Rare risk variants associate with epigenetic dysregulation in migraine

Techlo, Tanya Ramdal; Chalmer, Mona Ameri; Møller, Peter L.; Kogelman, Lisette J. A.; Olofsson, Isa A.; DBDS Genomic Consortium; Banasik, Karina; Nyegaard, Mette; Olesen, Jes; Hansen, Thomas Folkmann

DOI (link to publication from Publisher):
10.1101/2021.12.20.21268001

Publication date:
2021

Document Version
Publisher's PDF, also known as Version of record

Title: Rare risk variants associate with epigenetic dysregulation in migraine

Tanya Ramdal Techlo,[1] Mona Ameri Chalmer,[1] Peter L. Møller,[2] Lisette J. A. Kogelman,[1] Isa A. Olofsson,[1] DBDS Genomic Consortium, Karina Banasik,[3] Mette Nyegaard,[4] Jes Olesen,[1] Thomas Folkmann Hansen*[1,3]

[1]Danish Headache Center, Department of Neurology, Copenhagen University Hospital, Rigshospitalet-Glostrup, DK-2600 Glostrup, Denmark

[2]Department of Biomedicine, University of Aarhus, DK-8000 Aarhus C, Denmark

[3]Novo Nordisk Foundation Center for Protein Research, University of Copenhagen, DK-2200 Copenhagen N, Denmark

[4]Department of Health Science and Technology, University of Aalborg, DK-9220 Aalborg E, Denmark

***Corresponding author:**

Thomas Folkmann Hansen, Danish Headache Center, Department of Neurology, Copenhagen University Hospital, Rigshospitalet-Glostrup, Valdemar Hansens Vej 5, DK-2600 Glostrup, Denmark.

Phone: +45 38633051, Email: thomas.folkmann.hansen@regionh.dk

ABSTRACT

Migraine has a heritability of up to 65%. Genome-wide association studies (GWAS) on migraine have identified 123 risk loci, explaining only 10.6% of migraine heritability. Thus, there is a

considerable genetic component not identified with GWAS. Further, the causality of the identified risk loci remains inconclusive. Rare variants contribute to the risk of migraine but GWAS are often underpowered to detect these. Whole genome sequencing is reliable for analyzing rare variants but is not frequently used in large-scale. We assessed if rare variants in the migraine risk loci associated with migraine. We used a large cohort of whole genome sequenced migraine patients (1,040 individuals from 155 families). The findings were replicated in an independent case-control cohort (2,027 migraine patients, 1,650 controls). We found rare variants (minor allele frequency<0.1%) associated with migraine in a Polycomb Response Element in the *ASTN2* locus. The association was independent of the GWAS lead risk variant in the locus. The findings place rare variants as risk factors for migraine. We propose a biological mechanism by which epigenetic regulation by Polycomb Response Elements plays a crucial role in migraine etiology.

INTRODUCTION

Migraine is a complex neurovascular and genetic disorder which affects 15-20% of the population.[1] It is characterized by episodes of severe headache, which is typically unilateral and pulsating, accompanied by nausea, vomiting, photophobia and/or phonophobia.[2] Identifying factors that predispose to migraine is highly relevant. The Global Burden of Disease Study (2019) ranks migraine the highest cause of disability worldwide in young women (aged 15-49) and the second highest cause of disability worldwide among both sexes.[3] Additionally, migraine has a large socioeconomic impact through treatment costs and loss of productivity estimated at €27 billion in Europe alone.[4]

Migraine has a significant genetic component and is polygenic. There is an increased familial risk and twin studies have shown a heritability of 34%–65%.[5,6] Multiple, rare variants, each with large effect sizes, may contribute to disease risk together with common variants.[7,8] The most recent meta-analysis of migraine genome-wide association studies (GWAS) identified 123 migraine risk loci, of which 121 were autosomal. Still, only 10.6% of the total heritability was explained by the GWAS.[9] GWAS are often underpowered to detect rare variants, as these are imputed and not measured.[10] Whole genome sequencing (WGS) provides a direct assessment of rare variants. However, it is a technology that is still not frequently used in large-scale and is costly compared to genotyping using arrays.

How the migraine risk loci affect migraine pathogenesis is unknown. Many reside in non-protein coding regions of the DNA and may alter the regulation of gene expression.[11] Increased risk of migraine has been linked to disruption in gene regulation[12] and epigenetic regulation.[12–14] While candidate genes have been identified for increased genetic risk in migraine, the findings have been reported with a high risk of false positive results.[15–18] This suggests that migraine pathogenesis is affected by mutations outside of protein-coding regions of the DNA, such as in regulatory regions.

We hypothesized that migraine is influenced by rare variants in regulatory regions. We tested if rare variants in the migraine risk loci increase the risk of migraine. We defined blocks of linkage disequilibrium (LD) around the 121 autosomal GWAS lead risk variants using a Danish cohort of 90,312 individuals.[9,19] Within the LD-blocks, we analyzed rare variants (minor allele frequency (MAF)<0.1%, MAF<1%, and MAF<5%) in genes and regulatory regions using a Sequence Kernel Association Test. For the analysis, we used WGS data from a large cohort of 155 families ($n_{individuals}$=1,040) with clustering of migraine. The findings were replicated in an independent case-control cohort of migraine patients with no familial history of migraine and controls.

