-6}	[13]	1.05 [0.91-1.21]	1
ADH4	rs1126671	A	2.33 [1.25-4.37]	0.006[b]	[30]	0.87 [0.76-0.99]	0.36
ADH4	rs1126671	A	-	0.03	[31]	0.87 [0.76-0.99]	0.36
ADH4	rs1800759	A	-	0.03	[31]	0.80 [0.70-0.90]	0.0035
CLOCK	rs12649507	A	1.29 [1.08-1.54]	0.02	[29]	0.92 [0.80-1.05]	1
CHRNA3-CHRNA5	rs578776	A	-	0.038	[27]	0.95 [0.83-1.09]	1
HCRTR2	rs2653349	G	6.79 [2.25-22.99]	< 0.0002	[8]	1.08 [0.93-1.26]	1
HCRTR2	rs2653349	G	1.97 [1.32-2.92]	0.0007	[10]	1.08 [0.93-1.26]	1
HCRTR2	rs3122156	G	0.82 [0.68-0.99]	0.0421 (*Pcorr* 0.126)	[9]	0.92 [0.80-1.06]	1
HCRTR2	rs10498801	G	0.69 [0.49-0.97]	0.030	[28]	1.03 [0.88-1.21]	1
MME	rs147564881	C	-	0.019	[13]	0.24 [0.42-41.51]	1

The SNPs previously reported to be associated with cluster headache and the corresponding OR (based on the same EA) and *p*-values for these SNPs in the discovery sample. [a] *P*-values were Bonferroni corrected for 9 tests. [b] *P*-value is based on the carriers with homozygous AA genotype compared with GG/GA genotypes. CH = cluster headache; SNP = single nucleotide polymorphism; EA = effect allele; OR = odds ratio; CI = confidence interval

Table 5 RNA-seq expression data

Gene	Expression in blood	Locus number	CH *P*-value	CH *Pcorr* value[a]	M *P*-value
POLR1B	Yes	2	7.50×10^{-5}	8.3×10^{-4}	0.50
TMEM87B	Yes	2	0.014	0.15	0.45
ZC3H8	Yes	2	0.084	0.92	-
DUSP10	Yes	1	0.725	1	-
MERTK	Yes	2	0.527	1	-
TTL	Yes	2	0.465	1	-
MMS22L	Yes	4	0.424	1	-
FBLN7	Yes	2	0.361	1	-
ZC3H6	Yes	2	0.285	1	-
FTCDNL1	Yes	3	0.123	1	-
UFL1	Yes	4	0.123	1	-
FHL5	No	4	-	-	-
FUT9	No	4	-	-	-
GPR63	No	4	-	-	-
RGPD8	No	2	-	-	-
SATB2	No	3	-	-	-

Genes were selected based in the eQTL mapping in FUMA.[22] [a] *P*-values were Bonferroni corrected for 11 tests. CH = cluster headache; M = migraine.

Discussion

We performed a GWAS in CH and identified four independent genetic risk loci, of which three replicated in an independent sample. The association effect sizes, with ORs around 1.5, are high compared to those usually observed in GWAS (https://www.ebi.ac.uk/gwas/).[38] Whereas this may indicate that the risk for CH is driven by a limited number of loci with strong associations with CH, it is likely to be expected that follow-up studies with larger sample sizes also will identify loci with smaller effect sizes. Except for the *MERTK* locus (rs6541998), all loci replicated in our replication sample, suggesting that the signals are genuine. Gene-based mapping additionally found that expression of the *ASZ1* gene may be influenced by one or more CH loci, providing a possible additional locus. RNA-seq results show altered expression in CH patients of *POLR1B* and *TMEM87B*, suggesting their involvement in CH. Although there seems to be a considerable SNP-based heritability for CH, a robust estimation of SNP-based heritability is not possible given the small sample size, hence heritability estimates should be interpreted with caution.

The main limitations of our study are that, (1) although we identified and replicated genome-wide significant loci, the relatively small number of cases in the discovery sample will leave loci with smaller effect sizes or lower allele frequencies hidden; (2) it is unclear to what extent the present results can be extrapolated to ancestries other than European ancestry; (3) although cases

and controls were genotyped using the same platform, genotyping was performed in different laboratories possibly introducing batch effects. Therefore, we made significant efforts to circumvent possible problems arising from our design by rigorous quality control. Overall, our case sample was representative of the general CH population with a ~2:1 male-female ratio, chronic CH of ~30% and without any familial confounder, as familial cases were removed in the QC steps.[1] The difference in the percentage of men and women for cases and controls was corrected for in the statistical analysis.

Among previously suggested loci to be involved in CH, we found evidence for significant association to the alcohol dehydrogenase 4 gene (*ADH4*) although the effect identified is opposite to what was previously reported and at the genome wide level it was not significant.[31] In previous studies, *ADH4* was investigated mainly because alcohol is considered both a trigger and possible risk factor for transformation from episodic to chronic CH.[30, 31] Of note, we did not find evidence for an association of *HCRTR2,* as reported previously,[8-10] nor for any of the other previously reported loci in CH.

A remarkable finding in our study was that one of the leading loci, represented by rs2499799, which covers both *FHL5* and *UFL*, has previously been identified as a migraine risk locus.[34] *FHL5* encodes a transcription factor that regulates cAMP-responsive elements CREB6 and CREM, which play a role in synaptic plasticity and memory formation.[39] *UFL1* codes for the ubiquitin-fold modifier 1 (UFM1)-specific ligase 1, an ubiquitin-like protein that allows UFL1 to conjugate to its substrates.[40] The ubiquitin protease system (UPS) has been associated as a pathway in neuropsychiatric and neurodegenerative disorders.[41] In the latest migraine GWAS meta-analysis, the *UFL1/FHL5* locus had an odds ratio (OR) of 1.09 [1.08-1.11] based on the primary signal (rs67338227).[34] The direction of the effect in the *UFL1/FHL5* locus in our dataset was the same in both migraine and CH and, the lead SNPs for migraine and CH were in LD (r^2 = 0.64). Our colocalization analysis suggests that CH and migraine are more likely caused by distinct variants at this locus. Admittedly this finding could also be a result of different LD patterns in the samples that were compared in the colocalization analysis. The other 36 independent loci implicated in migraine showed no association with CH. Our results suggest though that the *UFL1/FHL5* locus is 'specific' for CH and that the association is not due to the mere presence of comorbid migraine among patients with CH. This is further supported by the similar prevalence of migraine among cases in our discovery sample (13%) and the expected population prevalence (10-17%), although the number of migraine cases in controls was not collected.[42] While no other migraine locus reached significance in our study individually there was a moderate correlation between association effect sizes of CH and migraine for the 37 examined migraine loci. This may reflect a shared genetic architecture underlying both disorders, which is not surprising given that they share pathophysiological features such as the involvement of the trigeminovascular system and efficacy of calcitonin gene-related peptide (CGRP) monoclonal antibodies and triptans.[4] It is possible that future studies with larger sample sizes may identify the involvement of more migraine loci in CH.

With respect to the other replicated loci, rs11579212 and rs10184573, which mapped to *RP11-815M8.1* and *AC093590.1* respectively, they have not previously been related to disease, and their role in CH pathogenesis remains unclear. Although rs6541998 did not replicate in the small replication sample, two genes (*POLR1B*, *TMEM87B*) in the locus showed differential expression in CH compared to controls in RNA-seq analyses, whereas no such effect was seen in migraine data. *POLR1B,* encoding DNA-directed RNA polymerase I subunit RPA2, has been associated with Treacher Collins and *TMEM87B,* encoding transmembrane protein 87B may be involved in restrictive cardiomyopathy.[43, 44] *MERTK,* the nearest gene, encodes a receptor tyrosine kinase of the TAM (Tyro3, Axl, MERTK) family, is among other tissues expressed in oligodendrocytes, astrocytes and microglia in the brain and has an effect on the immune response.[45] Unfortunately, the number of associated genes with CH is not large enough to perform meaningful further downstream pathway analyses. Based on the regression coefficients, we found no evidence for a different effect for the lead SNPs between men and women.

In conclusion, this GWAS of CH reveals four genetic risk loci for CH with unusually high effect sizes for a complex disorder, of which three replicated in an independent sample. One of the loci has previously been identified as a migraine risk locus. Our results suggest several genes to be involved in the pathogenesis of CH and offer a starting point for future research to elucidate the molecular mechanisms of this severe disease.

Post-script paragraph

Two parallel manuscripts (Harder *et al.* and O'Connor *et al.*), submitted to the journal, report the first replicated genomic loci associated with CH. Whereas Harder *et al*. investigated Dutch CH cases (n = 840) and controls (n = 1,457) and Norwegian CH cases (n = 144) and controls (n= 1,800), O'Connor *et al*. investigated UK cases (n = 852) and controls (n = 5,614) as well as Swedish cases (n = 591) and controls n = 1,134). The four loci reported by Harder *et al*. correspond to four loci reported by O'Connor *et al*., with the index variants reported in the two studies being in linkage disequilibrium with each other (D' = 0.86 and r^2 = 0.36 for rs11579212 and rs12121134; D' = 0.98 and r^2 = 0.95 for rs6541998 and rs4519530; D' = 0.95 and r^2 = 0.34 for rs10184573 and rs113658130; and D' = 0.93 and r^2 = 0.38 for rs2499799 and rs11153082, in the 1000 Genomes data for European populations). The independent discovery of the four loci in the two studies provides additional support that they represent genuine risk loci for cluster headache.

Next, we combined the summary statistics from the four studies (Dutch, Norwegian, UK, Swedish) using inverse-variance weighted meta-analysis as implemented in METAL (with the 'STDERR' option), after harmonizing the datasets using EasyQC.[46, 47] In total, 8,039,373 variants were analyzed. The association to CH remained significant for all eight index variants (in the four loci) reported in the two papers: rs11579212 (effect allele, EA: C), OR 1.31 (95% CI 1.21-1.41),

p-value 8.98 x 10^{-13}; rs12121134 (EA: T), OR 1.40 (95% CI 1.29-1.53), *p*-value 9.18 x 10^{-15}; rs6541998 (EA: C), OR 1.40 (95% CI 1.30-1.51), *p*-value 2.37 x 10^{-19}; rs4519530 (EA: C), OR 1.41 (95% CI 1.31-1.52), *p*-value 4.18 x 10^{-29}; rs10184573 (EA: T), OR 1.38 (95% CI 1.28-1.50), *p*-value 3.35 x 10^{-16}; rs113658130 (EA: C), OR 1.54 (95% CI 1.41-1.69), *p*-value 1.28 x 10^{-21}; rs2499799 (EA: C), OR 0.77 (95% CI 0.70-0.84), *p*-value 2.73 x 10^{-8}; rs11153082 (EA: G), OR 1.33 (95% CI 1.23-1.43), *p*-value 2.98 x 10^{-14}. The eight index variants in the overlapping loci showed a consistent effect direction across the two studies. Colocalization analysis, to determine whether the reported loci of both manuscripts represent the same causal variants, identified a high posterior probability for three loci (those on chromosomes 1 and 2) to likely represent the same causal variant.[33] Rs12121134 and rs11579212 have a posterior probability that the causal variants are the same (H4) of 80.4%, for rs4519530 and rs6541998 H4 is 87.4% and for rs113658130 and rs10184573 H4 is 96.9%. For the locus on chromosome 6, the colocalization analysis shows a higher probability that the loci in the two studies represent distinct causal variants (H3: 78.7%) rather than the same causal variant (H4: 21.2%).

Finally, the meta-analysis resulted in three additional loci becoming genome-wide significant: (1) a locus on chromosome 7 with 31 significant (*p*-value < 5 x 10^{-8}) variants with index variant rs6966836 (chr7:117002998, EA: C), OR 1.25 (95% CI 1.16-1.35), *p*-value 2.06 x 10^{-9}; (2) a locus on chromosome 10 with two significant variants with index variant rs10786156 (chr10:96014622, EA: C), OR 1.24 (95% CI 1.15-1.33), *p*-value 7.61 x 10^{-9}; and (3) a locus on chromosome 19 with two significant variants with index variant rs60690598 (chr19:55052198, EA: T), OR 1.87 (95% CI 1.51-2.33), *p*-value 1.70 x 10^{-8}.

Acknowlegdements: We thank the patients for their participation in this project. The Nord-Trøndelag Health Study (The HUNT Study) is a collaboration between HUNT Research Centre (Faculty of Medicine and Health Sciences, NTNU, Norwegian University of Science and Technology), Trøndelag County Council, Central Norway Regional Health Authority, and the Norwegian Institute of Public Health. The genotyping was financed by the National Institute of health (NIH), University of Michigan, The Norwegian Research council, and Central Norway Regional Health Authority and the Faculty of Medicine and Health Sciences, Norwegian University of Science and Technology (NTNU). The genotype quality control and imputation has been conducted by the K.G. Jebsen center for genetic epidemiology, Department of public health and nursing, Faculty of medicine and health sciences, Norwegian University of Science and Technology (NTNU). We express our gratitude to all individuals who participate in the Netherlands Epidemiology in Obesity study. We are grateful to all participating general practitioners for inviting eligible participants. We also thank P. van Beelen and all research nurses for collecting the data and P. Noordijk and her team for sample handling and storage and I. de Jonge, MSc for data management of the NEO study.

This work was supported by the South-Eastern Norway Regional Health Authority (grant no. 2015089 to J.A.Z.), and grants of the Netherlands Organization for Scientific Research, i.e. the Center of Medical System Biology established by the Netherlands Genomics Initiative/ Netherlands Organisation for Scientific Research (to A.M.J.M.v.d.M.), Spinoza 2009 (to M.D.F.) and EU-funded FP7 "EUROHEADPAIN" (grant no. 6026337 to, M.D.F., G.M.T, and A.M.J.M.v.d.M.).

The genotyping of Dutch and Norwegian case samples was provided by the Genomics Core Facility (GCF), Norwegian University of Science and Technology (NTNU). GCF is funded by the Faculty of Medicine and Health Sciences at NTNU and Central Norway Regional Health Authority.

The NEO study is supported by the participating Departments, the Division and the Board of Directors of the Leiden University Medical Centre, and by the Leiden University, Research Profile Area 'Vascular and Regenerative Medicine'.

The funding organizations had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

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

Members of "Cluster Headache Genetics Working Group": All members of the "Cluster Headache Genetics Working Group" were actively involved in patient collection. The members of the Cluster Headache Genetics Working group include Roemer B. Brandt, MD, Ilse F. de Coo, MD, Patty G.G. Doesborg, MD all at the Leiden University Medical Center (LUMC), Leiden the Netherlands; Roser Corominas, PhD from the University of Barcelona, Barcelona, Spain; Institut de Biomedicina de la Universitat de Barcelona (IBUB), Barcelona, Spain; Instituto de Salud Carlos III, Spain; Institut de Recerca Sant Joan de Déu (IR-SJD), Esplugues de Llobregat, Spain; Victor J. Gallardo, MSc at the Headache Research Group, Vall d'Hebron Institute of Research (VHIR), Universitat Autònoma de Barcelona, Barcelona, Spain; Nunu Lund, MD, PhD from the Danish Headache Center, Rigshospitalet Glostrup, Glostrup, Denmark; Paavo Häppölä, MSc at University of Helsinki, Helsinki, Finland.

Potential conflicts of interest: The authors declared no conflict of interest.

Reference

1. Lund N, Barloese M, Petersen A, Haddock B, Jensen R. Chronobiology differs between men and women with cluster headache, clinical phenotype does not. *Neurology*. 2017;88(11):1069-1076.
2. Lund N, Petersen A, Snoer A, Jensen RH, Barloese M. Cluster headache is associated with unhealthy lifestyle and lifestyle-related comorbid diseases: Results from the Danish Cluster Headache Survey. *Cephalalgia*. 2019;39(2):254-263.
3. Brandt RB, Doesborg PGG, Haan J, Ferrari MD, Fronczek R. Pharmacotherapy for Cluster Headache. *CNS Drugs*. 2020;34(2):171-184.
4. Vollesen AL, Benemei S, Cortese F, et al. Migraine and cluster headache - the common link. *J Headache Pain*. 2018;19(1):89.
5. Headache Classification Committee of the International Headache Society (IHS) The International Classification of Headache Disorders, 3rd edition. *Cephalalgia* 2018;38(1):1-211.
6. May A, Schwedt TJ, Magis D, et al. Cluster headache. *Nat Rev Dis Primers*. 2018;4:18006.
7. Sjaastad O, Shen JM, Stovner LJ, Elsas T. Cluster headache in identical twins. *Headache*. 1993;33(4):214-217.
8. Rainero I, Gallone S, Valfrè W, et al. A polymorphism of the hypocretin receptor 2 gene is associated with cluster headache. *Neurology*. 2004;63(7):1286-1288.
9. Fourier C, Ran C, Steinberg A, et al. Analysis of HCRTR2 Gene Variants and Cluster Headache in Sweden. *Headache*. 2019;59(3):410-417.
10. Schürks M, Kurth T, Geissler I, et al. Cluster headache is associated with the G1246A polymorphism in the hypocretin receptor 2 gene. *Neurology*. 2006;66(12):1917-1919.
11. Weller CM, Wilbrink LA, Houwing-Duistermaat JJ, et al. Cluster headache and the hypocretin receptor 2 reconsidered: a genetic association study and meta-analysis. *Cephalalgia*. 2015;35(9):741-747.
12. Gibson KF, Santos AD, Lund N, Jensen R, Stylianou IM. Genetics of cluster headache. *Cephalalgia*. 2019;39(10):1298-1312.
13. Bacchelli E, Cainazzo MM, Cameli C, et al. A genome-wide analysis in cluster headache points to neprilysin and PACAP receptor gene variants. *J Headache Pain*. 2016;17(1):114.
14. Ran C, Fourier C, Michalska JM, et al. Screening of genetic variants in ADCYAP1R1, MME and 14q21 in a Swedish cluster headache cohort. *J Headache Pain*. 2017;18(1):88.
15. Wilbrink LA, Weller CM, Cheung C, et al. Stepwise web-based questionnaires for diagnosing cluster headache: LUCA and QATCH. *Cephalalgia*. 2013;33(11):924-931.
16. de Mutsert R, den Heijer M, Rabelink TJ, et al. The Netherlands Epidemiology of Obesity (NEO) study: study design and data collection. *Eur J Epidemiol*. 2013;28(6):513-523.
17. Guo Y, He J, Zhao S, et al. Illumina human exome genotyping array clustering and quality control. *Nat Protoc*. 2014;9(11):2643-2662.
18. Grove ML, Yu B, Cochran BJ, et al. Best practices and joint calling of the HumanExome BeadChip: the CHARGE Consortium. *PLoS One*. 2013;8(7):e68095.
19. Anderson CA, Pettersson FH, Clarke GM, et al. Data quality control in genetic case-control association studies. *Nat Protoc*. 2010;5(9):1564-1573.
20. Loh PR, Danecek P, Palamara PF, et al. Reference-based phasing using the Haplotype Reference Consortium panel. *Nat Genet*. 2016;48(11):1443-1448.
21. Farh KK, Marson A, Zhu J, et al. Genetic and epigenetic fine mapping of causal autoimmune disease variants. *Nature*. 2015;518(7539):337-343.
22. Watanabe K, Taskesen E, van Bochoven A, Posthuma D. Functional mapping and annotation of genetic associations with FUMA. *Nat Commun*. 2017;8(1):1826.
23. Pruim RJ, Welch RP, Sanna S, et al. LocusZoom: regional visualization of genome-wide association scan results. *Bioinformatics*. 2010;26(18):2336-2337.

24. Pelzer N, Haan J, Stam AH, et al. Clinical spectrum of hemiplegic migraine and chances of finding a pathogenic mutation. *Neurology*. 2018;90(7):e575-e582.

25. Krokstad S, Langhammer A, Hveem K, et al. Cohort Profile: the HUNT Study, Norway. *Int J Epidemiol*. 2013;42(4):968-977.

26. Ferreira MA, Vonk JM, Baurecht H, et al. Shared genetic origin of asthma, hay fever and eczema elucidates allergic disease biology. *Nat Genet*. 2017;49(12):1752-1757.

27. Cainazzo MM, Tiraferri I, Ciccarese M, et al. O015. Evaluation of the genetic polymorphism of the α3 (CHRNA3) and α5 (CHRNA5) nicotinic receptor subunits, in patients with cluster headache. *J Headache Pain*. 2015;16(Suppl 1):A88.

28. Fan Z, Hou L, Wan D, et al. Genetic association of HCRTR2, ADH4 and CLOCK genes with cluster headache: a Chinese population-based case-control study. *J Headache Pain*. 2018;19(1):1.

29. Fourier C, Ran C, Zinnegger M, et al. A genetic CLOCK variant associated with cluster headache causing increased mRNA levels. *Cephalalgia*. 2018;38(3):496-502.

30. Rainero I, Rubino E, Gallone S, et al. Cluster headache is associated with the alcohol dehydrogenase 4 (ADH4) gene. *Headache*. 2010;50(1):92-98.

31. Zarrilli F, Tomaiuolo R, Ceglia C, et al. Molecular analysis of cluster headache. *Clin J Pain*. 2015;31(1):52-57.

32. Bulik-Sullivan BK, Loh PR, Finucane HK, et al. LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat Genet*. 2015;47(3):291-295.

33. Giambartolomei C, Vukcevic D, Schadt EE, et al. Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. *PLoS Genet*. 2014;10(5):e1004383.

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

35. Eising E, Pelzer N, Vijfhuizen LS, et al. Identifying a gene expression signature of cluster headache in blood. *Sci Rep*. 2017;7:40218.

36. Kogelman LJ, Falkenberg K, Halldorsson GH, et al. Comparing migraine with and without aura to healthy controls using RNA sequencing. *Cephalalgia*. 2019;39(11):1435-1444.

37. Clogg CC, Petkova E, Haritou A. Statistical Methods for Comparing Regression Coefficients Between Models. 1995;100(5):1261-1293.

38. Visscher PM, Wray NR, Zhang Q, et al. 10 Years of GWAS Discovery: Biology, Function, and Translation. *Am J Hum Genet*. 2017;101(1):5-22.

39. Anttila V, Winsvold BS, Gormley P, et al. Genome-wide meta-analysis identifies new susceptibility loci for migraine. *Nat Genet*. 2013;45(8):912-917.

40. Komatsu M, Chiba T, Tatsumi K, et al. A novel protein-conjugating system for Ufm1, a ubiquitin-fold modifier. *EMBO J*. 2004;23(9):1977-1986.

41. Cheon S, Dean M, Chahrour M. The ubiquitin proteasome pathway in neuropsychiatric disorders. *Neurobiol Learn Mem*. 2019;165:106791.

42. Lipton RB, Stewart WF. Prevalence and impact of migraine. *Neurol Clin*. 1997;15(1):1-13.

43. Sanchez E, Laplace-Builhe B, Mau-Them FT, et al. POLR1B and neural crest cell anomalies in Treacher Collins syndrome type 4. *Genet Med*. 2020;22(3):547-556.

44. Yu HC, Coughlin CR, Geiger EA, et al. Discovery of a potentially deleterious variant in TMEM87B in a patient with a hemizygous 2q13 microdeletion suggests a recessive condition characterized by congenital heart disease and restrictive cardiomyopathy. *Cold Spring Harb Mol Case Stud*. 2016;2(3):a000844.

45. Tondo G, Perani D, Comi C. TAM Receptor Pathways at the Crossroads of Neuroinflammation and Neurodegeneration. *Dis Markers*. 2019;2019:2387614.

46. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics*. 2010;26(17):2190-2191.

47. Winkler TW, Day FR, Croteau-Chonka DC, et al. Quality control and conduct of genome-wide association meta-analyses. *Nat Protoc*. 2014;9(5):1192-1212.

7

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

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

* shared authorship

Ann Neurol 2023; 10.1002/ana.26743

Abstract

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

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

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

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

Introduction

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

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

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

Methods

Cohorts and phenotyping

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

GWAS and meta-analysis

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

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

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

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

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

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

Candidate gene mapping

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

eQTL analysis

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

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

Transcriptome-wide association study analysis (TWAS-FUSION)

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

Fine-mapping of causal gene sets (FOCUS)

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

Genetically driven DNA methylation scan (MetaMeth)

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

Protein-altering variants (VEP-Ensembl)

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

Gene set and tissue enrichment analyses

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

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

Drug target identification

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

Genetic risk score analysis

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

Genetic correlation

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

Colocalization analysis

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

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

Mendelian randomization analysis

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

Results

European-ancestry GWAS meta-analysis

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

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

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

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

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

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

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

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

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

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

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

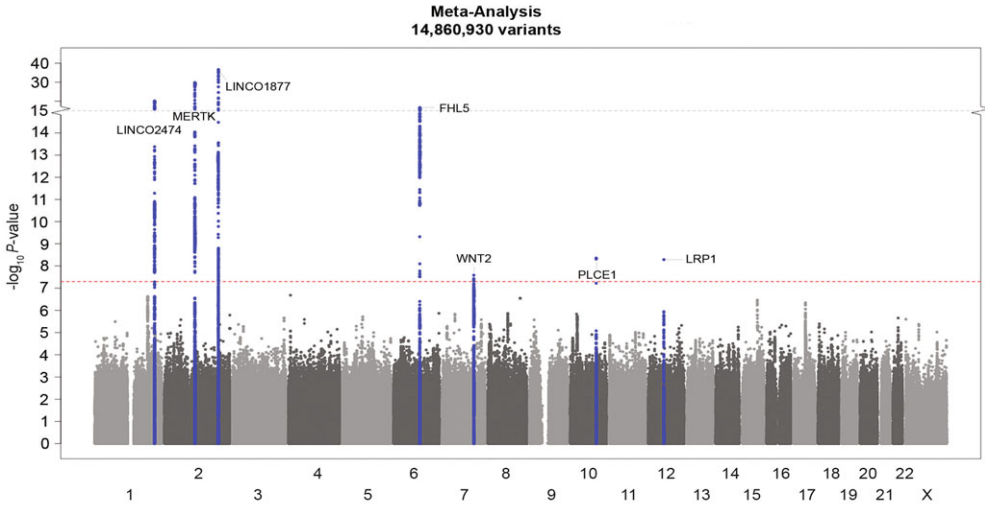
Locus name	Lead variant (Chr:Pos)	EA/ NEA (EAF)	OR (95% CI)	*p* value (Het *p*)	Variant type [Prioritized genes]
DUSP10	rs17011182 (1:222164327)	A/G (0.793)	1.38 (1.29-1.48)	7.76×10^{-21} (0.58)	regulatory region [*DUSP10*]
MERTK	rs13399108 (2:112747123)	A/G (0.373)	1.41 (1.33-1.50)	1.74×10^{-30} (0.16)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

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

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

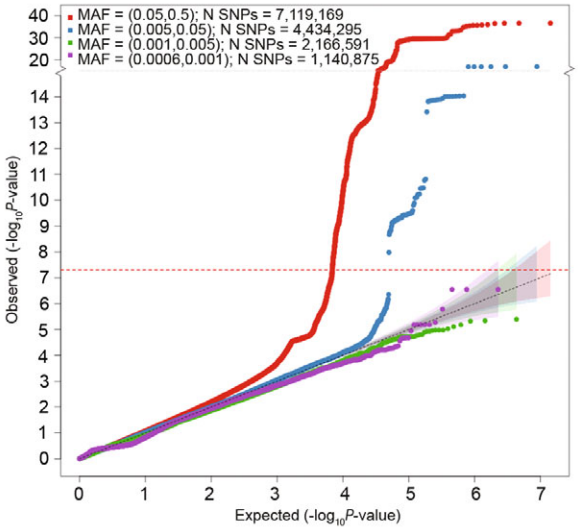
When considering the 20 prioritized genes, FUMA[24] found a significant enrichment for genes differentially expressed in artery (tibial artery) and brain (substantia nigra) (**Figure 5** and **Table S20**), and a significant overlap with genes reported in the GWAS catalog for 10 traits, most significantly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

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



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

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

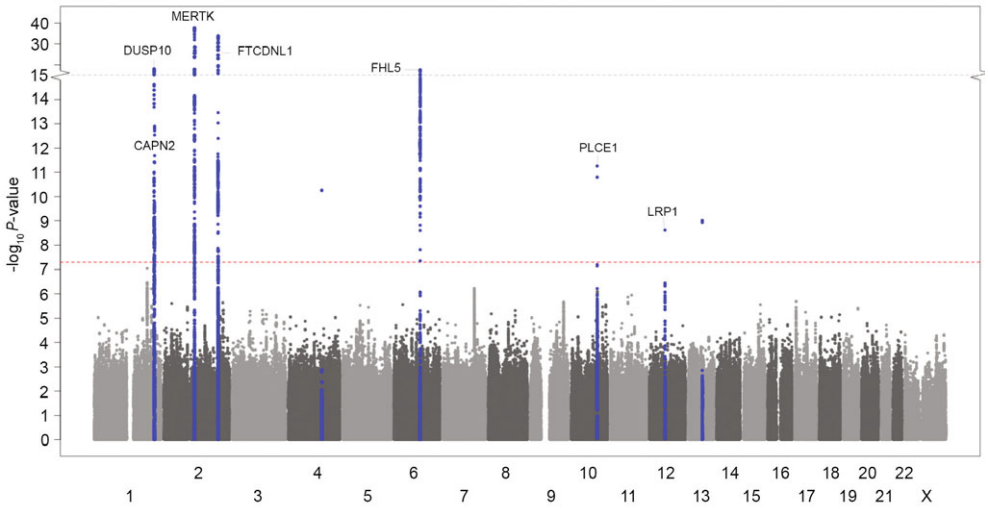


The horizontal axis shows -log10 p values expected under the null distribution. The vertical axis shows observed -log10 p values. Genomic inflation factor (λ) = 1.086. Red = common variants (MAF ≥ 5%), blue = low frequency variants (MAF = 0.5 - 5%), green = rare variants (MAF = 0.1 - 0.5%), purple =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

Genetic risk score analysis

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

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



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

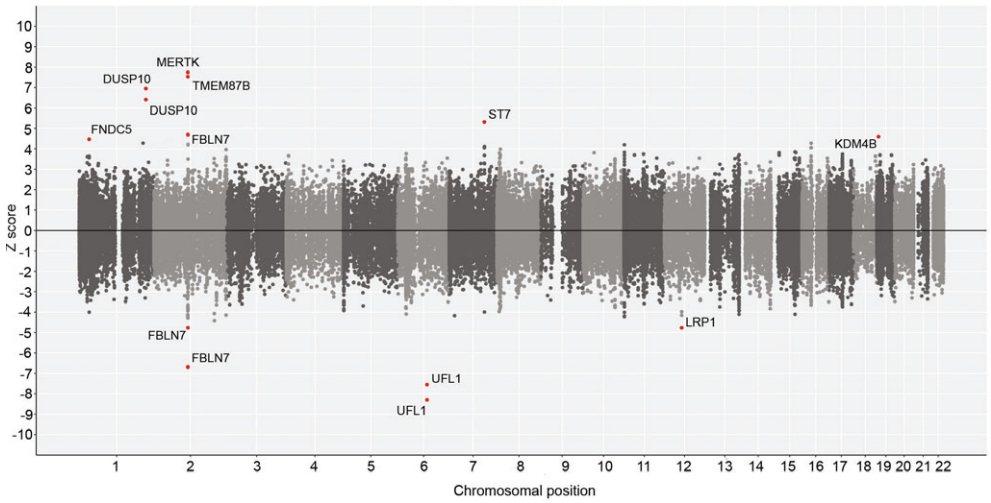
Genetic correlation

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

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

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

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

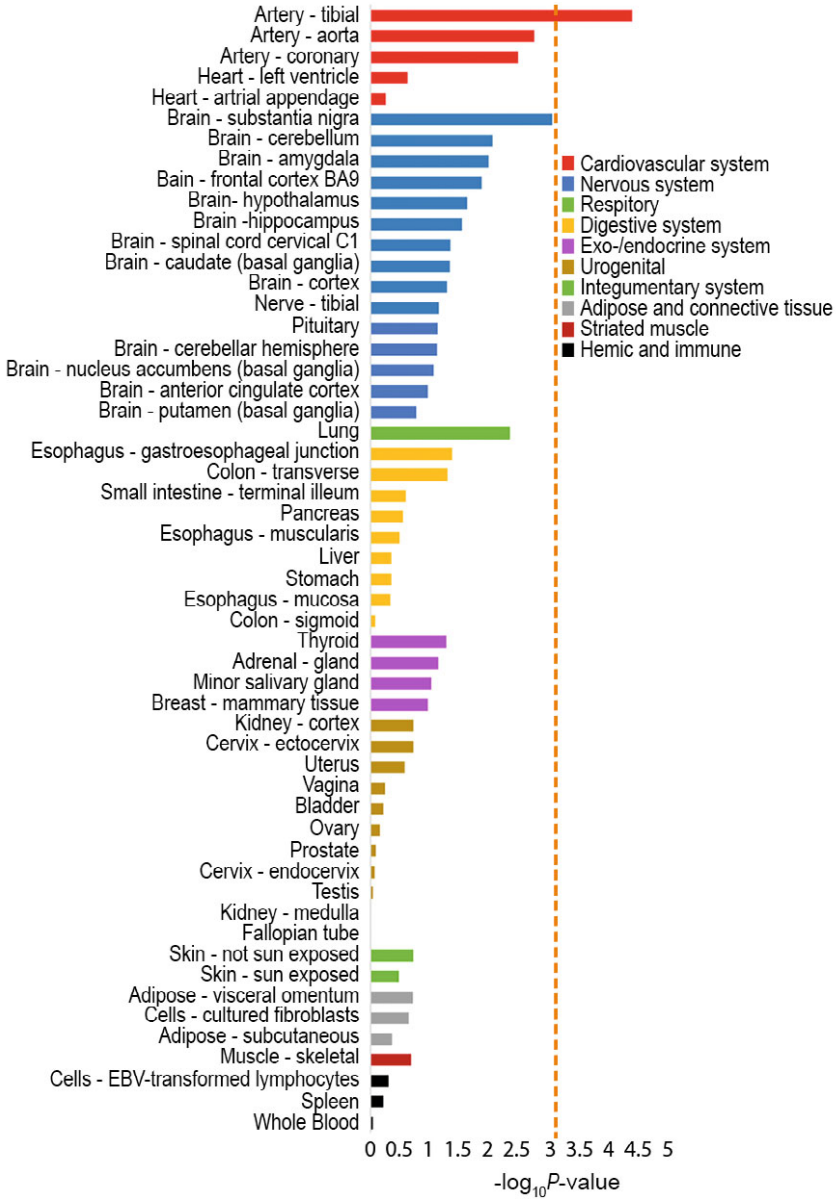


Computational prediction of genetically driven CpG methylation associated with cluster headache, using MetaMeth. Genes annotated to significant CpGs are shown (FDR-corrected p value < 0.05). Horizontal axis shows the chromosomal position and the vertical axis shows significance (-log10 pvalue). The top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

Mendelian randomization analysis

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

Figure 5 Tissue enrichment for the putative causal genes.



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

Discussion

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Supplementary Information

Supplementary Tables

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

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

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

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

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

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

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

Potential conflicts of interest: Nothing to report.

