Common Variant Burden Contributes to the Familial Aggregation of Migraine in 1,589 Families

Gormley, P.; Kurki, M.I.; Hiekkala, M.E.; Veerapen, K.; Happola, P.; Mitchell, A.A.; ... ; Int Headache Genetics Consortium

HHS Public Access

Author manuscript

Neuron. Author manuscript; available in PMC 2019 May 16.

Published in final edited form as:

Neuron. 2018 May 16; 98(4): 743–753.e4. doi:10.1016/j.neuron.2018.04.014.

Common variant burden contributes to the familial aggregation of migraine in 1,589 families

Padhraig Gormley[1,2,3,27], **Mitja I. Kurki**[1,2,3,27], **Marjo Eveliina Hiekkala**[4,27], **Kumar Veerapen**[1,2,3], **Paavo Häppölä**[5], **Adele A. Mitchell**[6,25], **Dennis Lal**[1,2,3,7], **Priit Palta**[5], **Ida Surakka**[5], **Mari Anneli Kaunisto**[5], **Eija Hämäläinen**[5], **Salli Vepsäläinen**[8], **Hannele Havanka**[9], **Hanna Harno**[8,10], **Matti Ilmavirta**[11], **Markku Nissilä**[12], **Erkki Säkö**[13], **Marja-Liisa Sumelahti**[14], **Jarmo Liukkonen**[15], **Matti Sillanpää**[16], **Liisa Metsähonkala**[17], **Seppo Koskinen**[18], **Terho Lehtimäki**[19], **Olli Raitakari**[20,21], **Minna Männikkö**[22], **Caroline Ran**[23], **Andrea Carmine Belin**[23], **Pekka Jousilahti**[18], **Verneri Anttila**[1,2,3,5], **Veikko Salomaa**[18], **Ville Artto**[8], **Markus Färkkilä**[8], **23andMe Research Team**[24], **International Headache Genetics Consortium (IHGC)**, **Heiko Runz**[6,26], **Mark J. Daly**[1,2,3], **Benjamin M. Neale**[1,2,3], **Samuli Ripatti**[5], **Mikko Kallela**[8], **Maija Wessman**[4,5], and **Aarno Palotie**[1,2,3,5,28,*]

[*]Correspondence: aarno.palotie@helsinki.fi (A.P.).
[25]Current address: Eisai Inc., 4 Corporate Drive, Andover, Massachusetts 01810, USA.
[26]Current address: Biogen, Research and Early Development, Cambridge, MA, USA.
[27]These authors contributed equally.
[28]Lead Contact

CONSORTIA

The members of the 23andMe Research Team are Michelle Agee, Babak Alipanahi, Adam Auton, Robert K. Bell, Katarzyna Bryc, Sarah L. Elson, Pierre Fontanillas, Nicholas A. Furlotte, Karen E. Huber, Aaron Kleinman, Nadia K. Litterman, Jennifer C. McCreight, Matthew H. McIntyre, Joanna L. Mountain, Carrie A.M. Northover, Steven J. Pitts, J. Fah Sathirapongsasuti, Olga V. Sazonova, Janie F. Shelton, Suyash Shringarpure, Chao Tian, Joyce Y. Tung, Vladimir Vacic, and Catherine H. Wilson.

The members of the IHGC are Verneri Anttila, Ville Artto, Andrea Carmine Belin, Dorret I. Boomsma, Sigrid Børte, Daniel I. Chasman, Lynn Cherkas, Anne Francke Christensen, Bru Cormand, Ester Cuenca-Leon, George Davey-Smith, Martin Dichgans, Cornelia van Duijn, Tonu Esko, Ann-Louise Esserlind, Michel Ferrari, Rune R. Frants, Tobias Freilinger, Nick Furlotte, Padhraig Gormley, Lyn Griffiths, Eija Hamalainen, Thomas Folkmann Hansen, Marjo Hiekkala, M Arfan Ikram, Andres Ingason, Marjo-Riitta Järvelin, Risto Kajanne, Mikko Kallela, Jaakko Kaprio, Mari Kaunisto, Christian Kubisch, Mitja Kurki, Tobias Kurth, Lenore Launer, Terho Lehtimaki, Davor Lessel, Lannie Ligthart, Nadia Litterman, Arn van den Maagdenberg, Alfons Macaya, Rainer Malik, Massimo Mangino, George McMahon, Bertram Muller-Myhsok, Benjamin M. Neale, Carrie Northover, Dale R. Nyholt, Jes Olesen, Aarno Palotie, Priit Palta, Linda Pedersen, Nancy Pedersen, Danielle Posthuma, Patricia Pozo-Rosich, Alice Pressman, Lydia Quaye, Olli Raitakari, Markus Schürks, Celia Sintas, Kari Stefansson, Hreinn Stefansson, Stacy Steinberg, David Strachan, Gisela Terwindt, Marta Vila-Pueyo, Maija Wessman, Bendik S. Winsvold, William Wrenthal, Huiying Zhao, and John-Anker Zwart.

AUTHOR CONTRIBUTIONS

Conceptualization, P.G., M.I.K., M.E.H., H.R., M.J.D., B.M.N., S.R., M.K., M.W., and A.P.; Methodology, P.G., M.I.K., M.E.H., K.V., P.H., A.A.M., M.J.D., B.M.N., S.R., M.K., M.W., and A.P.; Software, P.G., M.I.K., K.V., and P.H.; Validation, S.K., T.L., O.R., M.M., C.R., and A.C.B.; Formal Analysis, P.G., M.I.K., M.E.H., K.V., and P.H.; Investigation, S.V., H.Hav., H.Har., M.I., M.N., E.S., L-M.S., J.L., M.S., L.M., V.Ar., M.F., and M.K.; Resources, S.V., H.Hav., H.Har., M.I., M.N., E.S., L-M.S., J.L., M.S., L.M., M.K., 23andMe, and IHGC; Data Curation, M.E.H., P.H., P.P., I.S., M.A.K., E.H., and M.W.; Writing – Original Draft, P.G., M.I.K., M.E.H., M.W., and A.P.; Writing – Review & Editing, P.G., M.I.K., M.E.H., K.V., P.H., A.A.M., D.L., P.P., I.S., M.A.K., E.H., S.V., H.Hav., H.Har., M.I., M.N., E.S., L-M.S., J.L., M.S., L.M., S.K., T.L., O.R., M.M., C.R., A.C.B., P.J., V.An., V.S., V.Ar., M.F., 23andMe, IHGC, H.R., M.J.D., B.M.N., S.R., M.K., M.W., and A.P.; Visualization, P.G., M.I.K., M.E.H., and K.V.; Supervision, M.A.K., M.J.D., B.M.N., S.R., M.K., M.W., and A.P.; Funding Acquisition, H.R., M.W., and A.P.

DECLARATION OF INTERESTS

The study was partially funded by Merck and Co., Kenilworth, NJ, USA. Members of the 23andMe Research Team are current or former employees of 23andMe, Inc., and hold stock or stock options in 23andMe.

[1]Massachusetts General Hospital, Boston, MA, USA [2]Harvard Medical School, Boston, MA, USA [3]Broad Institute of MIT and Harvard, Cambridge, MA, USA [4]Institute of Genetics, Folkhälsan Research Center, Helsinki, Finland [5]Institute for Molecular Medicine Finland (FIMM), HiLIFE, University of Helsinki, Helsinki, Finland [6]Merck Research Laboratories, Merck and Co., Kenilworth, NJ, USA [7]Cologne Center for Genomics, University of Cologne, Cologne, Germany [8]Department of Neurology, Helsinki University Central Hospital, Helsinki, Finland [9]Regional State Administrative Agency for Northern Finland, Oulu, Finland [10]Division of Pain Medicine, Department of Anesthesiology, Intensive Care and Pain Medicine, University of Helsinki and Helsinki University Hospital, Finland [11]Department of Neurology, Central Hospital Central Finland, Jyväskylä [12]Terveystalo Clinical Research, Turku, Finland [13]Turku Headache Center, Turku, Finland [14]Terveystalo, Tampere, Finland [15]Lääkärikeskus Ikioma, Mikkeli, Finland [16]Departments of Child Neurology and General Practice, University of Turku, and Turku University Hospital, Turku, Finland [17]Epilepsy Unit for Children and Adolescents, Helsinki University Hospital, Helsinki, Finland [18]National Institute for Health and Welfare, Helsinki, Finland [19]Department of Clinical Chemistry, Fimlab Laboratories, Faculty of Medicine and Life Sciences, University of Tampere, Finland [20]Research Centre of Applied and Preventive Cardiovascular Medicine, University of Turku, Turku, Finland [21]Department of Clinical Physiology and Nuclear Medicine, Turku University Hospital, Turku, Finland [22]Northern Finland Birth Cohorts, Faculty of Medicine, University of Oulu, Oulu, Finland [23]Department of Neuroscience, Karolinska Institutet, Stockholm, Sweden [24]23andMe, Inc., Mountain View, CA, USA

SUMMARY

Complex traits, including migraine, often aggregate in families, but the underlying genetic architecture behind this is not well understood. The aggregation could be explained by rare, penetrant variants that segregate according to Mendelian inheritance or by the sufficient polygenic accumulation of common variants, each with an individually small effect, or a combination of the two hypotheses. In 8,319 individuals across 1,589 migraine families we calculated migraine polygenic risk scores (PRS) and found a significantly higher common variant burden in familial cases (n=5,317, OR=1.76, 95%CI=1.71–1.81, P=1.7×10^{-109}) compared to population cases from the FINRISK cohort (n=1,101, OR=1.32, 95%CI=1.25–1.38, P=7.2×10^{-17}). The PRS explained 1.6% of the phenotypic variance in the population cases and 3.5% in the familial cases (including 2.9% for migraine without aura, 5.5% for migraine with typical aura, and 8.2% for hemiplegic migraine). The results demonstrate a significant contribution of common polygenic variation to the familial aggregation of migraine.

INTRODUCTION

Familial aggregation in chronic diseases is well known but its background is not well understood (Agarwala et al., 2013). One hypothesis has been based on the Mendelian viewpoint that segregating, highly penetrant variants would strongly contribute to the familial nature of the disease. Linkage studies have had modest success in identifying highly penetrant disease variants; on the other hand, the numerous established genetic loci from genome-wide association studies (GWAS) rarely co-reside within linkage peaks. So far,

most of the whole-exome sequencing (WES) and whole-genome sequencing (WGS) studies have been underpowered to shed light on the question of why common diseases aggregate in families. Studies in familial dyslipidemias have shed some light on this, by demonstrating that both rare (penetrant) and common (less penetrant) variants have been associated to specific lipid traits (Khera et al., 2017a; Ripatti et al., 2016).