MATERIALS AND METHODS

Migraine family cohort

The migraine family cohort consisted of 155 families with a Mendelian-like inheritance pattern of migraine. The participants were recruited from the Danish Headache Center, Rigshospitalet-Glostrup, Denmark, as described elsewhere.[20,21] All participants were examined by a neurology resident or a senior medical student specifically trained in headache diagnostics. The participants were examined using a semi-structured interview[22,23] based on the International Classification of Headache Disorders.[2] The families consisted of 1,040 participants, including 746 individuals diagnosed with migraine and 294 without migraine. The mean number of participants per family was 6.7 and the mean number of migraine patients per family was 4.8. Most of the families were extended beyond the nuclear family. The smallest families consisted of at least two individuals with at least one migraine patient.

Replication cohort

The replication cohort consisted of 2,027 migraine patients with no familial history of migraine, i.e. unrelated migraine patients, and 1,650 controls. The unrelated migraine patients were recruited and assessed at the Danish Headache Center using the same procedures as described for the family cohort. The controls were recruited for the Danish study of Non-Invasive testing in Coronary Artery Disease (Dan-NICAD) according to procedures described by Nissen *et al*..[24]

96 **Whole genome sequencing**

97

98 Blood samples were taken from all participants and genomic DNA was extracted from whole blood.
99 The sequencing was performed on an Illumina NovaSeq 6000 sequencing platform with S4 flow
100 cells. The WGS data was subjected to quality control by deCODE genetics, as described
101 elsewhere.[25] All samples from both the migraine family cohort and the replication cohort were
102 subjected to the same sequencing and quality control procedures.

103

104 **Analytical approach**

105

106 **Selecting genomic regions for analysis**

107

108 We calculated blocks of linkage disequilibrium (LD) around the 121 autosomal migraine risk loci,[9]
109 using the genomic positions of the GWAS lead risk variants, i.e. the variants with the smallest
110 GWAS *p*-value. For this we used 90,312 participants from the Danish Blood Donor Study
111 (DBDS).[19] The genomic positions of the GWAS lead risk variants were based on SNP positions
112 from the Database of Single Nucleotide Polymorphisms (dbSNP) build 153 in the GRCh38.p12
113 (hg38) assembly. The LD-blocks were calculated using a LD-block recognition algorithm proposed
114 by Gabriel *et al*.[26] with PLINK2.[27] Here, we assessed variant-pairs within 1000 kilobases of each
115 other with a MAF>5% using a 90% D' confidence interval threshold >0.70. Subsequently, we
116 annotated genes and regulatory regions within the LD-blocks and analyzed these. If a LD-block
117 partially spanned a gene or regulatory region, we included the entire gene or regulatory region in
118 the analysis. If a LD-block was located in an intergenic region, we extended the genomic region for
119 analysis to include the nearest protein- or RNA-coding gene. No LD-block was generated for 22

loci. In such case, the nearest protein- or RNA-coding gene was included in the analysis. The genomic positions of the regions we analyzed are found in Table S1.

Annotating genes and regulatory regions for analysis

The gene transcripts were derived from GENCODE release 36 (basic gene-set only). They included protein-coding genes, non-coding RNA-genes, and pseudogenes. We analyzed the genes in their full length. In addition, we analyzed the coding exons, introns, 5' untranslated regions (UTR) exons, and 3' UTR exons by themselves. For regulatory regions, we analyzed promoters, enhancers, CpG islands, insulators, Polycomb Response Elements (PREs), and Transcription Factor Binding Sites (TFBSs). We included promoters from The Eukaryotic Promoter Database (EPD) version 6 and EPD non-coding promoters version 1.[28] The enhancers were included from the GeneHancer project,[29] using only enhancers that are supported by multiple studies. The genomic positions of the CpG islands were based on predictions by Gardiner-Garden *et al.*.[30] The insulators and PREs were included from the Broad Institute Chromatin State Segmentation by HMM[31] from The Encyclopedia of DNA Elements (ENCODE) Consortium[32] in H1-hESC cells in the GRCh37 (hg19) assembly. The TFBS were included from ENCODE[32] in the GRCh37 (hg19) assembly. All genomic positions in hg19 were converted to the CRCh38 (hg38) assembly. Table S2 displays the number of genes and regulatory regions that were analyzed in this study.