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

Reference

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

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

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

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

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

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

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

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

8

Genome-wide analysis of 102,084 migraine cases identifies 123 risk loci and subtype-specific risk alleles

Heidi Hautakangas, Bendik S. Winsvold, Sanni E. Ruotsalainen, Gyda Bjornsdottir, Aster V.E. Harder, Lisette J.A. Kogelman, Laurent F. Thomas, Raymond Noordam, Christian Benner, Padhraig Gormley, Ville Artto, Karina Banasik, Anna Bjornsdottir, Dorret I. Boomsma, Ben M. Brumpton, Kristoffer Sølvsten Burgdorf, Julie E. Buring, Mona Ameri Chalmer, Irene de Boer, Martin Dichgans, Christian Erikstrup, Markus Färkkilä, Maiken Elvestad Garbrielsen, Mohsen Ghanbari, Knut Hagen, Paavo Häppölä, Jouke-Jan Hottenga, Maria G. Hrafnsdottir, Kristian Hveem, Marianne Bakke Johnsen, Mika Kähönen, Espen S. Kristoffersen, Tobias Kurth, Terho Lehtimäki, Lannie Lighart, Sigurdur H. Magnusson, Rainer Malik, Ole Birger Pedersen, Nadine Pelzer, Brenda W.J.H. Penninx, Caroline Ran, P aul M. Ridker, Frits R. Rosendaal, Gudrun R. Sigurdardottir, Anne Heidi Skogholt, Olafur A. Sveinsson, Thorgeir E. Thorgeirsson, Henrik Ullum, Lisanne S. Vijfhuizen, Elisabeth Widén, Ko Willems van Dijk, International Headache Genetics Consortium, HUNT All-in Headache, Danish Blood Donor Study Genomic Cohort, Arpo Aromaa, Andrea Carmine Belin, Tobias Freilinger, M. Arfan Ikram, Marjo-Riitta Järvelin, Olli T. Raitakari, Gisela M. Terwindt, Mikko Kallela, Maija Wessman, Jes Olesen, Daniel I. Chasman, Dale R. Nyholt, Hreinn Stefánsson, Kari Stefansson, Arn M.M. van den Maagdenberg, Thomas Folkmann Hansen, Samuli Ripatti, John-Anker Zwart, Aarno Palotie and Matti Pirinen

Nat Genet. 2022;54(2):152160

Abstract

Migraine affects over a billion individuals worldwide but its genetic underpinning remains largely unknown. Here we performed a genome-wide association study (GWAS) of 102,084 migraine cases and 771,257 controls and identified 123 loci, of which 86 are novel. The loci provide an opportunity to evaluate shared and distinct genetic components in the two main migraine subtypes: migraine with aura and migraine without aura. Stratification of the risk loci using 29,679 cases with subtype information indicated three risk variants that appear specific for migraine with aura (in *HMOX2*, *CACNA1A* and *MPPED2*), two that appear specific for migraine without aura (near *SPINK2* and near *FECH*), and nine that increase susceptibility for migraine regardless of subtype. The new risk loci include genes encoding recent migraine-specific drug targets, namely calcitonin gene-related peptide (*CALCA/CALCB*) and serotonin 1F receptor (*HTR1F*). Overall, genomic annotations among migraine-associated variants were enriched in both vascular and central nervous system tissue/cell types, supporting unequivocally that neurovascular mechanisms underlie migraine pathophysiology.

Introduction

Migraine is a highly prevalent brain disorder characterized by disabling attacks of moderate to severe pulsating and usually one-sided headache that may be aggravated by physical activity and can be associated with symptoms such as a hypersensitivity to light and sound, nausea and vomiting.[1] Migraine has a lifetime prevalence of 15-20% and is ranked as the second most disabling condition in terms of years lived with disability.[2, 3] Migraine is three times more prevalent in females than in males. For about one-third of patients, migraine attacks often include an aura phase[4] characterized by transient neurological symptoms such as scintillations. Hence, the two main migraine subtypes are defined as migraine with aura (MA) and migraine without aura (MO).

It has been debated for decades whether or not the migraine subtypes are in fact two separate disorders,[5-7] and if so, what the underlying causes are. Prevailing theories about migraine pathophysiology emphasize neuronal and/or vascular dysfunction.[8, 9] Current knowledge on disease mechanisms largely comes from studies of a rare monogenic sub-form of MA, familial hemiplegic migraine, for which three ion transporter genes (*CACNA1A*, *ATP1A2* and *SCN1A*) have been identified.[10] The common forms of migraine, MA and MO, instead have a complex polygenic architecture with an increased familial relative risk,[5] increased concordance in monozygotic twins,[11] and a heritability of 40-60%.[12] The largest GWAS thus far, with 59,674 cases and 316,078 controls, reported 38 genomic loci that confer migraine risk.[13] Subsequent analyses of these GWAS data showed enrichment of migraine signals near activating histone marks specific to cardiovascular and central nervous system tissues,[14] as well as for genes expressed in vascular and smooth muscle tissues.[13] Other smaller GWAS[15-21] have suggested 10 additional loci. Of note, the previous datasets were too small to perform a meaningful comparison of the genetic background between migraine subtypes.

As migraine is globally the second largest contributor to years lived with disability,[2, 3] there is clearly a large need for new treatments. Triptans, i.e., serotonin 5-$HT_{1B/1D}$ receptor agonists, are migraine-specific acute treatments for the headache phase but are not effective in every patient, whereas preventive medication is far from satisfactory alltogether.[22] Recent promising alternatives for acute treatment are serotonin 5-HT_{1F} receptor agonists ('ditans')[23] and small-molecule calcitonin-gene related peptide (CGRP) receptor antagonists ('gepants').[24, 25] For preventive treatment, monoclonal antibodies (mAbs) targeting CGRP or its receptor have recently proven effective,[26] and new gepants are under development for migraine prevention.[27] Still, there remains an urgent need for treatment options for patients who do not respond to the existing treatments. Genetics has proven a promising way to develop novel therapeutic hypotheses in other prevalent complex diseases, such as cardiovascular disease[28] and type 2 diabetes,[29] and we anticipate that large genetic studies of migraine could also yield similar insights.

We conducted a GWAS meta-analysis of migraine by adding to the previous meta-analysis[13] 42,410 new migraine cases from four study collections (**Table 1**). This increased the number of migraine cases by 71% for a total sample of 102,084 cases and 771,257 controls. Furthermore,

we assessed the subtype specificity of the risk loci in 8,292 new MA and 6,707 new MO cases in addition to the 6,332 MA and 8,348 MO cases used previously[13] (**Table 2**). Here we report 123 genomic loci, of which 86 are novel, and include the first four loci that reach genome-wide significance ($P < 5 \times 10^{-8}$) in MA. Our subtype data compellingly show that migraine risk is conferred both by risk loci that appear specific for only one subtype as well as by loci that are shared by both subtypes. Our findings also include new risk loci containing target genes of recent migraine drugs acting on the CGRP pathway and the serotonin 5-HT_{1F} receptor. Finally, our data support the concept that migraine is brought about by both neuronal and vascular genetic factors, strengthening the view that migraine truly is a neurovascular disorder.

Methods

Cohorts and phenotyping

All participating studies were approved by local research ethics committees, and written informed consent was obtained from all study participants. For all the participating studies, an approval was received to use the data in the present work. Study-specific ethics statements are provided in the **Supplementary Note**.

First, we performed a genome-wide meta-analysis on migraine including five study collections listed in **Table 1** and **Supplementary Table 1**. Second, we performed subtype-specific meta-analyses on MA and on MO, both including five study collections listed in **Table 2**, for the 123 independent risk variants identified in the migraine analysis. A description of the study collections is given in the **Supplementary Note**. In particular, the migraine phenotype has been self-reported in other cohorts except in IHGC2016, where a subset of patients were phenotyped in specialized headache centers, as previously explained.[13]

Table 1 Five migraine study collections included in the meta-analysis

Abbreviation	Full name	Ancestry	Cases	Controls	Case %	Migraine definition
IHGC2016*	Gormley et al. 2016 (no 23andMe)	European descent	29,209	172,931	14.4	Self-reported and ICHD-II
23andMe**	23andMe, Inc. (23andMe.com)	European descent	53,109	230,876	18.7	Self-reported
UKBB	UK Biobank (ukbiobank.ac.uk)	European, British	10,881	330,170	3.2	Self-reported
GeneRISK	GeneRISK (generisk.fi)	European, Finnish	1,084	4,857	18.2	Self-reported
HUNT	Nord-Trøndelag Health Study (ntnu.edu/hunt)	European, Norwegian	7,801	32,423	19.4	Self-reported migraine or fulfilling modified ICHD-II criteria

*IHGC2016 is a meta-analysis of 21 studies listed in Supplementary Table 1 and does not include data from 23andMe. Some studies of IHGC2016 determined migraine status through clinical phenotyping, while migraine status in other studies is based on self-reported information. **23andMe includes 30,465 cases from Gormley et al. (2016) meta-analysis and 22,644 new cases. ICHD-II, the International Classification of Headache Disorders 2nd edition.

Table 2 Study collections included in MO and MA subtype analyses

Abbreviation	Full name	Ancestry	Subtype	Cases	Controls
IHGC2016*	Gormley et al. 2016	European descent	MO	8,348	139,622
			MA	6,332	144,883
UKBB	UK Biobank (ukbiobank.ac.uk)	European, British	MO	187	320,139
			MA	1,333	320,139
deCODE	deCODE Genetics Inc.	European, Icelandic	MO	1,648	193,050
			MA	2,297	209,338
DBDS	Danish Blood Donor Study	European, Danish	MO	3,756	28,045
			MA	3,938	28,045
LUMINA	LUMINA migraine without aura or with aura	European, Dutch	MO	1,116	1,445
			MA	724	1,447

*IHGC2016 MO is a meta-analysis of 11 studies and IHGC2016 MA is a meta-analysis of 12 studies listed in Gormley et al. 2016. MO, migraine without aura; MA, migraine with aura.

Quality control

Before the meta-analysis, a standard quality control protocol was applied to each individual GWAS. Related individuals were removed from all other cohorts except HUNT (which modeled relatedness via a logistic mixed model) by using an IBD cut-off of 0.185 or smaller. Multi-allelic variants were excluded from all studies, and only variants that satisfied the following thresholds were kept for further analysis: minor allele frequency (MAF) > 0.01, IMPUTE2 info or MACH $r^2 > 0.6$, and, when available, Hardy-Weinberg equilibrium (HWE) P-value > 1×10^{-6} and missingness < 0.05. Variants were matched by chromosome, position and alleles to the UK Biobank data. Indels were recoded as insertions (I) and deletions (D). For each study, the SNPs with an effect allele frequency (EAF) discrepancy of > 0.30 and indels with EAF discrepancy of > 0.20 to UK Biobank were excluded. MAF and EAF plots of cohorts against the reference cohort are shown in **Supplementary Data 7**. We conducted a sensitivity analysis on strand-ambiguous SNPs (with alleles A/T or G/C), by counting, for each pair of studies, how often the same allele of A/T or G/C SNP was coded as the minor allele in both cohorts, as a function of MAF threshold (**Supplementary Table 17**). Minor alleles were same at least in 97.39% of the SNPs without MAF threshold and the corresponding proportions were 99.96% and 79.58% when MAF < 0.25 and when MAF > 0.4, respectively. The very high concordance for SNPs with MAF < 0.25 suggests that the strand-ambiguous SNPs were consistently labeled for almost every SNP. Therefore, we did not exclude any SNPs based on possible labeling mismatches due to strand ambiguity.

Statistical analysis

All statistical tests conducted were two-sided unless otherwise indicated. The GWAS for the individual study cohorts were performed by logistic regression with an additive model of imputed dosage of the effect allele on the log-odds of migraine. The analyses for IHGC2016[13] and 23andMe[19] have been described before. For UKBB data and GeneRISK data, we used PLINK

v2.0.[30] For HUNT data, we used a logistic mixed model with the saddlepoint approximation as implemented in SAIGE v0.20[31] that accounts for the genetic relatedness. All models were adjusted for sex and at least for the four leading principal components of the genetic population structure (**Supplementary Table 18**). Age was used as a covariate when available. A detailed description is provided in **Supplementary Note**. For the chromosome X meta-analysis, male genotypes were coded as (0,2) in all cohorts, and the GWAS were conducted with an X chromosome inactivation model that treats hemizygous males as equivalent to homozygous females.[32]

We performed an inverse-variance weighted fixed-effect meta-analysis on the five study collections by using GWAMA.[33] After the meta-analysis, we excluded the variants with effective sample size *N*eff < 5,000 to remove results with very low precision compared to the majority of variants and were left with 10,843,197 variants surpassing the QC thresholds. We estimated the effective sample size for variant *i* as $N_{\mathrm{eff}(i)} = \frac{1}{fi(1-fi)\,sei^2}$, where fi is the effect allele frequency for variant *i* and *sei* is the standard error estimated by the GWAS software. This quantity approximates the value 2 *N t (1-t) I*, where *N* is the total sample size (cases + controls), *t* is the proportion of cases and *I* is the imputation info (derivation in **Supplementary Note**).

Risk loci

There were 8,117 genome-wide significant (GWS) variants with the meta-analysis *P*-value < 5×10^{-8}. For 8,067 of them that were available in UK Biobank, an LD matrix was obtained from UK Biobank using a random sample of 10,000 individuals included in the UKBB GWAS. We defined the index variants as the LD-independent GWS variants at LD threshold of $r^2 < 0.1$ in the following way. First, the GWS variant with the lowest *P*-value was chosen, and subsequently all GWS variants that were in LD with the chosen variant ($r^2 > 0.1$) were excluded. Next, out of the remaining GWS variants, the variant with the lowest *P*-value was chosen and the GWS variants in LD with that variant were excluded. This procedure was repeated until there were no GWS variants left. Out of the 8,067 variants with LD information, 170 were LD-independent (at $r^2 < 0.1$). For 18/50 variants that were not found in UK Biobank, LD information was available from the 23andMe data, and all 18 variants were in LD ($r^2 > 0.1$) with some index variant. Two of the 18 variants (rs111404218 and rs12149936) had lower *P*-value than the original index variant they were in LD with and hence they replaced the original index variants. For 32 GWS variants, LD remained unknown. Thus, at this stage, the GWS associations were represented by 202 = 168 + 2 + 32 index variants.

Next, to define the risk loci and their lead variants, an LD block around each index variant was formed by the interval spanning all GWS variants that were in high LD ($r^2 > 0.6$) with the index variant. Sizes of these regions ranged from 1 bp (only the variant itself, e.g., the variants with unknown LD) to 1,089 kb. Sets of regions that were less than 250 kb away from each other were merged (distance from the end of the first region to the beginning of the second region). This definition resulted in 126 loci. All other GWS variants were included in their nearest locus based

on their position and the locus boundaries were updated, and finally loci within 250 kb from each other were merged. This resulted in our final list of 123 risk loci. Each risk locus was represented by its lead variant defined as the variant with the lowest *P*-value and named by the nearest protein-coding gene to the lead variant or by the nearest non-coding gene if there was no protein-coding gene within 250 kb. The term "*Near*" was added to the locus name if the lead variant did not overlap with a gene transcript. We note that the nearest gene to the lead variant need not be a causal gene. None of the 32 variants without LD information became a lead variant of a risk locus because all had a variant in the vicinity with a smaller *P*-value.

We annotated and mapped these loci by their physical position to genes by using the Ensembl Variant Effect Predictor (VEP, GRCh37).[34] We used two different thresholds for annotating the nearest genes: a distance of 20 kb and 250 kb to the nearest transcript of a gene. The filtered results including all variants within a gene or a regulatory element are in **Supplementary Table 7B**.

Stepwise conditional analysis

We performed a stepwise conditional analysis (CA) on each risk locus by using FINEMAP v1.4.[35] FINEMAP uses GWAS summary statistics together with an LD reference panel and does not require individual-level data. When the reference LD does not accurately match the GWAS data, full fine-mapping is prone to false positives.[36] A simpler stepwise CA is more robust to inaccuracy in reference LD because CA has a much smaller search space than full fine-mapping, and therefore CA is less likely to run into most problematic variant combinations where LD is very inaccurate. Since we did not have the full in-sample LD from our GWAS data, we only carried out the CA and not the full fine-mapping. For the CA, we included only the SNPs, but no indels, and we used the same reference LD from the UK Biobank data as we used to define the risk loci. We restricted the CA only to the variants with a similar effective sample size (*N*eff) by using a threshold of ±10% of the *N*eff of the lead SNP of the risk locus, because our summary statistics came from the meta-analysis where sample sizes per variant vary greatly. This filter excluded approximately 17% of all GWS variants and was necessary since otherwise CA led to spurious conditional *P*-values, such as $P < 10^{-250}$, for some loci. Consequently, for two of the loci where the lead variant was an indel, the lead variant was not included in the CA. For such regions, we checked that the new lead variant from the CA output was in LD ($r^2 > 0.3$) with the original lead variant. For one locus (rs111404218) where the lead variant does not have LD information in the UK Biobank data, there were no GWS variants left in the CA after filtering by *N*eff. We used the standard GWS ($P < 5 \times 10^{-8}$) threshold to define the secondary variants that were conditionally independent from the lead variant. The CA results are in **Supplementary Tables 6A**,**B**.

eQTL mapping to genes and tissues

We used two data sources to map the risk variants to genes via eQTL associations. From GTEx v8 database (https://gtexportal.org), we downloaded the data of 49 tissues. We first mapped

all 123 lead variants to all significant *cis*-eQTLs across tissues using the FDR cut-off of 5% as provided by the GTEx project.[37] Next, we also mapped the variants in high LD ($r^2 > 0.6$) with the lead variants to all significant *cis*-eQTLs. Finally, we filtered the results to include only the new significant gene-tissue pairs that were not implicated by the lead variants. Results are shown in **Supplementary Tables 9** and **10**.

With FUMA v1.3.6,[38] we mapped the 123 lead variants, and the variants in high LD ($r^2 > 0.6$) with the lead variants, to the other eQTL data repositories provided by FUMA except GTEx, i.e., Blood eQTL Browser,[39] BIOS QTL browser,[40] BRAINEAC,[41] MuTHER,[42] xQTLServer,[43] CommonMind Consortium,[44] eQTLGen,[45] eQTL Cataloque,[46] DICE,[47] scRNA eQTLs,[48] and PsychENCODE.[49] Results are shown in **Supplementary Tables 9** and **10**.

To study whether the lead variants were enriched in any of the 49 tissues from GTEx v8, we fitted a linear regression model where the number of lead variants that are significant *cis*-eQTLs for a specific tissue was used as the outcome, and the overall number of genes with at least one significant *cis*-eQTL reported by GTEx for the tissue was the predictor.[37] We did a separate regression model for each tissue type by leaving the tissue of interest out from the model, and we used the model fitted on the other tissues for predicting the outcome variable for the tissue type of interest. Finally, we checked in which tissues the true observed number of migraine lead variants was outside of the 95% prediction intervals as given by the function 'predict.lm(, interval="prediction")' in R software. Details of the procedure are in the **Supplementary Note**.

LD Score regression

We estimated both the SNP-heritability (h^2_{SNP}) of migraine and pairwise genetic correlations (r_G) between each pair of study collections using LDSC v1.0.0.[50, 51] SNP-heritability and genetic correlations were estimated using European LD scores from the 1000 Genomes Project Phase 3 data for the HapMap3 SNPs, downloaded from https://data.broadinstitute.org/alkesgroup/LDSCORE/. We reformatted the meta-analysis association statistics to LDSC format with munge-tool that excluded variants that did not match with the HapMap3 SNPs, had strand ambiguity (i.e., A/T or G/C SNPs), MAF < 0.01 or missingness more than two-thirds of the 90th percentile of the total sample size, or resided in long-range LD regions,[52] in centromere regions or in the major histocompatibility locus (MHC) of chromosome 6, leaving 1,165,201 SNPs for the LDSC analyses. We used a migraine population prevalence of 16% and a sample proportion of cases of 11.7% = 102,084/(102,084 + 771,257) to turn the LDSC slope into the estimate of h^2_{SNP} on the liability scale.[53] Pairwise genetic correlation results are listed in **Supplementary Table 2**. We note that in the previous migraine meta-analysis,[13] LDSC reported h^2_{SNP} value of 14.6% (13.8–15.5%), which was considerably larger than the value 11.2% (10.8–11.6%) that we report in our analysis. When we ran our LDSC pipeline on the data of Gormley et al.[13], we estimated h^2_{SNP} value of 10.6% (10.1–11.1%). Thus, it seems that our liability transformation estimates lower values of heritability than the transformation used by Gormley et al..[13]

Stratified LD Score regression

We used stratified LD Score regression (S-LDSC) to partition the SNP heritability by functional genomic annotations.[54] We used the baseline-LD model[55] that contains 75 annotations, including conserved, coding and regulatory regions of the genome and different histone modifications. Baseline-LD model adjusts for MAF- and LD-related annotations, such as recombination rate and predicted allele age, which decreases the risk of model misspecification.[54-56] We used the same QC as with the univariate LDSC, and the baseline LDv1.1 European LD scores estimated from the 1000 Genomes Project Phase 3, downloaded from https://data.broadinstitute.org/alkesgroup/LDSCORE/. We set the significance threshold for enrichment of individual binary functional annotations to $\alpha = 0.05/24$, as we considered only 24 unique functional annotations without the flanking regions. Results are listed in **Supplementary Table 8**.

Subtype analyses of migraine with and without aura

First, we combined new MA and MO data (**Table 2**) with the previously used migraine subtype-specific meta-analysis data,[13] and estimated migraine subtype-specific effect sizes for the 123 lead variants from the migraine meta-analysis. We tested how often the direction of allelic effects was similar between the IHGC MA/MO and the new cohorts using a binomial test (**Supplementary Table 12B**). Next, we stratified the lead variants by using the information from the migraine subtype-specific analyses. For each of the variants, we estimated probabilities between four possible explanations of the observed data that we call 'NULL', 'MO', 'MA' and 'BOTH'. Under model NULL, the effect is not present in either of the migraine subtypes (i.e., the effect is zero); under model MO or MA, the effect is present only in MO or only in MA but not in both; and under model BOTH, a non-zero effect is shared by both MO and MA. We used a Bayesian approach for model comparison that combines a bivariate Gaussian prior distribution on the two effect sizes with a bivariate Gaussian approximation to the likelihood using GWAS summary statistics.[57] Across all models, the prior standard deviation for the effect is 0.2 on the log-odds scale for non-zero effects and 0 for a zero effect. The bivariate priors for the four models are as follows: NULL assumes a zero effect in both migraine subtypes, MO and MA assume a non-zero effect for one subtype and a zero effect for the other subtype, and BOTH combines the fixed-effect model (exactly the same effect in both subtypes) with the independent-effects model (the two effect sizes are non-zero but uncorrelated with each other) with equal weights. Finally, we assumed that each of the four models (NULL, MO, MA, BOTH) is equally probable *a priori*, which we considered an appropriate assumption since all these variants show a convincing association to overall migraine ($P < 5 \times 10^{-8}$). Then we used the Bayes formula to work out the posterior probability on each model. The results are shown in **Figure 3A**, thresholded by a probability cut-off of 95% and in **Supplementary Table 12A**. The correlation parameter between MO and MA GWAS statistics needed in the bivariate likelihood approximation was estimated to be 0.148 using the empirical Pearson correlation of the effect size estimates of the common variants (MAF > 0.05) that did not show a strong association to either of the migraine subtypes ($P > 1 \times 10^{-4}$).[58]

We tested whether the effect sizes between MA and MO were equal at a Bonferroni corrected significance threshold of $\alpha = 0.05/123$ by using a normal approximation and accounting for the correlation in effect size estimators.

We note that the amount of information in the data ("statistical power") is taken automatically into account in this model comparison, which we consider an advantage compared to a comparison of the raw *P*-values between the subtype analyses that does not automatically account for statistical power. In particular, observing a GWS *P*-value ($P < 5 \times 10^{-8}$) in one subtype but not in the other subtype is not yet evidence for a subtype-specific locus, because the effect could still be non-zero also for the other subtype but simply lack power to reach the stringent GWS threshold. Finally, we point out that the inference in the model comparison approach is conditional on the particular set of models being included in the comparison as well as on the particular choice of the prior distributions.

PheWAS with NHGRI GWAS Catalog and FinnGen R4

We performed phenome-wide association studies (PheWAS) for the 123 lead variants using the NHGRI GWAS Catalog and the FinnGen R4 GWAS summary statistics. In addition, we performed the same lookups for the 123 risk loci including all variants in high LD ($r^2 > 0.6$) with the lead variants. With the GWAS Catalog, we first downloaded all the available results (4,314 traits) from the GWAS Catalog webpage (accessed 6.4.2020). Next, we obtained all the associations for the 123 risk loci with all the high LD variants included using *P*-value thresholds of $P < 1 \times 10^{-5}$, $P < 1 \times 10^{-6}$ and $P < 1 \times 10^{-4}$ (**Supplementary Table 13A-C**). Because the GWAS Catalog includes results from several different GWAS for the same phenotype or for a very similar phenotype with a different name, we divided the phenotype associations into broader categories. The new categories are listed in **Supplementary Table 19**. The same approach was used for the PheWAS of FinnGen R4. We first downloaded all the available summary statistics (2,263 endpoints), and next, obtained all the associations for the 123 risk loci using the same three *P*-value thresholds as with the GWAS Catalog (**Supplementary Table 13A-C**). We also divided similar endpoints into broader categories that are listed in **Supplementary Table 20**.

We tested the direction of allelic effects between migraine and the following three traits that shared multiple associated variants with migraine: coronary artery disease (CAD),[59] diastolic blood pressure,[60] and systolic blood pressure[51]. We first took all migraine lead variants that were available also in the summary statistics of the other trait without any *P*-value threshold and used a binomial test to test whether the proportion of variants with same direction of effects was 0.5. Next, we used a *P*-value threshold of 1×10^{-5} for the association with the other trait. Results are in **Supplementary Table 13D**.

LD Score regression applied to specifically expressed genes

We used LD Score regression applied to specifically expressed genes (LDSC-SEG)[14] to identify tissues and cell types implicated by the migraine GWAS results. LDSC-SEG uses gene expression

data and GWAS results from all variants together with an LD reference panel. For our analyses, we used the same QC as for the other LDSC analyses and six different sets of readily constructed annotation-specific LD scores downloaded from https://data.broadinstitute.org/alkesgroup/LDSCORE/LDSC_SEG_ldscores/: multi-tissue gene expression, multi-tissue chromatin, GTEx brain, Cahoy, Corces ATAC and ImmGen LD Scores. FDR was controlled by the Benjamini-Hochberg method. The results are in **Supplementary Table 14A-F**. There were no significant results with the Cahoy, Corces ATAC and ImmGen data at FDR 5%.

Multi-marker Analysis of GenoMic Annotation (MAGMA)

We applied MAGMA v1.09[61] to identify genes and gene sets associated with the migraine meta-analysis results. First, we mapped the meta-analysis SNPs to 18,985 protein-coding genes based on their physical position in the NCBI 37 build by using default settings of MAGMA. Next, we performed a gene-based analysis using the default SNPwise-mean model and the same UK Biobank LD reference as for the other analyses. We applied a Bonferroni correction ($\alpha = 0.05/18{,}985$) to identify significantly associated genes for migraine with the results listed in **Supplementary Table 16A**. Finally, we used the results from the gene-based analysis to perform a gene-set analysis by using two different gene-set collections from the Molecular Signature Database v.7.0[62, 63]: the curated gene sets containing 5,500 gene sets and the GO gene sets containing 9,988 gene sets. The gene-set analysis was performed using the competitive gene set model and one-sided test that tests whether the genes in the gene-set are more strongly associated with the phenotype compared to the other genes. To correct for multiple testing, we used a Bonferroni correction ($\alpha = 0.05/(5{,}500 + 9{,}988)$). Results are in **Supplementary Table 16B,C** and in **Supplementary Figure 7**.

DEPICT

DEPICT[64] is an integrative tool to identify the most likely causal genes at associated loci, and enriched pathways and tissues or cell types in which the genes from the associated loci are highly expressed. As an input, DEPICT takes a set of trait-associated SNPs. First, DEPICT uses co-regulation data from 77,840 microarrays to predict biological functions of genes and to construct 14,461 reconstituted gene sets. Next, information of similar predicted gene functions is used to identify and prioritize gene sets that are enriched for genes in the associated loci. For the tissue and cell type enrichment analysis, DEPICT uses a set of 37,427 human gene expression microarrays. We used DEPICT v1.194 and ran the analyses twice for each of the *P*-value thresholds for clumping, as recommended,[64] and using the default settings of 500 permutations for bias adjustment and 50 replications for the FDR estimation and for the *P*-value calculation. As an input, we used only the autosomal SNPs and the same UK Biobank LD reference data as for the other analyses. First, we ran the analysis using a clumping *P*-value threshold of 5×10^{-8} that resulted in 165 clumps formed from 7,672 variants (**Supplementary Table 15D-F**). Second, we used a *P*-value threshold of 1×10^{-5} leading to 612 clumps formed from 22,480 variants (**Supplementary Table 15A-C**).

Transcriptome-wide association study and colocalization

We performed a transcriptome-wide association study (TWAS) by S-PrediXcan[65] v0.7.5 using GTEx v8 multivariate adaptive shrinkage models (MASHR-M) for 49 tissues downloaded from predictdb.org and the European 1000 Genomes v3 LD reference panel (hg38). We followed the recommended QC protocol, and first harmonized and imputed the migraine summary statistics to ensure an optimal overlap with the GTEx v8 expression weights. After the harmonization and summary statistic imputation, 8,909,736 variants were available for the TWAS. We performed the analysis with default settings to identify significant gene-tissue pairs. We applied a Bonferroni corrected significance level of $\alpha = 0.05/662{,}726$, corresponding to the number of unique gene-tissue pairs tested.

Next, we performed colocalization analysis with COLOCv4.0.4[66] R package for the 1,844 significant gene-tissue pairs to indicate pairs that could be due to LD contamination. COLOC compares five hypotheses where the null hypothesis (H0) corresponds to no association to either eQTL or GWAS, H1 and H2 correspond to associations with only one of the traits, H3 corresponds to association with both eQTL and GWAS but at distinct causal variants, and H4 corresponds to association with both eQTL and GWAS at a shared causal variant. We set a prior probability for colocalization as $p_{12} = 5 \times 10^{-6}$ for all tested regions and restricted the analysis to variants that had $N_{\text{eff}} \pm 10\%$ of the N_{eff} of the lead variant of the region. Results are presented in **Supplementary Table 11B**.

8

Fine-mapping of causal gene sets (FOCUS)

To prioritize genes for the migraine loci, we applied a gene-based fine-mapping approach using FOCUS v0.7.[67] FOCUS is a Bayesian approach that models predicted expression correlations among TWAS signals to estimate posterior probabilities for all genes within a tested region.

We used the European 1000 Genomes v3 LD reference panel and same GTEx v8 predicted expression weights for the 49 tissues as with S-PrediXcan. First, we mapped the migraine summary statistics from hg37 to hg38 with UCSC liftOver.[68] Next, we followed the suggested QC protocol and applied the modified munge-tool to obtain cleaned summary statistics. After the QC steps, we had 6,237,177 variants left for the analysis. We performed tissue-prioritized fine-mapping of gene-sets for the 49 tissues with otherwise default settings except that we increased the *P*-value threshold to 1×10^{-4} so that the fine-mapping would cover most of the same regions that contained at least one significant gene-tissue pair by S-PrediXcan. Posterior inclusion probability (PIP) from FOCUS is reported for all available significant S-PrediXcan gene-tissue pairs in **Supplementary Table 11B**, and all prioritized genes by FOCUS with PIP > 0.9 are reported in **Supplementary Table 11A**.

Data Availability

Results for 8,117 genome-wide significant SNP associations ($P < 5 \times 10^{-8}$) from the meta-analysis including 23andMe data are available on the International Headache Genetics Consortium website (http://www.headachegenetics.org/content/datasets-and-cohorts). Genome-wide summary

statistics for the other study collections except 23andMe are available for bona fide researchers (contact Dale Nyholt, d.nyholt@qut.edu.au) within two weeks from the request. The full GWAS summary statistics for the 23andMe discovery data set will be made available through 23andMe to qualified researchers under an agreement with 23andMe that protects the privacy of the 23andMe participants. Please visit research.23andme.com/collaborate/#publication for more information and to apply to access the data.

Code Availability

R code for the subtype specificity analysis: https://github.com/mjpirinen/migraine-meta.

Results

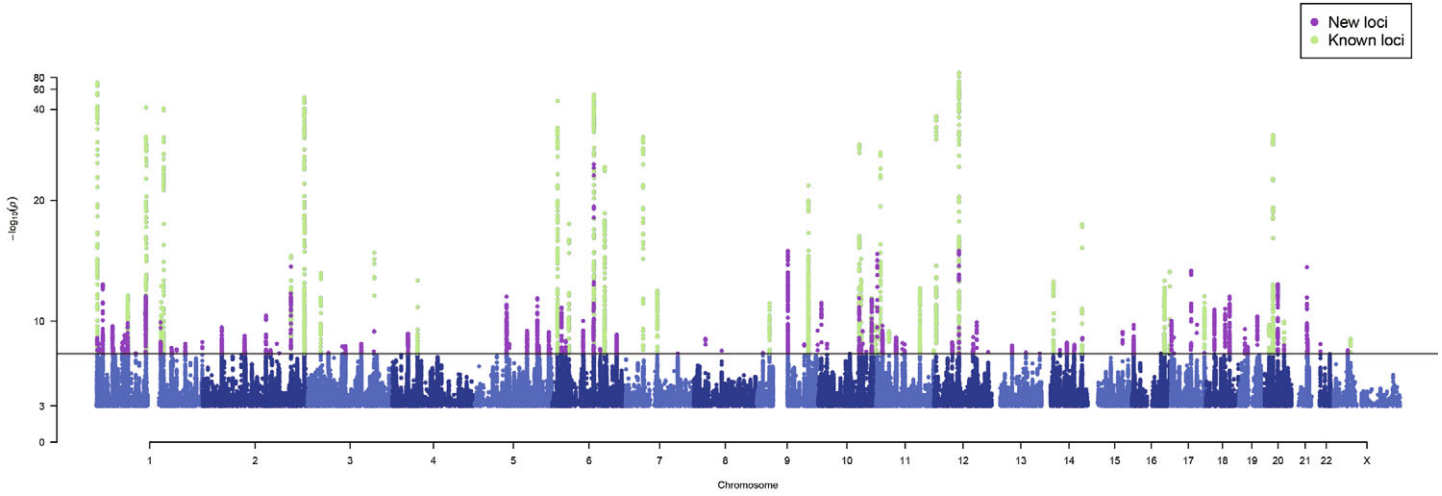
Genome-wide meta-analysis

We combined data on 873,341 individuals of European ancestry (102,084 cases and 771,257 controls) from five study collections (**Table 1** and **Supplementary Table 1**) and analyzed 10,843,197 common variants (Methods). Despite different approaches to the ascertainment of migraine cases across the studies, the pairwise genetic correlations were all near 1 (**Supplementary Table 2**), as determined by LD Score (LDSC) regression,[50] showing high genetic and phenotypic similarity across the studies, justifying their meta-analysis. Pairwise LDSC intercepts were all near 0, indicating little or no sample overlap (**Supplementary Table 2**).

The genomic inflation factor (λ_{GC}) of the fixed-effect meta-analysis results was 1.33 (**Supplementary Figure 1**), which is in line with other large meta-analyses[69-71]and is as expected for a polygenic trait.[72] The univariate LDSC[51] intercept was 1.05 (s.e. 0.01), which, being close to 1.0, suggests that most of the genome-wide elevation of the association statistics comes from true additive polygenic effects rather than from a confounding bias such as population stratification. The LDSC analysis showed a linear trend between the variant's LD-score and its association with migraine, as expected from a highly polygenic phenotype such as migraine (**Supplementary Figure 2**). The SNP-heritability estimate from LDSC was 11.2% (95%CI 10.8-11.6%) on a liability scale when assuming a population prevalence of 16%.

We identified 8,117 genome-wide significant (GWS; $P < 5 \times 10^{-8}$) variants represented by 170 LD-independent index variants ($r^2 < 0.1$). We defined the risk loci by including all variants in high LD ($r^2 > 0.6$) with the index variants and merged loci that were closer than 250 kb (Methods). This resulted in 123 independent risk loci (**Figure 1**, **Supplementary Table 3A**, and **Supplementary Data 1** and **2**). Of the 123 loci, 86 are novel whereas 36 overlap with the previously reported 47 autosomal risk loci (**Supplementary Table 4**) and one with the previously reported X chromosome risk locus. Of the 11 previously reported migraine risk loci that were not GWS in our study, six were GWS in Gormley et al.[13] and had $P < 3.50 \times 10^{-5}$ in our data, one had $P = 2.37 \times 10^{-3}$,

Figure 1 Manhattan plot of migraine GWAS meta-analysis (n = 873,341; 102,084 cases and 771,257 controls)



On the x-axis, variants are plotted along the 22 autosomes and the X chromosome. y-axis shows the statistical strength of the association from the inverse-variance weighted fixed-effect meta-analysis as the negative $\log_{10}$ of the uncorrected two-sided P-value. Horizontal line is the genome-wide significance threshold ($P = 5 \times 10^{-8}$). The 123 risk loci passing the threshold are divided into 86 new loci (purple) and 37 previously known loci (green). Adjacent chromosomes are colored in different shades of blue. Plotted are variants with $P < 0.001$.

three had $P > 0.14$, and one was not available in our data (**Supplementary Data 3**). When we represented each risk locus by its lead variant, i.e., the variant with the smallest *P*-value, 47 GWS variants were LD-independent ($r^2 < 0.1$) of the 123 lead variants, and with a more stringent threshold ($r^2 < 0.01$), 15 GWS variants remained LD independent of the 123 lead variants (**Supplementary Table 5**).

