Epigenetic DNA methylation changes associated with headache chronification: A retrospective case-control study

SCHOLARONE™
Manuscripts

Epigenetic DNA methylation changes associated with headache chronification: A retrospective case-control study

Bendik S. Winsvold[1,2*], Priit Palta[3,4*], Else Eising[5*], Christian M. Page[2,6*], *The International Headache Genetics Consortium*, Arn M. J. M. van den Maagdenberg[5,7], Aarno Palotie[3, 8-12,], John-Anker Zwart[1,2]

1 FORMI and Department of Neurology, Oslo University Hospital, PO Box 4956 Nydalen, 0424 Oslo, Norway.

2 Institute of Clinical Medicine, University of Oslo, 0316, Oslo, Norway

3 Institute for Molecular Medicine Finland (FIMM), University of Helsinki, Tukholmankatu 8, Helsinki, Finland.

4 Estonian Genome Center, University of Tartu, Tartu, Estonia

5 Department of Human Genetics, Leiden University Medical Center, Leiden, The Netherlands

6 Norwegian Institute of Public Health, PO Box 4404 Nydalen, 0403 Oslo, Norway.

7 Department of Neurology, Leiden University Medical Center, Leiden, The Netherlands

8 Psychiatric and Neurodevelopmental Genetics Unit, Massachusetts General Hospital and Harvard Medical School, Boston, USA.

9 Medical and Population Genetics Program, Broad Institute of MIT and Harvard, Cambridge, USA.

10 Stanley Center for Psychiatric Research, Broad Institute of MIT and Harvard, Cambridge, USA.

11 Analytic and Translational Genetics Unit, Massachusetts General Hospital and Harvard Medical School,

12 Department of Neurology, Massachusetts General Hospital, Boston, USA

* The authors contributed equally to this work.

Supplemental Information: Supplemental Methods; Supplemental Figure 1;

Supplemental Table 1.

Correspondence should be addressed to:

Bendik S. Winsvold, Department of Neurology, Oslo University Hospital, Pb 4956

Nydalen, 0424 Oslo, Norway

Tel: +47 97604500

email: bendik.s.winsvold@gmail.com

Abstract

Background: The biological mechanisms of headache chronification are poorly

understood. We aimed to identify changes in DNA methylation associated with the

transformation from episodic to chronic headache.

Methods: Participants were recruited from the population-based Norwegian HUNT-

Study. Thirty-six female headache patients who transformed from episodic to chronic

headache between baseline and follow-up 11 years later were matched against 35

controls with episodic headache. DNA methylation was quantified at 485,000 CpG

sites, and changes in methylation level at these sites were compared between cases

and controls by linear regression analysis. Data were analyzed in two stages (Stages 1

and 2) and in a combined meta-analysis.

Results: None of the top 20 CpG sites identified in Stage 1 replicated in Stage 2 after

multiple testing correction. In the combined meta-analysis the strongest associated

CpG sites were related to *SH2D5* and *NPTX2*, two brain-expressed genes involved in

the regulation of synaptic plasticity. Functional enrichment analysis pointed to

processes including calcium ion binding and estrogen receptor pathways.

Conclusion: In this first genome-wide study of DNA methylation in headache

chronification several potentially implicated loci and processes were identified. The

study exemplifies the use of prospectively collected population cohorts to search for

epigenetic mechanisms of disease.

Key words: chronic headache; epigenetics; cohort studies

Introduction

In a number of patients who suffer from episodic headache attacks, their headache

transforms into a chronic form (1, 2). Chronic headache, defined as headache

occurring on at least 15 days per month, affects 3-5% of the general population and

40% of patients in headache clinics, and is associated with severe disability and

impaired quality of life (1, 3). Approximately half of these patients have chronic

migraine (1). Established risk factors for headache chronification include the overuse

of acute headache medications and a high baseline attack frequency (4). The

biological mechanisms involved in the transition from episodic to chronic headache

are however unknown.