Preparing VCF files for analysis

The VCF files of each participant in the study were merged into one file. Multiallelic sites were converted to a biallelic representation. Subsequently, we annotated genes and regulatory regions to

the merged VCF file. We isolated variants, including SNVs, insertions, and deletions, with a MAF<0.1%, MAF<1%, or MAF<5%, using BCFtools.[33]

Rare variant association analysis

To test for an association between the rare variants in genes or regulatory regions and migraine, we used the software of Family Sequence Kernel Association Test (F-SKAT).[34] Age and gender were used as covariates. The age was defined as the age at the time of the interview (time of diagnosis). To avoid association merely given that longer genes and regulatory regions can contain more rare variants, we used modified beta-values.[35] The beta-values were modified by dividing them with the length of the genes or regulatory regions. The *p*-values were controlled for multiple testing using the Bonferroni method based on the number of genes and regulatory regions tested. A Bonferroni-corrected *p*-value<0.05 was considered significant.

To assess if the results were dependent of the GWAS lead risk variants, the number of risk alleles of the variants (0, 1, or 2) were included as covariates in a separate rare variant association analysis with F-SKAT.

Replication

The findings were replicated using an independent cohort of unrelated migraine patients and controls. Aggregated data were available on rare variants (MAF<0.1%, MAF<1%, MAF<5%) in the genes and regulatory regions for the controls. We tested if the frequency of the rare variants was increased in the migraine patients compared with the controls, using a Fisher's Exact Test.[36] We used the null hypothesis of no association with migraine under a dominant model of penetrance.

The resulting *p*-values were controlled for multiple testing using the Bonferroni method. A Bonferroni-corrected *p*-value<0.05 was considered significant.

Characterization of findings

To explore the effect of the rare variants of the replicated results, we used Ensembl Variant Effect Predictor.[37] To discover the target genes of the replicated results, we used *cis*-eQTLs data from the Genotype-Tissue Expression (GTEx) version 8.0 dataset (dbGaP Accession phs000424.v8.p2 release).

Ethics

Oral and written consent was obtained from all participants. All procedures performed in studies involving human participants were in accordance with the ethical standards of the National Committee on Health Research Ethics (file no. H-2-2010-122) and with the 1964 Helsinki declaration and its later amendments. The study was approved by the data-protection agency (01080/GLO-2010-10).

RESULTS

Rare variants in regulatory regions associate with migraine

We found an association between migraine and rare variants with a MAF<0.1% in 7 regulatory regions and an intron of the *MACF1* gene (see Table 1 for genomic regions and *p*-values). No

192 association was found for rare variants with a MAF<1% or a MAF<5% nor in coding regions (full
193 gene transcripts, coding exons, 5' UTR exons, 3' UTR exons). In line with what is expected when
194 analyzing rare variants for polygenic traits,[9,38,39] we found that the distribution of the test statistics
195 deviated from the null with inflated genomic inflation factors (λ=1.89, λ=1.63, and λ=1.68) (Figure
196 S1).
197
198 Subsequently, we replicated the Polycomb Response Element (PRE) located at chr9:117,411,901-
199 117,412,101 in the *ASTN2* locus ($p=1.1\cdot10^{-15}$) using an independent case-control cohort, henceforth
200 referred to as the replicated PRE.
201
202 We found that the findings were independent of the GWAS lead risk variants, with unaffected
203 results ($p=1.3\cdot10^{-4}$ in the original analysis (see Table 1) and $p=1.2\cdot10^{-4}$ with the allele count of the
204 risk variants as covariates).
205
206 **Table 1. Summary of the results obtained from the rare variant association analysis.** The
207 results from the rare variant association analysis together with the locus ID, as given in the meta-
208 analysis of migraine GWAS by Hautakangas *et al.*[9], the genomic positions in GRCh38.p12 (hg38),
209 and type of genomic feature.
210

Locus ID	Genomic positions	Type	*p*-value	Corrected *p*-value
MACF1	chr1:39,303,078-39,309,569	Intron	$1.8\cdot10^{-5}$	0.019
THADA	chr2:43,400,466-43,400,944	Transcription Factor Binding Site	$5.3\cdot10^{-6}$	0.013

PHACTR1	chr6:12,748,985-12,750,439	Enhancer	$1.0 \cdot 10^{-4}$	0.021
ASTN2	chr9:117,411,901-117,412,101	Polycomb Response Element	$1.3 \cdot 10^{-4}$	0.012
near *SPATA19*	chr11:133,939,661-133,940,767	CpG island	$1.9 \cdot 10^{-4}$	0.027
	chr14:26,917,154-26,917,554	Insulator	$1.2 \cdot 10^{-4}$	0.014
near *LRFN5*	chr14:41,610,647-41,610,847	Polycomb Response Element	$6.7 \cdot 10^{-5}$	$6.2 \cdot 10^{-3}$
	chr14:41,610,658-41,610,787	Enhancer	$6.7 \cdot 10^{-5}$	0.028

Rare variants in PRE map to CTCF binding sites

The replicated PRE was located in the first intron of the *ASTN2* gene and overlapped the TFBS of CTCF, SIN3A, ZNF444, GATA2, RCOR1, and MYC (Figure 1). There were 12 rare variants with a MAF<0.1% located in the replicated PRE. These rare variants consisted of 1 SNV, 9 deletions, and 2 insertions (Table 2). We found that all rare variants mapped to CTCF binding sites (Table 2) using the Ensembl Variant Effect Predictor (VEP).