In addition, we conducted an approximate stepwise conditional analysis for the 123 risk loci (Methods). Since sample sizes per variant varied considerably, we restricted the conditional analysis to variants with similar effective sample sizes to the lead variant. The conditional analysis returned 6 SNPs within the 123 risk loci that remained GWS after conditioning on the lead variants (**Supplementary Table 6A**,**B**).

Characterization of migraine risk loci

We mapped the 123 risk loci to genes by their physical location using the Ensembl Variant Effect Predictor (VEP).[34] Of the lead variants, 59% (72/123) were within a transcript of a protein-coding gene, and 80% (99/123) of the loci contained at least one protein-coding gene within 20 kb, and 93% (114/123) within 250 kb (**Supplementary Table 3**). Five of the 123 lead variants were missense variants (in genes *PLCE1*, *MRGPRE*, *SERPINA1*, *ZBTB4* and *ZNF462*), and 40 more missense variants were in high LD ($r^2 > 0.6$) with the lead variants (**Supplementary Table 7A**). Of note, three variants with a predicted high impact consequence on protein function were in high LD with the lead variants: (i) a stop gained variant (rs34358) with lead variant rs42854 ($r^2 = 0.85$) in gene *ANKDD1B*, (ii) a splice donor variant (rs66880209) with lead variant rs1472662 ($r^2 = 0.71$) in *RP11-420K8.1*, and (iii) a splice acceptor variant (rs11042902) with lead variant rs4910165 ($r^2 = 0.69$) in *MRVI1* (**Supplementary Table 7B**).

We used stratified LDSC (S-LDSC) to partition migraine heritability by 24 functional genomic annotations.[54, 55] We observed enrichment for 10 categories (**Supplementary Figure 3** and **Supplementary Table 8**), with conserved regions showing the highest enrichment (11.2-fold; $P = 1.95 \times 10^{-10}$), followed by coding regions (8.1-fold; $P = 1.36 \times 10^{-3}$) and enhancers (4.2-fold; $P = 3.64 \times 10^{-4}$).

Prioritization of candidate genes

We mapped the 123 lead variants to genes via expression quantitative trait locus (eQTL) association using the GTEx v8[37] and data repositories included in FUMA[38] at a false discovery rate (FDR) of 5% (Methods). The lead variants were *cis*-eQTLs for 589 genes (**Supplementary Table 9**), and variants in high LD with the lead variants were *cis*-eQTLs for an additional 624 genes (**Supplementary Table 10**). In total, 84% (103/123) of lead variants were *cis*-eQTLs for at least one gene. Tibial artery had the highest number (47/123) of lead variants as *cis*-eQTLs in GTEx v8, and it was the only tissue type where the enrichment was statistically higher ($P = 6.37 \times 10^{-6}$)

than expected based on the overall number of *cis*-eQTLs per tissue reported by GTEx (**Supplementary Figure 4** and **Supplementary Note**).

To prioritize candidate genes for the risk loci, we applied two approaches based on GTEx v8 expression data: fine-mapping of causal gene-sets by FOCUS[67] (**Supplementary Table 11A**) and a transcriptome-wide association study (TWAS) by S-PrediXcan[65] combined with colocalization analysis using COLOC[66] (**Supplementary Table 11B**).

With posterior probability (PP) > 0.5, FOCUS found candidate genes for 82 loci and S-PrediXcan + COLOC supported colocalization for 52 loci (**Supplementary Table 11C**). In total, 73 genes in 46 loci were prioritized by both methods. *MRC2* and *PHACTR1* were the only genes that both methods prioritized with strong evidence (PP > 0.99 for same tissue) and without any other gene prioritized within their loci.

Table 3 LDSC-SEG results that are significant at FDR 5%

Tissue/cell type and histone mark	Tissue category	*P*-value	FDR adjusted *P*-value
Multi-tissue gene expression data			
Aorta	Cardiovascular	1.78E-04	0.029
Tibial artery	Cardiovascular	3.60E-04	0.029
Coronary artery	Cardiovascular	4.29E-04	0.029
Gene expression data of 13 brain regions from GTEx			
Caudate (basal ganglia)	Central nervous system	6.00E-04	0.008
Multi-tissue chromatin annotation data			
Fetal brain female, H3K4me3	Central nervous system	2.49E-05	0.012
Brain dorsolateral prefrontal cortex, H3K27ac	Central nervous system	8.43E-05	0.018
Brain dorsolateral prefrontal cortex, H3K4me3	Central nervous system	1.11E-04	0.018
Aorta, H3K4me1	Cardiovascular	2.57E-04	0.031
Stomach mucosa, H3K36me3	Digestive	3.36E-04	0.032
Aorta, H3K27ac	Cardiovascular	4.40E-04	0.032
Artery-tibial ENTEX, H3K4me1	Cardiovascular	4.53E-04	0.032
Ganglion eminence derived primary cultured neurospheres, H3K4me3	Central nervous system	6.53E-04	0.04
Brain germinal matrix, H3K4me3	Central nervous system	8.42E-04	0.043
Aorta ENTEX, H3K27ac	Cardiovascular	1.11E-03	0.043
Artery-coronary ENTEX, H3K4me3	Cardiovascular	1.13E-03	0.043
Cortex derived primary cultured neurospheres, H3K36me3	Central nervous system	1.14E-03	0.043
Ovary, H3K27ac	Other	1.15E-03	0.043
Cortex derived primary cultured neurospheres, H3K4me3	Central nervous system	1.29E-03	0.045
Aorta ENTEX, H3K4me1	Cardiovascular	1.39E-03	0.045
Stomach smooth muscle, H3K4me3	Musculoskeletal/connective	1.55E-03	0.047

One-sided *P*-value from testing whether the regression coefficient is positive. FDR, false discovery rate based on Benjamini-Hochberg method. Full results are in Supplementary Table 14A-F.

Two of the new risk loci contain genes (*CALCA*/*CALCB* and *HTR1F*) whose protein products are closely related to targets of two migraine-specific drug therapies.[73] We observe a convincing association at the chromosome 11 locus that contains the *CALCA* and *CALCB* genes encoding CGRP itself (lead SNP rs1003194, $P = 2.43 \times 10^{-10}$; **Figure 2A**), while none of the genes encoding CGRP receptor proteins (*CALCRL*, *RAMP1* or *RCP*) show a statistically comparable association (all $P > 10^{-4}$; **Supplementary Figure 5**). Variant rs1003194 is a *cis*-eQTL for *CALCB*, but also for *COPB1*, *PDE3B* and *INSC* (**Supplementary Table 9**) and FOCUS prioritizes *CALCA*, *CALCB* and *INSC* (**Supplementary Table 11C**). In addition, a new locus on chromosome 3 contains *HTR1F* (lead SNP rs6795209, $P = 1.23 \times 10^{-8}$; **Figure 2B**), which encodes the serotonin 5-HT_{1F} receptor. Variant rs6795209 is a significant *cis*-eQTL for *HTR1F*, as well as for three other genes (*CGGBP1*, *ZNF654*, *C3orf38*) in the same locus (**Supplementary Table 9**). FOCUS or S-PrediXcan + COLOC did not prioritize *HTR1F* based on gene expression data (**Supplementary Table 11C**).

Migraine subtypes with aura and without aura

Previously, Gormley et al.[13] conducted subtype-specific GWAS with 6,332 MA cases against 144,883 controls and 8,348 MO cases against 139,622 controls, and reported that 7 loci were GWS in MO but none were GWS in MA. Here we added to the previous data 8,292 new MA and 6,707 new MO cases from headache specialist centers in Denmark and the Netherlands as well as from study collections in Iceland and UK Biobank (**Table 2**), for total sample sizes of 14,624 MA cases and 703,852 controls, and 15,055 MO cases and 682,301 controls. We estimated the effect size for each subtype at the 123 lead variants of the migraine GWAS (**Supplementary Table 3B**,**C** and **Supplementary Data 4** and **5**) and detected four GWS variants in the MA meta-analysis and 15 GWS variants in the MO meta-analysis. We also estimated a probability that the lead variant is either subtype-specific (i.e., associated only with MO or with MA but not with both), shared by both subtypes, or not associated with either subtype (**Methods**, **Supplementary Table 12A**, and **Supplementary Data 6**). With a probability above 95%, three lead variants (i.e., rs12598836 in the *HMOX2* locus, rs10405121 in the *CACNA1A* locus, and rs11031122 in the *MPPED2* locus) are MA-specific, while two lead variants (i.e., rs7684253 in the locus near *SPINK2* and rs8087942 in the locus near *FECH*) are MO-specific at a similar threshold. Nine lead variants were shared by MA and MO with > 95% probability (**Figure 3A**). In addition to the five subtype-specific lead variants, four other lead variants also showed differences in effect size between the subtypes ($P < 0.05/123$) (**Figure 3B**).

PheWAS with NHGRI GWAS Catalog and FinnGen R4

Next, we conducted phenome-wide association scans (PheWAS) for the lead variants for 4,314 traits with reported associations in the NHGRI GWAS Catalog (https://www.ebi.ac.uk/gwas/) and for the GWAS summary statistics of 2,263 disease traits in the FinnGen release 4 data. We identified 25 lead variants that were reported to be associated with 23 different phenotype

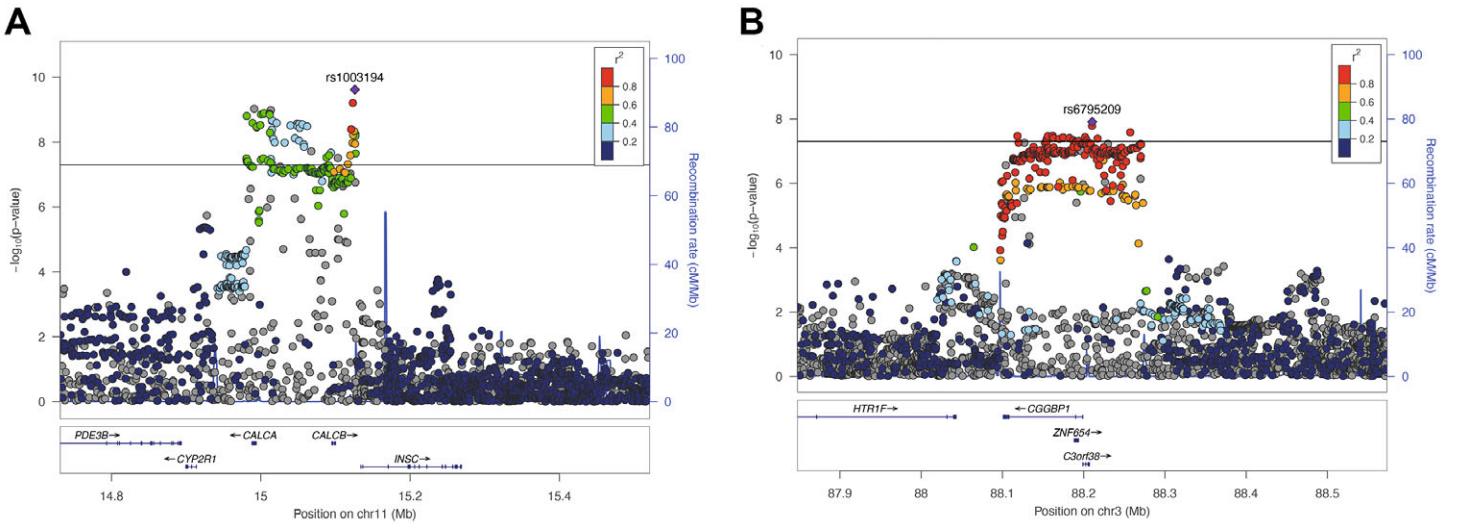
categories (Methods) in the GWAS Catalog, and 17 lead variants with 26 defined disease categories in FinnGen at $P < 1 \times 10^{-5}$. The categories with the highest number of reported associations were cardiovascular disease (7 lead variants) and blood pressure (6 lead variants) in the GWAS Catalog, and diseases of the circulatory system (11 lead variants) in FinnGen. When we performed PheWAS for all variants in high LD ($r^2 > 0.6$) with the lead variants, we observed associations for 79 loci with 54 different phenotype categories in the GWAS Catalog, and for 41 loci with 26 disease categories in FinnGen (**Supplementary Table 13A** and **Supplementary Figure 6**).

These findings are consistent with previous results that migraine is a risk factor for multiple cardiovascular traits[74-76], and genetically correlated with blood pressure.[77, 78] However, we did not observe a trend in the direction of the allelic effects between migraine and coronary artery disease (CAD) or migraine and blood pressure traits (**Supplementary Table 13D**) using the latest meta-analysis of CARDIoGRAMplusCD4 Consortium[59] (n = 336,924) and blood pressure GWAS from UK Biobank[60] (n = 422,771).

Enrichment in tissue or cell types and gene sets. We used LDSC applied to specifically expressed genes (LDSC-SEG)[14] (Methods) to evaluate whether the polygenic migraine signal was enriched near genes that were particularly active in certain tissue or cell types as determined by gene expression or activating histone marks. Using multi-tissue gene expression data, we found enrichment at FDR 5% in three cardiovascular tissue/cell types, i.e., aorta artery ($P = 1.78 \times 10^{-4}$), tibial artery ($P = 3.60 \times 10^{-4}$) and coronary artery ($P = 4.29 \times 10^{-4}$) (**Table 3** and **Supplementary Table 14A**), all of which have previously been reported enriched in migraine without aura.[14] The fine-scale brain expression data from GTEx, since recently including 13 brain regions, showed enrichment in the caudate nucleus of striatum, a component of basal ganglia ($P = 6.02 \times 10^{-4}$; **Table 3** and **Supplementary Table 14B**). With chromatin-based annotations, we found enrichment in five central nervous system (CNS) cell types, three cardiovascular cell types, one cell type of the digestive system, one musculoskeletal/connective cell type, and ovary tissue (**Table 3** and **Supplementary Table 14C**). In addition to replicating previous findings,[13, 14] the signal linking to ovary tissue has not been reported before.

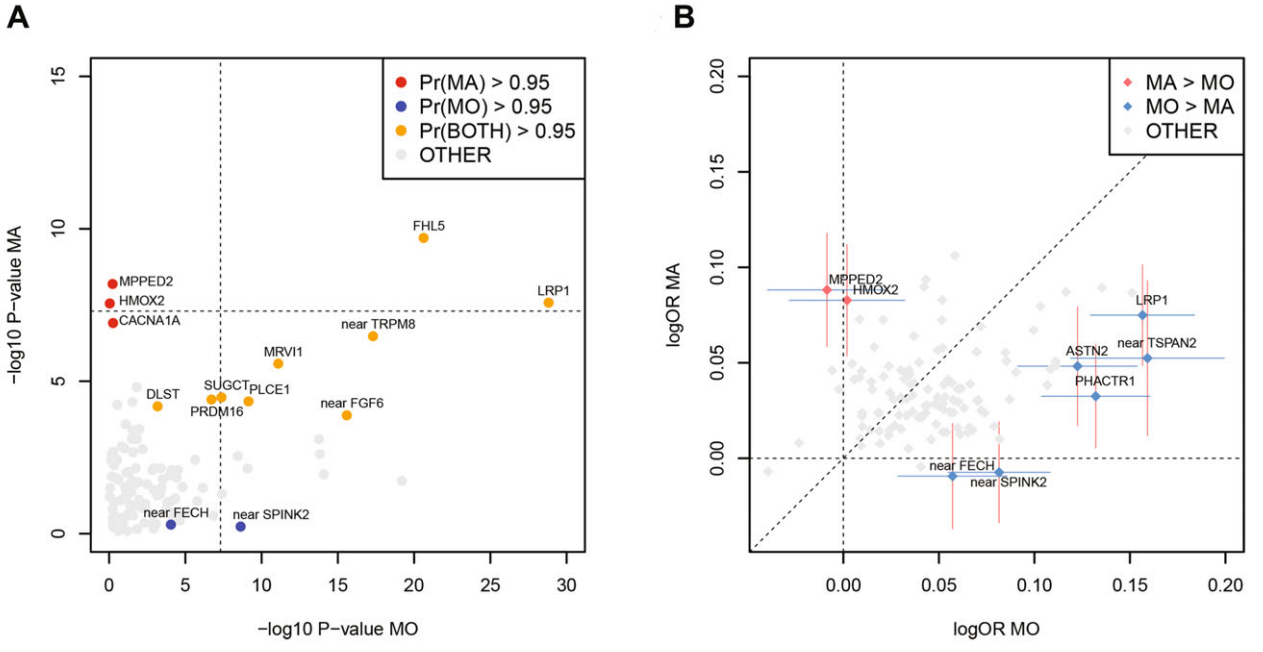
Finally, we used DEPICT[64] to identify tissues whose eQTLs were enriched for migraine-associated variants. The tissue enrichment analysis replicated three previously reported tissues[13]: arteries (nominal $P = 1.03 \times 10^{-3}$), stomach (nominal $P = 1.04 \times 10^{-3}$) and upper gastrointestinal tract (nominal $P = 1.29 \times 10^{-3}$) (**Supplementary Table 14A**). Results of gene set analyses using DEPICT[64] and MAGMA[61] are presented in **Supplementary Tables 15** and **16**.

Figure 2 LocusZoom plots of two novel migraine loci with genes that are targets of recent migraine specific drugs



A) Locus containing *CALCA* and *CALCB* genes, encoding CGRP, which is the target of preventive and acute therapies via monoclonal antibodies and gepants. **B**) Locus containing the *HTR1F* gene, which encodes a serotonin 5-HT_{1F} receptor that is the target of acute therapies via ditans. *x*-axis shows the chromosomal location, and *y*-axis shows the uncorrected two-sided negative $\log_{10}$ *P*-value from the inverse-variance weighted fixed-effects meta-analysis with 102,084 cases and 771,257 controls. The squared correlation to the lead variant is shown by colors based on the UK Biobank data for variants that have an effective sample size ± 20% of the lead variant's effective sample size. Horizontal line corresponds to $P = 5 \times 10^{-8}$. Blue graph shows the recombination rate.

Figure 3 Lead variants stratified by migraine subtype for risk loci with minor allele frequency > 5%



A) Axes show the negative $\log_{10}$ *P*-value of MO (*x*-axis; *n* = 697,356, 15,055 MO cases and 682,301 controls) and MA (*y*-axis; *n* = 718,476, 14,624 MA cases and 703,852 controls) analyses. Two-sided *P*-values are derived from inverse-variance weighted fixed-effect meta-analyses and are uncorrected. Symbols that are colored and annotated indicate > 95% posterior probability that a non-zero effect is present in both MO and MA (model BOTH), or that the effect is present only in MO or only in MA but not both (models MO and MA, respectively). Variants with a probability less than 95% for each of the three models are shown as gray. Dashed lines show the genome-wide significance threshold ($P = 5 \times 10^{-8}$). **B**) Axes show logarithm of odds ratios for MO (*x*-axis; *n* = 697,356, 15,055 MO cases and 682,301 controls) and MA (*y*-axis; *n* = 718,476, 14,624 MA cases and 703,852 controls) calculated for the migraine risk allele. The effects at variants that have been colored and annotated differ between the subtypes at significance level of 0.0004 = 0.05/123. The 95% confidence intervals for the logarithm of odds ratios are shown for the annotated variants. Dashed lines show the coordinate axes and the diagonal. MO, migraine without aura; MA, migraine with aura.

Discussion

We conducted the largest GWAS meta-analysis on migraine thus far by combining genetic data on 102,084 cases and 771,257 controls. We identified 123 migraine risk loci, of which 86 are novel since the previous migraine meta-analysis that yielded 38 loci.[13] This shows that we have now reached the statistical power for rapid accumulation of new risk loci for migraine, in line with the progress of GWAS seen with other common diseases[79], and as expected for a highly polygenic disorder like migraine.[80]

Migraine subtypes MO and MA were defined as separate disease entities some 30 years ago, and since then, the debate has continued as to what extent they are biologically similar. Over the years, arguments in favor[6] and against[5] have been presented, but convincing genetic evidence to support subtype-specific risk alleles has been lacking in genetic studies with smaller sample sizes.[18, 81, 82] Here we increased considerably the evidence for subtype specificity of some risk alleles by including new migraine subtype data at the 123 migraine risk variants. We observed that, with a probability of > 95%, three lead variants (in *HMOX2*, in *CACNA1A* and in *MPPED2*) are associated with MA but not MO. Of them, *CACNA1A* is a well-known gene linked to familial hemiplegic migraine, a rare subform of MA.[83, 84] The observation that *CACNA1A* seems involved in both monogenic and polygenic forms of migraine provides the first gene-based support for the increased sharing of common variants between the two disorders.[80] We find no evidence that any of the seven loci, previously reported as GWS in MO but not in MA,[13] would be specific for MO, while four of them (*LRP1*, *FHL5*, near *FGF6* and near *TRPM8*) are among the nine loci shared by both subtypes with a probability over 95%. Loci (e.g., *LRP1* and *FHL5*) that are strongly associated with both subtypes provide convincing evidence for a previous hypothesis that the subtypes partly share a genetic background.[13, 85] In accordance with our analysis, effects in both subtypes were suggested before at the *TRPM8* and *TSPAN2* loci while, in contrast to our results, the *LRP1* locus was previously reported to be specific for MO.[81] Finally, we also detected four lead variants (including *LRP1*) that do not appear specific for MO but do confer a higher risk for MO than for MA.

It has been long debated whether migraine has a vascular or a neuronal origin, or whether it is a combination of both.[8, 9, 86, 87] Here we found genetic evidence for the role of both vascular and central nervous tissue types in migraine from several tissue enrichment analyses, which refined earlier analyses based on smaller sample sizes.[13, 14]

With respect to a vascular involvement in the pathophysiology of migraine, both gene expression and chromatin annotation data from LDSC-SEG showed that migraine signals are enriched for genes and cell type-specific annotations that are highly expressed in aorta and tibial and coronary arteries. The involvement of arteries was also proposed by our DEPICT tissue enrichment analysis. In addition, cardiovascular disease and blood pressure phenotypes were among the top

categories in the PheWAS analyses. These results are consistent with previous reports of a shared etiology and some genetic correlation between migraine and cardiovascular and cerebrovascular endpoints.[76-78, 88-92] However, in our analysis, the migraine risk alleles neither consistently increased nor consistently decreased the risk of coronary artery disease or the risk of hypertension.

A key role of the central nervous system (CNS) in migraine pathophysiology has emerged from animal models, human imaging, and neurophysiological studies,[10, 93] while support for CNS involvement from genetic studies has been more difficult to obtain. A likely reason is the paucity of gene expression data from CNS tissue types, but recently more data have become available, making such studies feasible. Our LDSC-SEG analysis using gene expression data from 13 brain regions showed an enrichment for caudate nucleus in the basal ganglia, and with chromatin-based annotations for five CNS tissue types: dorsolateral prefrontal cortex, neurospheres derived from cortex, fetal brain, germinal matrix and neurospheres derived from ganglion eminence. Alterations in the structure and/or function of several brain regions,[93-95] including basal ganglia, cortex, hypothalamus, thalamus, brainstem, amygdala and cerebellum, have been reported for individuals who suffer from migraine, but the cause of these changes is not known.

In addition to the support for the hypothesis that both vascular and CNS are important in migraine pathogenesis,[8, 93, 96] the tissue enrichment analyses also reported some tissue types of the digestive system as well as ovary at FDR 5%. Given the female preponderance and suggested influence of sex hormones (e.g. menstrual related migraine) in migraine,[97-99] the involvement of the ovary is an interesting finding, although the statistical evidence for it currently remains weaker compared to that for the vascular and central nervous systems.

A particularly interesting finding in our GWAS was the identification of risk loci containing genes that encode targets for migraine-specific therapeutics. One new locus contains the *CALCA* and *CALCB* genes on chromosome 11 that encode calcitonin gene-related peptide (CGRP). CGRP-related monoclonal antibodies have been successful for the preventive treatment of migraine,[100] and they are considered as a major breakthrough in migraine-specific treatments since the development of the triptans for acute migraine over two decades ago. Another new locus contains the *HTR1F* gene that encodes serotonin 5-HT_{1F} receptor, which is the target of another recent migraine drug class called ditans.[101] Ditans provide a promising acute treatment especially for those migraine patients that cannot use triptans because of cardiovascular risk factors.[23] These two new GWAS associations near genes that are already targeted by effective migraine drugs suggest that there could be other potential drug targets among the new loci and provide a clear rationale for future GWAS efforts to increase the number of loci by increasing sample sizes further. In addition, GWAS data with migraine subtype information can help prioritize treatment targets for particular migraine symptomatology, such as aura symptoms, that currently lack treatment options. More generally, utilizing genetic evidence when selecting new drug targets is estimated to double the success rate in clinical development.[102, 103]

Even though we observed links between our new risk loci and known target genes of effective migraine drugs, the accurate gene prioritization at risk loci remains challenging. First, robust fine-mapping would require accurate LD information,[36] which is typically lacking in meta-analyses and further distorted from reference panels by variation in effective sample size across variants. Second, computational approaches to gene prioritization require further methodological work[104] and extension to additional sources of functional data in order to provide more robust and comprehensive gene prioritization results. Another limitation of our study is that a large proportion of migraine diagnoses are self-reported. Therefore, we cannot rule out misdiagnosis, such as, e.g., tension headache being reported as migraine, which could overemphasize genetic factors related to general pain mechanisms and not migraine *per se*. Regardless, the high genetic correlation that we observed supports a strong phenotypic concordance between the study collections that also included deeply phenotyped clinical cohorts from headache specialist centers, which were instrumental for the migraine subtype analyses. While the subtype data provided convincing evidence of both loci with genetic differences and other loci with genetic overlap between subtypes, larger samples are still needed to achieve a more accurate picture of the similarities and differences in genetic architecture behind the subtypes.

To conclude, we report the largest GWAS meta-analysis of migraine to date, detecting 123 risk loci. We demonstrated that both vascular and central nervous systems are involved in migraine pathophysiology, supporting the notion that migraine is a neurovascular disease. Our subtype analysis of migraine with aura and migraine without aura shows that these migraine subtypes have both shared risk alleles and risk alleles that appear specific to one subtype. In addition, new loci include two targets of recently developed and effective migraine treatments. Therefore, we expect that these and future GWAS data will reveal more of the heterogeneous biology of migraine and potentially point to new therapies against migraine that currently is a leading burden for population health throughout the world.

Supplementary Materials

Supplementary Note and Figs. 1–8.
https://static-content.springer.com/esm/art%3A10.1038%2Fs41588-021-00990-0/MediaObjects/41588_2021_990_MOESM1_ESM.pdf

Supplementary Tables 1–20.
https://static-content.springer.com/esm/art%3A10.1038%2Fs41588-021-00990-0/MediaObjects/41588_2021_990_MOESM4_ESM.xlsx

Supplementary Data 1: Regional LocusZoom plots of the 123 independent migraine risk loci identified from the meta-analysis.
https://static-content.springer.com/esm/art%3A10.1038%2Fs41588-021-00990-0/MediaObjects/41588_2021_990_MOESM5_ESM.pdf

Supplementary Data 2: Forest plots of the 123 lead migraine variants.
https://static-content.springer.com/esm/art%3A10.1038%2Fs41588-021-00990-0/MediaObjects/41588_2021_990_MOESM6_ESM.pdf

Supplementary Data 3: Forest plots of 10 variants that have been previously reported to associate with migraine but failed to replicate in our study.
https://static-content.springer.com/esm/art%3A10.1038%2Fs41588-021-00990-0/MediaObjects/41588_2021_990_MOESM7_ESM.pdf

Supplementary Data 4: Forest plots of the 123 lead migraine variants from the MO meta-analysis.
https://static-content.springer.com/esm/art%3A10.1038%2Fs41588-021-00990-0/MediaObjects/41588_2021_990_MOESM8_ESM.pdf

Supplementary Data 5: Forest plots of the 123 lead migraine variants from the MA meta-analysis.
https://static-content.springer.com/esm/art%3A10.1038%2Fs41588-021-00990-0/MediaObjects/41588_2021_990_MOESM9_ESM.pdf

Supplementary Data 6: Subtype-specific combined log-odds-ratio estimates and posterior probabilities from subtype-specificity analysis for the 123 lead migraine variants.
https://static-content.springer.com/esm/art%3A10.1038%2Fs41588-021-00990-0/MediaObjects/41588_2021_990_MOESM10_ESM.pdf

Supplementary Data 7: Pairwise EAF and MAF plots against the reference cohort (UKBB).
https://static-content.springer.com/esm/art%3A10.1038%2Fs41588-021-00990-0/MediaObjects/41588_2021_990_MOESM11_ESM.pdf

Acknowledgements: We thank the study participants for their contribution to this research. We also thank the numerous individuals who contributed to sample collection, storage, handling, phenotyping and genotyping for each of the individual cohorts. We acknowledge the participants and investigators of the FinnGen study. This research has been conducted using the UK Biobank Resource under Application Number 22627. We are supported by following grants: the US National Institute of Neurological Disorders and Stroke (NINDS) of the US National Institutes of Health (NIH) (grant numbers R21NS09296 and R21NS104398 (D.I.C.)), the Finnish innovation fund Sitra and Finska Läkaresällskapet (E.W.), the Academy of Finland (grant nos. 288509, 312076, 336825 (M.P.)), the Sigrid Juselius Foundation (M.P. and S.R.), the Academy of Finland Center of Excellence in Complex Disease Genetics (grant no. 312062 (S.R.)), the Finnish Foundation for Cardiovascular Research (S.R.), University of Helsinki HiLIFE Fellow and Grand Challenge grants (S.R.), The Novo Nordisk Foundation (NNF14CC0001 and NNF17OC0027594 (T.F.H. and K.B.)), CANDY foundation (CEHEAD) (T.F.H.), and the South-Eastern Norway Regional Health Authority (grant no. 2020034 (B.S.W.)). A list of study-specific acknowledgements and funding information can be found in the Supplementary Note.

Author's contributions: H.S., A.M.J.M.v.d.M., T.F.H. and M.P. conceived the study. H.H., B.S.W., S.E.R., G.B., A.V.E.H., L.J.A.K., L.F.T., R.N. and L.S.V. performed analyses in their respective cohorts. K. Hveem, H.S., K.S., A.M.J.M.v.d.M., T.F.H., S.R., J.-A.Z. and M.P. supervised analyses in their respective cohorts. H.H., B.S.W., D.I.C., D.R.N., A.M.J.M.v.d.M., T.F.H., J.-A.Z., A.P. and M.P. contributed to writing the manuscript. H.H. performed meta-analysis, and created figures and tables. H.H. and M.P. performed downstream analyses and drafted the manuscript. M.P. supervised project. All authors interpreted the results and reviewed and commented on the manuscript.

References

1. Classification Committee of The International Headache Society (IHS). The International Classification of Headache Disorders, 3rd edition. *Cephalalgia*. 2018;38(1):1-211.
2. Global burden of 369 diseases and injuries in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. *Lancet*. 2020;396(10258):1204-1222.
3. Steiner TJ, Stovner LJ, Jensen R, Uluduz D, Katsarava Z. Migraine remains second among the world's causes of disability, and first among young women: findings from GBD2019. *J Headache Pain*. 2020;21(1):137.
4. Russell MB, Rasmussen BK, Thorvaldsen P, Olesen J. Prevalence and sex-ratio of the subtypes of migraine. *Int J Epidemiol*. 1995;24(3):612-618.
5. Russell MB, Olesen J. Increased familial risk and evidence of genetic factor in migraine. *BMJ*. 1995;311(7004):541-544.
6. Kallela M, Wessman M, Havanka H, Palotie A, Färkkilä M. Familial migraine with and without aura: clinical characteristics and co-occurrence. *Eur J Neurol*. 2001;8(5):441-449.
7. de Boer I, van den Maagdenberg A, Terwindt GM. Advance in genetics of migraine. *Curr Opin Neurol*. 2019;32(3):413-421.
8. Tfelt-Hansen PC, Koehler PJ. One hundred years of migraine research: major clinical and scientific observations from 1910 to 2010. *Headache*. 2011;51(5):752-778.
9. Anttila V, Wessman M, Kallela M, Palotie A. Genetics of migraine. *Handb Clin Neurol*. 2018;148:493-503.
10. Ferrari MD, Klever RR, Terwindt GM, Ayata C, van den Maagdenberg AM. Migraine pathophysiology: lessons from mouse models and human genetics. *Lancet Neurol*. 2015;14(1):65-80.
11. Ulrich V, Gervil M, Kyvik KO, Olesen J, Russell MB. Evidence of a genetic factor in migraine with aura: a population-based Danish twin study. *Ann Neurol*. 1999;45(2):242-246.
12. Gervil M, Ulrich V, Kaprio J, Olesen J, Russell MB. The relative role of genetic and environmental factors in migraine without aura. *Neurology*. 1999;53(5):995-999.
13. Gormley P, Anttila V, Winsvold BS, et al. Meta-analysis of 375,000 individuals identifies 38 susceptibility loci for migraine. *Nat Genet*. 2016;48(8):856-866.
14. Finucane HK, Reshef YA, Anttila V, et al. Heritability enrichment of specifically expressed genes identifies disease-relevant tissues and cell types. *Nat Genet*. 2018;50(4):621-629.
15. Anttila V, Stefansson H, Kallela M, et al. Genome-wide association study of migraine implicates a common susceptibility variant on 8q22.1. *Nat Genet*. 2010;42(10):869-873.
16. Chasman DI, Schürks M, Anttila V, et al. Genome-wide association study reveals three susceptibility loci for common migraine in the general population. *Nat Genet*. 2011;43(7):695-698.
17. Freilinger T, Anttila V, de Vries B, et al. Genome-wide association analysis identifies susceptibility loci for migraine without aura. *Nat Genet*. 2012;44(7):777-782.
18. Anttila V, Winsvold BS, Gormley P, et al. Genome-wide meta-analysis identifies new susceptibility loci for migraine. *Nat Genet*. 2013;45(8):912-917.
19. Pickrell JK, Berisa T, Liu JZ, et al. Detection and interpretation of shared genetic influences on 42 human traits. *Nat Genet*. 2016;48(7):709-717.
20. Chen SP, Fuh JL, Chung MY, et al. Genome-wide association study identifies novel susceptibility loci for migraine in Han Chinese resided in Taiwan. *Cephalalgia*. 2018;38(3):466-475.
21. Chang X, Pellegrino R, Garifallou J, et al. Common variants at 5q33.1 predispose to migraine in African-American children. *J Med Genet*. 2018;55(12):831-836.
22. Tfelt-Hansen P, Olesen J. Taking the negative view of current migraine treatments: the unmet needs. *CNS Drugs*. 2012;26(5):375-382.