Migraine is an example of a common disease that can aggregate in families (Stewart et al., 1997). It is one of the most common brain disorders worldwide, affecting approximately 15–20% of the adult population in developed countries (Global Burden of Disease Study 2013 Collaborators, 2015). Therefore migraine studies may facilitate collection of the large sample sizes required to reveal some of the mechanisms of familial aggregation.

One third of migraine patients experience additional neurological symptoms during attacks, called aura (migraine with aura, MA, ICHD-3 code: 1.2). These can occur in rare forms called hemiplegic migraine (HM, ICHD-3 code: 1.2.3), typically accompanied by severe symptoms of motor weakness that can be either familial (FHM, ICHD-3 code: 1.2.3.1) or sporadic (SHM, ICHD-3 code: 1.2.3.2). Alternatively, the more common form is usually accompanied by a visual aura, called migraine with typical aura (ICHD-3 code: 1.2.1). Migraine that occurs without any aura symptoms is the most common subtype and is called migraine without aura (MO, ICHD-3 code: 1.1). Each subtype is diagnosed according to the third edition of the International Classification of Headache Disorders (ICHD-3) criteria (Headache Classification Committee of the International Headache Society (IHS), 2013).

A Mendelian inheritance model in familial migraine has been supported by mutations in three ion-transporter genes (*CACNA1A* (Ophoff et al., 1996), *ATP1A2* (De Fusco et al., 2003), and *SCN1A* (Dichgans et al., 2005)) identified by linkage studies and positional cloning of FHM families. However, mutations in these genes explain only a fraction of FHM/SHM cases (Thomsen et al., 2007, 2008) and none of the more common forms of migraine. Even in FHM these mutations vary in their penetrance, therefore, it is unlikely that penetrant mutations would entirely explain the observation that migraine is enriched in some families. Linkage studies in common forms of migraine have suggested several loci but no specific genes have been identified (Chasman et al., 2016).

The polygenic nature of migraine is well documented by GWAS that have identified over 40 loci associated to common forms of migraine (Anttila et al., 2010, 2013; Chasman et al., 2011; Freilinger et al., 2012; Gormley et al., 2016). Basic understanding of differences in the pathophysiology of these common forms (MA and MO) is limited. GWAS have identified many more common variant loci in MO than in MA (Anttila et al., 2010, 2013; Chasman et al., 2011; Freilinger et al., 2012; Gormley et al., 2016), likely due to the larger sample sizes collected for MO, but clear differences in prevalence (MA = 5%, MO = 12%) could instead point towards differences in the genetic architecture and heterogeneity of these diseases.

We hypothesize that in addition to some rare, highly penetrant variants, accumulation of common variants with small individual effect sizes contribute to the familial forms of migraine. To study this, we constructed a polygenic risk score (PRS) from the most recent migraine GWAS consisting of approximately 59,000 cases and 316,000 controls after

excluding all Finnish samples (Gormley et al., 2016). We then investigated the contribution of common polygenic and rare variation to migraine in our large migraine family cohort consisting of 1,589 families from Finland totaling 8,319 individuals, including 540 HM, 2,420 migraine with typical aura, 2,357 MO, and 3,002 family members with no migraine. We observed an overall increased PRS in familial migraine cases compared to population-based cases and controls and clear differences of the common variant load across different migraine subtypes.

RESULTS

In this study, we assessed the contribution of migraine-associated common genetic variation to a range of migraine subtypes in a collection of 1,589 families from Finland (Table S1 and Figure S1). We calculated the PRS for all 8,319 family members and 14,470 individuals from the FINRISK population cohort by combining each individual's genotypes with association summary statistics from a previously reported migraine GWAS (**STAR★METHODS**). The distributions of the PRS observed in each sample are shown in Figure 1. We then used the migraine PRS to assess the relative polygenic load contributing to both prevalent subtypes of migraine (migraine without aura and migraine with typical aura), to the rare subtype (HM), and to other available subtypes (Table 1).

Common polygenic load is enriched across migraine subtypes

Compared to 13,369 FINRISK population-based controls, we found that the burden of common migraine-associated variation, measured via the PRS, was enriched across all of the migraine subtypes in the family collection, including rare forms of the disease. Using a logistic mixed model (adjusted for sex, age, and genetic relatedness) to test for association between the PRS as a continuous variable and each of the migraine subtypes we found that the PRS was associated with all migraine cases combined (n = 5,317, OR = 1.76, 95% CI = 1.71–1.81, $P = 1.7\times10^{-109}$, Table 1 and Figure 2). We found the lowest enrichment for MO (n = 2,357, OR = 1.57, 95% CI = 1.51–1.63, $P = 1.1\times10^{-48}$), compared to significantly higher enrichment for migraine with typical aura (n = 2,420, OR = 1.85, 95% CI = 1.79–1.91, $P = 1.4\times10^{-86}$) and HM (n = 540, OR = 1.96, 95% CI = 1.86–2.07, $P = 8.7\times10^{-36}$). From additional analyses using only cases to compare between migraine subtypes, there was no significant difference in common variant burden between the migraine with aura subtypes, HM and migraine with typical aura (OR = 1.09, 95% CI = 0.99–1.19, $P = 0.09$), but both showed significantly higher enrichment compared to MO (OR = 1.28, 95% CI = 1.17–1.38, $P = 3.8\times10^{-6}$, and OR = 1.17, 95% CI = 1.11–1.23, $P = 7.3\times10^{-7}$, respectively, Table S2).

Investigating migraine with aura subtypes

We next looked at the four deeper-level subtypes of MA according to the ICHD-3 criteria. Two are subtypes of migraine with typical aura (ICHD-3 code: 1.2.1) called typical aura with headache (ICHD-3 code: 1.2.1.1) and typical aura without headache (ICHD-3 code: 1.2.1.2), and the other two are subtypes of HM, called FHM and SHM, where FHM is defined as those HM cases with at least one first- or second-degree relative that has also been diagnosed with HM (Table 1). We investigated if these sub-subtypes were any different

in terms of their polygenic burden. We found that there was no significant difference between FHM and SHM (OR = 0.94, 95% CI = 0.72–1.16, $P = 0.58$, Table S2 and Figure S2). However, for migraine with typical aura we found that typical aura with headache cases had significantly higher polygenic burden than typical aura without headache (OR = 1.85, 95% CI = 1.45–2.36, $P = 6.3\times10^{-7}$, Table S2 and Figure S2). In fact, the PRS burden in cases of typical aura without headache was observed to be equivalently as low as the FINRISK population controls (n = 70, OR = 0.85, 95% CI = 0.62 – 1.17, $P = 0.20$, Figure S2).

Variance explained by the PRS in familial and population-based cases

To quantify how much the common variation currently captured by the PRS contributes to migraine phenotypes, we calculated the variance explained by the PRS using a model with the PRS included compared to a model without the PRS. While we chose a *P*-value threshold of $P < 0.1$ for the PRS used throughout this study, we also calculated the variance explained across a range of different migraine GWAS *P*-value thresholds to confirm that the results would not be qualitatively different if we had chosen a higher *P*-value threshold that explained more of the variance (Figure S3 and Table S3). For the general category of any migraine, we found that the variance explained by the PRS ($P < 0.1$ threshold) in the familial cases was 3.5%. This finding was compared to only 1.6% variance explained by the same PRS in the FINRISK population cases. Additionally, for the migraine subtypes, the lowest variance explained by the PRS was 2.9% for migraine without aura (MO), 5.5% for migraine with typical aura, and 8.2% for hemiplegic migraine (HM).

Increased risk between upper and lower PRS quartiles

To quantify the effect for individuals carrying the highest burden of risk alleles relative to the FINRISK population distribution, we separated individuals from the family collection into population-level quartiles of the PRS. We used the FINRISK population sample to calculate the cut-off values for individuals in the upper and lower quartiles of PRS and tested for enrichment between individuals in the highest and lowest quartiles of the distribution (Table S4). Again, we observed the lowest enrichment in the MO subtype, where the mean PRS was estimated to be 2.2 times significantly higher than the mean PRS of the FINRISK population controls (OR = 2.2, 95% CI = 2.03–2.37, $P = 1.4\times10^{-19}$). The enrichment for migraine with typical aura was even higher, with mean PRS that was 3.0 times larger than the population mean (OR = 3.02, 95% CI = 2.85–3.20, $P = 9.0\times10^{-35}$). As before, in HM we observed the highest enrichment of common variation, with mean PRS that was 3.8 times significantly higher than the population mean (OR = 3.84, 95% CI = 3.52–4.15, $P = 2.5\times10^{-17}$).

Comparing familial cases to population cases identified from the national health-registry

We observed that common variant burden measured via the PRS was higher in familial cases of any migraine subtype (n = 5,317) compared to population-based cases (n = 1,101) identified from FINRISK via health-registry data (OR = 1.26, 95% CI = 1.18–1.34, $P = 3.2\times10^{-8}$, Table S5). Splitting this result by subtype, we found that the PRS in familial MO was modestly enriched compared to the population-based cases (OR = 1.13, 95% CI = 1.04–1.22, $P = 0.0075$), whereas migraine with typical aura and HM both showed higher

enrichment (OR = 1.31, 95% CI = 1.22–1.40, $P = 4.6\times10^{-9}$, and OR = 1.51, 95% CI = 1.37–1.64, $P = 1.2\times10^{-9}$ respectively).

We confirmed that the FINRISK sample was representative of other population-based migraine samples by also estimating enrichment of the PRS in population cases from four other migraine case-control studies that were included in the original GWAS of migraine (Gormley et al., 2016), including the Young Finns (OR = 1.21, 95% CI = 1.08–1.36, P = 8.7×10^{-4}), Swedish Twins (OR = 1.24, 95% CI = 1.17–1.32, P = 1.8×10^{-11}), Northern Finland Birth Cohort (OR = 1.20, 95% CI = 1.10–1.29, P = 1.1×10^{-5}), and Health 2000 (OR = 1.40, 95% CI = 1.18–1.29, P = 1.6×10^{-4}). These effect-sizes were in line with the enrichment found for the FINRISK population cases when comparing the PRS between cases and controls (OR = 1.32, 95% CI = 1.25–1.38, P = 7.2×10^{-17}) and all population-based studies showed lower enrichment for the PRS than in the familial migraine cases (Figure S4).