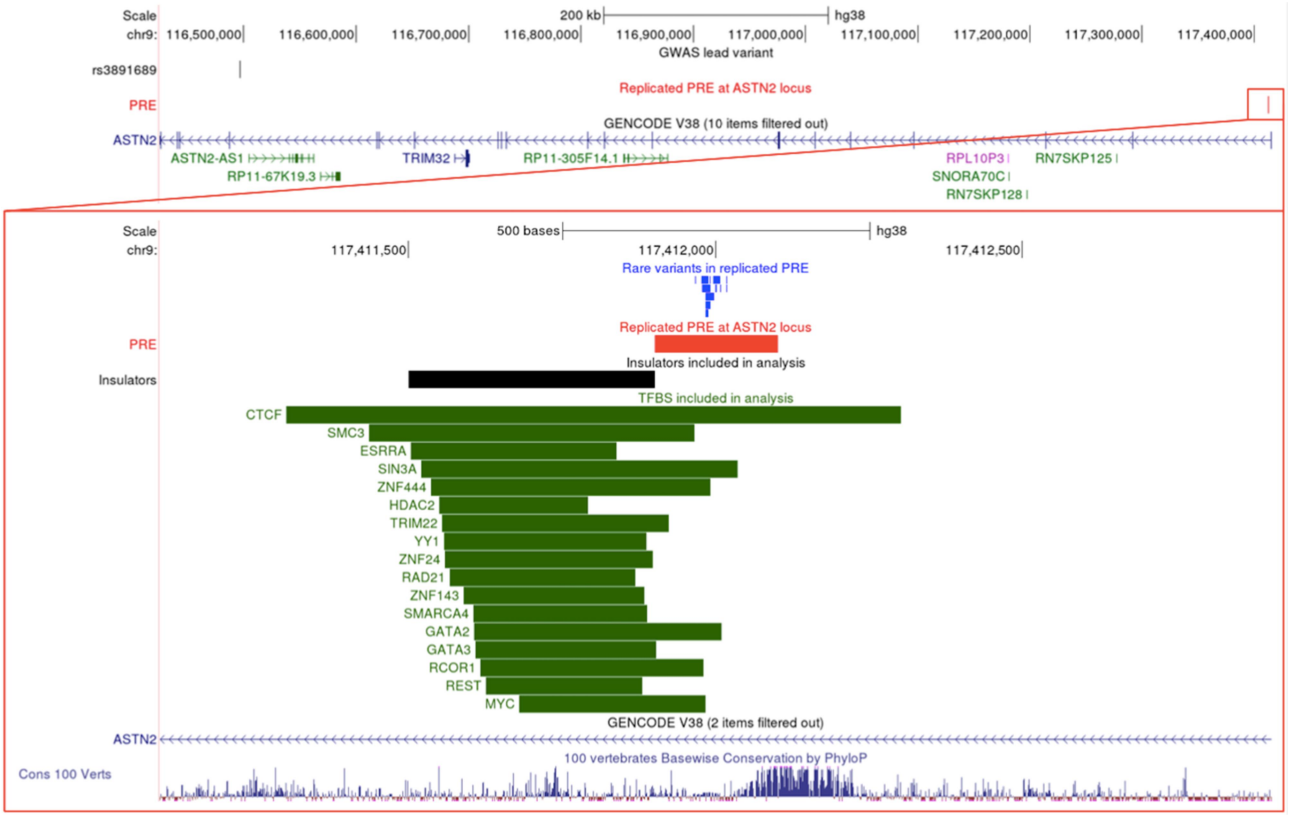


Figure 1. Overview of the genomic region around the *ASTN2* migraine risk locus. Top: The genomic region that was included for analysis (see Table S1 for genomic positions). Displayed is the GWAS lead risk variant (rs3891689), the replicated Polycomb Response Element, and the genes included for analysis. **Bottom:** A close-up of the replicated Polycomb Response Element and its surrounding genomic region. Displayed are the rare variants with a MAF<0.1%, that associated with migraine, the replicated Polycomb Response Element, near-by regulatory regions (an insulator and 17 transcription factor binding sites), the first intron of the *ASTN2* gene, and the evolutionary conservation across 100 vertebrates.

Table 2. Information on the rare variants in the replicated Polycomb Response Element. The table presents the genomic positions in GRCh38.p12 (hg38), the reference allele, the alternative

allele, the variant type, the rs number, and the Ensembl regulatory feature of the rare variants (MAF<0.1%) in the replicated Polycomb Response Element.