23. Kuca B, Silberstein SD, Wietecha L, et al. Lasmiditan is an effective acute treatment for migraine: A phase 3 randomized study. *Neurology*. 2018;91(24):e2222-e2232.
24. Dodick DW. Migraine. *Lancet*. 2018;391(10127):1315-1330.
25. Lipton RB, Dodick DW, Ailani J, et al. Effect of Ubrogepant vs Placebo on Pain and the Most Bothersome Associated Symptom in the Acute Treatment of Migraine: The ACHIEVE II Randomized Clinical Trial. *JAMA*. 2019;322(19):1887-1898.
26. Charles A, Pozo-Rosich P. Targeting calcitonin gene-related peptide: a new era in migraine therapy. *Lancet*. 2019;394(10210):1765-1774.
27. Goadsby PJ, Dodick DW, Ailani J, et al. Safety, tolerability, and efficacy of orally administered atogepant for the prevention of episodic migraine in adults: a double-blind, randomised phase 2b/3 trial. *Lancet Neurol*. 2020;19(9):727-737.
28. Cohen J, Pertsemlidis A, Kotowski IK, et al. Low LDL cholesterol in individuals of African descent resulting from frequent nonsense mutations in PCSK9. *Nat Genet*. 2005;37(2):161-165.
29. Flannick J, Thorleifsson G, Beer NL, et al. Loss-of-function mutations in SLC30A8 protect against type 2 diabetes. *Nat Genet*. 2014;46(4):357-363.
30. Chang CC, Chow CC, Tellier LC, et al. Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience*. 2015;4:7.
31. Zhou W, Nielsen JB, Fritsche LG, et al. Efficiently controlling for case-control imbalance and sample relatedness in large-scale genetic association studies. *Nat Genet*. 2018;50(9):1335-1341.
32. Tukiainen T, Pirinen M, Sarin AP, et al. Chromosome X-wide association study identifies Loci for fasting insulin and height and evidence for incomplete dosage compensation. *PLoS Genet*. 2014;10(2):e1004127.
33. Mägi R, Morris AP. GWAMA: software for genome-wide association meta-analysis. *BMC Bioinformatics*. 2010;11:288.
34. McLaren W, Gil L, Hunt SE, et al. The Ensembl Variant Effect Predictor. *Genome Biol*. 2016;17(1):122.
35. Benner C, Spencer CC, Havulinna AS, et al. FINEMAP: efficient variable selection using summary data from genome-wide association studies. *Bioinformatics*. 2016;32(10):1493-501.
36. Benner C, Havulinna AS, Järvelin M-R, et al. Prospects of Fine-Mapping Trait-Associated Genomic Regions by Using Summary Statistics from Genome-wide Association Studies. *Am J Hum Genet*. 2017;101(4):539-551.
37. The GTEx Consortium atlas of genetic regulatory effects across human tissues. *Science*. 2020;369(6509):1318-1330.
38. Watanabe K, Taskesen E, van Bochoven A, Posthuma D. Functional mapping and annotation of genetic associations with FUMA. *Nat Commun*. 2017;8(1):1826.
39. Westra H-J, Peters MJ, Esko T, et al. Systematic identification of trans eQTLs as putative drivers of known disease associations. *Nat Genet*. 2013;45(10):1238-1243.
40. Zhernakova DV, Deelen P, Vermaat M, et al. Identification of context-dependent expression quantitative trait loci in whole blood. *Nat Genet*. 2017;49(1):139-145.
41. Ramasamy A, Trabzuni D, Guelfi S, et al. Genetic variability in the regulation of gene expression in ten regions of the human brain. *Nat Neurosci*. 2014;17(10):1418-1428.
42. Grundberg E, Small KS, Hedman ÅK, et al. Mapping cis- and trans-regulatory effects across multiple tissues in twins. *Nat Genet*. 2012;44(10):1084-1089.
43. Ng B, White CC, Klein H-U, et al. An xQTL map integrates the genetic architecture of the human brain's transcriptome and epigenome. *Nat Neurosci*. 2017;20(10):1418-1426.
44. Fromer M, Roussos P, Sieberts SK, et al. Gene expression elucidates functional impact of polygenic risk for schizophrenia. *Nat Neurosci*. 2016;19(11):1442-1453.
45. Võsa U, Claringbould A, Westra HJ, et al. Large-scale cis- and trans-eQTL analyses identify thousands of genetic loci and polygenic scores that regulate blood gene expression. *Nat Genet*. 2021;53(9):1300-1310.

46. Kerimov N, Hayhurst JD, Peikova K, et al. A compendium of uniformly processed human gene expression and splicing quantitative trait loci. *Nat Genet*. 2021;53(9):1290-1299.

47. Schmiedel BJ, Singh D, Madrigal A, et al. Impact of Genetic Polymorphisms on Human Immune Cell Gene Expression. *Cell*. 2018;175(6):1701-1715.e16.

48. van der Wijst MGP, Brugge H, de Vries DH, et al. Single-cell RNA sequencing identifies celltype-specific cis-eQTLs and co-expression QTLs. *Nat Genet*. 2018;50(4):493-497.

49. Wang D, Liu S, Warrell J, et al. Comprehensive functional genomic resource and integrative model for the human brain. *Science (New York, NY)*. 2018;362(6420):eaat8464.

50. Bulik-Sullivan B, Finucane HK, Anttila V, et al. An atlas of genetic correlations across human diseases and traits. *Nat Genet*. 2015;47(11):1236-1241.

51. Bulik-Sullivan BK, Loh PR, Finucane HK, et al. LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat Genet*. 2015;47(3):291-295.

52. Price AL, Weale ME, Patterson N, et al. Long-range LD can confound genome scans in admixed populations. Elsevier; 2008. p. 132-139.

53. Lee SH, Wray NR, Goddard ME, Visscher PM. Estimating missing heritability for disease from genome-wide association studies. *Am J Hum Genet*. 2011;88(3):294-305.

54. Finucane HK, Bulik-Sullivan B, Gusev A, et al. Partitioning heritability by functional annotation using genome-wide association summary statistics. *Nat Genet*. 2015;47(11):1228-1235.

55. Gazal S, Finucane HK, Furlotte NA, et al. Linkage disequilibrium-dependent architecture of human complex traits shows action of negative selection. *Nat Genet*. 2017;49(10):1421-1427.

56. Hujoel MLA, Gazal S, Hormozdiari F, van de Geijn B, Price AL. Disease Heritability Enrichment of Regulatory Elements Is Concentrated in Elements with Ancient Sequence Age and Conserved Function across Species. *Am J Hum Genet*. 2019;104(4):611-624.

57. Trochet H, Pirinen M, Band G, et al. Bayesian meta-analysis across genome-wide association studies of diverse phenotypes. *Genet Epidemiol*. 2019;43(5):532-547.

58. Cichonska A, Rousu J, Marttinen P, et al. metaCCA: summary statistics-based multivariate meta-analysis of genome-wide association studies using canonical correlation analysis. *Bioinformatics*. 2016;32(13):1981-1989.

59. Nelson CP, Goel A, Butterworth AS, et al. Association analyses based on false discovery rate implicate new loci for coronary artery disease. *Nat Genet*. 2017;49(9):1385-1391.

60. Loh P-R, Kichaev G, Gazal S, Schoech AP, Price AL. Mixed-model association for biobank-scale datasets. *Nat Genet*. 2018;50(7):906-908.

61. de Leeuw CA, Mooij JM, Heskes T, Posthuma D. MAGMA: generalized gene-set analysis of GWAS data. *PLoS Comput Biol*. 2015;11(4):e1004219.

62. Liberzon A, Subramanian A, Pinchback R, et al. Molecular signatures database (MSigDB) 3.0. *Bioinformatics*. 2011;27(12):1739-1740.

63. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences*. 2005;102(43):15545-15550.

64. Pers TH, Karjalainen JM, Chan Y, et al. Biological interpretation of genome-wide association studies using predicted gene functions. *Nature Communications*. 2015;6 :5890

65. Barbeira AN, Dickinson SP, Bonazzola R, et al. Exploring the phenotypic consequences of tissue specific gene expression variation inferred from GWAS summary statistics. *Nat Commun*. 2018;9(1):1825.

66. Giambartolomei C, Vukcevic D, Schadt EE, et al. Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. *PLoS Genet*. 2014;10(5):e1004383.

67. Mancuso N, Freund MK, Johnson R, et al. Probabilistic fine-mapping of transcriptome-wide association studies. *Nat Genet*. 2019;51(4):675-682.

68. Hinrichs AS, Karolchik D, Baertsch R, et al. The UCSC Genome Browser Database: update 2006. *Nucleic Acids Res*. 2006;34:D590-8.

69. Nagel M, Jansen PR, Stringer S, et al. Meta-analysis of genome-wide association studies for neuroticism in 449,484 individuals identifies novel genetic loci and pathways. Nat Genet. 2018;50(7):920-927.

70. Pardiñas AF, Holmans P, Pocklington AJ, et al. Common schizophrenia alleles are enriched in mutation-intolerant genes and in regions under strong background selection. *Nat Genet*. 2018;50(3):381-389.

71. Howard DM, Adams MJ, Clarke T-K, et al. Genome-wide meta-analysis of depression identifies 102 independent variants and highlights the importance of the prefrontal brain regions. *Nat Neurosci*. 2019;22(3):343-352.

72. Yang J, Weedon MN, Purcell S, et al. Genomic inflation factors under polygenic inheritance. *Eur J Hum Genet*. 2011;19(7):807-812.

73. Do TP, Guo S, Ashina M. Therapeutic novelties in migraine: new drugs, new hope? *J Headache Pain*. 2019;20(1):37.

74. Kurth T, Winter AC, Eliassen AH, et al. Migraine and risk of cardiovascular disease in women: prospective cohort study. *BMJ*. 2016;353:i2610.

75. Hippisley-Cox J, Coupland C, Brindle P. Development and validation of QRISK3 risk prediction algorithms to estimate future risk of cardiovascular disease: prospective cohort study. *BMJ*. 2017;357:j2099.

76. Adelborg K, Szépligeti SK, Holland-Bill L, et al. Migraine and risk of cardiovascular diseases: Danish population based matched cohort study. *BMJ*. 2018;360:k96.

77. Siewert KM, Klarin D, Damrauer SM, et al. Cross-trait analyses with migraine reveal widespread pleiotropy and suggest a vascular component to migraine headache. *Int J Epidemiol*. 2020;49(3):1022-1031.

78. Guo Y, Rist PM, Daghlas I, et al. A genome-wide cross-phenotype meta-analysis of the association of blood pressure with migraine. *Nat Commun*. 2020;11(1):3368.

79. Need AC, Goldstein DB. Schizophrenia Genetics Comes of Age. *Neuron*. 2014;83(4):760-763.

80. Gormley P, Kurki M, Kurki MI, et al. Common Variant Burden Contributes to the Familial Aggregation of Migraine in 1,589 Families. *Neuron*. 2018;98(4):743-753.e4.

81. Chasman DI, Anttila V, Buring JE, et al. Selectivity in genetic association with sub-classified migraine in women. *PLoS Genet*. 2014;10(5):e1004366.

82. Nyholt DR, Anttila V, Winsvold BS, et al. Concordance of genetic risk across migraine subgroups: Impact on current and future genetic association studies. *Cephalalgia*. 2015;35(6):489-99.

83. Ophoff RA, Terwindt GM, Vergouwe MN, et al. Familial hemiplegic migraine and episodic ataxia type-2 are caused by mutations in the Ca2+ channel gene CACNL1A4. *Cell*. 1996;87(3):543-552.

84. de Vries B, Frants RR, Ferrari MD, van den Maagdenberg AMJM. Molecular genetics of migraine. *Hum Genet*. 2009;126(1):115-132.

85. Zhao H, Eising E, de Vries B, et al. Gene-based pleiotropy across migraine with aura and migraine without aura patient groups. *Cephalalgia*. 2016;36(7):648-57.

86. Jacobs B, Dussor G. Neurovascular contributions to migraine: Moving beyond vasodilation. *Neuroscience*. 2016;338:130-144.

87. Hoffmann J, Baca SM, Akerman S. Neurovascular mechanisms of migraine and cluster headache. *J Cereb Blood Flow Metab*. 2019;39(4):573-594.

88. Bigal ME, Kurth T, Hu H, Santanello N, Lipton RB. Migraine and cardiovascular disease: possible mechanisms of interaction. *Neurology*. 2009;72(21):1864-1871.

89. Malik R, Freilinger T, Winsvold BS, et al. Shared genetic basis for migraine and ischemic stroke: A genome-wide analysis of common variants. *Neurology*. 2015;84(21):2132-2145.

90. Winsvold BS, Nelson CP, Malik R, et al. Genetic analysis for a shared biological basis between migraine and coronary artery disease. *NeurologyGenetics*. 2015;1(1):e10.

91. Mahmoud AN, Mentias A, Elgendy AY, et al. Migraine and the risk of cardiovascular and cerebrovascular events: a meta-analysis of 16 cohort studies including 1 152 407 subjects. *BMJ open*. 2018;8(3):e020498.
92. Daghlas I, Guo Y, Chasman DI. Effect of genetic liability to migraine on coronary artery disease and atrial fibrillation: a Mendelian randomization study. *Eur J Neurol*. 2020;27(3):550-556.
93. Charles A. The pathophysiology of migraine: implications for clinical management. *The Lancet Neurology*. 2018;17(2):174-182.
94. Burstein R, Noseda R, Borsook D. Migraine: multiple processes, complex pathophysiology. *The Journal of neuroscience : the official journal of the Society for Neuroscience*. 2015;35(17):6619-6629.
95. Andreou AP, Edvinsson L. Mechanisms of migraine as a chronic evolutive condition. *J Headache Pain*. 2019;20(1):117.
96. Olesen J, Burstein R, Ashina M, Tfelt-Hansen P. Origin of pain in migraine: evidence for peripheral sensitisation. *The Lancet Neurology*. 2009;8(7):679-690.
97. Brandes JL. The influence of estrogen on migraine: a systematic review. *JAMA*. 2006;295(15):1824-1830.
98. Borsook D, Erpelding N, Lebel A, et al. Sex and the migraine brain. *Neurobiol Dis*. 2014;68:200-14.
99. Delaruelle Z, Ivanova TA, Khan S, et al. Male and female sex hormones in primary headaches. *J Headache Pain*. 2018;19(1):117.
100. Diener H-C. CGRP antibodies for migraine prevention — new kids on the block. *Nature Reviews Neurology*. 2019;15(3):129-130.
101. de Vries T, Villalón CM, Maassen VanDenBrink A. Pharmacological treatment of migraine: CGRP and 5-HT beyond the triptans. *Pharmacol Ther*. 2020:107528.
102. Nelson MR, Tipney H, Painter JL, et al. The support of human genetic evidence for approved drug indications. *Nat Genet*. 2015;47(8):856-860.
103. King EA, Davis JW, Degner JF. Are drug targets with genetic support twice as likely to be approved? Revised estimates of the impact of genetic support for drug mechanisms on the probability of drug approval. *PLoS Genet*. 2019;15(12):e1008489.
104. Wainberg M, Sinnott-Armstrong N, Mancuso N, et al. Opportunities and challenges for transcriptome-wide association studies. *Nat Genet*. 2019;51(4):592-599.

Whole exome sequencing of hemiplegic migraine patients shows an increased burden of missense variants in *CACNA1H* and *CACNA1I* genes

Neven Maksemous*, Aster V.E. Harder*, Omar Ibrahim*, Lisanne S. Vijfhuizen, Heidi Sutherland, Nadine Pelzer, Irene de Boer, Gisela M. Terwindt, Rodney A. Lea, Arn M.J.M. van den Maagdenberg* and Lyn R. Griffiths*

Abstract

Hemiplegic migraine (HM) is a rare subtype of migraine with aura. Given that causal missense mutations in the voltage-gated calcium channel α1A subunit gene *CACNA1A* have been identified in a subset of HM patients, we investigated whether HM patients without a mutation have an increased burden of such variants in the "*CACNA1x* gene family". Whole exome sequencing data of an Australian cohort of unrelated HM patients (n = 184), along with public data from gnomAD, as controls, was used to assess the burden of missense variants in *CACNA1x* genes. We performed both a variant and a subject burden test. We found a significant burden for the number of variants in *CACNA1E* ($p = 1.3 \times 10^{-4}$), *CACNA1H* ($p < 2.2 \times 10^{-16}$) and *CACNA1I* ($p < 2.2 \times 10^{-16}$). There was also a significant burden of subjects with missense variants in *CACNA1E* ($p = 6.2 \times 10^{-3}$), *CACNA1H* ($p < 2.2 \times 10^{-16}$) and *CACNA1I* ($p < 2.2 \times 10^{-16}$). Both the number of variants and number of subjects were replicated for *CACNA1H* ($p = 3.5 \times 10^{-8}$; $p = 0.012$) and *CACNA1I* ($p = 0.019$, $p = 0.044$), respectively in a Dutch clinical HM cohort (n = 32), albeit that *CACNA1I* did not remain significant after multiple testing correction. Our data suggest that HM, in the absence of a single causal mutation, is a complex trait, in which an increased burden of missense variants in *CACNA1H* and *CACNA1I* may contribute to the risk of disease.

KEYWORDS: Hemiplegic Migraine, *CACNA1x*, headache, burden-testing

Introduction

Hemiplegic migraine (HM) is a rare subtype of migraine with aura with attacks that are associated with motor weakness or hemiplegia during the aura phase.[1] HM is clinically and genetically heterogeneous[2-4] and can be subdivided in familial hemiplegic migraine (FHM) and sporadic hemiplegic migraine (SHM), distinguished by having a positive or negative family history for HM, respectively.[1]

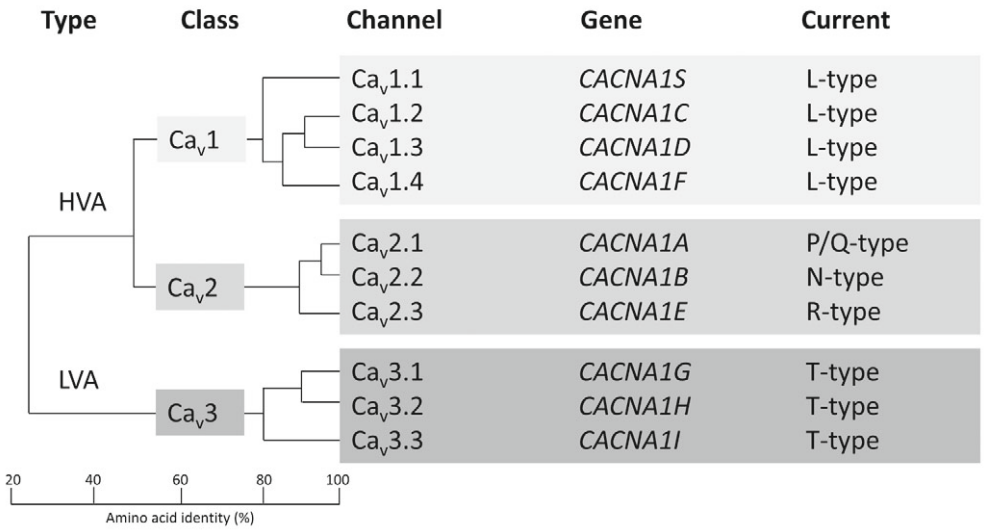
A subset of HM patients exhibits an autosomal dominant phenotype with single high-penetrant causal mutations present in ion transport genes *CACNA1A*, *ATP1A2* or *SCN1A*.[5-7] However, in many HM patients no such pathogenic mutation has been detected.[8, 9] Whereas evidence is accumulating that loss-of-function mutations in *PRRT2*,[10] a key component of the Ca^{2+}-dependent neurotransmitter release machinery, [11] are involved in HM, the gene more likely acts as a modifier of disease.[12] This suggests that HM, in a set of patients, may be regarded a complex disorder with multiple genetic factors contributing to the phenotype. Most relevant, a Finnish polygenic risk score study of genome-wide association study (GWAS) data has shown that HM patients without a high-penetrant disease-causing mutation in a known HM gene carry an excess of common (frequency > 1%) variants compared to patients suffering from common (complex) migraine subtypes.[13]

Following along this line of evidence, it has been hypothesized that complex disorders can be the result of an accumulation of genetic variants in a disease pathway, where the crossing of a certain threshold leads to disease.[14] Moreover, current evidence indicates that complex traits are likely to be underpinned by a combination of multiple common and rare variants.[15-17] Here we set out to investigate the contribution of modulatory genetic effects that can be studied through testing the synergistic burden of (functional) variants, best annotated as missense variants, rather than a single causative mutation. Burden can be regarded as an accumulation of variants that are more often present in cases compared to controls. We hypothesise here that the burden of missense variants in certain ion channel genes might be involved in the disease pathology of HM.

CACNA1A was the first HM gene discovered and encodes the pore-forming α_{1A} subunit of the neuronal voltage-gated calcium channel (VGCC) $Ca_V2.1$ (P/Q-type).[5, 18] $Ca_V2.1$ channels are predominantly localized at presynaptic terminals and play a prominent role in controlling neurotransmitter release at most synapses of the nervous system.[19-21] *CACNA1A* is a member of a family of rather conserved α1 subunit genes, hereafter referred to as "*CACNA1x*", which are part of VGCCs that are classified as either high-voltage-activated (HVA) or low-voltage-activated (LVA) channels that are present on the membranes of excitable cells (**Figure 1**).[22, 23] Ca_V channels are typically composed of multiple subunits namely an α_1, a β, an α_2/δ, and a γ subunit. An α_1 subunit has 24 transmembrane segments and forms the pore through which calcium ions pass into the cell. The main characteristic of the various Ca_V channel types is primarily determined by the type of α_1 subunit, so the presence of either α_{1A}, α_{1B}, α_{1C}, α_{1D}, α_{1E}, α_{1F}, α_{1G}, α_{1H}, α_{1I} or α_{1S}. Given the

important functions of Ca_V channels, it is not surprising that genetic variation in *CACNA1x* genes is not well-tolerated; the residual variation intolerance scores for these genes are high (**Table 1**).[24]

Figure 1 The voltage-gated calcium channel (VGCC) family of proteins



The α_1 subunits can be divided into three subclasses according to their amino acid sequence identity, as shown in the dendrogram. Ca_V1 and Ca_V2 channels are high-voltage-activated (HVA), whereas Ca_V3 channels are low-voltage-activated (LVA). The genes encoding the respective α_1 subunit are provided as well as the type of current the respective channel type produces. The Schematic is based upon Perez-Reyes and Dolphin.[22,23]

Table 1 Intolerance scores *CACNA1x* genes.

Gene	RVIS	%ExAC RVIS	%ExAC Lof FDR	%ExAC v2 RVIS	Edge case (%OE-ratio)
CACNA1A	-1.78 (2.27%)	1.68%	1.35×10^{-10}	-2.20 (2.45%)	N (4.41%)
CACNA1B	NA (NA)	0.79%	1.48×10^{-5}	-2.34 (2.17%)	N (10.98%)
CACNA1C	-2.09 (1.57%)	3.51%	9.23×10^{-8}	-1.53 (5.77%)	N (4.20%)
CACNA1D	-3.51 (0.32%)	0.39%	1.64×10^{-13}	-3.37 (0.72%)	N (11.785)
CACNA1E	-2.71 (0.71%)	3.34%	8.29×10^{-18}	-1.88 (3.59%)	N (5.73%)
CACNA1F	-0.83 (11.53%)	12.28%	1.03×10^{-4}	0.02 (52.82%)	N (45.93%)
CACNA1G	-2.37 (1.11%)	1%	2.45×10^{-8}	-2.08 (2.85%0	N (13.56%)
CACNA1H	-2.06 (1.63%)	25.99%	1.79×10^{-9}	-0.03 (48.18%)	N (13.68%)
CACNA1I	-0.83 (11.55%)	3.96%	4.29×10^{-6}	-1.90 (3.52%)	N (7.79%)
CACNA1S	-1.21 (5.71%)	61.36%	8.21×10^{-5}	0.52 (72.65%)	N (41.68%)

RVIS = Residual variation intolerance score (is a gene score based module intended to help in the interpretation of human sequence data); %ExAC RVIS = RVIS v3, constructed on the ExAC data release; ExAC LoF FDR = FDR-adjusted *p*-value reflects the significance of the departure from the expected rate of LoF variants; %ExAC v2 RVIS = RVIS v4, constructed on the ExAC v2 data release; Edge Case = Edge case genes. The OE-ratio is based on ExaC v2 data release.

The expression of *CACNA1x* genes varies considerably and, with the exception of *CACNA1S*, all are expressed in the brain.[25] In addition to *CACNA1A* being a well-known HM gene, there have been rare reports on other *CACNA1x* genes possibly involved in HM-relevant phenotypes. For instance, a link between hemiplegic migraine and brain stem aura migraine has been suggested for *CACNA1E* [26], and headache with neurological deficits and cerebrospinal fluid lymphocytosis (HaNDL), a headache syndrome with much phenotypic resemblance to HM, has been linked to the occurrence of antibodies against *CACNA1H*,[27] a gene implicated in childhood epilepsy although this has recently been debated.[28] Furthermore, using a systems genetics approach Rasmussen et al.[29] identified *CACNA1B* as one of the genes commonly mutated in migraine families. Finally, *CACNA1A* was identified as a risk locus for common migraine, as well as being one of the three genes specific for migraine with aura.[30] Although no definite proof for a causal link was provided in any of these cases, the existing data can be regarded as supportive evidence for a spectrum ranging from rare to common variants contributing to certain extent to the risk for both common and hemiplegic migraine. This variety of observed variants makes the family of *CACNA1x* genes an interesting candidate for burden testing in HM, with relevance, foremost, to patients with a complex genetic basis.

Whole exome sequencing (WES) enables comprehensive exploration of missense variants and investigation of their role in complex traits. When considering that these missense variants are unlikely to be causing HM as monogenic factors as occurs for patients with specific *CACNA1A*, *ATP1A2* and *SCN1A* mutations,[5-7] burden testing is a potential way to explore their potential synergistic effect on increasing HM disease risk. Burden testing typically requires a set of qualifying variants, often rare, protein-altering variants in the case of a monogenic condition. However, following on from the hypothesis that HM seems not a monogenic disorder in *all* patients, the accumulation of both rare and common protein-altering variants may be relevant in terms of disease susceptibility. The use of large publicly available WES datasets from general population controls can be incorporated in burden testing to gain more reliable estimates of gene-wide susceptibility.

We hypothesize that the burden of multiple missense variants in *CACNA1x* genes increases the risk for HM. Burden being the aggregation of both rare and common variants as well as the increased presence of a variant in cases compared to controls. To this end, we here used WES data from a large Australian HM patient cohort to identify missense variants in eight calcium channel genes (*CACNA1A, -B, -C, -D, -E, -G, -H,* and *-I)* and determined whether the aggregated effect of the variants across the genes was higher than observed in general population controls. Results were validated in an independent Dutch clinical HM cohort.

Methods

Study cohorts

The study consisted of two cohorts of HM patients: an Australian cohort of 184 patients (discovery cohort) and a Dutch cohort of 32 patients (replication cohort). Importantly, patients were *a priori*

excluded in case a pathogenic mutation was present in one of the three HM genes (*CACNA1A*, *ATP1A2* and *SCN1A*) or HM-related genes with mutations confirmed by Sanger sequencing.[9, 31]

Australian cohort. The Australian cohort was selected out of over 300 patients that had been referred to the Genomics Research Centre (GRC) Diagnostic Clinic for genetic diagnostic testing after a suspected diagnosis of HM from the referring neurologist. From this cohort, a subset of 184 (122 females and 62 males) unrelated individuals tested negative for known HM gene mutations (*CACNA1A*, *ATP1A2*, and *SCN1A*) and HM-related genes.[9, 31] All cases consented to genetic testing with their doctors, as required under current regulations. Positive family history was reported for 25% of the cases, 5% were reported as SHM, while family information was not available for the remainder of cases.

DNA was extracted from blood samples using QIAGEN QIAamp DNA Mini Kit as per the manufacturers' instructions. Next generation sequencing (NGS) libraries for WES were constructed using the Ion AmpliSeq™ Exome RDY library kits (ThermoFisher Scientific, Waltham, MA, USA) according to the manufacturer's protocol. The Ion Chef was used to load sample libraries (barcoded fragments of 200 bp). WES was performed in the Genomics Research Centre (GRC), Australia via the Ion Proton and GeneStudio S5 plus (ThermoFisher Scientific) instruments using default settings for Ion AmpliSeq Exome RDY Kit 4x2 (ThermoFisher Scientific). The study was conducted in accordance with the Declaration of Helsinki, and the protocol was approved by the Human Research Ethics Committee of the Queensland University of Technology (approval number: 1800000611).

Dutch cohort. The cohort consisted of 32 patients (22 females and 10 males) with FHM/SHM according to ICHD-3 criteria.[32] Patients were selected from the Leiden Headache Centre at the Leiden University Medical Centre (LUMC), and contained patients (i) seen in person by experienced headache clinicians or research physicians or (ii) referred from elsewhere for clinical genetic research with records being evaluated and clinical diagnosis confirmed by GMT, NP and IdB.[4] All patients were from different families and did not have a known pathogenic mutation in one of the three HM genes. The study was approved by the Medical Ethics Committee of LUMC and all participants provided informed consent.

Genomic DNA was extracted from peripheral blood leukocytes according to standard salting out protocol.[33] WES was performed using in-house sequencing facility (Leiden Genome Technology Centre; URL: lgtc.nl) or outsourced to the Beijing Genomics Institute sequencing facility (URL: bgi.com). In brief, for the LGCT, coding sequences in the DNA were enriched using the SureSelect Human All Exon 50 Mb kit (Agilent Technologies, Santa Clara, CA, USA). Following sequence capture and amplification, fragments were sequenced using the Illumina HiSeq2000 platform (San Diego, CA, USA).

Controls. As a control dataset, we used summary statistics from gnomAD (see below in the paragraph on TRAPD methods). The gnomAD database was chosen as it consists of a large

number of individuals and contains a detailed catalogue of exome-wide genetic variation. Furthermore, gnomAD provided ancestry information. The gnomAD database contains exome variant summary statistics for 56,885 non-Finnish Europeans, with a female-to-male ratio of ~1.27:1, depending on the available genotypes at each specific locus. As HM is a very rare disorder with a prevalence of 0.01%,[34] confounding effects due to the presence of HM patients in the control group were deemed to be negligible.

Whole exome sequencing and quality control

Australian cohort. Following WES, the Ion Torrent Server was used to generate quality metrics, align reads to the Human Genome 19 (Hg19), and the Ion Torrent Variant Caller (TVC) was used to call sequence variants and produce variant calling format (VCF) files.

Dutch cohort. Following sequencing, the sequence reads were aligned to the UCSC Genome Browser hg19 reference sequence using the Burrows-Wheeler Alignment tool.[35] The generated BAM files were subsequently converted to VCF files using BCFtools.[36]

Single-variant analysis

Prior to performing burden testing all variants were assessed to determine whether there were obvious, high-penetrant disease-causing mutations detected outside of the known HM genes that could cause HM in patients of either cohort. In the absence of such pathogenic mutation, individual missense variants in all the *CACNA1x* genes were assessed for patients of the Australian cohort. For the Dutch cohort, only those missense variants present in TRAPD-associated *CACNA1x* genes were investigated.

Burden testing

Variants pre-processing all cohorts. As VCFs were exported from different platforms, the respective analyses had to be unified. The first step for both cohorts was to normalize VCFs using BCFtools, this ensures that any platform-specific formatting differences are removed, and also expands multi-allelic variants.[36] VCFs were merged for each cohort using vcftools, and variants with average read depth coverage below 10X were excluded using either BCFtools or the snpEff program.[36, 37] For both cohorts, the coding exons of the *CACNA1x* genes were included with a 5-bp pad on either site of the exon. New VCFs (one merged for each cohort) were annotated with VEP Ensembl.[38] For the Dutch cohort as an extra quality control step only those variants with a quality-by-depth (QD) score > 4 were taken forward.

Selection of qualifying variants. To determine the number of variants, we selected "qualifying variants" being variants that meet the criteria of inclusion. Only those variants classified (annotated) as missense variants were considered as "qualifying variants". The number of individuals in the case cohort who carried at least one "qualifying variant" in that gene as well as the total number of

variants were used in the analysis. For the gnomAD control dataset, only summary statistics were available. Therefore, to approximate the number of control subjects carrying at least one qualifying variant in a given gene, the allele counts for all qualifying variants in that gene were summed. This summation based approximation probably is an overestimation as it is likely that some individuals carry multiple variants in the same gene. Contrary to rare variant analysis where only the locations of the qualifying variant in cases are used for controls, we selected all variants across the entire gene in controls, in the same way as what was done for the cases. As a result, we had a total number of all missense variants per *CACNA1x* gene in both cases and controls. These "qualifying variants" for both the case and the control cohort were compared. Insertions and deletions (Indels) were not included in the analysis due to their higher percentage of sequencing artefacts, especially given the differing sequencing platforms used across cohorts.

Multiple-variant burden testing of CACNA1x genes. Gene-based burden testing was performed for all variants that met the quality filters, which are referred to as "qualifying variants", using adaptation on the TRAPD test (Testing Rare vAriants using Public Data).[39] TRAPD was chosen because the control dataset consisted of summary data rather than individual-level genotype data as well as for its approach to collate variants which mitigates the statistical drawbacks of burden testing per variant or per individual. The TRAPD test was implemented to determine whether *CACNA1x* genes and subjects carried a significant burden of missense variants in cases. TRAPD produces counts of "collapsed" variant groups across each gene and for the respective case or control cohort. To conduct the test, a group file with the qualifying variants was created for each of the eight genes (*CACNA1A*, *CACNA1B*, *CACNA1C*, *CACNA1D*, *CACNA1E*, *CACNA1G*, *CACNA1H* and *CACNA1I)*. Of note, *CACNA1S* was excluded from the analysis as it encodes the pore-forming $Ca_V1.1$ α_{1S} subunit that is exclusively expressed in skeletal muscle, so not in the brain and *CACNA1F* was excluded as this gene is located on the X-chromosome and TRAPD is currently not configured to test non-autosomal chromosomes.

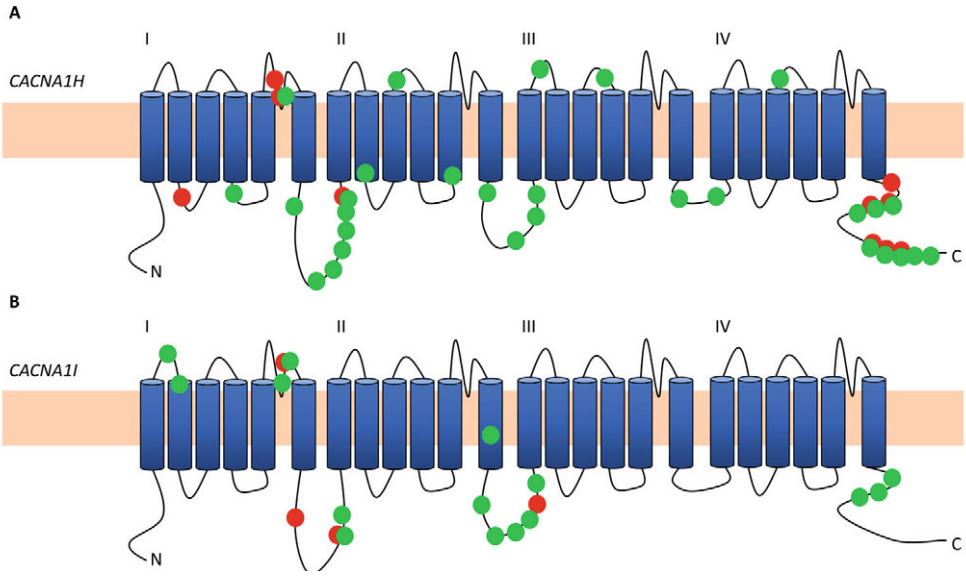
We performed gene-based burden testing for all single-point variants in each cohort. The following steps, in brief, were performed: (1) variants for each *CACNA1x* gene in the case group were compiled into a "SNP file", 2) a python script was used to interrogate the VCFs and count the occurrence of variants in each gene in both the case and the controls cohorts independently. This generated variant count data for each gene and (3) the one sided Fisher's exact test was used on the allele count tables to identify the probability of excess in the number of allele counts in cases relative to controls (i.e. the statistical significance of the burden). (4) the one sided Fisher's exact test was used on the subject count tables to identify the probability of excess in the number of subjects with variants in cases relative to controls (i.e. the statistical significance of the burden). *P*-values < 6.25×10^{-3} were considered significant (Bonferroni corrected for testing 8 genes). Odds ratios were calculated to assess the magnitude of the burden effect. Genes exhibiting statistically significant burden in HM from the Australian discovery cohort were also tested in the Dutch replication cohort.