In recent years it has become evident that epigenetic processes play an important role

in a range of multifactorial diseases (5). Epigenetics entail changes in gene expression

that are not due to variation in the DNA sequence. A main epigenetic mechanism is

DNA methylation, the covalent addition of a methyl group to the 5th carbon of

cytosine residues, which is typically associated with gene silencing (5, 6). Epigenetic

modifications can be stably inherited through cell divisions but can also be dynamic.

They enable cell and developmental stage-specific regulation of gene expression, and

play an important role in programming lasting responses to environmental cues (6, 7).

Due to the ability of DNA methylation changes to shift a biological system from one

stable state to another, their implication in headache chronification is of particular

interest. It has been hypothesized that frequent headache attacks may lower the

threshold for subsequent headache attacks through epigenetic mechanisms resulting in

a feed-forward loop (6). Similarly, psychological stress and female sex hormones,

5

which have been implicated in migraine genesis, are known to exert their

physiological effects partly through epigenetic mechanisms (6, 8).

In this study, subjects were recruited from the Norwegian population-based Nord-

Trøndelag Health (HUNT) Study, where they gave blood samples and were assessed

for headache at two time points 11 years apart. In order to identify changes in DNA

methylation associated with headache chronification we compared cases, who

transformed from episodic to chronic headache during the follow-up period, with

controls who did not develop chronic headache. For all selected subjects, genome-

wide DNA methylation levels were characterized at baseline and follow-up.

Materials and Methods

Participants

In the HUNT2 (1995–1997) and HUNT3 (2006–2008) studies, the entire population

of the Norwegian county of Nord-Trøndelag, 20 years of age or older, was invited to

complete extensive questionnaires concerning health-related issues, participate in a

clinical examination and donate blood samples. Among 92,566 individuals in HUNT2

and 94,194 in HUNT3 invited to participate, 51,383 (56%) and 39,701 (42%),

respectively, answered the headache questionnaire (9, 10).

The headache classification used in this study has been described in detail elsewhere

(9-11). Briefly, in each survey those who answered 'yes' to the screening question

'Have you suffered from headache during the last year?' were defined as headache

sufferers and asked to fill in the other headache questions. Participants indicated the

number of days with headache per month, with the options < 7, 7-14 or > 14

6

days/month in HUNT2 and the options < 1, 1-6, 7-14 or > 14 days/month in HUNT3.

Chronic daily headache (CDH) was defined as headache occurring on ≥ 15

days/month. In HUNT2, acute medication overuse was defined as 'analgesic use daily

or almost daily for ≥ 1 month during the last 12 months', and in HUNT3 as 'over-the-

counter-medication use for headache ≥ 4 days/week during the last month'.

Medication-overuse headache was defined as the combination of CDH and acute

medication overuse (9, 11, 12). This headache classification has been validated by

interview diagnoses (9, 10). In HUNT2, the sensitivity and specificity were,

respectively 85% and 83% for headache suffering (Cohen's kappa κ = 0.57), and 91%

and 56% for headache frequency < 7 days per month (κ = 0.50). In HUNT3, the

sensitivity and specificity were, respectively 88% and 86% for headache suffering (κ

= 0.70), 69% and 99% for CDH (κ = 0.75) and 75% and 100% for medication-

overuse headache (κ = 0.75).

Cases for the current study were selected among participating women who at the time

of HUNT2 (baseline) were ≤ 35 years of age and reported headache on < 7

days/month, and at the time of HUNT3 (follow-up) reported ≥ 15 headache

days/month (CDH). To avoid sample heterogeneity resulting from different DNA

methylation patterns in men and women, only women were included. Likewise, to

avoid DNA methylation changes resulting from the onset of menopause during

follow-up only women ≤ 35 years of age at baseline were included. In total 43

subjects fulfilled these criteria. After excluding samples with missing DNA from

either baseline or follow-up, 40 cases were included in the study. We selected an

equal number of controls, who were matched with cases on their headache frequency

and age at baseline, but at follow-up either did not have headache anymore or had

headache on < 7 days/month. Average time between baseline and follow-up

examinations was 11.4 years (range 10-12.6 years).

Ethics

Signed informed consent was obtained from all participants in the HUNT2 and

HUNT3 surveys, and the present study was approved by the Regional Committee for

Medical Research Ethics (2012/2316/REK midt).