Genomic position	Reference allele	Alternative allele	Variant type	rs number	Ensembl regulatory feature
chr9:117411967-117411967	A	C	SNV	rs866789329	CTCF binding site
chr9:117411977-117411988	CCACCCCTACCA	C	Deletion		CTCF binding site
chr9:117411978-117411991	CACCCCTACCACCA	C	Deletion	rs1588023872	CTCF binding site
chr9:117411983-117411997	CTACCA	C	Deletion	rs1446477137	CTCF binding site
chr9:117411983-117411991	CTACCACCA	C	Deletion		CTCF binding site
chr9:117411983-117411988	CTACCACCACCCCCA	C	Deletion		CTCF binding site
chr9:117411990-117411991	CA	C	Deletion		CTCF binding site
chr9:117411996-117412007	CACCTCACCCCT	C	Deletion		CTCF binding site
chr9:117411999-117412002	CTCA	C	Deletion	rs1383312195	CTCF binding site
chr9:117412007-117412008	TC	T	Deletion	rs145255224	CTCF binding site
chr9:117412018-117412018	A	ACCCCCCCCCC	Insertion		CTCF binding site
chr9:117412018-117412018	A	ACCCCCCCCCCC	Insertion		CTCF binding site

235

236 **Possible regulatory targets unidentified**

237

238 Finally, we assessed the possible regulatory targets of the replicated PRE. We found no significant
239 *cis*-eQTLs in the PRE using the GTEx database for the identification of regulatory targets.

240

241 **DISCUSSION**

242

243 Our study is, to our knowledge, the first to assess if rare variants in 121 autosomal migraine risk
244 loci increase the risk of migraine. We assessed if rare variants in genes and regulatory regions
245 associated with migraine, using WGS data from families with clustering of migraine. We replicated
246 a Polycomb Response Element (PRE) in the *ASTN2* locus using an independent case-control cohort.

247

248 **Regulatory regions harbor rare variants associated with migraine**

249

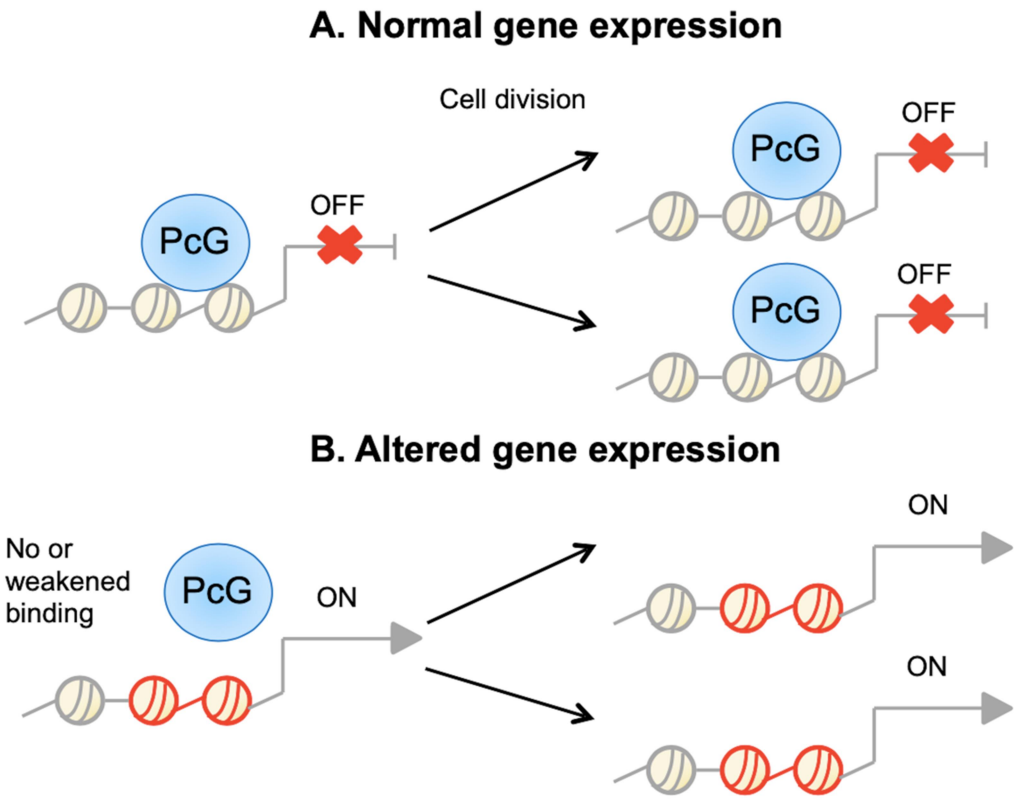
250 We found an association between migraine and rare variants with a MAF<0.1% in 7 regulatory
251 regions and an intron of the *MACF1* gene. Our results indicate that migraine can be associated with
252 rare variants in regulatory regions, since introns are likely to harbor elements with a regulatory
253 function.[40] We replicated a PRE in the *ASTN2* locus using an independent case-control cohort.
254 Interestingly, we have previously reported a rare variant association with migraine in a PRE in the
255 same locus, using a different study design.[12] Importantly, we show that the effect of the replicated
256 PRE was independent of the GWAS lead risk variant. Thus, our results show that migraine GWAS
257 loci also contain independent rare risk variants.