To ensure that the increased enrichment of the PRS in familial cases compared to population cases was not due to some systematic bias between the FINRISK and Finnish Migraine Families sample, we also calculated two additional PRS scores, this time based on single-nucleotide polymorphism (SNP) weights taken from GWAS of Intelligence Quotient (IQ) (Sniekers et al., 2017) and Schizophrenia (Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014). As expected, we found no difference in enrichment of these two PRS scores when comparing migraine cases and controls from either the FINRISK population sample or the Finnish Migraine Family sample (Figure S5).

We additionally found that migraine individuals within the family collection who had self-reported use of triptan medication had on average higher common variant burden compared with individuals who did not report use of triptans (OR = 1.12, 95% CI =1.06–1.19, $P = 5.7\times10^{-4}$), Table S6). Among the FINRISK population cases, we did not observe any difference in PRS when comparing cases that were either triptan users (defined as purchasing triptans at least twice) or individuals who had visited a migraine outpatient clinic (OR = 1.06, 95% CI = 0.86–1.26, $P = 0.56$). However, individuals that had both visited a specialist outpatient clinic and used triptans had significantly higher PRS enrichment (n = 131, OR = 1.70, 95% CI = 1.53–1.88, P = 3.9×10^{-9}). In fact, this was similar enrichment to the PRS profile observed in the familial migraine cases (n = 5,317, OR = 1.76, 95% CI = 1.71–1.81, P = 1.7×10^{-109}, Figure S6).

Migraine-associated alleles are over-transmitted to affected offspring

As an additional method for investigating the contribution of polygenic load that is robust to relatedness in a family sample, we used the polygenic transmission disequilibrium test (pTDT, **STAR★METHODS**). Here we extracted nuclear trios from the family collection for four phenotypes (any migraine, MO, migraine with typical aura, and HM) and tested separately whether common migraine-associated alleles (as measured by the PRS) were disproportionately over-transmitted from parents to both affected and unaffected offspring. In support of the results from the mixed model, we found that affected offspring received a significantly higher transmission of common polygenic load from their parents than would be expected by chance alone (Figure 3). The over-transmission was observed in every

migraine phenotype; any migraine, (n = 1,486 trios, P = 1.7×10^{-13}, pTDT deviation = 0.16, 95% CI = 0.12 – 0.20), MO (n = 727 trios, P = 4.9×10^{-4}, pTDT deviation = 0.10, 95%CI = 0.046 – 0.163), migraine with typical aura (n = 571 trios, P = 1.5×10^{-6}, pTDT deviation = 0.17, 95%CI = 0.10 – 0.24), and HM (n = 188 trios, P = 3.7×10^{-7}, pTDT deviation = 0.33, 95%CI = 0.21 – 0.45), see Table S7. As expected, no over-transmission was observed for unaffected offspring (n = 734 trios, P>0.05). While the observed over-transmission of the PRS was higher for migraine with typical aura and HM compared to MO (consistent with the association results from the mixed-model above), the difference was not significant between migraine groups.

Contribution of Mendelian variants and polygenic load to FHM

We examined the relative contribution of known pathogenic variants (**STAR★METHODS**) and polygenic load to FHM, using a subset of 74 families where familial aggregation of cases had been confirmed. From sequencing data on 101 FHM cases from 45 of these families, we have identified four families that carried a pathogenic, rare mutation in one of the three known FHM genes (Hiekkala et al., 2018; Kaunisto et al., 2004). Therefore out of the 45 sequenced families, 8.9% (4/45) could be potentially explained by a rare pathogenic mutation in one of the known FHM genes (Table 2 and Table S8). We have found no likely pathogenic mutations in these genes for any of the 201 (of 343) SHM individuals that were sequenced (Hiekkala et al., 2018). We next investigated what proportion of cases from each migraine subtype was found in the extreme tails of the distribution of polygenic risk from the FINRISK population. The expected proportion of individuals in the upper quartile of risk is 25%, which was approximately what we observed for individuals with no migraine (26.7%), but found large deviations from expectation for the other phenotypes, including MO (36.2%), migraine with typical aura (41.4%), and HM (43.0%). Furthermore, of the 197 FHM cases from 74 families, 80 of these cases (40.6%) were in the highest quartile of polygenic risk (Table 2 and Figure S7).

To assess the polygenic burden per family, we inspected the distribution of the median PRS from each of the 74 families with confirmed FHM cases (Figure S8). We found that over 44% of FHM families (33 out of 74) carried a common variant burden that was in the highest quartile of population risk. Additionally, only five of these 74 FHM families (6.8%) were found to be in the lowest quartile of population risk. Interestingly, the four apparently Mendelian families that carry a pathogenic rare mutation were not among the five families in the lowest quartile of risk and were instead spread across the distribution (in fact two families were in the highest PRS quartile, Figure S8), potentially indicating that other genetic and/or environmental factors also play a role in these families.

Relationship between age of onset and polygenic load

For a subset of individuals (n = 4,930) in the family collection we had information on age of onset of the migraine headache. We used this information to assess whether polygenic load was associated with age of onset. We grouped these 4,930 individuals with onset data into age of onset bins (0 to 10 years old [n = 1,295], over 10 to 15 years old [n = 1,402], over 15 to 20 years old [n = 990], and over 20 years old [n = 1,243]) and estimated whether the mean PRS in each bin was significantly different (Figure 4). Additionally, data from all cases

(n=5,317) was available on whether onset of headaches occurred before or after 20 years old, so we compared the PRS between these two groups using a logistic mixed model. We found that the mean PRS was significantly higher in migraine cases where headache onset occurred before 20 years of age compared to individuals with later onset (OR = 1.11, 95% CI = 1.05–1.18, $P = 8.2\times10^{-4}$, Table S6).

Higher PRS is associated with higher rate of clinical diagnostic symptoms

We used nine diagnostic criteria for migraine (attack length > 4 hours, unilaterality, pulsation, moderate/severe intensity, aggravation by physical exercise, nausea, vomiting, phonophobia, and photophobia, Figure S9) to test if the PRS was specifically associated with any of these individual criteria. As these data were obtained from some family members by questionnaire, only strict yes answers were interpreted as cases, whereas missing answers were interpreted as no-answers. We found that increased PRS predicted a higher rate of 8 out of 9 diagnostic symptoms (Table S9), with severity of headache showing the largest effect size (OR = 1.29, 95% CI = 1.20–1.39, $P = 1.1\times10^{-7}$). Also strongly associated were ‘attack length greater than 4 hours’ (OR = 1.23, 95% CI = 1.16–1.29, $P = 4.7\times10^{-9}$) and photophobia (OR = 1.23, 95% CI = 1.14–1.32, $P = 5.8\times10^{-6}$). Notably, the only diagnostic criteria not associated with the polygenic risk score was headache pulsation (OR = 1.04, 95% CI = 0.98–1.10, $P = 0.17$).

DISCUSSION

Our results show that common polygenic variation, as measured via the PRS, significantly contributes to the familial aggregation of migraine. PRS enrichment in families was observed in both common and rare subtypes of familial migraine compared to both population controls and to population cases of migraine. There were relatively large differences in the PRS burden observed between different migraine sub-categories. The polygenic burden was higher for MA compared to those individuals that do not suffer any migraine aura symptoms. Our results confirm that common variants identified by GWAS in populations play a considerable role in rare forms of migraine with aura (both FHM and SHM) and suggest that a large proportion of the disease risk in HM cases can be significantly explained by common polygenic variation, rather than solely by highly penetrant, rare variation.

In interpreting these results, we have also considered the possibility (particularly related to the familial cases) that the common variation encapsulated by the PRS could merely be tagging rare variants of large effect that are segregating with cases in the families. However, this is highly unlikely, as these so-called “synthetic associations” by rare variants have been shown to not explain most of the loci found by GWAS (Wray et al., 2011). Therefore we would expect that vast majority of genomic associations captured by the PRS are tagging only causal common variation.

We showed that the proportion of variance explained by the PRS was higher in the familial cases (3.5%) compared with only 1.6% in the population-based cases. These findings are in line with emerging evidence from other complex traits, where familial forms of disease that were once thought to be mostly explained by rare variants have been found to also have a

strong contribution from common variation. Recent examples include familial forms of dyslipidemia (Ripatti et al., 2016) and Alzheimer's disease (Tosto et al., 2017), and the observation for type 2 diabetes that rare variants of large effect explain only a small amount of disease heritability (Jun et al., 2018). Further support for the important contribution of common variation to disease risk comes from recent work on coronary artery disease (CAD) where it was shown that for some individuals, their common variant load contributes equivalent risk to that of a rare, monogenic variant of large effect (Khera et al., 2017b). It may be noted, however, that the proportion of variance explained so far by PRSs for migraine and other complex diseases is still small (typically less than 10%). For example, in Schizophrenia, a similar approach explains 3.4% of the variance on the liability scale using only genome-wide significant variants, which then rises to 7.0% of the variance when including more variants (PRS P-value threshold < 0.05) (Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014). Similarly, in a recent GWAS of amyotrophic lateral sclerosis (ALS), genome-wide significant variants captured 0.2% of the SNP-heritability, rising to 8.3% when including all common variants (van Rheenen et al., 2016). Even for traits such as inflammatory bowel disease (IBD) where GWAS have successfully been able to identify over 200 significant loci, the total variance explained to date by all common variation remains low for IBD and its subtypes; 13.1% for Crohn's disease and 8.2% for ulcerative colitis (Liu et al., 2015). However, the predictive power of PRS will improve with increasing GWAS sample sizes as the effect sizes (the weights used in the PRS) of common variants can be more reliably estimated. For migraine, based on the most recent GWAS sample size, the total SNP-heritability (i.e. from common, additive variation) was estimated previously by LD-Score regression to be 14.6% (Gormley et al., 2016). This sets an upper limit on the maximum variance we can currently explain with common variation; therefore with the PRS results that we present here we are capturing around one quarter of the possible phenotypic variance. Future studies using a PRS based on weights from a larger GWAS sample should be able to capture even more of the phenotypic variance.