Genome-wide quantification of DNA methylation

A detailed description of sample handling, quantification of DNA methylation, data

normalization and quality control is given in **Supplemental Methods**. Briefly, we

performed array quantification of DNA methylation on genomic DNA isolated from

leukocytes, after bisulfite conversion, using the Illumina Infinium

HumanMethylation450 BeadChip (Illumina, San Diego, CA, USA). For Stage 1, 96

samples from 24 cases and 24 controls were analyzed. For Stage 2, 64 samples from

16 cases and 16 controls were analyzed. The methylation level at each CpG site for

each sample was calculated as a β-value, scaling from 0 (not methylated) to 1 (fully

methylated). After sample-wise quality control steps, 36 cases and 35 controls

contributed to the final analysis. An overview of the study design and participants in

the final analysis are given in **Figure 1**. After marker-wise quality control steps, out

of 485,775 CpG markers initially interrogated 431,282 and 436,517 markers were

available for analysis in Stage 1 and Stage 2, respectively.

8

Analytic approach

<u>Baseline characteristics</u>

Baseline characteristics were compared between cases and controls by Fisher's exact

test (for binary variables) or independent samples *t*-test (for continuous variables).

The examined variables included known or suggested risk factors for the development

of chronic daily headache (4, 13); age (continuous); low educational level (upper

secondary school or less as highest education); self-reported daily use of tranquilizers

last year (yes, no); current daily smoking (yes, no); low physical activity (< 1 hour

light exercise per week); body-mass index (continuous); Hospital Anxiety and

Depression Scale (HADS) for each of anxiety and depression (continuous) (14),

chronic musculoskeletal complaints (yes, no) (15); gastrointestinal complaints (yes,

no) (16), self-reported diabetes mellitus (yes, no) and self-reported rheumatoid or

osteoarthritis (yes, no). Of the participants included in the study, 3 had missing

information on physical activity and 1 on chronic musculoskeletal complaints.

<u>Association analysis and replication</u>

The effect of headache chronification (case-control status) was modeled on changes in

DNA methylation levels. For each individual, normalized methylation level estimate

(β-value) at baseline was subtracted from methylation level estimate at follow-up for

each CpG site. Resulting methylation level changes at each CpG site were then

compared between cases and controls using linear regression model adjusted for age,

using the *lm* package in R software (17). Data from each stage was analyzed

separately and the 20 CpG sites showing strongest association with chronification in

Stage 1 were taken forward for replication in Stage 2. A Bonferroni-corrected p-value

threshold $(0.05/20 = 0.0025)$ was used to consider statistically significant replication.

9

Meta-analysis

Subsequently, a fixed-effects meta-analysis was performed on DNA methylation data

from Stage 1 and Stage 2, weighted by the inverse variance, using the GWAMA

software (18). In total 436,549 CpG markers were present in either stage after quality

control and could be used in the analysis, resulting in a Bonferroni-corrected p-value

threshold of 1.15×10^{-7} (0.05/436,549).

Characterization of top-ranking CpG sites

Functional profiling of the meta-analysis top-ranking CpG sites was carried out using

g:Profiler web-based software (19), see **Supplemental Methods** for a full description.

Briefly, g:Profiler performs a statistical gene set enrichment analysis to find

functional gene groups or biological pathways that are significantly over-represented

in user-provided input. The results presented are based on the 200 top-ranking CpG

sites from the meta-analysis, and results were considered significant if the multiple

testing corrected p-value for the enriched group or pathway was ≤ 0.05.

Characterization of gene expression in various human tissues was retrieved from the

GTEx database (20). Although DNA methylation is largely tissue-specific, inter-

individual variation in DNA methylation has been shown to correlate between blood

and brain for a subset of CpG sites (21). Correlation of DNA methylation levels

between blood and brain at the top CpG sites was analyzed using the Blood Brain

DNA Methylation Comparison tool (http://epigenetics.essex.ac.uk/bloodbrain) that

utilizes DNA methylation data of blood and four different brain regions from matched

samples (21).