258

Disruption in epigenetic regulation as a possible biological mechanism for migraine

The function of mammalian PREs is still largely unknown.[41] Mammalian PREs are likely capable of binding Polycomb Group (PcG) proteins as a part of transcriptional silencing during cell division, where the silencing is maintained in subsequent daughter cells.[41–43] Associated with the silencing is remodeling of the chromatin.[44,45] Possibly, the rare variants we find associated with migraine in the replicated PRE can cause a disruption in the binding of PcG proteins and chromatin remodeling. The disruption may lead to transcriptional silencing not being maintained after cell division. This ultimately leads to an altered gene expression (Figure 2). Epigenetic mechanisms have been linked to brain plasticity in the transition from episodic to chronic migraine.[13,14] However, we are, to our knowledge, the first to report an indication of an actual biological mechanism for migraine, where chromatin remodeling is disrupted by rare variants in PREs.

We found that the rare mutations in the replicated PRE were situated in binding sites of the transcription factor CTCF. CTCF is a transcriptional repressor, with many functions including regulation of gene expression through chromatin modifications.[46] Again, our findings indicate a link to a disruption in chromatin remodeling as a biological mechanism for migraine.



277

278 **Figure 2. Schematic overview of transcriptional silencing by PcG proteins in absence or**
279 **presence of rare variants in the DNA. A.** The recruitment of PcG proteins to the DNA results in
280 transcriptional silencing that is maintained after cell division. **B.** A disruption of PcG binding, e.g.
281 by rare variants, results in no binding or a weakened binding of the PcG proteins. A disrupted
282 binding leads to an altered gene expression where gene silencing is disturbed. The altered gene
283 expression is maintained after cell division.

284

285 ***ASTN2* as a possible regulatory target gene**

286

287 We were not able to identify regulatory targets of the replicated PRE based on significant *cis*-
288 eQTLs. Thus, the target gene could be assigned to the nearest gene, as done by Ringrose *et al.*.[47]
289 The nearest gene is *ASTN2*, of which the PRE is overlapping. The *ASTN2* gene encodes the
290 Astrotactin 2 protein. The Astrotactin 2 protein is expressed in the brain and is suggested to play a

part in neuronal migration. Thus, our findings support the hypothesis of migraine being a neuronal disorder.[48]

Implications for genetic studies on migraine

The most recent meta-analysis of migraine GWAS only explains 10.6% of the total heritability of migraine.[9] GWAS are often underpowered to detect rare variants with low-penetrant alleles.[10] We found that migraine associated with very rare variants, i.e. variants with a MAF<0.1%, in the migraine risk loci. Our results indicate that a part of the unexplained heritability of migraine could arise from rare variants not detected by GWAS. Also, our results suggest that the unexplained heritability could be hidden in epigenetics. For discovering the unexplained heritability of migraine, future studies focusing on the identification of heritable epigenetic marks for migraine can be highly relevant.

CONCLUSION

We found that rare variants (MAF<0.1%) in a Polycomb Response Element in the *ASTN2* locus associated with migraine. The association was independent of the original, common variant in the locus. Thus, migraine GWAS loci also contain independent rare risk variants. We propose a biological mechanism for migraine by which chromatin remodeling is disrupted by rare variants in Polycomb Response Elements.

ABBREVIATIONS

GWAS	Genome-Wide Association Study
LD	Linkage disequilibrium
MAF	Minor allele frequency
PcG	Polycomb Group protein
PRE	Polycomb Response Element
TFBS	Transcription factor binding site
UTR	Untranslated region
WGS	Whole genome sequencing

FUNDING

The study was funded by a grant from Candys Foundation "CEHEAD" and a grant from Jascha Fonden (web site: https://jaschafonden.dk/). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

COMPETING INTERESTS

The authors have declared that no competing interests exist.

DATA AVAILABILITY

The data supporting the conclusions of this article are unavailable, as they contain information that may be used for identifying study participants. However, summary data supporting the conclusions of this article are available on request to the corresponding author of this article.

ACKNOWLEDGEMENTS

We thank deCODE genetics for whole genome sequencing and quality-controlling our samples. We thank the Danish blood donors for their valuable participation in the Danish Blood Donor Study as well as the staff at the blood centers for making this study possible. Finally, we thank all of the staff and participating patients at the Danish Headache Center for their willing assistance.

CONSORTIA

Members list of the DBDS Genomic Consortium:

Steffen Andersen[1], Karina Banasik[2], Søren Brunak[2], Kristoffer Burgdorf[3], Christian Erikstrup[4], Thomas Folkmann Hansen[5], Henrik Hjalgrim[6], Gregor Jemec[7], Poul Jennum[8], Per Ingemar Johansson[3], Kasper Rene Nielsen[9], Mette Nyegaard[10], Mie Topholm Bruun[11], Ole Birger Pedersen[12], Susan Mikkelsen[4], Khoa Manh Dinh[4], Erik Sørensen[3], Henrik Ullum[3], Sisse Ostrowski[3], Thomas Werge[13], Daniel Gudbjartsson[14], Kari Stefansson[14], Hreinn Stefánsson[14], Unnur Þorsteinsdóttir[14], Margit Anita Hørup Larsen[3], Maria Didriksen[3], Susanne Sækmose[12].