Results

Single-variant analysis

No clear pathogenic mutations in *CACNA1x* genes were identified from the WES data in patients from either the Australian or the Dutch cohorts. However, the number of variants in *CACNA1x* genes prompted us to perform burden testing. In the Australian cohort we identified 79 different missense variants in the eight *CACNA1x* genes examined in the 184 HM patient group from Australia (**Supplementary Table 1**). All but seven of the variants had been previously identified (i.e. they have an rs number in dbSNP). The seven novel variants were all single-case across multiple different *CACNA1x* genes. In the Dutch cohort four different variants were identified in *CACNA1I* and ten in *CACNA1H*, all of which had been previously identified (**Supplementary Table 2**). Although some missense variants in *CACNA1x* genes were predicted to have a pathogenic potential there was not enough evidence for causality in a monogenic manner such as has been shown for the three well known HM genes. The results of the individual variant analyses indicate the existence of many variants across *CACNA1x* genes that in combination could plausibly confer increased susceptibility to HM, especially when considered collectively using burden analysis.

Figure 2 Schematic representation of the α1 subunit



Schematic representation of the α1 subunit, with the position of identified variants in this study, of the $Ca_V3.2$ and the $Ca_V3.3$ channel, encoded by the *CACNA1H* (**A**) and *CACNA1I* (**B**), respectively. Variants identified in the Australian HM cohort are depicted with a green dot and variants identified in the Dutch HM cohort are depicted with a red dot.

Multiple-variant burden testing of *CACNA1x* genes

Considering the results from the individual variant analysis, a multiple-variant burden analysis was performed to test whether there is an over-representation of *CACNA1x* missense variants in HM compared to controls. As shown in **Table 2**, in the Australian HM cohort this analysis revealed a significantly increased burden of missense variants for those with HM in *CACNA1E* ($p = 1.3 \times 10^{-4}$), *CACNA1H* ($p < 2.2 \times 10^{-16}$) and *CACNA1I* ($p < 2.2 \times 10^{-16}$). For the Dutch cohort this replicated for *CACNA1H* ($p = 3.5 \times 10^{-8}$, $p_{cor} = 1.04 \times 10^{-7}$) and C*ACNA1I* ($p = 0.019$, $p_{cor} = 0.056$), but not *CACNA1E* ($p = 0.85$), albeit that *CACNA1I* did not remain significant after correction for multiple testing. In addition, the number of subjects carrying a variant was also higher cases in *CACNA1E* ($p = 6.2 \times 10^{-3}$), *CACNA1H* ($p < 2.2 \times 10^{-16}$) and *CACNA1I* ($p < 2.2 \times 10^{-16}$) (**Table 2**). The results showed evidence of replication in the Dutch cohort for *CACNA1H* ($p = 1.2 \times 10^{-2}$; $p_{cor} = 3.6 \times 10^{-2}$) and *CACNA1I* ($p = 4.4 \times 10^{-2}$; $p_{cor} = 0.13$), but not for *CACNA1E* ($p = 0.88$), albeit *CACNA1I* did not remain significant after correction for multiple testing. All but four variants were outside the transmembrane domains that are typically affected in case of a pathogenic mutation in either Ca_V channel (**Figure 2**)

Table 2 Missense variant burden in the Discovery cohort consisting of Australian cases and GnomAD controls**.**

Gene	Gene length (bp)	Case count HET	Case count HOM	Control count	Case allele count	Control allele count	*P*-value Subjects[a]	*P*-value Allele Counts[a]
CACNA1A	8392	54	1	23977	60	26841	1	1
CACNA1B	9790	6	0	1815	7	1819	0.54	0.39
CACNA1C	8425	4	0	1221	4	1225	0.56	0.57
CACNA1D	7636	4	0	2014	4	2024	0.89	0.90
CACNA1E	7067	109	41	29644	228	53792	6.21×10^{-3}*	1.30×10^{-4}*
CACNA1G	7648	12	0	1768	12	1770	1.33×10^{-2}	1.65×10^{-2}
CACNA1H	8084	155	88	20456	637	22942	$<2.2 \times 10^{-16}$*	$< 2.2 \times 10^{-16}$*
CACNA1I	10004	126	44	4652	277	4694	$<2.2 \times 10^{-16}$*	$< 2.2 \times 10^{-16}$*

Case count = number of cases with at least one variant (Het = heterozygous variant Hom = homozygous variant), Control count = number of controls derived from gnomAD with a variant, Case allele count = Total allele count in cases, Control allele count = Total allele count in controls derived from gnomAD. * Significant results; [a] p-values < 6.25×10^{-3} were considered significant (Bonferroni corrected for testing 8 genes)

Discussion

Here we used WES data from 184 suspected HM patients from an Australian clinically-referred cohort, and compared these to the publicly available gnomAD control dataset using TRAPD, finding that *CACNA1E*, *CACNA1I* and *CACNA1H* missense variants were more prevalent in cases. Furthermore, we show evidence for replication of these findings for *CACNA1H* and *CACNA1I* in a Dutch clinical HM cohort. This finding emphasizes that although the cohorts differ in terms of inclusion criteria, the results are transferable to both groups.

In the general population, females are overrepresented in most forms of migraine including hemiplegic migraine. The overall female:male sex ratio in our HM cohort was ~1.97:1. This observed difference in prevalence will in part be explained by the fact that females are more inclined to consult a physician and thus are diagnosed earlier and more often than males.[40] We cannot rule out that there is any sexual dimorphic effect at any of the *CACNA1x* genes (i.e. a sex bias in gene function), but we consider this a minor factor compared to the ascertainment bias.

We have hypothesised that HM may not be autosomal dominant in a substantial proportion of cases, but rather is genetically a more complex trait. The difficulty in confirming this hypothesis lies in how to identify such variants, as they are neither identified by gene association approaches, nor in genome-wide association studies (GWAS). In order to identify such variants, we have used a methodology adapted from a TRAPD analysis. The methodology has proven itself by the identification of functional genetic variants in idiopathic hypogonadotropic hypogonadism.[39] By slightly adapting the method, we have been able to investigate all missense variance and thereby determine the variant and subject burden. Similarly, our results show that the accumulation of missense variants in *CACNA1H* and *CACNA1I* plays a role in HM.

CACNA1H and *CACNA1I* encode the α1 subunits of $Ca_V3.2$ and $Ca_V3.3$ LVA T-type calcium channels, respectively (**Figure 1**).[22, 23] *CACNA1H* is expressed ubiquitously, whereas *CACNA1I* is predominantly expressed in the brain.[25] $Ca_V3.2$ and $Ca_V3.3$ LVA T-type calcium channels,[41] which open by only a small membrane depolarization, coupled with their tonic inactivation near resting membrane potential, underlie the spike/rebound bursting phenomenon seen with many types of neurons [42, 43]. These channels are localized at presynaptic nerve terminals[44] where they control synaptic transmission by directly triggering the release of neurotransmitters.[45-47] Inactivation of *Cacna1h* in mice led to decreased nociceptive signalling[48, 49] and several neurological symptoms,[50, 51] whereas *Cacna1i* knock-out mice, and also *Cacna1i*/*Cacna1h* double knockout mutants, show implications for sleep rhythmogenesis.[52] T-type channels are important for human physiology, so mutations in these channels may lead, at least in theory, to channelopathies with clinical manifestations resulting from aberrant biophysical characteristics and/or cell surface trafficking issues of channels due to a gain or loss of channel function. Indeed, specific missense variants in *CACNA1H* have been implicated in a range of human conditions,[50] including autism spectrum disorders[53] and amyotrophic lateral sclerosis.[54] Many missense variants in the human *CACNA1H* gene have been reported in patients presenting with a range of epilepsy syndromes,[50] so the gene was labelled a risk gene for idiopathic generalised epilepsies.[55] Functional analyses in embryonic kidney cells, however, revealed that the variants in *CACNA1H* generally produce mild biophysical changes and in some cases do not alter the gating of the channel and variants do not segregate with the phenotype.[50] Hence their contribution to human epilepsies should be debated, as was recently suggested.[28] In line with this suggestion, it is not unexpected that *CACNA1H* variants identified in HM patients also not solely cause disease, although a burden of variants in this gene can still contribute to HM risk. Similarly, *CACNA1I* loss-of-function variants were identified that alter the

gating properties of $Ca_V3.3$ channels, disrupt neuronal excitability and network activity, and have been associated with risk of developing schizophrenia and a range of neurodevelopmental disorders featuring developmental delay and epilepsy.[56, 57] Moreover, using patch-clamp electrophysiology, we have shown various functional alterations of channel activity for selected $Ca_v3.3$ rare variants, providing further evidence that *CACNA1I* may play a role in the development of HM.[58]

Hence, the most likely scenario is that an increased burden of missense variants in *CACNA1H* and *CACNA1I* acts as a genetic modifier of disease risk. Such a modification of risk is not different when reviewing mutations that have been identified in some HM patients in a number of genes, including *PRRT2*,[12] *PNKD*,[59] *SLC4A4*,[60] *SLC1A3*,[61] and *SLC2A1*,[62] that are primarily associated with movement or solute transport disorders.

Our study has some limitations. First of all, contrary to what is commonly undertaken in genetics, we considered both rare and common variants as an overarching burden of missense variants in this study. To support the validity of this approach we used the Dutch replication cohort to validate findings from the Australian cohort. Further replication efforts in other independent cohorts would be of benefit in future studies of these genes. Secondly, as this study is the first of its kind, we narrowed the genes targeted to *CACNA1x* ion channels, due to known association of genes of this family with HM. However, the burden of variants in additional genes is likely to play a role in determining HM disease risk. Thirdly, we have used summary statistics for the controls that prevented us to compare ancestry of cases and controls together, although we ensured that both cases and controls were of European ancestry. Finally, the use of the gnomAD population as a control cohort means that we are not comparing truly matched populations. Both in our case cohorts and the gnomAD cohort there are slightly more female than male participants, that is the female: male ratio for the cases is ~1.97:1 and for the controls it is ~1.27:1, which may result in a slight confounding effect, as does differences in environmental and cultural differences that could not be controlled in our study.

Conclusion

This study provides evidence that increased burden of missense variants in the amount of variants and the number of subjects carrying a variant in *CACNA1H* and *CACNA1I* exists for HM, and that these genes can modify HM disease risk, supporting more complex types of heritability for HM, in addition to the strictly monogenic forms.

Supplementary Information

Supplementary Table 1
https://static-content.springer.com/esm/art%3A10.1007%2Fs12035-023-03255-5/MediaObjects/12035_2023_3255_MOESM1_ESM.xlsx

Supplementary Table 2
https://static-content.springer.com/esm/art%3A10.1007%2Fs12035-023-03255-5/MediaObjects/12035_2023_3255_MOESM2_ESM.xlsx

Acknowledgements: We greatly acknowledge the subjects who participated in this study.

Author's Contributions: Omar Ibrahim, Aster VE Harder, Lisanne S Vijfhuizen, Rodney A Lea, Arn MJM van den Maagdenberg and Lyn R Griffiths contributed to the study conception and design. Whole exome sequencing and data collection were performed by Neven Maksemous, Heidi Sutherland, Nadine Pelzer, Irene de Boer and Gisela M Terwindt. Material preparation and analysis were performed by Omar Ibrahim, Aster VE Harder, Neven Maksemous, Lisanne S Vijfhuizen and Rodney A Lea. The first draft of the manuscript was written by Omar Ibrahim and Aster VE Harder with additional input from all co-authors. The final manuscript was read and approved by all authors.

Funding: The authors disclose receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by: the Australian National Health and Medical Research Council (NHMRC-APP1122387) (LRG); a Migraine Research Foundation grant (2016), NY, USA (LRG); an Australian International Science Linkages grant (LRG); by infrastructure purchased with Australian Government EIF Super Science Funds as part of the Therapeutic Innovation Australia – Queensland Node project (LRG); the Centre of Medical System Biology (CMSB) in the framework of the Netherlands Genomics Initiative (NGI) 050-060-409 (AMJMvdM); and the European Community (EC) FP7-EUROHEADPAIN (no. 602633; AMJMvdM).

Data availability: The data used and/or analysed during the current study are available from the corresponding author on reasonable request.

Competing interests: GMT reports consultancy support from Novartis, Allergan/Abbvie, Lilly,

and Teva, Lundbeck and independent support from Dutch Research Council, the Dutch Heart & Brain Foundations, IRRF and Dioraphte. AMJMvdM reports research support from Praxis Precision Medicine and Schedule 1 Therapeutics and consultancy support from AbbVie. LRG reports recent consultancy support from Teva and research support from the Australian National Health and Medical Research Council and the US Migraine Research Foundation. IdB reports independent support from the International Retinal Research Foundation and Dutch Heart Foundation. AVEH, LSV, OI, NP, NM, HS, and RAL declare no conflict of interest.

Ethics approval: The study was conducted in accordance with the Declaration of Helsinki, and the protocol was approved by the Human Research Ethics Committee of the Queensland University of Technology (approval number: 1800000611) and by the Medical Ethics Committee of LUMC.

Consent to participate: Appropriate consents for the patient cohort, are already in place.

Consent for publication: Not applicable

References

1. Headache Classification Committee of the International Headache Society (IHS) The International Classification of Headache Disorders, 3rd edition. *Cephalalgia*. 2018;38(1):1-211.
2. Ducros A, Denier C, Joutel A, et al. The clinical spectrum of familial hemiplegic migraine associated with mutations in a neuronal calcium channel. *N Engl J Med*. 2001;345(1):17-24.
3. Russell MB, Ducros A. Sporadic and familial hemiplegic migraine: pathophysiological mechanisms, clinical characteristics, diagnosis, and management. *Lancet Neurol*. 2011;10(5):457-70.
4. Pelzer N, Haan J, Stam AH, et al. Clinical spectrum of hemiplegic migraine and chances of finding a pathogenic mutation. *Neurology*. 2018;90(7):e575-e582.
5. Ophoff RA, Terwindt GM, Vergouwe MN, et al. Familial hemiplegic migraine and episodic ataxia type-2 are caused by mutations in the Ca2+ channel gene CACNL1A4. *Cell*. 1996;87(3):543-52.
6. Dichgans M, Freilinger T, Eckstein G, et al. Mutation in the neuronal voltage-gated sodium channel SCN1A in familial hemiplegic migraine. *Lancet*. 2005;366(9483):371-377.
7. De Fusco M, Marconi R, Silvestri L, et al. Haploinsufficiency of ATP1A2 encoding the Na+/K+ pump alpha2 subunit associated with familial hemiplegic migraine type 2. *Nat Genet*. 2003;33(2):192-196.
8. Hiekkala ME, Vuola P, Artto V, et al. The contribution of CACNA1A, ATP1A2 and SCN1A mutations in hemiplegic migraine: A clinical and genetic study in Finnish migraine families. *Cephalalgia*. 2018;38(12):1849-1863.
9. Sutherland HG, Maksemous N, Albury CL, et al. Comprehensive Exonic Sequencing of Hemiplegic Migraine-Related Genes in a Cohort of Suspected Probands Identifies Known and Potential Pathogenic Variants. *Cells*. 2020;9(11) :2368.
10. Riant F, Roos C, Roubertie A, et al. Hemiplegic Migraine Associated With PRRT2 Variations: A Clinical and Genetic Study. *Neurology*. 2022;98(1):e51-e61.
11. Valente P, Castroflorio E, Rossi P, et al. PRRT2 Is a Key Component of the Ca(2+)-Dependent Neurotransmitter Release Machinery. *Cell Rep*. 2016;15(1):117-131.
12. Pelzer N, de Vries B, Kamphorst JT, et al. PRRT2 and hemiplegic migraine: a complex association. *Neurology*. 2014;83(3):288-290.
13. Gormley P, Kurki M, Kurki MI, et al. Common Variant Burden Contributes to the Familial Aggregation of Migraine in 1,589 Families. *Neuron*. 2018;98(4):743-753.e4.
14. Wray NR, Goddard ME. Multi-locus models of genetic risk of disease. *Genome Med*. 2010;2(2):10.
15. Gudmundsson J, Sulem P, Gudbjartsson DF, et al. A study based on whole-genome sequencing yields a rare variant at 8q24 associated with prostate cancer. *Nat Genet*. 2012;44(12):1326-1329.
16. Rivas MA, Beaudoin M, Gardet A, et al. Deep resequencing of GWAS loci identifies independent rare variants associated with inflammatory bowel disease. *Nat Genet*. 2011;43(11):1066-1073.
17. Panoutsopoulou K, Tachmazidou I, Zeggini E. In search of low-frequency and rare variants affecting complex traits. *Hum Mol Genet*. 2013;22(R1):R16-21.
18. Ferrari MDP, Klever RRM, Terwindt GMMD, Ayata CMD, van den Maagdenberg AMJMP. Migraine pathophysiology: lessons from mouse models and human genetics. *Lancet Neurology, The*. 2015;14(1):65-80.
19. Westenbroek RE, Sakurai T, Elliott EM, et al. Immunochemical identification and subcellular distribution of the alpha 1A subunits of brain calcium channels. *J Neurosci*. 1995;15(10):6403-6418.

20. Catterall WA. Structure and function of neuronal Ca2+ channels and their role in neurotransmitter release. *Cell Calcium*. 1998;24(5-6):307-323.
21. Dunlap K, Luebke JI, Turner TJ. Exocytotic Ca2+ channels in mammalian central neurons. *Trends Neurosci*. 1995;18(2):89-98.
22. Perez-Reyes E. Molecular physiology of low-voltage-activated t-type calcium channels. *Physiol Rev*. 2003;83(1):117-161.
23. Dolphin AC. Calcium channel auxiliary α2δ and β subunits: trafficking and one step beyond. *Nat Rev Neurosci*. 2012;13(8):542-555.
24. Petrovski S, Gussow AB, Wang Q, et al. The Intolerance of Regulatory Sequence to Genetic Variation Predicts Gene Dosage Sensitivity. *PLoS Genet*. 2015;11(9):e1005492.
25. The Genotype-Tissue Expression (GTEx) project. *Nat Genet*. 2013;45(6):580-585.
26. Ambrosini A, D'Onofrio M, Buzzi MG, et al. Possible Involvement of the CACNA1E Gene in Migraine: A Search for Single Nucleotide Polymorphism in Different Clinical Phenotypes. *Headache*. 2017;57(7):1136-1144.
27. Kurtuncu M, Kaya D, Zuliani L, et al. CACNA1H antibodies associated with headache with neurological deficits and cerebrospinal fluid lymphocytosis (HaNDL). *Cephalalgia : an international journal of headache*. 2013;33(2):123-129.
28. Calhoun JD, Huffman AM, Bellinski I, et al. CACNA1H variants are not a cause of monogenic epilepsy. *Hum Mutat*. 2020;41(6):1138-1144.
29. Rasmussen AH, Kogelman LJA, Kristensen DM, et al. Functional gene networks reveal distinct mechanisms segregating in migraine families. *Brain : a journal of neurology*. 2020;143(10):2945-2956.
30. Hautakangas H, Winsvold BS, Ruotsalainen SE, et al. Genome-wide analysis of 102,084 migraine cases identifies 123 risk loci and subtype-specific risk alleles. *Nat Genet*. 2022; 54(2):152-160
31. Maksemous N, Smith RA, Sutherland HG, et al. Targeted next generation sequencing identifies a genetic spectrum of DNA variants in patients with hemiplegic migraine. *Cephalalgia Reports*. 2019;2:2515816319881630.
32. The International Classification of Headache Disorders, 3rd edition (beta version). *Cephalalgia*. 2013;33(9):629-808.
33. Miller SA, Dykes DD, Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res*. 1988;16(3):1215.
34. Thomsen LL, Eriksen MK, Roemer SF, et al. A population-based study of familial hemiplegic migraine suggests revised diagnostic criteria. *Brain*. 2002;125(Pt 6):1379-1391.
35. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2009;25(14):1754-60.
36. Danecek P, Bonfield JK, Liddle J, et al. Twelve years of SAMtools and BCFtools. *Gigascience*. 2021;10(2)
37. Cingolani P, Platts A, Wang le L, et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. *Fly (Austin)*. 2012;6(2):80-92.
38. McLaren W, Gil L, Hunt SE, et al. The Ensembl Variant Effect Predictor. *Genome Biol*. 2016;17(1):122.
39. Guo MH, Plummer L, Chan YM, Hirschhorn JN, Lippincott MF. Burden Testing of Rare Variants Identified through Exome Sequencing via Publicly Available Control Data. *Am J Hum Genet*. 2018;103(4):522-534.
40. Rasmussen BK, Jensen R, Olesen J. Impact of headache on sickness absence and utilisation of medical services: a Danish population study. *J Epidemiol Community Health*. 1992;46(4):443-446.
41. Ertel EA, Campbell KP, Harpold MM, et al. Nomenclature of voltage-gated calcium channels. *Neuron*. 2000;25(3):533-535.

42. Chemin J, Monteil A, Perez-Reyes E, et al. Specific contribution of human T-type calcium channel isotypes (alpha(1G), alpha(1H) and alpha(1I)) to neuronal excitability. *J Physiol*. 2002;540(Pt 1):3-14.

43. Cain SM, Snutch TP. Contributions of T-type calcium channel isoforms to neuronal firing. *Channels (Austin, Tex)*. 2010;4(6):475-482.

44. Huang Z, Lujan R, Kadurin I, et al. Presynaptic HCN1 channels regulate Cav3.2 activity and neurotransmission at select cortical synapses. *Nat Neurosci*. 2011;14(4):478-486.

45. Pan ZH, Hu HJ, Perring P, Andrade R. T-type Ca(2+) channels mediate neurotransmitter release in retinal bipolar cells. *Neuron*. 2001;32(1):89-98.

46. Weiss N, Zamponi GW. Control of low-threshold exocytosis by T-type calcium channels. *Biochim Biophys Acta*. 2013;1828(7):1579-1586.

47. Zamponi GW, Striessnig J, Koschak A, Dolphin AC. The Physiology, Pathology, and Pharmacology of Voltage-Gated Calcium Channels and Their Future Therapeutic Potential. *Pharmacol Rev*. 2015;67(4):821-870.

48. Choi S, Na HS, Kim J, et al. Attenuated pain responses in mice lacking Ca(V)3.2 T-type channels. *Genes Brain Behav*. 2007;6(5):425-431.

49. Bourinet E, Alloui A, Monteil A, et al. Silencing of the Cav3.2 T-type calcium channel gene in sensory neurons demonstrates its major role in nociception. *EMBO J*. 2005;24(2):315-324.

50. Weiss N, Zamponi GW. Genetic T-type calcium channelopathies. *J Med Genet*. 2020;57(1):1-10.

51. Gangarossa G, Laffray S, Bourinet E, Valjent E. T-type calcium channel Cav3.2 deficient mice show elevated anxiety, impaired memory and reduced sensitivity to psychostimulants. Original Research. 2014;8:92.

52. Astori S, Wimmer RD, Prosser HM, et al. The Ca(V)3.3 calcium channel is the major sleep spindle pacemaker in thalamus. *Proc Natl Acad Sci U S A*. 2011;108(33):13823-13828.

53. Splawski I, Yoo DS, Stotz SC, et al. CACNA1H Mutations in Autism Spectrum Disorders*. *J Biol Chem*. 2006;281(31):22085-22091.

54. Carter MT, McMillan HJ, Tomin A, Weiss N. Compound heterozygous CACNA1H mutations associated with severe congenital amyotrophy. *Channels (Austin, Tex)*. 2019;13(1):153-161.

55. Chen Y, Lu J, Pan H, et al. Association between genetic variation of CACNA1H and childhood absence epilepsy. *Ann Neurol*. 2003;54(2):239-243.

56. Gulsuner S, Walsh T, Watts AC, et al. Spatial and temporal mapping of de novo mutations in schizophrenia to a fetal prefrontal cortical network. *Cell*. 2013;154(3):518-529.

57. El Ghaleb Y, Schneeberger PE, Fernández-Quintero ML, et al. CACNA1I gain-of-function mutations differentially affect channel gating and cause neurodevelopmental disorders. *Brain*. 2021;144(7):2092-2106.

58. Maksemous N, Blayney CD, Sutherland HG, et al. Investigation of CACNA1I Cav3.3 Dysfunction in Hemiplegic Migraine. *Front Mol Neurosci*. 2022;15:892820.

59. Gardiner AR, Jaffer F, Dale RC, et al. The clinical and genetic heterogeneity of paroxysmal dyskinesias. *Brain*. 2015;138(Pt 12):3567-3580.

60. Suzuki M, Van Paesschen W, Stalmans I, et al. Defective membrane expression of the Na(+)-HCO(3)(-) cotransporter NBCe1 is associated with familial migraine. *Proc Natl Acad Sci U S A*. 2010;107(36):15963-15968.

61. Jen JC, Wan J, Palos TP, Howard BD, Baloh RW. Mutation in the glutamate transporter EAAT1 causes episodic ataxia, hemiplegia, and seizures. *Neurology*. 2005;65(4):529-534.

62. Weller CM, Leen WG, Neville BG, et al. A novel SLC2A1 mutation linking hemiplegic migraine with alternating hemiplegia of childhood. *Cephalalgia*. 2015;35(1):10-15.

10

General discussion

Introduction

In this thesis both migraine and cluster headache were investigated. For migraine, studies were performed in the fields of genomics, metabolomics but also specific molecules involved in a biological pathway were investigated. For cluster headache, the research focussed only on genomics. The aim of **Part I** of the thesis was to gain a better understanding of the biochemical profile of migraine. We hypothesized that individuals with migraine have a different biochemical profile, reflecting disturbances in molecular processes relevant to the pathophysiology of the disease. To this end, we investigated profiles of biochemical compounds interictally (i.e. outside an attack) and ictally (i.e. during an attack) in individuals with migraine and compared these with profiles obtained from control individuals to identify differences in metabolites and molecules relevant to the disease. Findings of the studies conducted for the thesis are discussed and put in perspective with other biochemical findings in migraine. In addition, it is shortly discussed what can be learned from such studies. The aim of **Part II** was to investigate the genetic underpinnings of migraine and cluster headache. The most recent findings in genome-wide association studies (GWAS) and next-generation sequencing analysis in migraine and cluster headache are presented. It is shown that cross-trait and causal analyses are beginning to identify and characterise specific biological factors that contribute to migraine and cluster headache risk, as well as their comorbid conditions. The combination of metabolomics and genomics data is discussed and followed by a short general reflection and opportunities for future research.

Biochemistry of migraine

Human health is dependent on the interaction of genes and environmental factors (e.g. medication and lifestyle). The metabolomic profile can be considered the endpoint of biological workings that include genomic, epigenomic, transcriptomic, proteomic, and environmental factors. A benefit of metabolomics is that it takes environmental factors into account and therefore represents a broader view of the phenotype at the molecular level than, for instance, genomics alone. In a way, metabolomics is the unbiased sum of all biochemical processes. This makes studying metabolomics very relevant when trying to elucidate the aetiology of a disease.

Within biochemical research, studies can focus on different aspects of the metabolic system, as was done by either looking at specific molecules, i.e. at amines (**Chapter 3**), endocannabinoids (**Chapter 4**), and prostaglandin E_2 (PGE_2) (**Chapter 5**), or by looking at a more unbiased biochemical profile (**Chapter 2**). By investigating the overall system and/or by combining different aspects of the architecture of a disease the results of the studies in the end might implicate similar biological mechanisms. One step further is to combine various omics approaches by which new insights in the disease pathology can be gained and through which biological systems can be interlinked.

Migraine in perspective, focussing on lipids and amines

In **Chapters 2 - 5** different metabolites in different body fluids obtained inside or outside migraine attacks were investigated. Because migraine is a multifactorial complex brain disorder there are a multitude of aspects that can be and have to be investigated in order to fully understand the mechanism of disease. An important question in the context of migraine is whether outside of attacks, migraine patients differ biochemically from controls or whether this is only during attacks. What we know from **Part II** of the thesis is that migraine patients differ genetically from controls. This makes it conceivable that there are also differences in metabolite concentrations interictally between migraine patients and controls. In this respect it is of interest to note that twin studies have shown that the heritability of blood metabolites is up to 80%.[1] Given that genetics are static, this means that when looking at metabolites in body fluids interictally a difference may be expected. This is also follows from the research in **Chapters 2 and 3**, were migraine patients differed in their metabolite profile from controls. Although we have investigated a variety of aspects in migraine patients (i.e. different biochemical and genetic aspects), we have not specifically investigated the interplay between the biochemical and the genetic system in this thesis. The correlation of metabolites and genetic data was combined with previous knowledge on the increased cardiovascular disease risk in migraine patients, by an Australian group.[2-5] The Australian researchers hypothesized that routine chemistry tests and serum markers for cardiovascular disease might be associated with migraine risk, directly as well as on a genetic basis. They showed that there was an association between clinical chemistry tests traditionally used to monitor cardiovascular disease and migraine risk.[6] More importantly, in cross-trait genetic analyses it was shown that migraine and the associated chemistry tests have an underlying shared genetic basis.[6] Notably, the association and genetic overlap between a lower level of HDL-C and an increased migraine risk are in part due to shared biology.[6] This observed link where genetic factors influence blood metabolite levels and risk for migraine indicates a change of metabolite concentrations in migraine patients. This example from literature illustrates the relevant interplay between the genetic and biochemical underpinnings of migraine and in addition confirms that a difference in metabolite concentration interictally is very probable and thus in line with the results from **Chapter 2 and 3**.

In a subsequent study by the same Australian group, the genetic underpinning and causality of even more blood metabolites (n = 316) and migraine risk were investigated.[7] Their study showed a significant correlation between migraine and 44 metabolites that were mostly associated with lipid metabolism.[7] In addition, they found a causal effect of some lipoproteins on migraine risk.[7] These findings in lipids were perhaps not that unexpected as epidemiological studies had already shown that obesity is a risk factor for migraine, as well as the comorbidity of cardiovascular and cerebrovascular disease in migraine.[2-5] Regardless, studies directly into lipid levels of migraine have not shown consistent results; in some studies differences in lipid levels were found in migraine patients,[8, 9] whereas in other studies this was not the case.[8, 10] A previous very thorough meta-

analysis from our group into lipid levels of 2,800 migraine patients and 7,353 controls using a ^{1}H-NMR metabolomics platform did show an indication for alterations in HDL metabolism.[9] Another, possibly less eloborate study looking at both the phenotypic and the genotypic associations independently between lipoproteins and migraine in only women with data from the Women's Genome Health study (over 22,000 subjects) found the association with triglyceride-rich lipoproteins (TRLP) subfractions to be more evident than LDL or HDL.[11] The authors from the USA claim that their large sample size, detailed information on migraine subtypes, consideration of important confounders including menopause status and postmenopausal hormone therapy in the phenotypic analysis may be a factor explaining the differences. All in all the studies support the view that it may be possible to use lipoprotein measures in blood to predict migraine status. As shown in **Chapter 2** it was indeed possible to produce a predictive metabolite profile that was largely dependent on lipids. Taking the results of the different studies together suggests that a limited number of lipids do play a role in migraine pathophysiology, but that there are many factors to take into account in the study design before meaningful conclusions can be drawn at the individual lipoprotein compound level. Regarding causality of these lipid metabolites and migraine the data are also contradictory, where one study,[11] reports no causality between migraine and lipoprotein subfractions, another study found a causally protective role for a longer length of fatty acids against migraine as well as a causal role of a higher level of a lysophosphatidylethanolamine on migraine.[7]

The second most associated type of metabolite in the above mentioned Australian study[7] was amino acids. A genome-wide genetic overlap was seen for higher levels of phenylalanine, isoleucine and gamma-glutamylisoleucine and lower levels of proline betaine and 2-hydroxyisobutyric acid in migraineurs. Amino acids are a subgroup of amines, which were also investigated in **Chapter 3**. Amino acids have been previously investigated in several biochemical studies in migraine,[12, 13] and one particular amino acid, namely glutamate/glutamine has been repeatedly linked to migraine.[14-16] In **Chapter 2** our study into the metabolomics profile of migraine patients did show that multiple amino acids were part of our prediction model able to predict active migraine status, namely: leucine, isoleucine, proline, methionine, valine and serine. Noteworthy is the overlap found for isoleucine with **Chapter 2**, but also with phenylalanine in **Chapter 3**. In **Chapter 3** six amine plasma/CFS ratios were found to be different in migraineurs compared to controls. In two of the ratios phenylalanine was present and the overall ratio correlation between $\frac{\text{L-Valine}_{plasma}}{\text{L-Phenylalanine}_{plasma}} / \frac{\text{L-Valine}_{CSF}}{\text{L-Phenylalanine}_{CSF}}$ and $\frac{\text{L-Leucine}_{plasma}}{\text{L-Phenylalanine}_{plasma}} / \frac{\text{L-Leucine}_{CSF}}{\text{L-Phenylalanine}_{CSF}}$ was different between individuals with migraine compared to controls. Whether the metabolomic ratio changes are the consequence of amino acid changes in the brain itself or malfunctioning of the blood brain transporter that co-regulates concentrations inside and outside the brain still needs to be investigated.

Taken together, there is ample evidence that lipids and amino acids and their related biomolecules remain interesting targets in migraine research. In addition it is demonstrated that when

combining both metabolomic and genomics data better insights into the metabolism and co-morbidity underlying disease can be provided, which will eventually lead to the identification of new diagnostic and therapeutic targets. Such approach has already been successful in cancer research, where the integration of metabolomics data with genomics data led to the identification of new drug targets.[17, 18]

Interplay between biochemical systems

In **Chapters 4** and **5** we investigated endocannabinoids and prostaglandin E_2 levels, respectively, in relation to migraine risk. Although we did not find any differences in concentrations between migraineurs and controls, these studies did show that both systems are very complex and influenced by many factors. Illustrative of the complexity of both the endocannabinoid and prostaglandin system is the role the transient receptor potential vanilloid 1 (TRPV1) cation channel has in *both* systems. TRPV1 is a nonselective cation channel, predominantly expressed in sensory neurons, about 40-50% of the trigeminal sensory neurons express TRPV1.[19-21] It is believed that TRPV1 has a role in nociception, which makes it a molecular target for pain treatment.[20, 22] The channel has a preference for calcium and can be activated by numerous stimuli, such as pH < 5.9, heat, but TRPV1 is *also* sensitive to endocannabinoids and prostaglandins.[20, 23] Activation of TRPV1 causes calcitonin-gene related peptide (CGRP) release.[24]

Two of the most well-known endocannabinoids, *N*-arachidonoylethanolamine (AEA) and 2-arachidonoylglycerol (2-AG) were investigated in this thesis. These endocannabinoids are synthesised on demand from lipid membrane phospholipids and have a half-life of only minutes due to rapid degradation by fatty acid amide hydrolase (FAAH) and monoacylglycerol lipase (MAGL), respectively.[25-27] AEA and 2-AG are partial agonist for the cannabinoid receptors (CB_1 and CB_2). The CB_1 receptor is strongly involved in inhibiting the release of gamma-aminobutyric acid (GABA), serotonin, glutamate, and CGRP, all neurotransmitters believed to play an important role in migraine.[28-31] It has been shown that AEA activates the TRPV1 channel on trigeminal ganglion neurons. Thus promoting the release of calcitonin-gene related peptide (CGRP) independent of any action at the CB_1.[32] Thus AEA can both inhibit and stimulate CGRP release via CB_1 and TRPV1, respectively.[30, 33] In addition, there are various studies that indicate that TRPV1 is co-expressed with either CB_1, CB_2 or both, in neuronal and non-neuronal cells.[33] Therefore, AEA may have a distinct role in nociception, both being pronociceptive at TRPV1 and antinociceptive at cannabinoid receptors.[34]

Similarly, prostaglandins work in a complex manner as they act on various receptors including EP1, EP2, EP3, EP4 and IP that activate different G protein-coupled signalling pathways.[35] It has been shown that the activation of EP1/ EP4 and IP by PGE_2 and PGI_2, respectively, is critical for TRPV1 sensitization.[36] In addition, PGE_2 has been shown to increase TRPV1 expression, cell surface and axonal trafficking.[37, 38] In addition, PGE_2 also has an interaction with CGRP, as it stimulates the release of CGRP in rat trigeminal neurons, trigeminal ganglia and trigeminal

nucleus caudalis.[39-41] Cyclooxygenases (COX-1 and COX-2) are responsible for the conversion of arachidonic acid to prostaglandins and thromboxane and the conversion of AEA and 2-AG into prostanoid-like derivatives.[42]

Together this shows that biological systems do not stand alone and that both AEA and PGE_2 are intertwined and can converge on one pathway via the TRPV1 channel and CGRP. Given this mutual convergence, it is a perhaps not surprising that in neither study (**Chapter 4** nor **Chapter 5**) we found a difference between migraineurs and cases. It was already known that the TRPV1 channel is an important contributor to migraine pathology.[34] Furthermore, it has been demonstrated that there is an increased expression of TRPV1 in chronic migraine patients.[43] Additional argumentation for the involvement of TRPV1 in migraine, are the expression of TRPV1 on the meningeal nociceptors and that their activation is known to promote the release of CGRP from sensory nerve endings.[44] A future step in research would be looking more holistically to our biological system.