In addition to showing that familial cases of migraine have higher polygenic burden on average compared to population controls, we also showed, using the pTDT approach, that offspring with migraine have inherited a higher burden of common polygenic variation associated with migraine than would be expected by chance alone. Together, these two methods produce results that are robust to genetic relatedness of individuals within the sample. Therefore, we have validated by two independent statistical methods (mixed-model and pTDT) our result that common polygenic variation associated with migraine significantly contributes to the familial aggregation of both prevalent and rare subtypes of migraine.

The National Health Register system provides data from every hospital- and out-patient visit and every prescription drug purchase of every citizen. This data was used to sub-categorize migraine cases in the population-based FINRISK sample. In addition to the formal ICD-code based migraine subtype definition, the registries enabled us to sub-categorize patients based on their use of the health care system. While a large fraction of migraine patients in the Finnish health care system are treated in primary care, the more complicated patients tend to be referred to secondary and tertiary treatment units, like hospital neurology outpatient clinics. Interestingly, migraine cases that had visited a specialist outpatient clinic

and additionally had purchased triptans had the highest PRS, which was as high as the mean value of familial migraine cases ascertained from the clinics. This finding is consistent with recent observations in other traits, for example hyperlipidemias, where familial dyslipidemic cases were observed to have similarly high PRS compared to hyperlipidemia cases in the FINRISK population cohort (Ripatti et al., 2016). These findings suggest that severe cases identified from a population cohort, in terms of their polygenic profile, can be genetically similar to familial cases, and that familial aggregation might just be a reflection of a cumulative effect of many common variants.

Furthermore, we attempted to characterize the proportion of FHM cases that could be explained by rare pathogenic variants in the three known FHM genes. We identified only four out of 45 sequenced FHM families (8.9%) with cases that carried one of these variants, and 0 out of 201 sequenced SHM individuals (Hiekkala et al., 2018). While it is possible that more pathogenic, rare variants for FHM are yet to be discovered, it is striking that so few known variants could be identified in our large family collection (n = 302 sequenced HM cases from 1,589 families). Together, with the observation of significant polygenic burden also seen in individuals with FHM and SHM (including over 40% of FHM cases from 74 families that were in the highest quartile of population polygenic risk), it is likely that a large proportion of the risk for these rare migraine phenotypes are explained by a higher burden of common polygenic variation, with HM falling on the high end of a spectrum of disease liability, possibly in some instances combined with rare variants of larger effect.

In the more prevalent forms of MA, we found that there was a significant difference in common polygenic burden between the migraine with typical aura subtypes - typical aura with headache (ICHD-3 code: 1.2.1.1) and typical aura without headache (ICHD-3 code: 1.2.1.2). While the typical aura with headache group showed a similarly high polygenic burden compared with the more rare forms of MA (i.e. both FHM and SHM), and was not significantly different from these groups, we observed that the typical aura without headache group looked very different and carried a substantially lower common polygenic burden relative to the other MA subtypes. The contribution of the PRS (in terms of effect size) to this migraine phenotype was in fact no different than controls from the population (Figure S2). As such, one might speculate that much of the common variation captured by the PRS is influencing genes involved in the etiology of the pain characteristics of migraine rather than the aura features, but more investigations are needed to determine this.

Interestingly, in the family sample, migraine cases who experienced earlier age of onset of headaches tended to carry a higher polygenic burden of common migraine risk alleles on average. This association was observed for all migraine types and is consistent with similar findings in other complex disorders (Tosto et al., 2017), as well as previous hypotheses that suggest that migraines that have earlier onset (before 15 – 20 years old) have a higher genetic burden compared to migraines that begin later in life, where some combination of genetic plus environmental factors (e.g. stress, diet, medications, general health, and other stimulants) may play a larger role.

Finally, when looking at nine diagnostic symptoms of migraine, we found that higher polygenic risk scores were associated with an increased risk for eight out of nine symptoms. The only symptom not associated with higher polygenic load was headache pulsation. While it is not unusual for a polygenic risk score based on migraine-associated variation to be associated with the diagnostic symptoms of migraine, the consistency of the associations and direction of effect across these symptoms suggest that migraine severity is positively correlated with higher burden of common polygenic variation. In addition to these diagnostic symptoms, we also found that migraine cases that had received treatment with triptan medication, both in the family sample and the population cohort, were associated with higher polygenic load. This again points to a higher polygenic load for individuals that are more likely to suffer from more disabling migraines since they have sought out specialist treatment.

In conclusion, our study supports the hypothesis that migraine subtypes are genetically heterogeneous diseases, and that regardless of whether they are common (i.e. MO and migraine with typical aura) or rare subtypes (i.e. FHM and SHM), common polygenic variation significantly contributes to the aggregation of the disease in families.

STAR★METHODS

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Aarno Palotie (aarno.palotie@helsinki.fi).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Study cohorts

Finnish Migraine Families collection: The families were collected over a period of 25 years from six headache clinics in Finland (Helsinki, Turku, Jyväskylä, Tampere, Kemi, and Kuopio) and through advertisements on the national migraine patient organization web page (www.migreeni.org). Geographically, family members are represented from across the entire country. The current collection consists of 1,589 families, which included a complete range of pedigree sizes from small to large (e.g. 1,023 families had 1–4 related individuals and 566 families had 5+ related individuals, see Table S1 and Figure S1). It should be noted here that 455 individuals in the sample were single probands (i.e. unrelated cases without available affected relatives for analysis) but since they were ascertained in the same way as the other migraine families we have included them. Currently, the collection consists of 8,319 family members, of whom 5,317 have a migraine diagnosis based on the third edition of the established International Classification for Headache Disorders (ICHD-3) criteria (Headache Classification Committee of the International Headache Society (IHS), 2013). In about 50% of these affected individuals, the migraine attack is preceded by an aura phase. Another 3,002 family members were classified as having no migraines, including 1,557 individuals with no headache, 427 individuals with headache, 755 individuals with probable migraine, and 263 individuals with unknown diagnosis.

Migraine phenotype data was collected with a combination of individual interviews and an extensively validated Finnish Migraine Specific Questionnaire for Family Studies (FMSQFS (Kallela et al., 2001)). All participants were also asked to donate a blood sample. Over 200 variables were recorded, including information on the ICHD-3 symptoms, typical attack features, age of onset, other diseases, place of birth, etc. For the index patient in each family a neurologist performed a physical examination and sometimes other family members were examined as well. In all cases where the diagnosis was not clear from the questionnaire, a neurologist specialized in headache disorders interviewed the study subject. A summary of the sample characteristics of this collection is shown in Table 1 and Table S10.

FINRISK population-based cohort: FINRISK is a series of population-based health examination surveys carried out every five years since 1972 to monitor the risk of chronic diseases in Finland, as detailed elsewhere (Borodulin et al., 2015). Individuals in these cohorts have been prospectively followed for cardiovascular events and cause-specific death until 31st December 2015 using annual record linkage with the Finnish National Hospital Discharge Register and the National Causes-of-Death Register. A total of 14,470 subjects from FINRISK were genotyped in five batches (Lim et al., 2014), having been randomly sampled from the full cohort, stratified by sex and cohort year (i.e. FINRISK 1992i.e. FINRISK 1997i.e. FINRISK 2002 or 2007 cohorts). A summary of the FINRISK sample characteristics is shown in Table 1.

Finnish National Health Registry data for population-based cases: The population-based migraine cases within the FINRISK cohort were identified using Finnish National Health Registry data by two means: 1) From a specialist outpatient registry (from 1998 onwards) if an individual had received a migraine diagnosis (ICD-10 code: G43 or ICD-9 code: 346) either during a hospital visit (hospital discharge registry) or a specialty outpatient clinic visit (outpatient discharge registry) or 2) From a prescription drug purchase registry (from 1995 onwards) if an individual had been prescribed triptans at least twice (ATC codes under the N02CC category). This approach is likely to underreport migraine cases, particularly those with a sufficiently mild form of the disease so as to require neither triptan use nor visits to a hospital outpatient specialty clinic. Additionally, this data does not have symptom level information, so ICHD-3 based classification into migraine subtypes was not possible. Altogether 1,101 individuals fulfilling one of the above criteria were identified among a total of 14,470 study participants, giving a frequency of 7.6% of migraine cases.

Population-based migraine studies used for replication: We used four population-based genome-wide association studies for replication of our PRS enrichment in the FINRISK population-based sample, including Health 2000 (Heistaro, 2008), Young Finns (Raitakari et al., 2008), Northern Finland Birth Cohort (Sovio et al., 2007), and the Swedish Twins (Ran et al., 2014), see Figure S4. The sample collections have been described in more detail in the original publications.

METHOD DETAILS

Genotyping and Quality Control—Genotyping was performed in seven batches on either the Illumina® CoreExome or Illumina® PsychArray, which share the Infinium®

HumanCore backbone including 480,000 variants in common. Samples from the migraine family collection were genotyped in two batches, one on the CoreExome and one on the PsychArray, with cases and controls distributed across both batches. Individual samples from the FINRISK cohort were genotyped in five batches, all on the CoreExome. A summary of genotyping batches is provided in Table S11.

Before merging any batches we performed standard quality control procedures on each dataset individually, according to established GWAS protocols (Anderson et al., 2010). Briefly, we excluded markers that exhibited high 'missingness' rates (> 5%), low minor allele frequency (< 1%), or failed a test of Hardy–Weinberg equilibrium ($P < 10^{-6}$). We also excluded individuals with high rates of heterozygosity (> 3 standard deviations from the mean), or a high proportion of missing genotypes (> 5%). To control for any possible population stratification, we merged the genotypes from individuals passing QC with HapMap III data from European (CEU), Asian (CHB+JPT), and African (YRI) populations. We then performed a principal-components analysis on this combined data and excluded any population outliers not clustering with the other Finnish samples. We also performed a second principal-components analysis within each batch to ensure that cases and/or controls were clustering evenly together.

We then merged genotyping batches one-by-one and repeated the QC procedures described above on the merged dataset. To prevent any potential batch effects in the merged data, we also excluded any markers that failed a test of differential missingness ($P < 10^{-5}$) between the merged batches. Furthermore, during each round of merging, we performed a pseudo-association analysis (using a logistic mixed-model for batches with related individuals) between samples from each batch to identify markers where the minor allele frequency deviated significantly between batches ($P < 10^{-5}$). Markers with significant deviation were subsequently removed.