[1] Department of Finance, Copenhagen Business School, Copenhagen, Denmark.

[2] Novo Nordisk Foundation Center for Protein Research, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark

[3] Department of Clinical Immunology, Copenhagen University Hospital, Copenhagen, Denmark

[4] Department of Clinical Immunology, Aarhus University Hospital, Aarhus, Denmark

[5] Danish Headache Center, Department of Neurology, Copenhagen University Hospital, Glostrup, Denmark

[6] Department of Epidemiology Research, Statens Serum Institut, Copenhagen, Denmark

[7] Department of Clinical Medicine, Sealand University Hospital, Roskilde, Denmark

[8] Department of Clinical Neurophysiology, University of Copenhagen, Copenhagen, Denmark

[9] Department of Clinical Immunology, Aalborg University Hospital, Aalborg, Denmark

[10] Department of Biomedicine, Aarhus University, Aarhus, Denmark

[11] Department of Clinical Immunology, Odense University Hospital, Odense, Denmark

[12] Department of Clinical Immunology, Næstved Hospital, Næstved, Denmark

[13] Institute of Biological Psychiatry, Mental Health Centre Sct. Hans, Copenhagen University Hospital, Roskilde, Denmark

[14] deCODE genetics, Reykjavik, Iceland

REFERENCES

1. Stovner, L. J. *et al.* Global, regional, and national burden of migraine and tension-type headache, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet Neurol.* **17**, 954–976 (2018).
2. Headache Classification Committee of the International Headache Society (IHS) The International Classification of Headache Disorders, 3rd edition. *Cephalalgia* **38**, 1–211 (2018).
3. 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 (London, England)* **396**, 1204–1222 (2020).

4. Stovner, L. J. & Andrée, C. Impact of headache in Europe: a review for the Eurolight project On behalf of the Eurolight Steering Committee. *J Headache Pain* **9**, 139–146 (2008).

5. Mulder, E. J. *et al.* Genetic and environmental influences on migraine: a twin study across six countries. *Twin Res* **6**, 422–431 (2003).

6. Ulrich, V., Gervil, M., Kyvik, K. O., Olesen, J. & Russell, M. B. The inheritance of migraine with aura estimated by means of structural equation modelling. *J. Med. Genet.* **36**, 225–227 (1999).

7. Hansen, R. D., Christensen, A. F. & Olesen, J. Family studies to find rare high risk variants in migraine. *J Headache Pain* **18**, 32 (2017).

8. Levy, D., Strassman, A. M. & Burstein, R. A critical view on the role of migraine triggers in the genesis of migraine pain. *Headache* **49**, 953–957 (2009).

9. Hautakangas, H. *et al.* Genome-wide analysis of 102,084 migraine cases identifies 123 risk loci and subtype-specific risk alleles. *medRxiv* 2021.01.20.21249647 (2021) doi:10.1101/2021.01.20.21249647.

10. Manolio, T. A. *et al.* Finding the missing heritability of complex diseases. *Nature* **461**, 747–753 (2009).

11. Nicolae, D. L. *et al.* Trait-associated SNPs are more likely to be eQTLs: annotation to enhance discovery from GWAS. *PLoS Genet* **6**, e1000888 (2010).

12. Techlo, T. R. *et al.* Familial analysis reveals rare risk variants for migraine in regulatory regions. *Neurogenetics* (2020) doi:10.1007/s10048-020-00606-5.

13. Eising, E., A Datson, N., van den Maagdenberg, A. M. J. M. & Ferrari, M. D. Epigenetic mechanisms in migraine: a promising avenue? *BMC Med.* **11**, 26 (2013).

14. Winsvold, B. S. *et al.* Epigenetic DNA methylation changes associated with headache chronification: A retrospective case-control study. *Cephalalgia* **38**, 312–322 (2018).

15. Corominas, R. *et al.* Association study of the serotoninergic system in migraine in the Spanish population. *Am. J. Med. Genet. Part B, Neuropsychiatr. Genet. Off. Publ. Int. Soc. Psychiatr. Genet.* **153B**, 177–184 (2010).

16. Lafreniere, R. G. *et al.* A dominant-negative mutation in the TRESK potassium channel is linked to familial migraine with aura. *Nat Med* **16**, 1157–1160 (2010).

17. Nyholt, D. R. *et al.* A high-density association screen of 155 ion transport genes for involvement with common migraine. *Hum. Mol. Genet.* **17**, 3318–3331 (2008).

18. de Vries, B. *et al.* Systematic re-evaluation of genes from candidate gene association studies in migraine using a large genome-wide association data set. *Cephalalgia* **36**, 604–614 (2016).

19. Hansen, T. F. *et al.* DBDS Genomic Cohort, a prospective and comprehensive resource for integrative and temporal analysis of genetic, environmental and lifestyle factors affecting health of blood donors. *BMJ Open* **9**, e028401 (2019).