Many components of these systems have already been investigated with the aim to develop targeted therapies. Currently migraine treatment consists of acute and preventive migraine medications that do not result in the desired effect for all patients. The efficacy (the achievement of freedom from pain or from the most bothersome symptom within 2 hours) of acute medications in clinical trials lies around 19-31%.[45] The efficacy (a 50% reduction in monthly migraine days) of preventive medications lies around 40-50%.[46,47] Capsaicin, the pungent ingredient in chili peppers, is a TRPV1 agonist and has been observed to provide pain relief when used in the appropriate amount and frequency. Research has shown that repeated intranasal capsaicin treatment can reduce migraine attacks by 50-80% in chronic migraine sufferers.[48] This is believed to be due to the desensitising effect of capsaicin via TRPV1, which leads to a decrease in CGRP release. SB-705498 is a chemical that acts as a competitive antagonist of the TRPV1 receptor, was found to be less effective than placebo in treating migraine headache, photophobia, and phonophobia in a phase II clinical trial.[49] While several TRPV1 antagonists have been studied in clinical trials, they have failed to advance to Phase III trials due to adverse side effects such as hyperthermia or loss of thermal pain sensitivity.[50] CGRP is one of the latest avenues in migraine treatment, it is known that the release of CGRP can be blocked by sumatriptan a conventional anti-migraine drug.[51] Sumatriptan binds to the 5-HT-1D/1F receptor in the brain and causes constriction of the extracerebral blood vessels within the cranial vasculature, to inhibit release of inflammatory mediators. Monoclonal antibodies that target the CGRP receptor, known as gepants, have been recently approved for migraine treatment.[52] Examples are, erenumab, fremanezumab en galcanezumab.

Non-steroidal anti-inflammatory drugs (NSAIDs), which inhibit prostaglandin synthesis by inhibiting COX enzymes, are a first line treatment for migraine headaches Investigating treatment response via the PGE_2 pathway was in migraine was done with the EP4 receptor antagonist BGC20-1531 in a randomised, double blind, placebo-controlled, three-way intra-individual crossover study. This studies showed that BGC20-1531 was not effective against PGE_2 induced

headache, which suggests that either the compound was not effective for full blockage of the receptor or that EP4 is not the only receptor involved in PGE_2 induced headache.[53] Regarding the endocannabinoid system, medical marijuana is used as a preventive treatment of migraine, although controlled research on the effectiveness is lacking,[47, 54] Chemical compounds targeting the endocannabinoid catabolic enzymes are new treatment targets and allosteric ligands of the CB_1 and CB_2 receptor are gaining interest and are currently being investigated in animal studies.[47] The newest treatment avenue entails incorporating the interplay of these systems is dual therapy, whereby components of different pathways are blocked/stimulated with one medicine. Currently, there is development of combined FAAH/TRPV1 and FAAH/COX-2 inhibitors.[55] Other upcoming targeted medication are the ditans and gepants. Ditans are selective 5-HT-1F receptor agonists, such as lasmiditan. Gepants are small molecules that work as antagonists of CGRP receptors, examples are ubrogepant and rimegepant. Both can be used as medication during an attack, ditans can also be used to prevent attacks.

Based on these pathways thus far no effective treatment has been found for migraine, a consecutive question would be: why that is and are we looking at the right targets, given that no difference in concentration was found in either **Chapter 4** or **Chapter 5**? The answer could be that the absence of significance does not prove the absence of an effect. In addition, a difference might be too subtle to be detected or only occur at the cellular level. There is still a lot of evidence pointing to the involvement of these systems in pain and migraine, this should not be disregarded. However, given the large clinical heterogeneity in migraine there might be an equally large heterogeneity in the components of the pathways that are involved. To truly understand the mechanisms involved, we might have to change the way we investigate migraine.

Finding biochemical differences in case-control studies

Often the central aim of metabolomics studies is to test for the association between (specific) metabolites and a trait of interest, in our case migraine. Once a difference is found this can provide biological insight into the disease, be it the cause or consequence of a disease process. In this thesis, we set out to investigate the biochemical aspects of migraine. For both **Chapters 3 and 5** we performed a robust analysis to compare plasma levels of endocannabinoids and PGE_2, respectively, between cases and controls. Contrary to what other studies have found, we did not detect differences in migraineurs compared to controls. Notably, our sample sizes were larger than in most previous studies that did find differences between migraineurs and controls. The lack of replicability and comparability is a major and challenging issue when investigating metabolites. One reason underlying the contradictory findings might be related to the study design. As stated above, both the endocannabinoid system and the pathways involving PGE_2 are complex and can be influenced by numerous possibly confounding factors apart from the presence or absence of migraine. Confounding occurs when there is a distortion of the association between the exposure and the outcome. The problem of confounding arises when the study groups differ with respect to variables that can influence the outcome, such as age, sex,

dietary factors or medication. Logically, when a study is subjected to confounding the results do not faithfully reflect the relationship between the variables under study. From a research perspective it is therefore sensible to account for such factors as much as possible, but it is impossible and perhaps even undesirable to take all into account. Factors that can influence the metabolome can be divided into two; 1) influenceable factors, such as storage of samples and time of day of sample collection; [56-59] 2) uninfluenceable factors, such as sex and age.[60, 61]

Sampling is an important step in research into metabolites, as discussed in the General Introduction (**Chapter 1**), and should be planned and carried out meticulously as changes in handling of a sample can have a large effect on the outcome, already because of the (sometimes) extreme variability in stability of certain metabolites and lipids.[58] A way to control the data for these changes is protocolized sampling, as we have done in our metabolomics studies. Unfortunately, not in all studies collection and storing of samples is done in such a protocolized fashion or even described in the respective papers[62-66]. Demographic variables such as age, sex, ethnicity, smoking behaviour and obesity are frequent confounders in omics association studies. It is known that sex and age are often the principal factors explaining metabolome variability.[60, 67]

Migraine prevalence differs between men and women and lifestyle characteristics, such as smoking and alcohol intake, differ from those without migraine.[68-70] Given these differences in prevalence they are important conceivable confounding factors that need to be considered in migraine research to avoid a false-positive (Type I) error. It is possible to either account for confounding factors through experimental design, by means of randomization, restriction and matching before data gathering or after the data gathering process by means of statistical analyses, via stratification or multivariate models.[71, 72] Endocannabinoids for instance, are involved in a large number of pathways and processes.[73] Several endocannabinoids have been associated with body mass index (BMI), sex, age, weekly alcohol consumption and smoking.[73-75] In addition, there is also a well-known relation between depression and endocannabinoids.[76-81] Energy homeostasis is one of the processes best known to be controlled by the endocannabinoid system.[73, 82] Thus, the endocannabinoid system affects the food intake and indirectly the BMI.[73, 83] Although the pathways underlying sex differences in the endocannabinoid system are not fully elucidated, it is generally accepted that differences exist between the male and female endocannabinoid system.[84-86] Moreover, studies in mice have found that chronic exposure to ethanol leads to increased levels of AEA and 2-AG.[87-89] Smoking is expected to also play a role in the endocannabinoid system as it is involved in the reward circuits of the brain. There is some evidence on a decrease in the expression of cannabinoid receptor 1 (CB_1) molecules and an increase in the AEA level in smokers.[90, 91] Given that these factors could influence both the exposure and the outcome, all these factors were taken into account when studying endocannabinoid levels in **Chapter 3** when we were solely interested on the endocannabinoid-migraine relationship. Comparing our study with similar studies there was a considerable clinical and methodological heterogeneity across studies. In previous studies on endocannabinoid levels in patients there is only mention of age-matching in the design of one

study,[62] and restriction to female participants in another study,[92] but no other correctional steps were taken in the design or statistical model for any of the other confounding factors (BMI, alcohol consumption, smoking and depression). In the case of PGE_2 less confounders are known, but it seems likely that age is an important factor, as PGE_2 plays a role in inflammation,[93-96] and aging is associated with dysregulation of the immune and inflammatory responses. Neuroinflammation has been implicated to contribute to neurodegeneration in normal aging and in age-associated neurological disorders.[97] A recent study showed, for the first time, *in vivo* compelling evidence of neuroinflammation in migraine.[98] It has previously been shown that peripheral blood mononuclear cells from elderly human subjects as well as macrophages and splenocytes from old mice make significantly more PGE_2 than their young counterparts.[99, 100] This increase was mainly mediated via an increase in COX-2 expression.[101] In **Chapter 5**, where PGE_2 levels were investigated in relation to a provoked migraine attack, our statistical model was adjusted for age. In none of the other studies on PGE_2 levels there was a mention of age being taken into account.[63-66] In the most recent study, age was not even reported for the individuals under investigation.[66] In **Chapter 4**, an extensive profile of metabolites was investigated, hence we corrected our regression model for the most commonly known metabolic confounders; age, sex, BMI and smoking status.

In conclusion, it is very important prior to starting a metabolomics study and prior to analysing data from such study to consider the role of confounders to be able to meaningfully quantify biological signals of interest. Conflicting results between studies can be the cause of a large heterogeneity in sample collection and analysis and therefore such studies need to be performed with caution.

Genetics of migraine

Since the last two decades, genome-wide association studies (GWAS) have been able to identify many thousands of low-effect risk DNA variants for numerous diseases.[102] An important challenge in GWAS is to determine the exact gene that is affected by a risk SNP, as this is often far from straightforward. A complicating factor is that the location of the associated SNP is often intronic or intergenic, were it functions merely as genomic marker at a distance of, so in linkage disequilibrium (LD) with, the true causal variant. One of the most used methods to link a risk locus/index SNP (i.e., the SNP with the lowest *p*-value) to a gene is to relate it to the nearest gene. However, in two-thirds of cases it was demonstrated that it is not the nearest gene that is affected by the risk SNP.[103, 104] Moreover, there can be multiple independent association signals at the same locus (secondary SNPs) that can influence other regulatory features of the same (or nearby) gene. Most risk SNPs do not directly affect amino acid changes, but appear to regulate gene expression indirectly by disrupting enhancer elements. A way to determine the causal gene more effectively is by integrating data from GWAS and expression data in a way that expression quantitative trait loci (eQTL) analyses is able to prioritize likely causal genes.[102] **Chapter 6, 7** and **8** illustrate that the road ahead to understanding the pathophysiology behind migraine and cluster headache risk loci is far from easy.

Overlap between migraine and cluster headache

Comorbidities between disorders can be the result of either sharing of genetic and/or environmental factors, or a causal relationship. GWAS data can be used to investigate a possible shared genetic architecture in correlation studies where the proportion of variance two trait share due to genetics is measured (the genetic correlation r_G). In **Chapter 6** and **Chapter 7** susceptibility loci for cluster headache were identified and in **Chapter 8** risk loci for migraine were identified. **Chapter 6** and a parallel GWAS paper on cluster headache independently of each other showed genetic overlap between migraine and cluster headache, as one (*UFL1*/*FHL5*) of the four genes associated with cluster headache was the same as in migraine.[105] This is not that surprising, as there was already a presumed link between cluster headache and migraine as they both share key clinical characteristics (unilateral headache and cranial autonomic symptoms), pathophysiological mechanisms (involvement of the trigeminal vascular system) and response to therapy (CGRP monoclonal antibodies).[106] This overlapping locus was reason to further investigate the genetic overlap with a correlation analysis, this is a parameter that describes the proportion of variance two traits share due to genetic causes. In **Chapter 6** we found a positive correlation of 33% between migraine and cluster headache. However, due to the relatively small sample size of the cluster headache GWAS this was not significant, therefore, it was important to perform this analysis in a larger cohort. In a meta-analysis of these cluster headache GWAS (**Chapter 7**), there was possibility to further investigate the possibility of an overlapping genetic architecture between both disorders. In the European meta-analysis a total of three (*FHL5*, *PLCE1*, *LRP1*) of all seven cluster headache risk loci were similar to migraine risk loci. Apart from these three overlapping loci, none of the other 119 known migraine risk loci showed an association with cluster headache. Genetic correlation analysis did show that there was an overall genetic correlation between cluster headache and migraine, suggesting a shared genetic architecture. It could be argued that this overlap would be due to "contamination" of migraine cases in our cluster headache dataset. However, there are findings that would contradict this. In case the signal in our cluster headache meta-analysis came from migraine patients, one would expect that the exact same loci would come up. *LRP1* is associated with the lowest *p*-value in migraine, but the second most prominent locus in migraine is *PRDM16*. This locus was already reported in 2011 in the second migraine GWAS, which only found three risk loci (*LRP1*, *PRDM16* and *TRPM8*).[107] This locus shows no signal in the cluster headache meta-analysis. Hence, the absence of *PRMD16*, argues that there is no contamination of migraine patients in the cluster headache sample. The third locus associated gene in migraine is *FHL5*, but all other top-associated genes up-until *PLEC1*, which is the 11th associated gene in migraine, show no signal in cluster headache. Thus, it seems that there is some sharing in genetics architecture between cluster headache and migraine, but that their exact genetic signature differs. In fact, investigation of the top 200 ranked gene-based *p*-values between both disorders found an overlap of only 10 genes (data not shown), which confirms the notion that the two disorders have an overall different genetic signature. The advantage of gene-based *p*-values is that it is only based on SNPs that are located within a gene, therefore, you have more power. Furthermore, we show

that the effect of these overlapping risk loci in cluster headache is much larger than it ever was in migraine, even though the sample size is smaller. This result also disputes that the overlapping loci are merely a result of contamination of migraine cases. In addition, cluster headache also showed a correlation with pain, depression, but above all with smoking and risk-taking behaviour traits, but this was different from what was seen in migraine. Migraine did also show a correlation with pain and depression, but not with smoking and risk-taking behaviour (**Chapter 7**). This suggest that migraine and cluster headache have a partly shared and a partly distinct genetic basis, which reflects the partly shared and partly distinct biological mechanisms.

Giving meaning to GWAS data

Over the last decade or so, GWAS have also been able to identify many low-effect risk DNA variants for common migraine. With each study, the sample size steadily increased as did the number of associated DNA variants.

In 2016 the at the time most recent migraine GWAS (59,674 cases and 316,078 controls) produced 44 LD-independent SNPs generating 38 distinct genomic loci associated with migraine.[108] Downstream bioinformatics analysis showed that the genes associated with these loci were primarily linked to arterial and smooth muscle function. This was in agreement with the shared polygenic genetic risk between migraine, stroke and cardiovascular diseases that had been described previously.[109, 110] In **Chapter 8** we performed the latest (2022) migraine GWAS with almost twice as many cases and controls (102,084 cases and 771,257 controls).[108] We were able to identify 123 distinct genomic regions associated with migraine, of which 86 loci were novel compared to the 2016 migraine GWAS.[108] Given that many of the loci replicate provides proof for the robustness of the loci. In addition, with downstream analysis an additional 44 genome-wide significant independent variants were identified resulting in a total of 167 independent SNPs in **Chapter 8**. Enrichment analyses suggested the involvement of vascular and central nervous system tissue/cell types, this is in concordance with the concept that migraine is a neurovascular disorder. In addition, risk loci that encode genes associated with migraine drug targets (calcitonin gene-related peptide; CGRP encoded by *CALCA/CALCB* and serotonin 1F receptor encoded by *HTR1F*) were identified.[111, 112] This clearly illustrates the potential of genetic studies to identify future drug targets. Additionally, migraine subtype analysis in clinically well diagnosed cases resulted in three risk variants (in *MPPED2, CACNA1A* and *HMOX2*) that were specific for migraine with aura and two (one near *FECH* and one near *SPINK2*) that were specific for migraine without aura. Most notably is *CACNA1A*, which is known as one of the three genes (*CACNA1A*, *ATP1A2* or *SCN1A*) in which a single high-penetrant mutation could lead to hemiplegic migraine.[113-115] Taken together mutations in *CACNA1A* can lead to hemiplegic migraine, but variants in *CACNA1A* contribute also to the risk for common migraine, linking *CACNA1A* to both monogenetic and polygenetic forms of migraine. In 2021 a migraine GWAS (28,852 cases and 525,717 controls) with more ethnically diverse individuals (East Asian, African American, and Hispanic/Latino

descent) was published by Choqet et al.,[116] reporting 95 single nucleotide polymorphisms (SNPs) to be associated with migraine. When putting **Chapter 8** and this new GWAS next to each other 99 SNPs were unique to our study, 11 were unique to Choquet et al. and 68 were found in both studies. Choquet et al. also identified 3 loci specific for woman, this means that there are now in total at least 181 SNPs robustly associated with migraine.[116] A benefit of finding as many risk loci as was done in migraine, is that these loci can be used to give biological meaning to the data, by detecting related genes or relevant functional groups, pathways and tissues with specific integrative analyses tools/software. There is a wide range of tools that can be used, all with their own specific goals and methodologies. Although pathway analysis have resulted in finding biomarkers for other diseases than migraine,[117, 118] proper application and interpretation is not straightforward.[119] In the cluster headache studies (**Chapters 6 and 7**) the number of loci was too few to perform these analysis.

Apart from giving biological meaning to samples, there are also bioinformatics approaches that can be used to add clinical meaning to genetic data or to better understand the relationships with other traits. Again proper usage is important, as with all statistical analyses assumptions are made regarding the data, which need to be understood by the researcher. In addition, sufficient power needs to be present to perform the analyses, and therefore a large enough sample sizes is necessary.

One of the methods uses polygenic risk scores (PRSs), which interrogates information of all the variants identified in GWAS and the combined effect of all these risk loci can be used to estimate a specific risk in a cohort. In breast cancer such a score is already being used for distinguishing women with a high and a low risk of breast cancer and advising sufficient preventative screening programs.[120] Multiple studies have been performed with calculating a PRS in migraine and relating it to a specific trait. However, to what extend such a score is of benefit in the current clinical setting is hard to envisage, as in contrast to breast cancer migraine is a disease with attacks and prevention has a different meaning. A possibility would be in assessing whether somebody is likely to respond to certain medications, as medications are expensive and reaching the desired concentration of effect also takes time. One study has investigated the migraine PRS in relation to treatment response.[121] However, there is still much to be improved on this study, indicating that a lot more work is to be done.

Another methods to infer clinical meaning is Mendelian Randomisation (MR) with MR analysis one can investigate causality between two traits with genetic data via a statistical method. In the last 5 years this has already resulted in over 20 MR studies investigating causal relationships with migraine for various traits. These studies focussed on traits that had already shown an association with migraine in epidemiological studies. Important traits that were investigated are again the cardiovascular and cerebrovascular diseases, given they are well known comorbidities of migraine.[2-5] The causal relationship of migraine with related traits such as stroke, blood pressure and coronary artery disease and migraine were investigated. Causality with migraine and stroke was examined based on GWAS data in three different studies by three different groups. Two studies, did not find a

causal relationship, and one found a protective influence of large artery stroke on migraine.[122-124] One study showed that an increase in blood pressure leads to an increase in migraine risk.[125] A similar result was observed in another study where both a raise in systolic blood pressure and diastolic blood pressure increased migraine susceptibility.[126] In addition, another study showed that migraine leads to an increased coronary artery disease risk.[127] Combining these results with what is already found in metabolomics studies supports a role of lipids in migraine pathophysiology. Apart from disease traits a some MR studies in migraine have investigated behavioural traits, such as smoking and coffee and alcohol consumption.[128, 129] **Chapter 6** was one of the first GWASs that resulted in risk loci for cluster headache. By the benefit of this study the genetic relationship of cluster headache with other traits and diseases could further be explored. The most notable, the long debated relationship between smoking and cluster headache, as a large proportion of cluster headache patients are known to smoke (70-90%).[130-132] Given that up to 40-60% of smoking addiction is due to genetic factors it may well be that there is a genetic link predisposing to both cluster headache and smoking.[133, 134] In addition, tobacco-smoking might cause cluster headache, for instance as a result of tobacco leading to excessive trigeminal autonomic pathway and hypothalamic activation.[135] This is supported by the observation that a majority of cluster headache patients (~80%) started smoking prior to cluster headache debut.[131] This could be explained by smoking causing cluster headache or by the sharing of genetic architecture and/or other causal risk factors. Contrary to this, is the fact that there is a long latency (>15 years) between smoking onset and cluster headache debut and additionally smoking cessation does not lead to changes in disease severity.[130-132] In **Chapter 6** the relationship between smoking and cluster headache was investigated with a genetic correlation analysis. We found a positive genetic correlation between smoking and cluster headache, indicating a genetic overlap. In **Chapter 7** the relationship was further unravelled, this time we analysed the causality of smoking on cluster headache. It was found that smoking and cluster headache not only have a shared genetic architecture, but also that smoking is one of the factors that causes cluster headache. As the positive correlation between two traits could lead to bias, an additional analysis, the latent causal variable (LCV) analysis was performed to verify the causality between smoking and cluster headache, this analysis reconfirmed the causality.[136] With GWAS data many genetic aspects of a disease can be unravelled. It is, however, important to note that MR studies are still in their infancy therefore it is critical and necessary to replicate the results of different studies to reach robust conclusions.

Combining genetic data: the full spectrum of migraine from polygenic to monogenic forms

In the search for a "fourth" hemiplegic migraine gene sequencing was applied to cohorts of hemiplegic migraine patients that are negative for mutations in *CACNA1A*, *ATP1A2*, and *SCN1A*. Thus far no "fourth" gene has been found, however, there are indications,[137] that the heritability of hemiplegic in mutation-negative patients is more complex (i.e. more oligogenic or even polygenic). A number of genes including *PRRT2*,[138] *PNKD*,[139] *SLC4A4*,[140] *SLC1A3*,[141] and *SLC2A1*,[142] are primarily associated with movement or solute transport disorders, but some patients with mutations in these genes were reported to display signs of hemiplegic migraine, suggesting an

effect of genetic modifiers. Furthermore, a Finnish GWAS demonstrated that hemiplegic migraine patients without a high-penetrant disease-causing mutation carry an excess of common (frequency > 1%) variants compared to patients suffering from common (complex) migraine subtypes.[143] In addition, in **Chapter 8** it has become clear that *CACNA1A* plays a role in monogenic as well as in complex forms of migraine. Thus there are multiple lines of evidence provided that show that common genetic variants contributes to the aetiology of hemiplegic migraine. Something that is also seen in developmental and epileptic encephalopathies (DEEs), where the aetiology of the DEEs is now regarded as complex, rather than monogenic.[144] This data suggests that the heritability of hemiplegic migraine can also be more complex. It is likely that the combination of both rare and common variants influence disease rick of complex disorders, however, the relative contribution of rare and common variants to disease risk is unknown. The difficulty in identifying these variants involved lies in the analysis method, where the GWA-approach (used in **Chapters 6 - 8**) is ideal in identifying common variants, next-generation sequencing is best in identifying rare variants. The inheritance of hemiplegic migraine might lie in a combination of both rare and common variants and/or an accumulation of variants in essential genes. It has already become clear that both common and rare variants can contribute to the same disease locus in other diseases.[145] Even though some analysis methods have been developed to look at both common and rare variants at once, the all require individual level data.[146] In **Chapter 9** we have adapted a method for summary level data to look at all missense variants, both rare and common, and there association with hemiplegic migraine. In other words we looked at the accumulation of variants or burden in a predefined number of genes, in order to have enough power. Therefore in **Chapter 9** the variant burden was investigated in next-generation sequencing data of mutation-negative hemiplegic migraine patients in genes of the *CACNA1x* family. From this study it can be concluded that a burden of likely not *per se* pathogenic missense variants in *CACNA1H* and *CACNA1I* is implicated as modifier and hence confers susceptibility to hemiplegic migraine. Burden in this case can be regarded as both the total number of variants as well as the amount of variants each subjects carried. The allele frequencies of the variants included, differed from 0.00003 – 0.59, thus including both common and rare variants. These results substantiate the hypothesis that all forms of migraine including hemiplegic migraine are explained by a spectrum of genetic variants ranging from rare high-risk mutations to common low-risk variants. However, these result still needs to be replicated in an independent cohort preferably based on individual level data.

The GWAS-era has left us with "missing heritability", where we know that the contribution of genetics is larger than thus far found with GWAS. Part of this "missing heritability" might be explained by the contribution of rare variants to common risk loci. Another way of looking at the contribution of both rare and common variants to disease risk is by combining data types. This approach is shown in a recent Danish study in extended migraine pedigrees (consisting of 1,040 individuals from 155 families) where they combined whole-genome sequencing (WGS) data with GWAS data.[147] Their aim was to fine-map the migraine risk loci with WGS data. It was found that there is an excess of rare segregating variants in regulatory regions (one CpG island and three polycomb group response

elements) near four migraine GWAS risk loci.[147] Different types of data can be combined as was shown by the same Danish group who combines RNA-seq data with WGS data of more than hundred migraine families in a so called 'systems genetics approach'.[148] With this approach they implicated a "visual cortex module" in migraine pathophysiology.[148] Their approach also suggested the involvement of glutamate, serotonin, G-protein signalling pathways and hormonal pathways. Other ways of integrating both common and rare variants is by deep resequencing of GWAS loci or performing whole-genome sequencing (WGS). However, thus far we have mainly focussed on the protein coding components of the genome which only make up about 1% of the whole genome. Thus analysis and interpretation of WGS data has still a long way to go.

Future of data integration

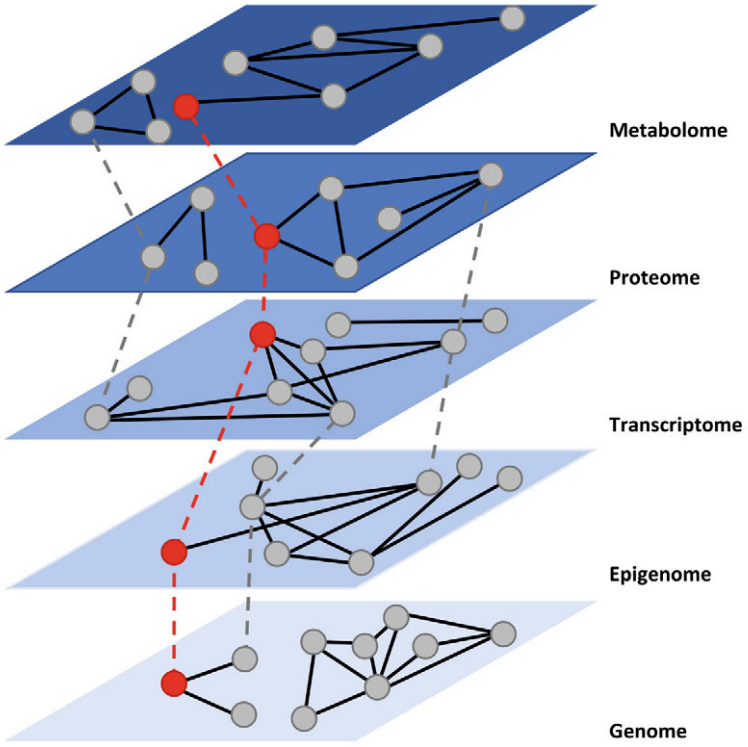
Multiple -omics fields have arisen over the last years, from genomics, epigenomics, transcriptomics to proteomics and metabolomics. GWAS is considered part of the most mature omics field, namely genomics and as presented its data can be utilised to assess the correlation between different traits, assess causality and calculate PRS. Large amounts of biological data are generated with analysing omics with high-throughput technologies. Each field of omics can be depicted as a layer, where there is a connection within different components of each layer (intra-layer relationship), for example different metabolites relate to each other within the metabolomics layer. In addition the different omics field layers also have a biological effect on each other (inter-layer relationships), so the epigenomics layer influences which genes are transcribed (**Figure 1**). All these omics relationships give new possibilities to further investigate the pathophysiology of disease. Studying isolated biological components/layers is not enough to understand biological systems. Integrating this multi-omics data can identify novel biological pathways that are not necessarily distinguishable in the individual omics layers. It has already been shown that integrating DNA copy number, loss of heterozygosity, and methylation led to an increase in explained gene expression changes in breast cancer that would otherwise have been overlooked in single layer analysis.[149] Integrating these layers has the potential to give a more comprehensive and deeper understanding of biological systems, thus a more holistic view on the systems biology.[150] The field of multi-omics integration is relatively new field with a lot of new developments.

From SNP to gene

Over the last two decades GWAS has identified thousands of loci for a large number of human diseases. However, a major short coming of GWAS is that the associated loci rarely reflect the causal variants, target genes, cell types and biological functions. The majority of risk loci in GWAS fall into non-coding regions of the genome. Even when they are located in a coding region it is often unclear whether this is a functional variant or a variant in LD with the functional variant. Genetic variants can modulate transcription of target genes up to several megabase pairs (Mbp)

away, when the variants are located in regulatory elements. Many statistical methods have been developed to prioritize the risk loci by integrating functional evidence. These computational tools often use the genomic structure, eQTLs, transcription factor binding sites, histone modifications and others for gene prioritization. In this thesis we have mapped the risk variants to causal genes and identified relevant gene sets using a variety of such tools, DEPICT, eQTL mapping, MAGMA, FOCUS, transcriptome-wide association studies (TWAS).[103, 151-155]

Figure 1 Multi-layer network



In this multi-layer network, each network represents an independent layer. A group of molecules with similar chemical properties, such as genome, transcriptome, proteome, and metabolome, is called an 'omic' layer, which can be measured by next-generation sequencers (NGS), microarray, mass spectrometry, and nuclear magnetic resonance (1H-NMR) spectroscopy. The common nodes, or identified metabolites, proteins, genes etc., are connected to each other across different layers by inter-layer edges. The edges within individual layers and between them can help to understand biological systems.

Pathway analysis

Pathway analysis is also known as functional enrichment analysis. The main purpose is to give biological context to the lists of loci, genes or other biological data generated by high-throughput pipelines. The vast amount of data generated makes it impossible to model by a naïve approach.

Therefore, powerful mathematical approaches are necessary to couple biological knowledge with biological omics data. There are two main statistical difficulties with integrating data: (1) the dimensionality, were the number of variables is greater than the number of samples and (2) the development of algorithms that are able to integrate and analyse biological information based on the most up-to-date information.[119] Currently there is an enormous number of pathway databases that store pathway topology information and that can be used for data integration, focussing on different networks such as databases on metabolic pathways, signalling pathways, transcription factor targets, gene regulatory networks, genetic interactions, protein-compound interactions, and protein-protein interactions.[156] Some well-known pathway databases are KEGG, Gene Ontology and REACTOME. Important to note is that not all database update their information as regularly as others. Pathway analysis software help to combine the information from the omics platforms with the pathway databases and are able to perform all statistical and mathematical computations. Although all software uses different statistical methods they fundamentally test the same thing, the possibility that any given pathway is represented by the high throughput data.[156] Choosing the best software platform depends on the hypothesis that needs answering and the user skills. In this thesis we have used DEPICT, FOCUS, TWAS, FUMA and LDSC-SEG for the identification of relevant tissues, cell types and pathways.[103, 153-155, 157, 158]

Multi-omics

Integrated approaches combine single-layer omics data to understand the interplay of molecules and help bridge the gap from genotype to phenotype. Integrative approaches can be more or less stringent on the types of omics considered as input, some methods are designed for a specific combination of datasets, while others are more general. In addition tools can differ in respect to sequential and simultaneous analysis of multiple layers. In addition different types of mathematics are used in current integration tools.[159] The most challenging part of the field are the mathematical methodologies and interpretation of data, because of the complexity of biological systems, the technological limits, the many biological variables and the relatively low number of biological samples.[159, 160] In addition, data quality remains crucial for each omics layer, hence the community should also focus on standardization of sample quality, sample analysis pipelines, data analysis pipelines and data formats. Given the vast amount of data on genomics of cluster headache and migraine in previous studies, and in this thesis (**Chapters 6**, **7 and 8**), together with the ever growing field of metabolites and metabolomics (**Chapters 2 - 5**) a promising future research field would be integration of omics data.

References

1. Hagenbeek FA, Pool R, van Dongen J, et al. Heritability estimates for 361 blood metabolites across 40 genome-wide association studies. *Nat Commun*. 2020;11(1):39.
2. Bigal ME, Lipton RB. Obesity is a risk factor for transformed migraine but not chronic tension-type headache. *Neurology*. 2006;67(2):252-257.
3. Linstra KM, Ibrahimi K, Terwindt GM, Wermer MJ, MaassenVanDenBrink A. Migraine and cardiovascular disease in women. *Maturitas*. 2017;97:28-31.
4. Kurth T, Rist PM, Ridker PM, et al. Association of Migraine With Aura and Other Risk Factors With Incident Cardiovascular Disease in Women. *JAMA*. 2020;323(22):2281-2289.
5. Kurth T, Chabriat H, Bousser MG. Migraine and stroke: a complex association with clinical implications. *Lancet Neurol*. 2012;11(1):92-100.
6. Tanha HM, Martin NG, Whitfield JB, Nyholt DR. Association and genetic overlap between clinical chemistry tests and migraine. *Cephalalgia*. 2021;41(11-12):1208-1221.
7. Tanha HM, Sathyanarayanan A, Nyholt DR. Genetic overlap and causality between blood metabolites and migraine. *Am J Hum Genet*. 2021;108(11):2086-2098.
8. Tietjen GE, Khubchandani J. Vascular biomarkers in migraine. *Cephalalgia*. 2015;35(2):95-117.
9. Onderwater GLJ, Ligthart L, Bot M, et al. Large-scale plasma metabolome analysis reveals alterations in HDL metabolism in migraine. *Neurology*. 2019;92(16):e1899-e1911.
10. Rubino E, Vacca A, Govone F, et al. Investigating the role of adipokines in chronic migraine. *Cephalalgia*. 2017;37(11):1067-1073.
11. Guo Y, Daghlas I, Gormley P, et al. Phenotypic and Genotypic Associations Between Migraine and Lipoprotein Subfractions. *Neurology*. 2021;97(22):e2223-e2235.
12. Ferrari MD, Odink J, Bos KD, Malessy MJ, Bruyn GW. Neuroexcitatory plasma amino acids are elevated in migraine. *Neurology*. 1990;40(10):1582-1586.
13. Martinez F, Castillo J, Rodriguez JR, Leira R, Noya M. Neuroexcitatory amino acid levels in plasma and cerebrospinal fluid during migraine attacks. *Cephalalgia*. 1993;13(2):89-93.
14. Cananzi AP, Dandrea G, Perini F, Zamberlan F, Welch KMA. Platelet and Plasma-Levels of Glutamate and Glutamine in Migraine with and without Aura. *Cephalalgia*. 1995;15(2):132-135.
15. Zielman R, Wijnen JP, Webb A, et al. Cortical glutamate in migraine. *Brain*. 2017;140(7):1859-1871.
16. Hoffmann J, Charles A. Glutamate and Its Receptors as Therapeutic Targets for Migraine. *Neurotherapeutics*. 2018;15(2):361-370.
17. Folger O, Jerby L, Frezza C, et al. Predicting selective drug targets in cancer through metabolic networks. *Mol Syst Biol*. 2011;7:501.
18. Li L, Zhou X, Ching WK, Wang P. Predicting enzyme targets for cancer drugs by profiling human metabolic reactions in NCI-60 cell lines. *BMC Bioinformatics*. 2010;11:501.
19. Caterina MJ, Schumacher MA, Tominaga M, et al. The capsaicin receptor: a heat-activated ion channel in the pain pathway. *Nature*. 1997;389(6653):816-824.
20. Vriens J, Appendino G, Nilius B. Pharmacology of vanilloid transient receptor potential cation channels. *Mol Pharmacol*. 2009;75(6):1262-79.
21. Saloman JL, Chung MK, Ro JY. $P2X_3$ and TRPV1 functionally interact and mediate sensitization of trigeminal sensory neurons. *Neuroscience*. 2013;232:226-238.
22. Caterina MJ, Leffler A, Malmberg AB, et al. Impaired nociception and pain sensation in mice lacking the capsaicin receptor. *Science*. 2000;288(5464):306-313.
23. Tominaga M, Caterina MJ, Malmberg AB, et al. The cloned capsaicin receptor integrates multiple pain-producing stimuli. *Neuron*. 1998;21(3):531-543.
24. Meents JE, Neeb L, Reuter U. TRPV1 in migraine pathophysiology. *Trends Mol Med*. 2010;16(4):153-159.