Finally, for the FINRISK samples we additionally used identity-by-descent (IBD) estimates to remove any closely related individuals (proportion IBD > 0.185), as the goal was to use them as a set of independent population controls. We further calculated kinship coefficients between all individuals using the software KING (Manichaikul et al., 2010) in order to estimate genetic relatedness and to correct or remove individuals causing clear pedigree errors in the family sample.

Reference panel for genotype imputation—To impute missing genotypes into the merged dataset (migraine families and FINRISK) we created a Finnish population-specific reference panel derived from sequencing data generated as part of the Sequencing Initiative Suomi (SISu) project (Chheda et al., 2017; Surakka et al., 2016). The reference panel combined low-coverage (mean depth 4.6x) WGS data and high coverage WES data described further below.

Finnish low-coverage (4.6x) WGS reference dataset: Sample- and variant-level quality control for the data was done at the Wellcome Trust Sanger Institute. Only 1,940 high quality unrelated individuals and polymorphic autosomal PASS SNP variants were included in the reference panel. Additionally, SNPs in low-complexity regions and those with Hardy-

Weinberg equilibrium (HWE) P-value < 10^{-5} (n=99,191) were removed leaving a total of 13,625,209 markers after quality control. The data was then phased using SHAPEIT2(Delaneau et al., 2011) with default options and effective population size of 11,418.

Finnish WES reference dataset: From the raw WES data, variants were filtered according to the following criteria: 1) Multi-allelic variants were removed, 2) Genotypes with QC < 20 were set to missing, 3) SNPs with call rate < 95% were removed, and 4) Monomorphic markers were removed. In addition, samples that were also in the WGS panel (n=7) were excluded together with individuals whose genotyping rate was < 95% (n=43). After filtering steps, the reference panel contained 1,540 individuals and 3,008,675 markers. The data was phased using SHAPEIT2(Delaneau et al., 2011) using default options and effective population size of 11,418.

Finally, the two reference panels (WGS and WES) were combined during imputation of the FINRISK and migraine family data using the software IMPUTE2 (Howie et al., 2012) and its option to merge reference panels (i.e. '-merge_ref_panels' option). We treated all available haplotypes from the two reference panels as informative (i.e. set total number of haplotypes as 6,962 with parameters: '-k_hap 3882 3080').

Imputation—Following genotyping QC, phased haplotypes were estimated for each individual using the program SHAPEIT2 (Delaneau et al., 2011) and its duoHMM method to improve accuracy by refining the estimation to haplotypes that are consistent with the pedigree structure. For phasing we chose an effective population size of 11,418, a window size of three, and 200 states for fitting the model. Missing genotypes were then imputed into these haplotypes using the program IMPUTE2 (Howie et al., 2012) and a manually created Finnish reference panel described above. We chose an effective population size of 20,000. While all samples (migraine family cases, controls, and FINRISK population) were imputed together, we split chromosomes into chunks of 3Mb with 500kb buffer to reduce the computation.

QUANTIFICATION AND STATISTICAL ANALYSIS

Calculation of polygenic risk scores (PRS)—To calculate the migraine PRS, we used the SNP effect sizes estimated for common variants from a previously published GWAS of migraine in 375,000 individuals (Gormley et al., 2016). To ensure there were no overlapping samples from our family collection or FINRISK cohort, we excluded all samples of Finnish descent (i.e. four cohorts; the Finnish MA, Health 2000, NFBC, and Young Finns) from the original 22 cohort GWAS (Table S12). We then recalculated the SNP effect size estimates for migraine from the remaining 18 studies that were of other European origin (i.e. 57,471 cases and 305,141 controls) using a fixed-effects meta-analysis. We then took the intersection of variants from the migraine GWAS dataset that overlapped with the imputed variants from our combined dataset of the migraine family collection and the FINRISK population. Next, we reduced the list of intersecting variants to an independent set by performing LD-clumping (r^2 < 0.1 within 500kb from the most significant variant in each locus) using PLINK (Purcell et al., 2007). Finally, we chose a subset of SNPs (n = 38,872)

from the list of independent variants with *P*-values below a threshold of 0.1 in order to capture most of the variation influencing migraine risk while excluding the remainder of variants that do not show even a modest association. We then calculated the PRS for each individual as a sum of these alleles, weighted by the effect size estimates from the migraine GWAS results. Our script for calculating the PRS for any trait with GWAS data has been made publicly available on the website Github (see **Software**).

Mixed-model association analyses—To account for the high degree of relatedness within our family sample, we used logistic mixed-models to adjust for the genetic relatedness matrix (GRM) as a random effect. We calculated the GRM after filtering to a set of independent LD-pruned common SNPs (minor allele frequency > 5% and SNP missingness < 3%) using the program PLINK (Purcell et al., 2007) (parameter options: '--maf 0.05 --geno 0.03 --make-rel square gz'). In addition to adjusting for the GRM as a random effect, we also adjusted for sex, age, age^2, and age^3 as fixed effects. We then tested if the PRS was associated with migraine phenotypes using a Wald test of one degree of freedom. All mixed models and Wald tests were implemented in the statistical software R using the GMMAT package (Chen et al., 2016). We adjusted for multiple testing using Bonferroni correction.

Estimation of variance explained—To estimate the variance explained by the PRS, we fitted a logistic mixed-model as described above, adjusted for relatedness using the GRM (random effect variable) and additionally adjusted for sex, age, age^2, and age^3 (fixed effect variables). We then compared the full model (including the PRS) with the null model (with PRS variable excluded) and estimated the variance explained using Nagelkerke's pseudo-R^2. We calculated the variance explained by the PRS across a range of GWAS *P*-value thresholds (Figure S3 and Table S3) and determined that while the optimal threshold for the most variance explained was at $P < 0.5$, the findings were not qualitatively different than at our chosen *P*-value threshold of $P < 0.1$.

Polygenic Transmission Disequilibrium Test (pTDT)—To assess polygenic burden of migraine risk alleles over-transmission from parents to affected offspring, we used the Polygenic Transmission Disequilibrium Test (pTDT) method (Weiner et al., 2017). The method is robust to the relatedness structure as it uses full trios within a family sample and calculates an expected distribution of PRS for offspring based on the average PRS of the parents. This expected distribution for the offspring is then used to test deviations from the null hypothesis in the observed mean parent PRS distribution. To calculate the expected distribution, we first separated the families into nuclear trios for any migraine (n = 1,486 trios) and then further performed subset analyses for MO (ICHD-3 code: 1.1, n = 737 trios), migraine with typical aura (ICHD-3 code: 1.2.1, n = 571 trios), and HM (ICHD-3 code: 1.2.3, n = 188 trios). We tested this hypothesis separately for offspring that were both cases and controls using a two-tailed, one-sample *t*-test.

Identification of known pathogenic variants—To identify any families/individuals that carry a known pathogenic variant, which we define as a rare, mutation contributing mechanistically to the disease (similar to guidelines provided elsewhere (MacArthur et al.,

2014)), we have screened 302 (101 FHM and 201 SHM) HM patients for these mutations in the three known genes for FHM (*CACNA1A*, *ATP1A2*, and *SCN1A*). To extract a list of possibly pathogenic variants in these genes, the selected 302 HM patients were either WES (293 HM cases from 243 families) (Hiekkala et al., 2018) or Sanger-sequenced (9 FHM cases from one family) (Hiekkala et al., 2018; Kaunisto et al., 2004). We then filtered out variants that were common in gnomAD (MAF > 1% in either all populations combined, or in the Finnish population alone MAF > 0.1%), that were predicted to be benign, or that did not segregate with cases in the family collection - as described in the original publication (Hiekkala et al., 2018).

DATA AND SOFTWARE AVAILABILITY

A list of the software tools, datasets and sample resources used in this work are given below. Access to the Finnish data can be arranged through application to the National Institute for Health and Welfare (THL) Biobank (https://thl.fi/en/web/thl-biobank/for-researchers). The migraine GWAS data used in calculating the PRS weights can be obtained via application to the International Headache Genetics Consortium (IHGC, contact: Risto.Kajanne@helsinki.fi). The script we used for calculating the PRS in each cohort has been made available on the Github repository (see link below).

Data resources and databases—Finnish Migraine Families Study: http://www.nationalbiobanks.fi/index.php/studies2/20-migraine-family-study

Finnish National Health Registry: https://thl.fi/fi/web/thlfi-en

FINRISK Study: http://www.nationalbiobanks.fi/index.php/studies2/7-finrisk

gnomAD browser: http://gnomad.broadinstitute.org/

HapMap data: http://www.sanger.ac.uk/resources/downloads/human/hapmap3.html

Health 2000 Survey: http://www.nationalbiobanks.fi/index.php/studies2/8-health2000

IHGC migraine data: http://www.headachegenetics.org/content/datasets-and-cohorts

Northern Finland Birth Cohort: http://www.nationalbiobanks.fi/index.php/studies2/11-nfbc

Sequencing Initiative Suomi: http://www.sisuproject.fi/

Swedish Twin Registry: https://ki.se/en/research/the-swedish-twin-registry

Young Finns Study: http://www.nationalbiobanks.fi/index.php/studies2/23-yfs

Software—GMMAT: https://content.sph.harvard.edu/xlin/software.html#gmmat

IMPUTE2: https://mathgen.stats.ox.ac.uk/impute/impute_v2.html

KING: http://people.virginia.edu/~wc9c/KING/

PLINK: https://www.cog-genomics.org/plink2

PRS script: https://github.com/pgormley/polygenic-risk-scores

pTDT: https://github.com/ypaialex/ptdt

R: https://www.r-project.org/

SHAPEIT2: https://mathgen.stats.ox.ac.uk/genetics_software/shapeit/shapeit.html

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Critical Commercial Assays		
Infinium CoreExome-24 Kit	Illumina	Cat# WG-331-1101
Infinium PsychArray-24 Kit	Illumina	Cat# WG-356-1101
Software and Algorithms		
GMMAT	(Chen et al., 2016)	https://content.sph.harvard.edu/xlin/software.html#gmmat
IMPUTE2	(Howie et al., 2012)	https://mathgen.stats.ox.ac.uk/impute/impute_v2.html
KING	(Manichaikul et al., 2010)	http://people.virginia.edu/~wc9c/KING/
PLINK	(Purcell et al., 2007)	https://www.cog-genomics.org/plink2
PRS script	Current paper	https://github.com/pgormley/polygenic-risk-scores
PTDT	(Weiner et al., 2017)	https://github.com/ypaialex/ptdt
R	R Core Team	https://www.r-project.org/
SHAPEIT2	(Delaneau et al., 2011)	https://mathgen.stats.ox.ac.uk/genetics_software/shapeit/shapeit.html
Other		
Migraine GWAS data	(Gormley et al., 2016)	https://www.ncbi.nlm.nih.gov/pubmed/27322543