20. Chalmer, M. A., Hansen, T. F. & Olesen, J. Nosographic analysis of osmophobia and field testing of diagnostic criteria including osmophobia. *Cephalalgia* **39**, 38–43 (2019).

21. Eriksen, M. K., Thomsen, L. L., Andersen, I., Nazim, F. & Olesen, J. Clinical characteristics of 362 patients with familial migraine with aura. *Cephalalgia* **24**, 564–575 (2004).

22. Gervil, M., Ulrich, V., Olesen, J. & Russell, M. B. Screening for migraine in the general population: validation of a simple questionnaire. *Cephalalgia* **18**, 342–348 (1998).

23. Rasmussen, B. K., Jensen, R. & Olesen, J. Questionnaire versus clinical interview in the diagnosis of headache. *Headache* **31**, 290–295 (1991).

24. Nissen, L. *et al.* Danish study of Non-Invasive testing in Coronary Artery Disease (Dan-NICAD): study protocol for a randomised controlled trial. *Trials* **17**, 262 (2016).

25. Jonsson, H. *et al.* Whole genome characterization of sequence diversity of 15,220 Icelanders.

Sci. data **4**, 170115 (2017).

26. Gabriel, S. B. *et al.* The structure of haplotype blocks in the human genome. *Science* **296**, 2225–2229 (2002).

27. Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* **81**, 559–575 (2007).

28. Dreos, R., Ambrosini, G., Cavin Perier, R. & Bucher, P. EPD and EPDnew, high-quality promoter resources in the next-generation sequencing era. *Nucleic Acids Res* **41**, D157-64 (2013).

29. Fishilevich, S. *et al.* GeneHancer: genome-wide integration of enhancers and target genes in GeneCards. *Database (Oxford)* **2017**, (2017).

30. Gardiner-Garden, M. & Frommer, M. CpG islands in vertebrate genomes. *J Mol Biol* **196**, 261–282 (1987).

31. Ernst, J. & Kellis, M. Discovery and characterization of chromatin states for systematic annotation of the human genome. *Nat Biotechnol* **28**, 817–825 (2010).

32. Gerstein, M. B. *et al.* Architecture of the human regulatory network derived from ENCODE data. *Nature* **489**, 91–100 (2012).

33. Li, H. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* **27**, 2987–2993 (2011).

34. Yan, Q. *et al.* A Sequence Kernel Association Test for Dichotomous Traits in Family Samples under a Generalized Linear Mixed Model. *Hum Hered* **79**, 60–68 (2015).

35. Wu, M. C. *et al.* Rare-variant association testing for sequencing data with the sequence kernel association test. *Am J Hum Genet* **89**, 82–93 (2011).

36. Fisher, R. A. *Statistical methods for research workers*. (Oliver and Boyd) (1934).

458 37. McLaren, W. *et al.* The Ensembl Variant Effect Predictor. *Genome Biol* **17**, 122 (2016).
459 38. Mathieson, I. & McVean, G. Differential confounding of rare and common variants in
460 spatially structured populations. *Nat Genet* **44**, 243–246 (2012).
461 39. Pirie, A., Wood, A., Lush, M., Tyrer, J. & Pharoah, P. D. The effect of rare variants on
462 inflation of the test statistics in case-control analyses. *BMC Bioinformatics* **16**, 53 (2015).
463 40. Maston, G. A., Evans, S. K. & Green, M. R. Transcriptional regulatory elements in the
464 human genome. *Annu Rev Genomics Hum Genet* **7**, 29–59 (2006).
465 41. Klymenko, T. & Müller, J. The histone methyltransferases Trithorax and Ash1 prevent
466 transcriptional silencing by Polycomb group proteins. *EMBO Rep.* **5**, 373–377 (2004).
467 42. Reinig, J., Ruge, F., Howard, M. & Ringrose, L. A theoretical model of Polycomb/Trithorax
468 action unites stable epigenetic memory and dynamic regulation. *Nat. Commun.* **11**, 4782
469 (2020).
470 43. Chan, C. S., Rastelli, L. & Pirrotta, V. A Polycomb response element in the Ubx gene that
471 determines an epigenetically inherited state of repression. *EMBO J.* **13**, 2553–2564 (1994).
472 44. Kalb, R. *et al.* Histone H2A monoubiquitination promotes histone H3 methylation in
473 Polycomb repression. *Nat. Struct. Mol. Biol.* **21**, 569–571 (2014).
474 45. Cao, R. *et al.* Role of histone H3 lysine 27 methylation in Polycomb-group silencing. *Science*
475 **298**, 1039–1043 (2002).
476 46. Phillips, J. E. & Corces, V. G. CTCF: master weaver of the genome. *Cell* **137**, 1194–1211
477 (2009).
478 47. Ringrose, L., Rehmsmeier, M., Dura, J.-M. & Paro, R. Genome-wide prediction of
479 Polycomb/Trithorax response elements in Drosophila melanogaster. *Dev. Cell* **5**, 759–771
480 (2003).
481 48. Goadsby, P. J. Pathophysiology of migraine. *Neurol Clin* **27**, 335–360 (2009).

482