25. Marsicano G, Goodenough S, Monory K, et al. CB1 cannabinoid receptors and on-demand defense against excitotoxicity. *Science*. 2003;302(5642):84-88.

26. Rouzer CA, Ghebreselasie K, Marnett LJ. Chemical stability of 2-arachidonylglycerol under biological conditions. *Chem Phys Lipids*. 2002;119(1-2):69-82.

27. Willoughby KA, Moore SF, Martin BR, Ellis EF. The biodisposition and metabolism of anandamide in mice. *J Pharmacol Exp Ther*. 1997;282(1):243-247.

28. Rea K, Roche M, Finn DP. Supraspinal modulation of pain by cannabinoids: the role of GABA and glutamate. *Br J Pharmacol*. 2007;152(5):633-648.

29. Szabo B, Schlicker E. Effects of Cannabinoids on Neurotransmission. In: Pertwee RG, ed. *Cannabinoids*. Springer Berlin Heidelberg; 2005:327-365.

30. Engel MA, Izydorczyk I, Mueller-Tribbensee SM, et al. Inhibitory CB1 and activating/desensitizing TRPV1-mediated cannabinoid actions on CGRP release in rodent skin. *Neuropeptides*. 2011;45(3):229-237.

31. Akerman S, Kaube H, Goadsby PJ. Anandamide is able to inhibit trigeminal neurons using an in vivo model of trigeminovascular-mediated nociception. *J Pharmacol Exp Ther*. 2004;309(1):56-63.

32. Akerman S, Kaube H, Goadsby PJ. Anandamide acts as a vasodilator of dural blood vessels in vivo by activating TRPV1 receptors. *Br J Pharmacol*. 2004;142(8):1354-1360.

33. Di Marzo V, De Petrocellis L. Why do cannabinoid receptors have more than one endogenous ligand? *Philos Trans R Soc Lond B Biol Sci*. 2012;367(1607):3216-3228.

34. Starowicz K, Przewlocka B. Modulation of neuropathic-pain-related behaviour by the spinal endocannabinoid/endovanilloid system. *Philos Trans R Soc Lond B Biol Sci*. 2012;367(1607):3286-3299.

35. Breyer RM, Bagdassarian CK, Myers SA, Breyer MD. Prostanoid receptors: subtypes and signaling. *Annu Rev Pharmacol Toxicol*. 2001;41:661-690.

36. Moriyama T, Higashi T, Togashi K, et al. Sensitization of TRPV1 by EP1 and IP reveals peripheral nociceptive mechanism of prostaglandins. *Mol Pain*. 2005;1:3.

37. Ma W, St-Jacques B, Rudakou U, Kim YN. Stimulating TRPV1 externalization and synthesis in dorsal root ganglion neurons contributes to PGE2 potentiation of TRPV1 activity and nociceptor sensitization. *Eur J Pain*. 2017;21(4):575-593.

38. Jang Y, Kim M, Hwang SW. Molecular mechanisms underlying the actions of arachidonic acid-derived prostaglandins on peripheral nociception. *J Neuroinflammation*. 2020;17(1):30.

39. Jenkins DW, Feniuk W, Humphrey PP. Characterization of the prostanoid receptor types involved in mediating calcitonin gene-related peptide release from cultured rat trigeminal neurones. *Br J Pharmacol*. 2001;134(6):1296-1302.

40. Neeb L, Hellen P, Boehnke C, et al. IL-1beta stimulates COX-2 dependent PGE(2) synthesis and CGRP release in rat trigeminal ganglia cells. *PLoS One*. 2011;6(3):e17360.

41. Jenkins DW, Langmead CJ, Parsons AA, Strijbos PJ. Regulation of calcitonin gene-related peptide release from rat trigeminal nucleus caudalis slices in vitro. *Neurosci Lett*. 2004;366(3):241-244.

42. Woodward DF, Carling RW, Cornell CL, et al. The pharmacology and therapeutic relevance of endocannabinoid derived cyclo-oxygenase (COX)-2 products. *Pharmacol Ther*. 2008;120(1):71-80.

43. Del Fiacco M, Quartu M, Boi M, et al. TRPV1, CGRP and SP in scalp arteries of patients suffering from chronic migraine. *J Neurol Neurosurg Psychiatry*. 2015;86(4):393-397.

44. Benemei S, Dussor G. TRP Channels and Migraine: Recent Developments and New Therapeutic Opportunities. *Pharmaceuticals (Basel)*. 2019;12(2)

45. Tfelt-Hansen P, Messlinger K. Why is the therapeutic effect of acute antimigraine drugs delayed? A review of controlled trials and hypotheses about the delay of effect. *Br J Clin Pharmacol*. 2019;85(11):2487-2498.

46. Schoenen J, Manise M, Nonis R, Gérard P, Timmermans G. Monoclonal antibodies blocking CGRP transmission: An update on their added value in migraine prevention. *Rev Neurol (Paris)*. 2020;176(10):788-803.

47. Greco R, Demartini C, Zanaboni AM, et al. The endocannabinoid system and related lipids as potential targets for the treatment of migraine-related pain. *Headache*. 2022;62(3):227-240.

48. Fusco BM, Barzoi G, Agrò F. Repeated intranasal capsaicin applications to treat chronic migraine. *Br J Anaesth*. 2003;90(6):812.

49. Lambert GA, Davis JB, Appleby JM, et al. The effects of the TRPV1 receptor antagonist SB-705498 on trigeminovascular sensitisation and neurotransmission. *Naunyn Schmiedebergs Arch Pharmacol*. 2009;380(4):311-325.

50. Aghazadeh Tabrizi M, Baraldi PG, Baraldi S, et al. Medicinal Chemistry, Pharmacology, and Clinical Implications of TRPV1 Receptor Antagonists. *Med Res Rev*. 2017;37(4):936-983.

51. Eltorp CT, Jansen-Olesen I, Hansen AJ. Release of calcitonin gene-related peptide (CGRP) from guinea pig dura mater in vitro is inhibited by sumatriptan but unaffected by nitric oxide. *Cephalalgia*. 2000;20(9):838-844.

52. Dodick DW, Ashina M, Brandes JL, et al. ARISE: A Phase 3 randomized trial of erenumab for episodic migraine. SAGE Publications Ltd STM; 2018. p. 1026-1037.

53. Antonova M, Wienecke T, Maubach K, et al. The pharmacological effect of BGC20-1531, a novel prostanoid EP4 receptor antagonist, in the prostaglandin E2 human model of headache. *J Headache Pain*. 2011;12(5):551-559.

54. Cuttler C, Spradlin A, Cleveland MJ, Craft RM. Short- and Long-Term Effects of Cannabis on Headache and Migraine. *J Pain*. 2020;21(5-6):722-730.

55. Malek N, Starowicz K. Dual-Acting Compounds Targeting Endocannabinoid and Endovanilloid Systems-A Novel Treatment Option for Chronic Pain Management. *Front Pharmacol*. 2016;7:257.

56. Dallmann R, Viola AU, Tarokh L, Cajochen C, Brown SA. The human circadian metabolome. *Proc Natl Acad Sci U S A*. 2012;109(7):2625-2629.

57. Kasukawa T, Sugimoto M, Hida A, et al. Human blood metabolite timetable indicates internal body time. *Proc Natl Acad Sci U S A*. 2012;109(37):15036-15041.

58. Lehmann R. From bedside to bench-practical considerations to avoid pre-analytical pitfalls and assess sample quality for high-resolution metabolomics and lipidomics analyses of body fluids. *Anal Bioanal Chem*. 2021;413(22):5567-5585.

59. Chua EC, Shui G, Lee IT, et al. Extensive diversity in circadian regulation of plasma lipids and evidence for different circadian metabolic phenotypes in humans. *Proc Natl Acad Sci U S A*. 2013;110(35):14468-14473.

60. Trabado S, Al-Salameh A, Croixmarie V, et al. The human plasma-metabolome: Reference values in 800 French healthy volunteers; impact of cholesterol, gender and age. *PLoS One*. 2017;12(3):e0173615.

61. Darst BF, Koscik RL, Hogan KJ, Johnson SC, Engelman CD. Longitudinal plasma metabolomics of aging and sex. *Aging (Albany N Y)*. 2019;11(4):1262-1282.

62. Sarchielli P, Pini LA, Coppola F, et al. Endocannabinoids in chronic migraine: CSF findings suggest a system failure. *Neuropsychopharmacology*. 2007;32(6):1384-1390.

63. Nattero G, Allais G, De Lorenzo C, et al. Relevance of prostaglandins in true menstrual migraine. *Headache*. 1989;29(4):233-238.

64. Tuca JO, Planas JM, Parellada PP. Increase in PGE2 and TXA2 in the saliva of common migraine patients. Action of calcium channel blockers. *Headache*. 1989;29(8):498-501.

65. Vardi J, Flechter S, Alguati A, Regev I, Ayalon D. Prostaglandin--E2 levels in the saliva of common migrainous women. *Headache*. 1983;23(2):59-61.

66. Sarchielli P, Alberti A, Codini M, Floridi A, Gallai V. Nitric oxide metabolites, prostaglandins and trigeminal vasoactive peptides in internal jugular vein blood during spontaneous migraine attacks. *Cephalalgia*. 2000;20(10):907-918.

67. Boheler KR, Volkova M, Morrell C, et al. Sex- and age-dependent human transcriptome variability: implications for chronic heart failure. *Proc Natl Acad Sci U S A*. 2003;100(5):2754-2759.

68. Scher AI, Terwindt GM, Picavet HS, et al. Cardiovascular risk factors and migraine: the GEM population-based study. *Neurology*. 2005;64(4):614-620.

69. Hagen K, Åsberg AN, Stovner L, et al. Lifestyle factors and risk of migraine and tension-type headache. Follow-up data from the Nord-Trøndelag Health Surveys 1995-1997 and 2006-2008. *Cephalalgia*. 2018;38(13):1919-1926.

70. Le H, Tfelt-Hansen P, Skytthe A, Kyvik KO, Olesen J. Association between migraine, lifestyle and socioeconomic factors: a population-based cross-sectional study. *J Headache Pain*. 2011;12(2):157-172.

71. Pourhoseingholi MA, Baghestani AR, Vahedi M. How to control confounding effects by statistical analysis. *Gastroenterol Hepatol Bed Bench*. 2012;5(2):79-83.

72. McNamee R. Regression modelling and other methods to control confounding. *Occup Environ Med*. 2005;62(7):500-6, 472.

73. Hillard CJ. Circulating Endocannabinoids: From Whence Do They Come and Where are They Going? *Neuropsychopharmacology*. 2018;43(1):155-172.

74. Engeli S, Böhnke J, Feldpausch M, et al. Activation of the peripheral endocannabinoid system in human obesity. *Diabetes*. 2005;54(10):2838-2843.

75. Piyanova A, Lomazzo E, Bindila L, et al. Age-related changes in the endocannabinoid system in the mouse hippocampus. *Mech Ageing Dev*. 2015;150:55-64.

76. Hill MN, Miller GE, Ho WS, Gorzalka BB, Hillard CJ. Serum endocannabinoid content is altered in females with depressive disorders: a preliminary report. *Pharmacopsychiatry*. 2008;41(2):48-53.

77. Kranaster L, Hoyer C, Aksay SS, et al. Electroconvulsive therapy enhances endocannabinoids in the cerebrospinal fluid of patients with major depression: a preliminary prospective study. *Eur Arch Psychiatry Clin Neurosci*. 2017;267(8):781-786.

78. Hill MN, Miller GE, Carrier EJ, Gorzalka BB, Hillard CJ. Circulating endocannabinoids and N-acyl ethanolamines are differentially regulated in major depression and following exposure to social stress. *Psychoneuroendocrinology*. 2009;34(8):1257-1262.

79. Patel S, Hillard CJ. Role of endocannabinoid signaling in anxiety and depression. *Curr Top Behav Neurosci*. 2009;1:347-371.

80. Gallego-Landin I, García-Baos A, Castro-Zavala A, Valverde O. Reviewing the Role of the Endocannabinoid System in the Pathophysiology of Depression. Review. *Front Pharmacol*. 2021;12

81. Smaga I, Bystrowska B, Gawliński D, Przegaliński E, Filip M. The endocannabinoid/endovanilloid system and depression. *Curr Neuropharmacol*. 2014;12(5):462-474.

82. Piazza PV, Cota D, Marsicano G. The CB1 Receptor as the Cornerstone of Exostasis. *Neuron*. 2017;93(6):1252-1274.

83. Blüher M, Engeli S, Klöting N, et al. Dysregulation of the peripheral and adipose tissue endocannabinoid system in human abdominal obesity. *Diabetes*. 2006;55(11):3053-3060.

84. Wagner EJ. Sex differences in cannabinoid-regulated biology: A focus on energy homeostasis. *Front Neuroendocrinol*. 2016;40:101-109.

85. Rubino T, Parolaro D. Sexually dimorphic effects of cannabinoid compounds on emotion and cognition. *Front Behav Neurosci*. 2011;5:64.

86. Craft RM, Marusich JA, Wiley JL. Sex differences in cannabinoid pharmacology: a reflection of differences in the endocannabinoid system? *Life Sci*. 2013;92(8-9):476-481.

87. Alvarez-Jaimes L, Stouffer DG, Parsons LH. Chronic ethanol treatment potentiates ethanol-induced increases in interstitial nucleus accumbens endocannabinoid levels in rats. *J Neurochem*. 2009;111(1):37-48.

88. Basavarajappa BS, Hungund BL. Chronic ethanol increases the cannabinoid receptor agonist anandamide and its precursor N-arachidonoylphosphatidylethanolamine in SK-N-SH cells. *J Neurochem*. 1999;72(2):522-528.

89. Vinod KY, Yalamanchili R, Xie S, Cooper TB, Hungund BL. Effect of chronic ethanol exposure and its withdrawal on the endocannabinoid system. *Neurochem Int*. 2006;49(6):619-625.

90. Scherma M, Fadda P, Le Foll B, et al. The endocannabinoid system: a new molecular target for the treatment of tobacco addiction. *CNS Neurol Disord Drug Targets*. 2008;7(5):468-481.

91. Hirvonen J, Zanotti-Fregonara P, Gorelick DA, et al. Decreased Cannabinoid CB1 Receptors in Male Tobacco Smokers Examined With Positron Emission Tomography. *Biol Psychiatry*. 2018;84(10):715-721.

92. Gouveia-Figueira S, Goldin K, Hashemian SA, et al. Plasma levels of the endocannabinoid anandamide, related N-acylethanolamines and linoleic acid-derived oxylipins in patients with migraine. *Prostaglandins Leukot Essent Fatty Acids*. 2017;120:15-24.

93. Tilley SL, Coffman TM, Koller BH. Mixed messages: modulation of inflammation and immune responses by prostaglandins and thromboxanes. *J Clin Invest*. 2001;108(1):15-23.

94. Pradhan SS, Salinas K, Garduno AC, et al. Anti-Inflammatory and Neuroprotective Effects of PGE(2) EP4 Signaling in Models of Parkinson's Disease. *J Neuroimmune Pharmacol*. 2017;12(2):292-304.

95. Liang X, Wang Q, Shi J, et al. The prostaglandin E2 EP2 receptor accelerates disease progression and inflammation in a model of amyotrophic lateral sclerosis. *Ann Neurol*. 2008;64(3):304-314.

96. Andreasson K. Emerging roles of PGE2 receptors in models of neurological disease. *Prostaglandins Other Lipid Mediat*. 2010;91(3-4):104-112.

97. Yin F, Sancheti H, Patil I, Cadenas E. Energy metabolism and inflammation in brain aging and Alzheimer's disease. *Free Radic Biol Med*. 2016;100:108-122.

98. Albrecht DS, Mainero C, Ichijo E, et al. Imaging of neuroinflammation in migraine with aura: A [(11)C]PBR28 PET/MRI study. *Neurology*. 2019;92(17):e2038-e2050.

99. Wu D, Mura C, Beharka AA, et al. Age-associated increase in PGE2 synthesis and COX activity in murine macrophages is reversed by vitamin E. *Am J Physiol*. 1998;275(3):C661-668.

100. Wu D, Meydani SN. Mechanism of age-associated up-regulation in macrophage PGE2 synthesis. *Brain Behav Immun*. 2004;18(6):487-494.

101. Hayek MG, Mura C, Wu D, et al. Enhanced expression of inducible cyclooxygenase with age in murine macrophages. *J Immunol*. 1997;159(5):2445-2451.

102. Visscher PM, Wray NR, Zhang Q, et al. 10 Years of GWAS Discovery: Biology, Function, and Translation. *Am J Hum Genet*. 2017;101(1):5-22.

103. Gusev A, Ko A, Shi H, et al. Integrative approaches for large-scale transcriptome-wide association studies. *Nat Genet*. 2016;48(3):245-252.

104. Zhu Z, Zhang F, Hu H, et al. Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. *Nat Genet*. 2016;48(5):481-487.

105. O'Connor E, Fourier C, Ran C, et al. Genome-Wide Association Study Identifies Risk Loci for Cluster Headache. *Ann Neurol*. 2021;90(2):193-202.

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

107. Chasman DI, Schürks M, Anttila V, et al. Genome-wide association study reveals three susceptibility loci for common migraine in the general population. *Nat Genet*. 2011;43(7):695-698.

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

109. Winsvold BS, Nelson CP, Malik R, et al. Genetic analysis for a shared biological basis between migraine and coronary artery disease. *Neurol Genet*. 2015;1(1):e10.

110. Malik R, Freilinger T, Winsvold BS, et al. Shared genetic basis for migraine and ischemic stroke: A genome-wide analysis of common variants. *Neurology*. 2015;84(21):2132-2145.

111. Diener H-C. CGRP antibodies for migraine prevention — new kids on the block. *Nature Reviews Neurology*. 2019;15(3):129-130.

112. de Vries T, Villalón CM, Maassen VanDenBrink A. Pharmacological treatment of migraine: CGRP and 5-HT beyond the triptans. *Pharmacol Ther*. 2020;211:107528.

113. Ophoff RA, Terwindt GM, Vergouwe MN, et al. Familial hemiplegic migraine and episodic ataxia type-2 are caused by mutations in the Ca2+ channel gene CACNL1A4. *Cell*. 1996;87(3):543-552.

114. Dichgans M, Freilinger T, Eckstein G, et al. Mutation in the neuronal voltage-gated sodium channel SCN1A in familial hemiplegic migraine. *Lancet*. 2005;366(9483):371-377.

115. De Fusco M, Marconi R, Silvestri L, et al. Haploinsufficiency of ATP1A2 encoding the Na+/K+ pump alpha2 subunit associated with familial hemiplegic migraine type 2. *Nat Genet*. 2003;33(2):192-196.

116. Choquet H, Yin J, Jacobson AS, et al. New and sex-specific migraine susceptibility loci identified from a multiethnic genome-wide meta-analysis. *Commun Biol*. 2021;4(1):864.

117. Santiago JA, Potashkin JA. Network-based metaanalysis identifies HNF4A and PTBP1 as longitudinally dynamic biomarkers for Parkinson's disease. *Proc Natl Acad Sci U S A*. 2015;112(7):2257-2262.

118. Kotelnikova E, Shkrob MA, Pyatnitskiy MA, Ferlini A, Daraselia N. Novel approach to meta-analysis of microarray datasets reveals muscle remodeling-related drug targets and biomarkers in Duchenne muscular dystrophy. *PLoS Comput Biol*. 2012;8(2):e1002365.

119. García-Campos MA, Espinal-Enríquez J, Hernández-Lemus E. Pathway Analysis: State of the Art. *Front Physiol*. 2015;6:383.

120. Mavaddat N, Michailidou K, Dennis J, et al. Polygenic Risk Scores for Prediction of Breast Cancer and Breast Cancer Subtypes. *Am J Hum Genet*. 2019;104(1):21-34.

121. Kogelman LJA, Esserlind A-L, Francke Christensen A, et al. Migraine polygenic risk score associates with efficacy of migraine-specific drugs. *Neurology Genetics*. 2019;5(6):e364.

122. Lee KJ, Lee SJ, Bae HJ, Sung J. Exploring the causal inference of migraine on stroke: A Mendelian randomization study. *Eur J Neurol*. 2022;29(1):335-338.

123. Shu MJ, Li JR, Zhu YC, Shen H. Migraine and Ischemic Stroke: A Mendelian Randomization Study. *Neurol Ther*. 2021;11(1):237-246.

124. Daghals I, Sargurupremraj M, Danning R, et al. Migraine, Stroke, and Cervical Arterial Dissection: Shared Genetics for a Triad of Brain Disorders With Vascular Involvement. *Neurol Genet*. 2022;8(1):e653.

125. Siewert KM, Klarin D, Damrauer SM, et al. Cross-trait analyses with migraine reveal widespread pleiotropy and suggest a vascular component to migraine headache. *Int J Epidemiol*. 2020;49(3):1022-1031.

126. Guo Y, Rist PM, Daghlas I, et al. A genome-wide cross-phenotype meta-analysis of the association of blood pressure with migraine. *Nat Commun*. 2020;11(1):3368.

127. Daghlas I, Guo Y, Chasman DI. Effect of genetic liability to migraine on coronary artery disease and atrial fibrillation: a Mendelian randomization study. *Eur J Neurol*. 2020;27(3):550-556.

128. Yuan S, Daghlas I, Larsson SC. Alcohol, coffee consumption, and smoking in relation to migraine: a bidirectional Mendelian randomization study. *Pain*. 2022;163(2):e342-e348.

129. Johnsen MB, Winsvold BS, Børte S, et al. The causal role of smoking on the risk of headache. A Mendelian randomization analysis in the HUNT study. *Eur J Neurol*. 2018;25(9):1148-e102.

130. Ferrari A, Zappaterra M, Righi F, et al. Impact of continuing or quitting smoking on episodic cluster headache: a pilot survey. *J Headache Pain*. 2013;14:48.

131. Lund N, Petersen A, Snoer A, Jensen RH, Barloese M. Cluster headache is associated with unhealthy lifestyle and lifestyle-related comorbid diseases: Results from the Danish Cluster Headache Survey. *Cephalalgia*. 2019;39(2):254-263.

132. May A, Schwedt TJ, Magis D, et al. Cluster headache. *Nat Rev Dis Primers*. 2018;4:18006.

133. Schurks M, Kurth T, de Jesus J, et al. Cluster headache: clinical presentation, lifestyle features, and medical treatment. *Headache*. 2006;46(8):1246-1254.

134. Agrawal A, Verweij KJ, Gillespie NA, et al. The genetics of addiction-a translational perspective. *Transl Psychiatry*. 2012;2:e140.

135. Rozen TD. Cluster Headache Clinical Phenotypes: Tobacco Nonexposed (Never Smoker and No Parental Secondary Smoke Exposure as a Child) versus Tobacco-Exposed: Results from the United States Cluster Headache Survey. *Headache*. 2018;58(5):688-699.

136. O'Connor LJ, Price AL. Distinguishing genetic correlation from causation across 52 diseases and complex traits. *Nat Genet*. 2018;50(12):1728-1734.

137. Gormley P, Kurki MI, Hiekkala ME, et al. Common Variant Burden Contributes to the Familial Aggregation of Migraine in 1,589 Families. *Neuron*. 2018;99(5):1098.

138. Pelzer N, de Vries B, Kamphorst JT, et al. PRRT2 and hemiplegic migraine: a complex association. *Neurology*. 2014;83(3):288-290.

139. Gardiner AR, Jaffer F, Dale RC, et al. The clinical and genetic heterogeneity of paroxysmal dyskinesias. *Brain*. 2015;138(Pt 12):3567-3580.

140. Suzuki M, Van Paesschen W, Stalmans I, et al. Defective membrane expression of the Na(+)-HCO(3)(-) cotransporter NBCe1 is associated with familial migraine. *Proc Natl Acad Sci U S A*. 2010;107(36):15963-15968.

141. Jen JC, Wan J, Palos TP, Howard BD, Baloh RW. Mutation in the glutamate transporter EAAT1 causes episodic ataxia, hemiplegia, and seizures. *Neurology*. 2005;65(4):529-534.

142. Weller CM, Leen WG, Neville BG, et al. A novel SLC2A1 mutation linking hemiplegic migraine with alternating hemiplegia of childhood. *Cephalalgia*. 2015;35(1):10-15.

143. Gormley P, Kurki M, Kurki MI, et al. Common Variant Burden Contributes to the Familial Aggregation of Migraine in 1,589 Families. *Neuron*. 2018;98(4):743-753.e4.

144. Campbell C, Leu C, Feng YA, et al. The role of common genetic variation in presumed monogenic epilepsies. *EBioMedicine*. 2022;81:104098.

145. Diogo D, Kurreeman F, Stahl EA, et al. Rare, low-frequency, and common variants in the protein-coding sequence of biological candidate genes from GWASs contribute to risk of rheumatoid arthritis. *Am J Hum Genet*. 2013;92(1):15-27.

146. Ionita-Laza I, Lee S, Makarov V, Buxbaum JD, Lin X. Sequence kernel association tests for the combined effect of rare and common variants. *Am J Hum Genet*. 2013;92(6):841-853.

147. Techlo TR, Rasmussen AH, Møller PL, et al. Familial analysis reveals rare risk variants for migraine in regulatory regions. *Neurogenetics*. 2020;21(3):149-157.

148. Rasmussen AH, Kogelman LJA, Kristensen DM, et al. Functional gene networks reveal distinct mechanisms segregating in migraine families. *Brain*. 2020;143(10):2945-2956.

149. Chari R, Coe BP, Vucic EA, Lockwood WW, Lam WL. An integrative multi-dimensional genetic and epigenetic strategy to identify aberrant genes and pathways in cancer. *BMC Syst Biol*. 2010;4:67.

150. Subramanian I, Verma S, Kumar S, Jere A, Anamika K. Multi-omics Data Integration, Interpretation, and Its Application. *Bioinform Biol Insights*. 2020;14:1177932219899051.

151. de Leeuw CA, Mooij JM, Heskes T, Posthuma D. MAGMA: generalized gene-set analysis of GWAS data. *PLoS Comput Biol*. 2015;11(4):e1004219.

152. Kerimov N, Hayhurst JD, Manning JR, et al. eQTL Catalogue: a compendium of uniformly processed human gene expression and splicing QTLs. *bioRxiv*. 2020:2020.01.29.924266.

153. Pers TH, Karjalainen JM, Chan Y, et al. Biological interpretation of genome-wide association studies using predicted gene functions. *Nature Communications*. 2015;6:5890

154. Mancuso N, Freund MK, Johnson R, et al. Probabilistic fine-mapping of transcriptome-wide association studies. *Nat Genet*. 2019;51(4):675-682.

155. Barbeira AN, Dickinson SP, Bonazzola R, et al. Exploring the phenotypic consequences of tissue specific gene expression variation inferred from GWAS summary statistics. *Nat Commun*. 2018;9(1):1825.

156. Bader GD, Cary MP, Sander C. Pathguide: a pathway resource list. *Nucleic Acids Res*. 2006;34(Database issue):D504-506.

157. Finucane HK, Reshef YA, Anttila V, et al. Heritability enrichment of specifically expressed genes identifies disease-relevant tissues and cell types. *Nat Genet*. 2018;50(4):621-629.

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

159. Bersanelli M, Mosca E, Remondini D, et al. Methods for the integration of multi-omics data: mathematical aspects. *BMC Bioinformatics*. 2016;17(2):S15.

160. Misra BB, Langefeld C, Olivier M, Cox LA. Integrated omics: tools, advances and future approaches. *J Mol Endocrinol*. 2019;62(1):R21-R45.

APPENDICES

Summary

This thesis explores biochemical and genetical changes associated with migraine and cluster headache. The research aimed to enhance our knowledge on the pathophysiology and pathways involved in these diseases and hopefully will provide first steps in the identification of new treatment targets. Uncovering the biological mechanisms on how patients differ from those without disease leads to a better understanding of the pathophysiology of primary headache disorders. The biological systems in our body are related to each other, and are based on the genomic blueprint and lead via epigenetics, transcription and translation to proteins and biomolecules. By investigating the various molecular levels, we hope to get a better insight in how headache disorders are brought about. The research is divided into two parts. **Part 1**, (**Chapters 2 - 5**) is aimed at the investigation of biomolecules in biofluids (i.e. blood and cerebrospinal fluid (CSF)) in migraine patients, whereas **Part 2** (**Chapters 6 - 9**) is focused on the genetical underpinnings of cluster headache migraine and related disorders.

Part I Biochemistry of migraine

Part 1 describes biochemical studies for which various molecules and metabolites were measured in people with migraine and compared to those in people without migraine. Migraine is a primary headache disorder that is characterized by recurrent episodes of severe often unilateral pulsating headache accompanied by nausea, vomiting and/or photo- and phonophobia. In one third of patients the attacks are preceded by an aura phase. Typically, a migraine attack consists of a preictal, ictal (aura and/or headache), and postictal (postdromal) phase. Although over the years much progress has been made with understanding the pathophysiological mechanisms involved in migraine, it is still unclear why some people get migraine and others do not and why those with migraine get attacks at certain timepoints, and which factors trigger attacks. The aim of **Part 1** was to uncover biochemical underpinnings of migraine to better understand the pathways involved. The studies focused on the measurement of a wide variety of biomolecules, either via a targeted or untargeted approach. The studies focused on both blood and CSF and investigated the comparability of the two biofluids specifically in relation to amine concentrations. Furthermore, in all but one chapter (**Chapters 2,3 and 4**) we investigated these biofluids outside of an attack and in the remaining chapter (**Chapter 5**) over the course of a provoked migraine attack. This way different aspects of migraine were put under investigation.

For **Chapter 2** the metabolic profile of migraine patients was compared to those without migraine using an untargeted approach based on high-throughput proton nuclear magnetic resonance (^{1}H-NMR) spectroscopy. In total, 100 signals representing 49 different metabolites in the fasting serum samples from 189 migraine patients and 1,360 controls were detected. Migraine status was divided in lifetime migraine (migraine diagnosis during lifetime) and active migraine (defined as having at least one severe migraine in the last 12 months). Using elastic net regression analysis

predictive profiles consisting of 6 metabolites for lifetime migraine status and 22 metabolites for active migraine status were identified. These predictive metabolites included lipids, amino acids, and metabolites of glucose metabolism. This data suggests an overall disturbance of the metabolic profile in migraine patients.

In **Chapter 3** the aim was to determine the relationship in amine levels between blood plasma and CSF. The CSF is the biofluid of choice in central nervous system disorders, such as migraine, as it is closest to the brain and thus best reflects the brains neurochemistry at the level of metabolites in it. Given that lumbar puncture is an invasive method it is not always feasible or ethical to perform. Blood measurements would be a much preferred to assess the biochemical aspects of a disease. In order to use amine levels in blood instead of CSF it is, however, essential to determine to what extent blood is representative of the amine levels in CSF. Therefore, the amine levels in blood were compared to those in CSF of 95 healthy people. What was observed is that the amine concentrations in plasma did not directly correlate with amine concentrations in CSF. This means that plasma concentrations are poor predictors of CSF concentrations for most amines, when studied as direct correlations ($\frac{\textbf{Amine x}_{plasma}}{\textbf{Amine y}_{CSF}} = r$). However, when studied as ratios (i.e., $\frac{\textbf{Amine x}_{plasma}}{\textbf{Amine y}_{plasma}} / \frac{\textbf{Amine x}_{CSF}}{\textbf{Amine y}_{CSF}} = r$) there was a significant correlation between plasma and CSF. The ratio correlations were significantly higher than the related single metabolite correlation for 308 of the 741 amine combinations. This indicates using ratio correlations is a superior method of comparing amine concentration of blood in relation to CSF. In addition, these results imply that the ratios are tightly regulated by blood-brain barrier transport systems most likely indicating the cotransport of amines. As a first proof-of-concept, the amine ratios in 95 healthy controls were also compared with those in 197 migraine patients and significant changes for some of the migraine ratios were observed, suggesting that amines seem to have a role in migraine pathophysiology.

In **Chapter 4** the biochemical alterations in endocannabinoids in the blood plasma of migraine patients outside of an attack were investigated. The endocannabinoid system is a relevant target for migraine research as endocanabinoids have a strong influence on neurotransmission, the neuroimmune and neuroendocrine systems, which are all implicated in the pathophysiology of migraine. Another interesting aspect is the role of the endocannabinoid system has in depression, an established comorbid condition of migraine. Multiple clinical studies in patients with (major) depressive disorders have identified a dysregulation of the endocannabinoid system. Several previous studies in migraine patients have investigated endocannabinoid levels but have shown inconsistent results. In our study the levels of endocannabinoids anandamide (AEA) and 2-arachidonoylglycerol (2-AG), and the endocannabinoid analogue docosahexaenoylethanolamine (DHEA) in the CSF could be reliably measured in 94 healthy volunteers and 97 migraine with aura patients and 97 migraine without aura patients outside of an attack. The endocannabinoid concentrations were measured via previously validated micro-liquid chromatography – tandem mass spectrometry (micro-LC-MS/MS). Given the number of possibly confounding factors on endocannabinoid levels, timing and processing of the samples was strictly protocolized and our

analysis was corrected for sex, age, BMI, weekly alcohol consumption, cigarettes per day, and comorbid lifetime depression. No differences were found in endocannabinoid levels between migraine subgroups and when compared to controls. However, several of the possibly confounding factors (sex, smoking, age, and BMI, and comorbid lifetime depression) did have a significant effect on endocannabinoid levels. Most striking is the effect on depression on endocannabinoid levels, given that depression is a comorbid disorder of migraine. This study is illustrative for the complexity of the endocannabinoid system, where it is important to take into consideration the contribution of confounding factors.

In **Chapter 5** the role of prostaglandin-E_2 (PGE_2) in relation to a provoked migraine attack was investigated. It has been shown that intravenous infusion of PGE_2 is able to provoke a migraine-like headache. Contrary to other provocation substances, such as glyceryl trinitrate (GTN), calcitonin gene-related peptide (CGRP), which show a delay of several hours in provoking a migraine attack, PGE_2 causes a very rapid onset of an attack, i.e. within 1 hour. The fact that PGE_2 causes such a rapid attack onset, suggests that PGE_2 may be at the end of the signalling pathway towards an attack, so PGE_2 might be closely upstream of GTN-induced migraine attacks. To reliable investigate the role of PGE_2 over the course of an attack, blood samples from 37 women with migraine and 25 age-matched female controls were obtained at three time points, i.e. (1) before provocation with GTN and (2) ~140 min and (3) ~320 min after migraine provocation with GTN. After analysis with a generalized linear mixed-effect model, no differences were found in PGE_2 levels between the interictal and preictal state nor between the interictal and ictal state. This suggests that rise in PGE_2 is not essential for the initiation of GTN-induced migraine-like attacks.