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by the Wellcome Trust (grant numbers WT089062, 098051 to A.P.); the Academy of Finland (grant numbers 286500, 293404 to A.P., and 139795 to M.W.); the Academy of Finland Center of Excellence for Complex Disease Genetics (grant number 312074); the EuroHead project (LSM-CT-2004-504837 to A.P.); FP7-EUROHEADPAIN (grant number 602633 to A.P.); ENGAGE Consortium (grant agreement HEALTH-F4-2007-201413 to A.P.); EU/SYNSYS-Synaptic Systems (grant number 242167 to A.P.); the Sigrid Juselius Foundation, Finland (to A.P.); the Nordic Information for Action eScience Center, NIASC (Grant No 62721 to A.P.); NIH/Genomic strategies to identify high-impact psychiatric risk variants (Grant No 1U01MH105666-01 to A.P.); the Folkhälsan Research Foundation, Finland (to M.W.); Medicinska Understödsföreningen Liv & Hälsa, Finland (to M.W.); the Helsinki University Central Hospital (to M.K., M.F., V.A.); the Swedish Research Foundation (grant number 2017-01096 to A.C.B); the Swedish Brain Foundation (to A.C.B. and C.R.); and Karolinska Institutet Research Funds (to A.C.B.). T.L. was supported by the Academy of Finland: grants 286284, 134309 (Eye), 126925, 121584, 124282, 129378 (Salve), 117787 (Gendi), and 41071 (Skidi); the Social Insurance Institution of Finland; Competitive State Research Financing of the Expert Responsibility area of Kuopio, Tampere and Turku University Hospitals (grant X51001); Juho Vainio Foundation; Paavo Nurmi Foundation; Finnish

Foundation for Cardiovascular Research; Finnish Cultural Foundation; Tampere Tuberculosis Foundation; Emil Aaltonen Foundation; Yrjö Jahnsson Foundation; Signe and Ane Gyllenberg Foundation; Diabetes Research Foundation of Finnish Diabetes Association; and EU Horizon 2020 (grant 755320 for TAXINOMISIS). The work was further supported by funding from Merck and Co., Kenilworth, NJ, USA. We would like to thank the participants of the NFBC1966 study and the NFBC project center. We are sincerely grateful to all members of the Finnish migraine families for participating in this research. We would also like to thank the research participants and employees of 23andMe, Inc., a personal genetics company, for making this work possible.

References

Agarwala V, Flannick J, Sunyaev S, Altshuler D. GoT2D Consortium. Evaluating empirical bounds on complex disease genetic architecture. Nat Genet. 2013; 45:1418–1427. [PubMed: 24141362]

Anderson CA, Pettersson FH, Clarke GM, Cardon LR, Morris AP, Zondervan KT. Data quality control in genetic case-control association studies. Nat Protoc. 2010; 5:1564–1573. [PubMed: 21085122]

Anttila V, Stefansson H, Kallela M, Todt U, Terwindt GM, Calafato MS, Nyholt DR, Dimas AS, Freilinger T, Müller-Myhsok B, et al. Genome-wide association study of migraine implicates a common susceptibility variant on 8q22.1. Nat Genet. 2010; 42:869–873. [PubMed: 20802479]

Anttila V, Winsvold BS, Gormley P, Kurth T, Bettella F, McMahon G, Kallela M, Malik R, de Vries B, Terwindt G, et al. Genome-wide meta-analysis identifies new susceptibility loci for migraine. Nat Genet. 2013; 45:912–917. [PubMed: 23793025]

Borodulin K, Vartiainen E, Peltonen M, Jousilahti P, Juolevi A, Laatikainen T, Männistö S, Salomaa V, Sundvall J, Puska P. Forty-year trends in cardiovascular risk factors in Finland. Eur J Public Health. 2015; 25:539–546. [PubMed: 25422363]

Chasman DI, Schürks M, Anttila V, de Vries B, Schminke U, Launer LJ, Terwindt GM, van den Maagdenberg AMJM, Fendrich K, Völzke H, et al. Genome-wide association study reveals three susceptibility loci for common migraine in the general population. Nat Genet. 2011; 43:695–698. [PubMed: 21666692]

Chasman DI, Schürks M, Kurth T. Population-based approaches to genetics of migraine. Cephalalgia. 2016; 36:692–703. [PubMed: 27013237]

Chen H, Wang C, Conomos MP, Stilp AM, Li Z, Sofer T, Szpiro AA, Chen W, Brehm JM, Celedón JC, et al. Control for population structure and relatedness for binary traits in genetic association studies via logistic mixed models. Am J Hum Genet. 2016; 98:653–666. [PubMed: 27018471]

Chheda H, Palta P, Pirinen M, McCarthy S, Walter K, Koskinen S, Salomaa V, Daly M, Durbin R, Palotie A, et al. Whole-genome view of the consequences of a population bottleneck using 2926 genome sequences from Finland and United Kingdom. Eur J Hum Genet. 2017; 25:477–484. [PubMed: 28145424]

De Fusco M, Marconi R, Silvestri L, Atorino L, Rampoldi L, Morgante L, Ballabio A, Aridon P, Casari G. Haploinsufficiency of ATP1A2 encoding the Na+/K+ pump alpha2 subunit associated with familial hemiplegic migraine type 2. Nat Genet. 2003; 33:192–196. [PubMed: 12539047]

Delaneau O, Marchini J, Zagury JF. A linear complexity phasing method for thousands of genomes. Nat Methods. 2011; 9:179–181. [PubMed: 22138821]

Dichgans M, Freilinger T, Eckstein G, Babini E, Lorenz-Depiereux B, Biskup S, Ferrari MD, Herzog J, van den Maagdenberg AMJM, Pusch M, et al. Mutation in the neuronal voltage-gated sodium channel SCN1A in familial hemiplegic migraine. Lancet. 2005; 366:371–377. [PubMed: 16054936]

Freilinger T, Anttila V, de Vries B, Malik R, Kallela M, Terwindt GM, Pozo-Rosich P, Winsvold B, Nyholt DR, van Oosterhout WPJ, et al. Genome-wide association analysis identifies susceptibility loci for migraine without aura. Nat Genet. 2012; 44:777–782. [PubMed: 22683712]

Global Burden of Disease Study 2013 Collaborators. 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. 2015; 386:743–800. [PubMed: 26063472]

Gormley P, Anttila V, Winsvold BS, Palta P, Esko T, Pers TH, Farh K-H, Cuenca-Leon E, Muona M, Furlotte NA, et al. Meta-analysis of 375,000 individuals identifies 38 susceptibility loci for migraine. Nat Genet. 2016; 48:856–866. [PubMed: 27322543]

Headache Classification Committee of the International Headache Society (IHS). The International Classification of Headache Disorders, 3rd edition (beta version). Cephalalgia. 2013; 33:629–808. [PubMed: 23771276]

Heistaro, S. Health 2000 survey. Finland: Publications of the National Public Health Institute; 2008. Methodology report.

Hiekkala ME, Vuola P, Artto V, Häppölä P, Häppölä E, Vepsäläinen S, Cuenca-León E, Lal D, Gormley P, Hämäläinen E, et al. The contribution of CACNA1A, ATP1A2 and SCN1A mutations in hemiplegic migraine: A clinical and genetic study in Finnish migraine families. Cephalalgia. 2018 333102418761041.

Howie B, Fuchsberger C, Stephens M, Marchini J, Abecasis GR. Fast and accurate genotype imputation in genome-wide association studies through pre-phasing. Nat Genet. 2012; 44:955–959. [PubMed: 22820512]

Jun G, Manning A, Almeida M, Zawistowski M, Wood AR, Teslovich TM, Fuchsberger C, Feng S, Cingolani P, Gaulton KJ, et al. Evaluating the contribution of rare variants to type 2 diabetes and related traits using pedigrees. Proc Natl Acad Sci USA. 2018; 115:379–384. [PubMed: 29279374]

Kallela M, Wessman M, Färkkilä M. Validation of a migraine-specific questionnaire for use in family studies. Eur J Neurol. 2001; 8:61–66. [PubMed: 11509082]

Kaunisto MA, Harno H, Vanmolkot KRJ, Gargus JJ, Sun G, Hämäläinen E, Liukkonen E, Kallela M, van den Maagdenberg AMJM, Frants RR, et al. A novel missense ATP1A2 mutation in a Finnish family with familial hemiplegic migraine type 2. Neurogenetics. 2004; 5:141–146. [PubMed: 15133718]

Khera AV, Won HH, Peloso GM, O'Dushlaine C, Liu D, Stitziel NO, Natarajan P, Nomura A, Emdin CA, Gupta N, et al. Association of rare and common variation in the lipoprotein lipase gene with coronary artery disease. JAMA. 2017a; 317:937–946. [PubMed: 28267856]

Khera AV, Chaffin M, Aragam K, Emdin CA, Klarin D, Haas M, Roselli C, Natarajan P, Kathiresan S. Genome-wide polygenic score to identify a monogenic risk-equivalent for coronary disease. BioRxiv. 2017b

Lim ET, Würtz P, Havulinna AS, Palta P, Tukiainen T, Rehnström K, Esko T, Mägi R, Inouye M, Lappalainen T, et al. Distribution and medical impact of loss-of-function variants in the Finnish founder population. PLoS Genet. 2014; 10:e1004494. [PubMed: 25078778]

Liu JZ, van Sommeren S, Huang H, Ng SC, Alberts R, Takahashi A, Ripke S, Lee JC, Jostins L, Shah T, et al. Association analyses identify 38 susceptibility loci for inflammatory bowel disease and highlight shared genetic risk across populations. Nat Genet. 2015; 47:979–986. [PubMed: 26192919]

MacArthur DG, Manolio TA, Dimmock DP, Rehm HL, Shendure J, Abecasis GR, Adams DR, Altman RB, Antonarakis SE, Ashley EA, et al. Guidelines for investigating causality of sequence variants in human disease. Nature. 2014; 508:469–476. [PubMed: 24759409]

Manichaikul A, Mychaleckyj JC, Rich SS, Daly K, Sale M, Chen WM. Robust relationship inference in genome-wide association studies. Bioinformatics. 2010; 26:2867–2873. [PubMed: 20926424]

Ophoff RA, Terwindt GM, Vergouwe MN, van Eijk R, Oefner PJ, Hoffman SM, Lamerdin JE, Mohrenweiser HW, Bulman DE, Ferrari M, et al. Familial hemiplegic migraine and episodic ataxia type-2 are caused by mutations in the Ca2+ channel gene CACNL1A4. Cell. 1996; 87:543–552. [PubMed: 8898206]

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklar P, de Bakker PIW, Daly MJ, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007; 81:559–575. [PubMed: 17701901]

Raitakari OT, Juonala M, Rönnemaa T, Keltikangas-Järvinen L, Räsänen L, Pietikäinen M, Hutri-Kähönen N, Taittonen L, Jokinen E, Marniemi J, et al. Cohort profile: the cardiovascular risk in Young Finns Study. Int J Epidemiol. 2008; 37:1220–1226. [PubMed: 18263651]

Ran C, Graae L, Magnusson PKE, Pedersen NL, Olson L, Belin AC. A replication study of GWAS findings in migraine identifies association in a Swedish case-control sample. BMC Med Genet. 2014; 15:38. [PubMed: 24674449]

van Rheenen W, Shatunov A, Dekker AM, McLaughlin RL, Diekstra FP, Pulit SL, van der Spek RAA, Võsa U, de Jong S, Robinson MR, et al. Genome-wide association analyses identify new risk

variants and the genetic architecture of amyotrophic lateral sclerosis. Nat Genet. 2016; 48:1043–1048. [PubMed: 27455348]

Ripatti P, Rämö JT, Söderlund S, Surakka I, Matikainen N, Pirinen M, Pajukanta P, Sarin A-P, Service SK, Laurila P-P, et al. The contribution of GWAS loci in familial dyslipidemias. PLoS Genet. 2016; 12:e1006078. [PubMed: 27227539]

Schizophrenia Working Group of the Psychiatric Genomics Consortium. Biological insights from 108 schizophrenia-associated genetic loci. Nature. 2014; 511:421–427. [PubMed: 25056061]

Sniekers S, Stringer S, Watanabe K, Jansen PR, Coleman JRI, Krapohl E, Taskesen E, Hammerschlag AR, Okbay A, Zabaneh D, et al. Genome-wide association meta-analysis of 78,308 individuals identifies new loci and genes influencing human intelligence. Nat Genet. 2017; 49:1107–1112. [PubMed: 28530673]

Sovio U, King V, Miettunen J, Ek E, Laitinen J, Joukamaa M, Veijola J, Järvelin M-R. Cloninger's Temperament dimensions, socio-economic and lifestyle factors and metabolic syndrome markers at age 31 years in the Northern Finland Birth Cohort 1966. J Health Psychol. 2007; 12:371–382. [PubMed: 17284500]

Stewart WF, Staffa J, Lipton RB, Ottman R. Familial risk of migraine: a population-based study. Ann Neurol. 1997; 41:166–172. [PubMed: 9029065]

Surakka I, Sarin A-P, Ruotsalainen SE, Durbin R, Salomaa V, Daly M, Palotie A, Ripatti S. The rate of false polymorphisms introduced when imputing genotypes from global imputation panels. BioRxiv. 2016

Thomsen LL, Kirchmann M, Bjornsson A, Stefansson H, Jensen RM, Fasquel AC, Petursson H, Stefansson M, Frigge ML, Kong A, et al. The genetic spectrum of a population-based sample of familial hemiplegic migraine. Brain. 2007; 130:346–356. [PubMed: 17142831]

Thomsen LL, Oestergaard E, Bjornsson A, Stefansson H, Fasquel AC, Gulcher J, Stefansson K, Olesen J. Screen for CACNA1A and ATP1A2 mutations in sporadic hemiplegic migraine patients. Cephalalgia. 2008; 28:914–921. [PubMed: 18513263]

Tosto G, Bird TD, Tsuang D, Bennett DA, Boeve BF, Cruchaga C, Faber K, Foroud TM, Farlow M, Goate AM, et al. Polygenic risk scores in familial Alzheimer disease. Neurology. 2017; 88:1180–1186. [PubMed: 28213371]

Weiner DJ, Wigdor EM, Ripke S, Walters RK, Kosmicki JA, Grove J, Samocha KE, Goldstein JI, Okbay A, Bybjerg-Grauholm J, et al. Polygenic transmission disequilibrium confirms that common and rare variation act additively to create risk for autism spectrum disorders. Nat Genet. 2017; 49:978–985. [PubMed: 28504703]

Wray NR, Purcell SM, Visscher PM. Synthetic associations created by rare variants do not explain most GWAS results. PLoS Biol. 2011; 9:e1000579. [PubMed: 21267061]

Highlights

- Polygenic risk scores (PRS) implicated in familial aggregation of complex disease
- PRS explains more phenotypic variance in familial cases than in population cases
- Evidence suggests greater role for polygenic load in aggregation than rare variants
- Higher PRS associated with symptoms of migraine severity and earlier age of onset

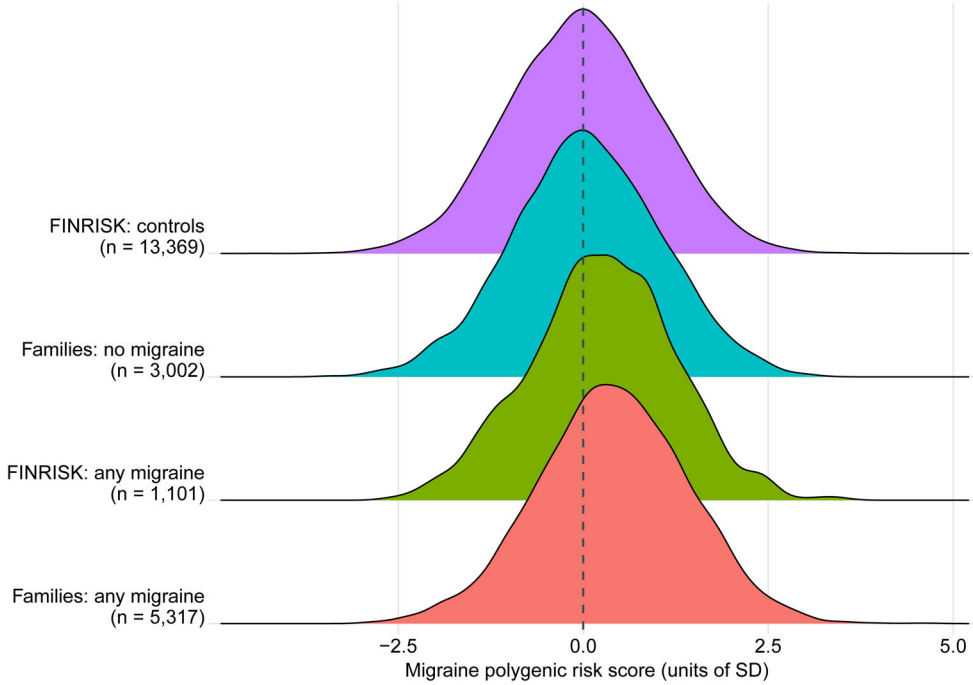

Figure 1. Distributions of the migraine polygenic risk scores (PRS) in the FINRISK population and the Finnish migraine families

For FINRISK, population controls and cases (any migraine subtype) are shown. For the families, family members with no migraine and familial cases (any migraine subtype) are shown. The vertical axis is density of individuals. SD is standard deviation. n is number of individuals.

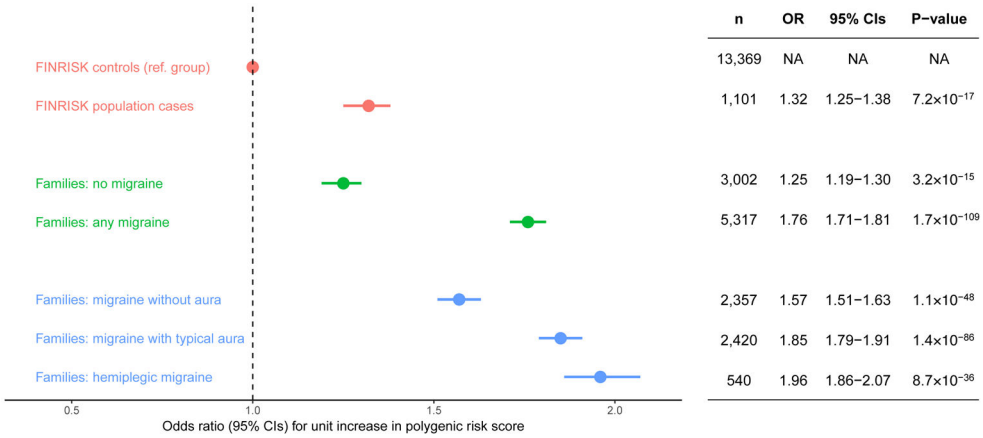

Figure 2. Enrichment of polygenic risk scores (PRS) in familial and population cases

Odds ratios (OR) given are for one standard deviation (SD) increase in PRS compared to 13,369 FINRISK population controls and calculated using a logistic mixed-model adjusted for genetic relatedness, sex, and age. The PRS was calculated using weights from a published migraine genome-wide association study (GWAS, n = 375,000) (Gormley et al., 2016) for an independent set of 38,872 SNPs (GWAS *P*-value threshold < 0.1). FINRISK population cases include any migraine cases identified from Finnish National Health Registry data. ‘Families’ are individuals from the Finnish Migraine Families collection. n is number of individuals. CIs are confidence intervals.