Part II Genetics of different headache forms

The aim of **Part 2** was to uncover the genetical underpinnings of multiple headache disorders. Apart from migraine we looked also at the genetic underpinnings of cluster headache and hemiplegic migraine. Cluster headache is characterized by excruciating unilateral headaches or facial pain accompanied by ipsilateral facial autonomic symptoms and/or restlessness. Hemiplegic migraine is a subtype of migraine with aura, where the aura phase is associated with motor weakness that can go into a hemiplegia. In most chapters (**Chapters 6,7 and 8**) of **Part 2** genome-wide association studies (GWAS) were used to investigate differences in the genetic architecture between cases and controls. In GWAS several millions of single nucleotide polymorphisms (SNPs) are tested for association with a trait by assessing differences in allele frequencies between large numbers of patients and controls. In addition, in **Chapter 9** we used a next-generation sequencing approach where the whole exome is sequenced and looked for differences between hemiplegic migraine patients and controls in a targeted approach by testing genes that belong to the family of voltage-gated calcium channel alpha subunits.

In **Chapter 6** the aim was to identify genetic susceptibility loci for cluster headache. To this end the DNA of 840 cluster headache patients and 1,457 controls was genotyped, quality controlled and imputated and compared to each other with a logistic regression model. An association was found with cluster headache for four independent loci rs11579212 (near *RP11-815M8.1*), rs6541998 (near *MERTK*), rs10184573 (near *AC093590.1*), and rs2499799 (near *UFL1/FHL5*), these loci collectively explain 7.2% of the variance of the phenotype. Three of the four loci replicated in an independent cohort of 144 cluster headache cases and 1,800 controls from Norway. Contrary to what is normally found in a GWAS, very large effect sizes were found in our study (all with an odds ratio of >1.2), indicating that the risk for cluster headache seems driven by a limited number of loci. In addition, the effect of the four loci on the expression of genes was investigated and this yielded 16 genes that seem involved in the pathophysiology of the disease. The expression of these genes was investigated in a previously performed RNA-sequencing dataset, by which a differential expression of *POLR1B* and *TMEM87B* in cluster headache patients compared to controls was found. In a separate correlation analysis, there was an indication of a genetical overlap of cluster headache with migraine. At a similar time, a UK-Swedish cohort also performed a GWAS in cluster headache patients based in 852 UK cases and 5,614 controls as well as 591 Swedish cases and 1,134 controls. This parallel study found four loci that were in linkage disequilibrium with our loci. This provides support that the four loci are genuine risk loci for cluster headache. Data of our study were also combined with the UK-Sweden data in a meta-analysis in a post-scriptum of the chapter. This meta-analysis resulted in three additional loci becoming genome-wide significant.

The research for **Chapter 7** is a further expansion on the meta-analysis of cluster headache performed in **Chapter 6**. In **Chapter 7** the genetic data of the two earlier studies (Netherlands/Norway and UK/Sweden) together with another previously published small Italian GWAS study and five novel European cohorts was jointly analysed. In total, the meta-analysis encompassed 4,043 cluster headache patients and 21,729 controls from ten cohorts, all of European ancestry. Seven genome-wide significant loci were identified, of which three are novel rs2402176 (*WNT2*), rs57866767 (*PLCE1*) and rs11172113 (*LRP1*) and four previously identified rs17011182, (*DUSP10*), rs13399108 (*MERTK*), rs6714578 (*FTCDNL1*) and rs9486725 (*FHL5*). Downstream bioinformatics analyses showed enrichment for artery and brain tissue. Furthermore, correlation analyses showed a genetic overlap of cluster headache with migraine, cigarette smoking, risk-taking behaviour, ADHD, depression, and musculoskeletal pain. An in-depth analysis of the overlap with migraine showed that cluster headache and migraine share three genetic risk loci. This suggests that the two disorders are genetically partially the same and partially distinct. Furthermore, the potential causality of smoking on cluster headache was investigated with a two-sample Mendelian randomization analysis based on the summary statistics. This analysis indicated a *causal* effect of smoking intensity on cluster headache. This effect of smoking on cluster headache has potential clinical implications. In a secondary trans-ancestry meta-analysis we added 734 cases and 9,846 controls of East Asian ancestry to our meta-analysis. One additional genome-wide significant cluster headache locus, in *CAPN2*, was identified when adding the East Asian cohort in an

ancestry-adjusted GWAS meta-analysis.

Previously, multiple genetic risk loci for migraine have been reported by multiple studies based on ever increasing sample sizes. In **Chapter 8** the sample size of the cohort was further increased to 102,084 migraine cases and 771,257 controls of European decent based on different international cohorts. In addition, the migraine subtype specificity was investigated in clinically diagnosed patients in 14,625 migraine with aura cases and 15,055 migraine without aura cases. The main meta-analysis identified 123 loci, of which 86 were novel compared to earlier studies. Remarkably the new risk loci included loci corresponding to genes transcribing calcitonin gene-related peptide (*CALCA*/*CALCB*) and serotonin 1F receptor (*HTR1F*). These are well-known migraine drug targets. This finding shows the clear potential for GWASs to identify (new) drug targets. Downstream analyses reconfirmed that neurovascular mechanisms underlie migraine pathophysiology as there was an enrichment for vascular and central nervous system tissue based on the risk loci. In a subtype analysis different risk loci were found for migraine with and without aura, and risk loci that increase disease susceptibility for both subtypes. Of note, *CACNA1A* a well-known hemiplegic migraine gene was identified as one of three susceptibility loci for migraine with aura, thus now also genetically linking monogenic and polygenic forms of migraine.

In continuation of the genetic studies of both monogenic and polygenic forms of migraine, in **Chapter 9**, genetic factors contributing to hemiplegic migraine were further investigated. For this study DNA was investigated of hemiplegic migraine patients in whom no mutation was detected in the known *CACNA1A*, *ATP1A2* and *SCN1A* disease genes. Given the previous association with *CACNA1A* that popped up in **Chapter 8** and associations of migraine with *CACNA1B*, *CACNA1E* and *CACNA1H* in other studies, made us consider the family of *CACNA1x* genes as an interesting target for burden testing (burden being the aggregation of both rare and common genetic variants as well as the increased presence of a variant in cases compared to controls). In total, the exome of 184 Australian hemiplegic migraine patients was sequenced. The subject and variant burden of missense variants in the *CACNA1x* genes of the hemiplegic migraine patients was compared to publicly available sequencing data from controls from gnomAD. An increase in burden in cases was found in *CACNA1H* and *CACNA1I* genes, a result that was replicated in an independent Dutch cohort of 32 patients with hemiplegic migraine. From this study it can be concluded that a burden of (non-pathogenic) missense variants in *CACNA1H* and *CACNA1I* is implicated as modifier genes that confer susceptibility to the hemiplegic migraine.

Finally, **Chapter 10** provides a general discussion of this thesis with considerations and suggestions for future research.

Nederlandse Samenvatting

Dit proefschrift onderzoekt welke biochemische en genetische veranderingen geassocieerd zijn met migraine en clusterhoofdpijn. Dit onderzoek heeft als doel om de kennis over de pathofysiologie en de signaalroutes die betrokken zijn bij migraine en clusterhoofdpijn te vergroten en daardoor nieuwe aangrijpingspunten voor behandeling te identificeren. Door te onderzoeken hoe patiënten verschillen van mensen zonder de ziekte worden de bijdragende biologische mechanismen blootgelegd, dit leidt vervolgens tot een nieuw begrip van de pathofysiologie van hoofdpijnaandoeningen. De biologische systemen in ons lichaam zijn allemaal aan elkaar gerelateerd, ze verhouden zich tot elkaar vanaf de genomische blauwdruk via epigenetica, transcriptie en translatie tot aan de eiwitten en biomoleculen. Door deze verschillende moleculaire niveaus te onderzoeken hopen we een beter inzicht te krijgen in de werking van hoofdpijnaandoeningen. Dit proefschrift bestaat uit twee delen: **Deel 1**, (**hoofdstukken 2 t/m 5**) is gericht op het onderzoek van biomoleculen in lichaamsvloeistoffen (bloed en hersenvocht) bij migrainepatiënten; en **Deel 2**, (**hoofdstuk 6 t/m 9**) gericht is op de genetische architectuur van clusterhoofdpijn, migraine en aanverwante aandoeningen.

Deel I Biochemie van migraine

Deel 1 beschrijft biochemische studies waarbij verschillende moleculen en metabolieten werden gemeten bij mensen met migraine en vergeleken met mensen zonder migraine. Migraine is een hoofdpijnaandoening die wordt gekenmerkt door terugkerende episoden van ernstige, vaak eenzijdig pulserende hoofdpijn die gepaard gaat met misselijkheid, braken en/of foto- en fonofobie. Bij een derde van de patiënten worden de aanvallen voorafgegaan door een aurafase. Typisch bestaat een migraineaanval uit een preictale, ictale (aura en/of hoofdpijn), en postictale (postdromale) fase. Hoewel in de loop der jaren veel vooruitgang is geboekt bij het begrijpen van de pathofysiologische mechanismen die betrokken zijn bij migraine, is het nog steeds onduidelijk waarom sommige mensen migraine krijgen en anderen niet en waarom degenen met migraine op bepaalde momenten aanvallen krijgen en welke factoren de aanvallen uitlokken. Het doel van **deel 1** was het verder identificeren van de biochemische werking van migraine om de betrokken signaalroutes te kunnen identificeren. De onderzoeken richtten zich op het meten van een groot aantal verschillende biomoleculen, hetzij via een hypothese-gerichte, hetzij via een hypothese-vrije aanpak. De studies waren gericht op zowel bloed als hersenvocht en daarnaast is onderzocht in hoeverre beide lichaamsvloeistoffen vergelijkbaar zijn met betrekking tot amineconcentraties. We onderzochten in op één na alle hoofdstukken (**hoofdstuk 2,3 en 4**) deze lichaamsvloeistoffen buiten een aanval en in één hoofdstuk in de loop van een uitgelokte migraineaanval. Op die manier werden verschillende aspecten van migraine onderzocht.

In **hoofdstuk 2** wordt het metabole profiel van migrainepatiënten vergeleken met dat van patiënten zonder migraine in een hypothese-vrije aanpak met "high-throughout" proton

nucleaire magnetische resonantie (^{1}H-NMR) spectroscopie. In totaal werden 100 signalen van 49 verschillende metabolieten van 189 migrainepatiënten en 1.360 controles in nuchtere serummonsters gedetecteerd. De migrainestatus werd onderverdeeld in "lifetime" migraine (migrainediagnose tijdens het leven) en actieve migraine (gedefinieerd als ten minste één ernstige migraine aanval in de afgelopen 12 maanden). Met behulp van "elastic net regression" analyses werden voorspellende profielen geïdentificeerd voor "lifetime" migrainestatus bestaande uit 6 metabolieten en voor actieve migrainestatus bestaande uit 22 metabolieten. Deze voorspellende metabolieten omvatten lipiden, aminozuren en metabolieten van het glucosemetabolisme. Deze resultaten wijzen op een algemene verstoring van het metabole profiel bij migrainepatiënten.

In **hoofdstuk 3** werd de relatie in aminegehaltes tussen bloedplasma en hersenvocht bepaald. Hersenvocht is een lichaamsvloeistof die bij uitstek veel kan vertellen over aandoeningen van het centrale zenuwstelsel, zoals migraine, omdat deze vloeistof het dichtst bij de hersenen ligt en daarmee het best de neurochemie van de hersenen weerspiegelt aan de hand van de metabolieten niveaus die aanwezig zijn. Bloedmetingen verdienen de voorkeur om de biochemische aspecten van een ziekte te beoordelen, omdat een lumbaalpunctie (ter verkrijging van hersenvocht) een invasieve en ingrijpende methode is, daarbij is het niet altijd haalbaar of ethisch verantwoord om deze uit te voeren. Om te bepalen of het mogelijk is om bloed te gebruiken voor aminen metingen in plaats van hersenvloeistof is het van essentieel belang om te bepalen of de aminegehaltes in bloed representatief zijn voor de aminegehaltes in hersenvocht. Daarom werden de aminegehaltes in het bloed vergeleken met die in het hersenvocht van 95 gezonde mensen. De resultaten lieten zien dat de amineconcentraties in plasma niet direct correleerden met de amineconcentraties in het hersenvocht. Dit betekent dat de plasmaconcentraties voor de meeste aminen slechte voorspellers zijn van de hersenvochtconcentraties, wanneer zij als directe correlaties worden bestudeerd ($\frac{\mathbf{Amine\ x}_{bloed}}{\mathbf{Amine\ y}_{hersenvocht}} = r$). Bij bestudering als ratio's (d.w.z. $\frac{\mathbf{Amine\ x}_{bloed}}{\mathbf{Amine\ y}_{bloed}} / \frac{\mathbf{Amine\ x}_{hersenvocht}}{\mathbf{Amine\ y}_{hersenvocht}} = r$) was er echter een significante correlatie tussen bloed en hersenvocht. Voor 308 van de 741 aminecombinaties waren de ratio correlaties significant hoger dan de gerelateerde enkelvoudige metaboliet correlatie. Dit wijst erop dat het gebruik van ratiocorrelaties een betere methode is om de amineconcentratie in het bloed te vergelijken met die in het hersenvocht. Bovendien impliceren deze resultaten dat de ratio's tussen bloed en hersenvocht strak gereguleerd worden door de transportsystemen van de bloed-hersenbarrière, hetgeen waarschijnlijk wijst op het co-transport van amines. Als "proof-of-concept" werden de amineratio's in 95 gezonde controles vergeleken met die in 197 migrainepatiënten en werden significante verschillen gezien tussen de twee groepen, hetgeen suggereert dat amines een rol spelen in de pathofysiologie van migraine.

In **hoofdstuk 4** werden biochemische veranderingen in endocannabinoïden niveaus van migrainepatiënten in bloedplasma buiten een aanval onderzocht. Het endocannabinoïde systeem is een interessant onderwerp voor onderzoek naar migraine, omdat bekend is dat het een sterke invloed heeft op neurotransmissie, het neuro-immuunsysteem en het neuro-endocriene systeem, en al die systemen zijn ook betrokken bij de pathofysiologie van migraine. Een ander interessant aspect is de

rol die het endocannabinoïde systeem speelt bij depressie, een bekende comorbiditeit van migraine. Meerdere klinische studies hebben een ontregeling van het endocannabinoïde systeem vastgesteld bij patiënten met ernstige depressieve stoornissen. Verschillende eerdere studies onderzochten de endocannabinoïden niveaus bij migrainepatiënten, maar de resultaten waren inconsistent. In onze studie werden in 94 gezonde vrijwilligers en 97 migraine met aura en 97 migraine zonder aura patiënten de niveaus van endocannabinoïden anandamide (AEA) en 2-arachidonoylglycerol (2-AG), en het endocannabinoïde-analoog docosahexaenoylethanolamine (DHEA) buiten een aanval betrouwbaar gemeten in hersenvocht. De endocannabinoïde concentraties werden gemeten via eerder gevalideerde microvloeistofchromatografie - tandem massaspectrometrie (micro-LC-MS/MS). Gezien het aantal mogelijke confounders op de endocannabinoïde concentraties, werden de timing en verwerking van de monsters strikt geprotocolleerd en werd onze analyse gecorrigeerd voor geslacht, leeftijd, BMI, wekelijks alcoholgebruik, sigaretten per dag en "lifetime" depressie. Er werden geen verschillen gevonden in endocannabinoïde niveaus tussen migraine patiënten en controles. Verschillende van de mogelijke confounders (geslacht, roken, leeftijd en BMI, en "lifetime" depressie) hadden wel een significant effect op de endocannabinoïde spiegels. Het meest opvallend was het effect van depressie op endocannabinoïde concentraties, aangezien depressie een comorbiditeit van migraine is. Deze studie is illustratief voor de complexiteit van het endocannabinoïde systeem, waarbij het belangrijk is gebleken rekening te houden met de rol van confounders op endocannabinoiden concentraties.

In **hoofdstuk 5** werd de rol van prostaglandine-E_2 (PGE_2) in relatie tot een uitgelokte migraineaanval onderzocht. Er is gebleken dat intraveneuze toediening van PGE_2 een migraine -achtige hoofdpijn aanval kan uitlokken. Toediening van PGE_2 veroorzaakt een zeer snel begin van een aanval namelijk binnen een uur, in tegenstelling tot andere provocerende stoffen, zoals glyceryltrinitraat (GTN), calcitonine-gen-gerelateerd peptide (calcitonin-gene-related peptide; CGRP), die een migraineaanval pas na enkele uren uitlokken. Het feit dat PGE_2 zo'n snelle aanval veroorzaakt, suggereert dat PGE_2 een laat signaal eiwit kan zijn in de signaal route naar een migraineaanval, zodanig dat PGE_2 nauw betrokken kan zijn bij GTN-geïnduceerde migraineaanvallen. Om de rol van PGE_2 in de aanloop naar een aanval betrouwbaar te onderzoeken, werden bloedmonsters van 37 vrouwen met migraine en 25 op leeftijd "gematchte" vrouwelijke controles verkregen op drie tijdsmomenten, namelijk (1) vóór provocatie met GTN en (2) ~140 min en (3) ~320 min na provocatie met GTN. Na analyse met een "generalized mixed-effect model", werden geen verschillen gevonden in PGE_2-niveaus tussen de interictale en de preictale toestand, noch tussen de interictale en de ictale toestand. Dit betekent dat er geen bewijs is dat een stijging van PGE_2 essentieel is voor het ontstaan van door GTN geïnduceerde migraine aanvallen.

Deel II Genetica van verschillende vormen van hoofdpijn

Het doel van **deel 2** was de genetische architectuur van meerdere hoofdpijnaandoeningen te onderzoeken. Naast migraine hebben we ook gekeken naar de genetische architectuur van

clusterhoofdpijn en hemiplegische migraine. Clusterhoofdpijn wordt gekenmerkt door ondraaglijke eenzijdige hoofdpijn of aangezichtspijn die gepaard gaat met ipsilaterale craniale autonome verschijnselen en/of rusteloosheid. Hemiplegische migraine is een subtype van migraine met aura, waarbij de aurafase gepaard gaat met motorische zwakte die kan overgaan in hemiplegie. In de meeste hoofdstukken (**hoofdstukken 6,7, en 8**) van **deel 2** werden genoomwijde associatiestudies (genome-wide asssociation studies; GWAS) gebruikt om het verschil in genetische architectuur tussen patiënten en controles te onderzoeken. In GWAS worden miljoenen "single nucleotide polymorfismen" (SNPs) getest op associatie met een aandoening door verschillen in allelfrequenties tussen grote aantallen patiënten en controles te vergelijken. Daarnaast hebben we in **hoofdstuk 9** een next generation sequencing aanpak gebruikt waarbij het hele exoom werd gesequenced en keken wij gericht naar verschillen tussen hemiplegische migrainepatiënten en controles.

In **hoofdstuk 6** was het doel om genetische risico loci voor clusterhoofdpijn te identificeren. Hiervoor werd het DNA van 840 clusterhoofdpijnpatiënten en 1.457 controles gegenotypeerd, op kwaliteit gecontroleerd en geïmputeerd en met elkaar vergeleken met een logistisch regressiemodel. Voor vier onafhankelijke loci rs11579212 (nabij *RP11-815M8.1*), rs6541998 (nabij *MERTK*), rs10184573 (nabij *AC093590.1*), en rs2499799 (nabij *UFL1*/*FHL5*) werd een associatie met clusterhoofdpijn gevonden, deze loci verklaarden gezamenlijk 7,2% van de variantie van het clusterhoofdpijn fenotype. De resultaten van drie van de vier loci werden gerepliceerd in een onafhankelijk Noors cohort bestaande uit 144 clusterhoofdpijn patiënten en 1800 controles. In tegenstelling tot wat gewoonlijk wordt gevonden in GWAS, werden in onze studie zeer grote effectgroottes gevonden (alle met een odds ratio van >1,2), wat erop wijst dat het risico voor clusterhoofdpijn mogelijk wordt bepaald door een beperkt aantal risico loci. Daarnaast werd het effect van deze vier risico loci op de expressie van genen onderzocht, dit leverde 16 genen op die betrokken lijken te zijn bij de pathofysiologie van deze ziekte. De expressie van deze genen werd onderzocht in een eerder uitgevoerde RNA-sequencing dataset, waarbij een verandering in expressie van *POLR1B* en *TMEM87B* in clusterhoofdpijn patiënten werd gevonden. In een aparte correlatieanalyse werden aanwijzingen gevonden voor een genetische overlap van clusterhoofdpijn met migraine. Op hetzelfde moment werd in een Brits-Zweeds cohort ook een GWAS uitgevoerd bij clusterhoofdpijnpatiënten op basis van 852 Britse clusterhoofdpijn patiënten en 5.614 controles en 591 Zweedse clusterhoofdpijn patiënten en 1.134 controles. Deze parallelle studie vond vier loci die in linkage disequilibrium waren met onze loci. Dit ondersteunt de bevinding dat deze loci ook daadwerkelijk risico loci voor clusterhoofdpijn zijn. Data van onze studie werden gecombineerd met de Brits-Zweedse data in een meta-analyse in een post-scriptum. Deze meta-analyse resulteerde in drie additionele genoombrede significante risico loci.

Hoofdstuk 7 is een verdere uitbreiding van de meta-analyse van clusterhoofdpijn uit **hoofdstuk 6**. In **hoofdstuk 7** zijn de genetische data van de twee eerdere studies (Nederland/Noorwegen en V-K/Zweden) samen met een andere eerder gepubliceerde Italiaanse GWAS-studie en vijf nieuwe Europese cohorten samen geanalyseerd. In het totaal werd een meta-analyse uitgevoerd op 4.043

clusterhoofdpijn patiënten en 21.729 controles uit tien cohorten allen van Europese achtergrond. Zeven genoombrede significante loci werden geïdentificeerd, waarvan drie nieuwe rs2402176 (*WNT2*), rs57866767 (*PLCE1*) en rs11172113 (*LRP1*) en vier eerder geïdentificeerde rs17011182, (*DUSP10*), rs13399108 (*MERTK*), rs6714578 (*FTCDNL1*) en rs9486725 (*FHL5*). Downstream bio-informatica analyses toonde een verrijking voor slagader- en hersenweefsel aan. Verder werden correlatieanalyses uitgevoerd, waaruit een genetische overlap bleek van clusterhoofdpijn met migraine, roken, risicogedrag, ADHD, depressie en pijn aan het bewegingsapparaat. Een additionele analyse naar overlap met migraine toonde aan dat clusterhoofdpijn en migraine drie genetische risico loci delen. Dit suggereert dat de twee aandoeningen genetisch deels hetzelfde en deels verschillend zijn. Verder werd het mogelijke causale verband van roken op clusterhoofdpijn onderzocht met een "two-sample" Mendeliaanse randomisatie analyse. Deze analyse liet een causaal effect van de intensiteit van roken op clusterhoofdpijn zien. Dit effect van roken op clusterhoofdpijn heeft mogelijke klinische implicaties. In een secundaire "trans-ancestry" meta-analyse voegden we 734 gevallen en 9.846 controles van Oost-Aziatische afkomst toe aan onze meta-analyse. Eén extra genoombreed significant cluster hoofdpijn locus, in *CAPN2*, werd geïdentificeerd bij het toevoegen van het Oost-Aziatische cohort in een voor afkomst gecorrigeerde GWAS meta-analyse.

Over de jaren zijn steeds meer genetische risico loci voor migraine vanuit verschillende studies gerapporteerd op basis van steeds grotere aantal patiënten dat word meegenomen. In **hoofdstuk 8** werd de studie omvang verder uitgebreid tot 102.084 migraine patiënten en 771.257 controles van verschillende internationale cohorten allen met een Europese achtergrond. Daarnaast werd de specificiteit van de migraine subtypes onderzocht bij klinisch gediagnosticeerde patiënten in 14.625 patiënten met migraine met aura en 15.055 patiënten met migraine zonder aura. De meta-analyse identificeerde 123 risico loci, waarvan 86 nieuwe ten opzichte van eerdere studies. Opmerkelijk is dat de nieuwe risico loci locaties omvatten die corresponderen met genen die calcitonine gen-gerelateerd peptide (*CALCA*/*CALCB*) en serotonine 1F receptor (*HTR1F*) transcriberen. Dit zijn bekende aangrijpingspunten voor migrainemedicatie. Deze bevinding toont duidelijk aan dat GWASen de mogelijkheid hebben om (nieuwe) doelwitten voor behandeling te identificeren. Downstream-analyse op basis van deze meta-analyse bevestigde nogmaals dat neurovasculaire mechanismen ten grondslag liggen aan de pathofysiologie van migraine, aangezien er een verrijking was voor vaatweefsel en weefsel van het centrale zenuwstelsel op basis van de risico loci. In de subtype-analyse werden verschillende risico loci gevonden voor migraine met en zonder aura, en risico loci die correspondeerden met beide subtypen. Opmerkelijk is dat *CACNA1A*, een bekend hemiplegische migraine gen, werd geïdentificeerd als één van de risico loci voor migraine met aura, op basis hiervan kunnen zowel monogene als polygene vormen van migraine met elkaar in verband worden gebracht.

In het verlengde van de koppeling van zowel monogene als polygene vormen van migraine, werd in **hoofdstuk 9** verder onderzocht welke genetische factoren bijdragen aan hemiplegische migraine. Dit werd onderzocht in hemiplegische migrainepatiënten bij wie geen mutatie werd aangetoond in *CACNA1A*, *ATP1A2* en *SCN1A*. Gezien de eerdere associatie met *CACNA1A* in **hoofdstuk 8** en andere associaties van migraine met *CACNA1B*, *CACNA1E* en *CACNA1H* in andere studies, is de familie van *CACNA1x*-genen een interessant onderwerp voor "burden" onderzoek ("burden" is de aggregatie van zowel zeldzame als veel voorkomende genetische varianten, en het frequenter voorkomen van een variant in patiënten in vergelijking met controles). In totaal werd van 184 Australische hemiplegische migrainepatiënten het exoom gesequenced. De subject- en variant "burden" van "missense"-varianten in de *CACNA1x*-genen van deze hemiplegische migrainepatiënten werd vergeleken met controles uit gnomAD. Er werd een toename in "burden" gevonden in *CACNA1H*-genen en *CACNA1I*-genen, dit werd gerepliceerd in een onafhankelijk Nederlands cohort bestaande uit 32 patiënten met hemiplegische migraine. Uit deze studie kan worden geconcludeerd dat een "burden" van (niet-pathogene) "missense" varianten in *CACNA1H* en *CACNA1I* werkt als modificerende factor die bijdraagt een het krijgen van hemiplegische migraine.

Tenslotte bevat **hoofdstuk 10** van dit proefschrift een algemene discussie van alle resultaten en worden overwegingen en suggesties voor toekomstig onderzoek gegeven.

List of Publications

Bendik S. Winsvold*, **Aster V.E. Harder***, Caroline Ran*, Mona A Chalmer*, Maria Carolina Dalmasso*, Egil Ferkingstad* et al. *Cluster headache genome-wide association study identifies eight loci and implicates smoking as causal risk factor*. Ann Neurol 2023 Jul; 10.1002/ana.26743

Aster V.E. Harder, Gisela M. Terwindt, Dale Nyholt, Arn M.J.M van den Maagdenberg. *Migraine genetics: status and road forward.* Cephalalgia. 2023;43(2):3331024221145962

Aster V.E. Harder*, Gerrit L.J. Onderwater*, Robin M. van Dongen, Marieke Heijink, Erik W. van Zwet, Martin Giera, Arn M.J.M. van den Maagdenberg and Gisela M. Terwindt. *Prostaglandin-E2 levels over the course of glyceryl trinitrate provoked migraine attacks*. Neurobiol. Pain. 2023 Jan;13:100112

Nike Z. Welander, Gull Rukh, Mathias Rask-Andersen, **Aster V.E. Harder**, Arn M.J.M. van den Maagdenberg, Helgi B. Schiöth, Jessica Mwinyi. *Is migraine causally linked to inflammatory bowel disease or coeliac disease? A Mendelian randomisation study*. Headache. 2023 May;63(5):642-651.

Neven Maksemous*, **Aster V.E. Harder***, Omar Ibrahim*, Lisanne S. Vijfhuizen, Heidi Sutherland, Nadine Pelzer, Irene de Boer, Gisela M. Terwindt, Rodney A. Lea, Arn M.J.M. van den Maagdenberg and Lyn R. Griffiths. *Whole exome sequencing of hemiplegic migraine patients shows an increased burden of missense variants in CACNA1H and CACNA1I genes.* Mol Neurobiol. 2023 Jun;60(6):3034-3043

Bingshu He, Xinyu Di, Faisa Guled, **Aster V.E. Harder**, Arn M.J.M. van den Maagdenberg, Gisela M. Terwindt, Elke H.J. Krekels, Isabelle Kohler, Amy Harms, Rawi Ramautar, Thomas Hankemeier. *Quantification of endocannabinoids in human cerebrospinal fluid using a novel micro-flow liquid chromatography-mass spectrometry method.* Analytica Chimica Acta 2022 Jun;1210:339888.

Heidi Hautakangas, Bendik S. Winsvold, Sanni E. Ruotsalainen, Gyda Bjornsdottir, **Aster V.E. Harder**, Lisette J. A. Kogelman, et al. *Genome-wide analysis of 102,084 migraine cases identifies 123 risk loci and subtype-specific risk alleles*. Nat Genet. 2022;54(2):152-160.

Aster V.E. Harder, Lisanne S. Vijfhuizen, Peter Henneman, Ko Willems van Dijk, Cornelia M. van Duijn, Gisela M. Terwindt, et al. *Metabolite changes in serum of migraine patients detected using* ^{1}*H-NMR spectroscopy*. J Headache Pain. 2021 Nov 24;22(1):142.

Aster V.E. Harder*, Bendik S. Winsvold*, Raymond Noordam*, Lisanne S. Vijfhuizen, Sigrid Børte, Lisette J.A. Kogelman et al. *Genetic susceptibility loci in genome-wide association study of cluster headac*he. Ann Neurol. 2021 Aug;90(2):203-216.

Emer O'Connor, Carmen Fourier, Caroline Ran, Prasanth Sivakumar, Franziska Liesecke, Laura Southgate , **Aster V.E. Harder** , Lisanne S. Vijfhuizen LS et al. *Genome-Wide Association Study Identifies Risk Loci for Cluster Headache*. Ann Neurol. 2021 Aug;90(2):193-202.

Katrine M. Johannesen, Yuanyuan Liu, Mahmoud Koko, Catherine E. Gjerulfsen, Lukas Sonnenberg, Julian Schubert, ….. **Aster V.E. Harder** ….. et al. *Genotype-phenotype correlations in SCN8A-related disorders reveal prognostic and therapeutic implications*. Brain. 2021 Aug 25:awab321.

Sarah L. Gardiner, **Aster V.E. Harder**, Yvonne J.M. Campman, Stella Trompet, Jacobijn Gussekloo, Martine J. van Belzen, Merel W. Boogaard, Raymund A.C. Roos, et al. *Repeat length variations in ATXN1 and AR modify disease expression in Alzheimer's disease*. Neurobiological Aging. 2019;73:230.e9-230.e17.

Eduardo Calpena, Alexia Hervieu, Teresa Kaserer, Sigrid M.A. Swagemakers, Jaqueline A.C. Goos, Olajumoke Popoola, ….. **Aster V.E. Harder** ….. et al. *De Novo Missense Substitutions in the Gene Encoding CDK8, a Regulator of the Mediator Complex, Cause a Syndromic Developmental Disorder*. American Journal of Human Genetics 2019;104(4):709-720

About the author

Aster Victorine Elsemieke Harder was born on the 3rd of November 1989 in Zoetermeer, she grew up in Bavel. She graduated secondary school in Breda (Stedelijk Gymnasium Breda) in 2008. That same year she went to the United Kingdom and followed courses at the English Language Centre in Bristol and St. Clare's College in Oxford. In 2009 she started her study Biomedical Sciences at Leiden University. In her second year she followed a semester of her studies at Karolinska Institutet in Stockholm, Sweden as part of an Erasmus exchange program. From 2012 she combined her study Biomedical Sciences with studying Medicine. After receiving a Bachelor in Biomedical Science (2012) and a Bachelor in Medicine (2013), Aster continued her education in Leiden. In 2017 she completed both a Master in Biomedical Science and a Master in Medicine.

During her Masters Aster did two research internships gaining experience in pre-clinical research. In her first internship Aster worked on '*Exon skipping*' in Duchenne muscular dystrophy and was supervised by Prof. dr. A. Aartsma-Rus at the Leiden University Medical Centre (LUMC). In her second internship she worked on '*CAG repeat polymorphisms*' in Parkinson's and Alzheimer's disease, supervised by dr. A. Aziz. Simultaneously Aster was enrolled in the Leiden Leadership Program from the Leiden University Honours Academy. She did her final clinical internship at the department of Clinical Genetics at the LUMC.

After graduation, she started working as a resident at the department of Clinical Genetics at University Medical Centre Utrecht (UMCU) in 2017. In 2018 Aster started her PhD at the department of Neurology and the department of Human Genetics, supervised by Prof. dr. A.M.J.M. van den Maagdenberg, Prof. dr. G.M. Terwindt and Prof. dr. M.D. Ferrari. As part of her PhD she conducted research at Oslo University Hospital under supervision of dr. B.S. Winsvold and Prof. dr. J.A.H. Zwart for which she received a grant from Leiden University Fund. Currently, Aster works as a resident at the department of Clinical Genetics in the LUMC.

Dankwoord

Dit proefschrift was niet tot stand gekomen zonder de hulp en steun van vele mensen, waarvan ik de volgende personen hieronder in het bijzonder wil bedanken.

In de eerste plaats wil ik alle mensen die hebben deelgenomen aan het onderzoek in dit proefschrift van harte bedanken. Zonder jullie medewerking zou dit proefschrift er niet zijn gekomen.

Daarnaast wil ik mijn promotoren bedanken. Arn, Gisela en Michel bedankt voor de mogelijkheid die jullie me hebben geboden om me wetenschappelijk te ontwikkelen. Ieder van jullie had een eigen rol binnen de totstandkoming van dit proefschrift. Beste Arn, je nam altijd de tijd om de voortgang van projecten te bespreken. Met je tactische adviezen heb je mij en menig project op de rails weten te houden. Onze wekelijkse overleggen heb ik als prettig, vol humor en stimulerend ervaren, je zette me altijd aan tot kritisch nadenken over de volgende stap binnen de vele verschillende projecten. Beste Gisela, ik wil mijn bewondering uitspreken voor je aanstekelijke enthousiasme voor de wetenschap en je drive om te streven naar degelijk en klinisch zinvol onderzoek. Beste Michel, ik wil mijn waardering uitspreken voor jouw wetenschappelijke kennis en de leerzame discussie momenten die we hebben gehad.

Furthermore, I would like to thank my colleagues from the Netherlands and from abroad. Thank you for your pleasant collaboration and support in the various genetics projects. In particular, I would like to thank Bendik and John-Anker from Oslo University Hospital for your trust in me and your help in completing the cluster headache projects from this thesis. Thank you for the possibility of performing research in Oslo, I sincerely hope we will see each other again in the future.

Ook wil ik de collega's van de Humane Genetica en mijn mede-promovendi en collega's van de Neurologie bedanken. Ontzettend bedankt voor de leuke en mooie tijd die ik met jullie heb beleefd. De congressen, borrels, spelletjesavonden, Sinterklaasavond, LUMINA-uitjes en onderzoekersweekenden maakte dit tot een onvergetelijke tijd. Daarnaast wil ik jullie bedanken voor de goede gesprekken en jullie bereidheid tot hulp, ook bij inhoudelijke vraagstukken. Het was heel fijn om deel uit te maken van deze groep enthousiaste collega's. Specifiek wil ik nog mijn dank betuigen aan Lisanne voor de altijd fijne samenwerking.

Lieve vrienden, van de middelbare school, vereniging, studie biomedische wetenschappen, studie geneeskunde en daarbuiten, oneindig veel dank voor jullie geduld, luisterend oor en bemoedigende woorden. Lieve familie, bedankt voor jullie onvoorwaardelijke steun, vertrouwen en dat jullie altijd voor mij klaar staan. Lieve David, je stimuleert me altijd om de beste versie van mezelf te zijn, bedankt voor je eeuwige geduld, rationele blik en adviezen. Zonder jou had ik dit proefschrift niet kunnen realiseren. Ik kijk uit naar alle toekomstige mijlpalen die we samen zullen bereiken.