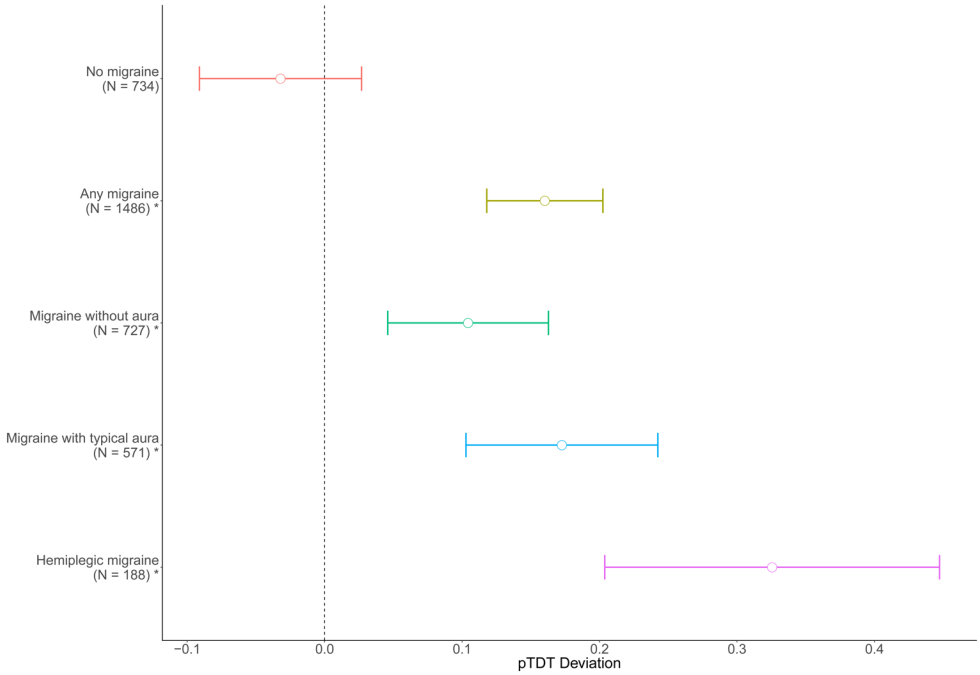

Figure 3. Polygenic transmission disequilibrium test (pTDT) in migraine subtypes

The Finnish migraine families were subset into trios and grouped by disease status of the offspring, making 734 trios for offspring with no migraine and 1,486 trios for offspring with any migraine. Trios were further divided into migraine subtypes, including 727 trios for migraine without aura, 571 trios for migraine with typical aura, and 188 trios for hemiplegic migraine. The horizontal axis shows the pTDT deviation (and 95% confidence intervals) from the mean PRS that would be expected to be transmitted by the parents under the null. N is the number of trios. Groups with significant over-transmission are marked with '*'.

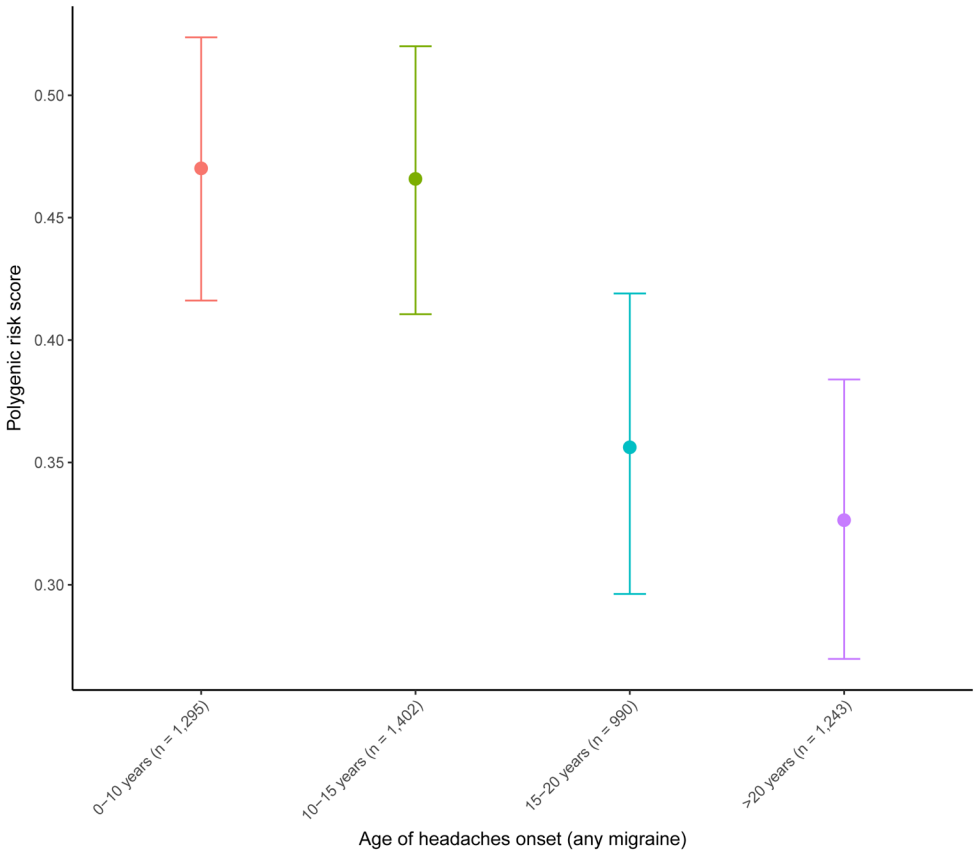

Figure 4. Mean polygenic risk score (PRS) stratified by age of onset of headaches in migraine cases

The data shows that higher PRS corresponds to earlier age of onset of migraine headache. Means and 95% confidence intervals (CIs) were estimated within the Finnish migraine families using bootstrap resampling (10,000 replicates) within each age of onset bin. n is number of individuals.

Table 1

Sample characteristics of the FINRISK population (n = 14,470) and the Finnish migraine families (n = 8,319 individuals from 1,589 families)

Classification codes shown are from the 3rd edition of the International Classification of Headache Disorders (ICHD-3) for each migraine subtype (note - classification criteria was not available for FINRISK). n is number of individuals. PRS is polygenic risk score. SD is standard deviation.

Sample group	ICHD-3 code	n	Mean age (SD)	Sex = Male (%)	Index case (%)	Mean PRS (SD)
FINRISK cohort (all)	-	14,470	46.44 (12.91)	6,356 (43.9)	-	0.02 (1.00)
• Controls	-	13,369	46.96 (12.91)	6,155 (46.0)	-	0.00 (1.00)
• Any migraine	-	1,101	40.09 (11.13)	201 (18.3)	-	0.26 (0.98)
→ triptans only	-	812	39.65 (10.77)	143 (17.6)	-	0.22 (0.98)
→ outpatient only	-	125	45.04 (13.26)	37 (29.6)	-	0.26 (0.98)
→ outpatient + triptans	-	164	38.47 (10.11)	21 (12.8)	-	0.45 (0.93)
Family collection (all)	-	8,319	44.56 (18.24)	2,982 (35.8)	1,568 (18.8)	0.28 (1.02)
• No migraine	-	3,002	44.67 (18.33)	1,687 (56.2)	85 (2.8)	0.06 (1.01)
• Any migraine	1	5,317	44.49 (18.19)	1,295 (24.4)	1,483 (27.9)	0.41 (1.01)
→ migraine without aura (MO)	1.1	2,357	44.85 (17.95)	600 (25.5)	561 (23.8)	0.32 (0.99)
→ migraine with aura (MA)	1.2	2,960	44.21 (18.38)	695 (23.5)	922 (31.1)	0.48 (1.03)
❑ migraine with typical aura	1.2.1	2,420	44.35 (18.52)	612 (25.3)	636 (26.3)	0.47 (1.03)
- typical aura with headache	1.2.1.1	2,350	44.32 (18.50)	575 (24.5)	633 (26.9)	0.47 (1.03)
- typical aura without headache	1.2.1.2	70	45.09 (19.48)	37 (52.9)	3 (4.3)	−0.13 (0.94)
❑ hemiplegic migraine (HM)	1.2.3	540	43.58 (17.72)	83 (15.4)	286 (53.0)	0.54 (0.99)
- familial hemiplegic migraine (FHM)	1.2.3.1	197	44.83 (17.11)	43 (21.8)	53 (26.9)	0.51 (1.00)
- sporadic hemiplegic migraine (SHM)	1.2.3.2	343	42.86 (18.05)	40 (11.7)	233 (67.9)	0.56 (0.98)

Table 2

Summary of family members with a known pathogenic variant or high polygenic risk scores (PRS)

Individuals were assigned to percentiles based on the PRS distribution of the FINRISK population controls. 14 individuals from four families were categorized as having a previously reported pathogenic variant for familial hemiplegic migraine (FHM). n is the number of individuals.

Phenotype	n	Pathogenic FHM variant	Very low PRS (<10th percentile)	Low PRS (<25th percentile)	High PRS (>75th percentile)	Very high PRS (>90th percentile)
• No migraine	3,002	-	272 (9.1%)	701 (23.4%)	803 (26.7%)	333 (11.1%)
• Any migraine	5,317	-	261 (4.9%)	752 (14.1%)	2088 (39.3%)	993 (18.7%)
→ migraine without aura (MO)	2,357	-	129 (5.5%)	374 (15.9%)	854 (36.2%)	373 (15.8%)
→ migraine with aura (MA)	2,960	-	132 (4.5%)	378 (12.8%)	1234 (41.7%)	620 (20.9%)
❑ migraine with typical aura	2,420	-	110 (4.5%)	319 (13.2%)	1002 (41.4%)	502 (20.7%)
- typical aura with headache	2,350	-	102 (4.3%)	301 (12.8%)	989 (42.1%)	497 (21.1%)
- typical aura without headache	70	-	8 (11.4%)	18 (25.7%)	13 (18.6%)	5 (7.1%)
❑ hemiplegic migraine (HM)	540	14 (2.6%) *	22 (4.1%)	59 (10.9%)	232 (43.0%)	118 (21.9%)
- familial hemiplegic migraine (FHM)	197	14 (7.1%) *	7 (3.6%)	23 (11.7%)	80 (40.6%)	45 (22.8%)
- sporadic hemiplegic migraine (SHM)	343	0 (0) *	15 (4.4%)	36 (10.5%)	152 (44.3%)	73 (21.3%)

* Note - the numbers given for individuals carrying a rare pathogenic variant in a known risk gene are incomplete since we have not yet sequenced all 540 HM cases but have instead screened 302 cases to date (101 FHM and 201 SHM).