Analysis of medication overuse

Overuse of acute headache medications is a known risk factor for headache

chronification. To evaluate to what extent medication overuse influenced the

identified associations, the association analysis and meta-analysis was repeated

stratified by medication overuse. Cases with or without medication overuse at follow-

up were analyzed separately against all controls. Results were visualized as forest

plots using the *metafor* R package (version 1.9-8) (22). Additionally, unsupervised

clustering of the samples was performed using the *hclust* R package, based on their

methylation levels at follow-up at each of the 200 top-ranking CpG sites from the

initial meta-analysis.

Results

Demographic variables and potential risk factors for CDH did not differ between

cases and controls at baseline (**Table 1**). However, four participants reported

medication overuse at baseline, and all four later developed CDH. Likewise no

difference between cases and controls was found for the listed variables when Stage 1

and Stage 2 were analyzed separately (data not shown). Average time between

baseline and follow-up examinations was 11.4 years (range 10-12.6 years). At the

time of follow-up 17 of 36 cases (47%) and 0 of 35 controls reported overuse of acute

headache medication.

The 20 CpG sites that demonstrated the strongest association with headache

chronification in Stage 1 are listed in **Table 2**. Using the Bonferroni-corrected

statistical significance level of 0.0025, none of these CpG sites reached statistical

significance for replication in Stage 2. When data from the two stages were combined

in a fixed-effects meta-analysis, no CpG sites reached the multiple-testing corrected

study-wide significance threshold of 1.15×10^{-7}. The most strongly associated CpG

marker was cg01010870 (p-value $= 6.17 \times 10^{-6}$), located near the transcription start

site of the gene *SH2D5* (**Table 3**) that is highly expressed in various regions of the

human brain (**Supplemental Figure 1**). The second most strongly associated CpG

marker cg11086695 (p-value $= 6.95 \times 10^{-6}$) is intergenic, located 76 kb downstream

from *NPTX2* and 141 kb downstream from *TMEM130*. For 6 of the top 10 CpG

markers (**Table 3**), methylation levels in blood showed high correlation with

methylation levels in one or several regions of the human brain at p-value < 0.05 (21),

exemplified for the most significantly associated CpG site (cg01010870) in **Figure 2**.

Genomic regions corresponding to the 200 top-ranking CpG sites from the meta-

analysis were significantly enriched for genes belonging to specific biological

functions (**Table 4**, **Supplemental Table 1**). Statistically strongest enrichment was

observed for "calcium ion binding" (p-value $= 1.21 \times 10^{-5}$), followed by "biological

adhesion" (p-value $= 4.23 \times 10^{-5}$), "intracellular estrogen receptor signaling" (p-value

$= 0.02$), and "negative regulation of homotypic cell-cell adhesion" (p-value $= 0.04$).

When stratifying the meta-analysis for overuse of acute headache medications among

cases, the associations for the 10 top-ranking CpG sites from the initial meta-analysis

had the same direction and were of comparable strength in each of the two groups

(**Figure 3**). Likewise, unbiased hierarchical cluster analysis based on the 200 top-

ranking CpG sites did not discriminate between cases with and without medication

overuse (**Figure 4**).

12

Discussion

This is, to our knowledge, the first study of epigenetic changes in headache

chronification. Using a retrospective case-control design, DNA methylation changes

were examined across the genome comparing patients who converted from episodic to

chronic headache during a 11 year follow-up period with matched controls. In the

discovery stage several CpG sites showed suggestive association to headache

chronification, but these failed to replicate in the second stage. Combining association

results from both stages in a fixed-effects meta-analysis we identified CpG sites that

may point to genes involved in headache chronification. An enrichment was found for

genes related to calcium ion binding, biological adhesion, and intracellular estrogen

receptor signaling.

Since cases in this study reported episodic headache at baseline and CDH at the time

of follow-up, it is likely that they represent *transformed* rather than *de-novo* chronic

headache. A limitation is however that the questionnaire-based assessment did not

allow distinction of chronic migraine from other types of chronic headache. Previous

epidemiological studies suggest that approximately half of those with CDH fulfill

criteria for chronic migraine (1), making it likely that migraineurs contribute

substantially to our case group.

Medication overuse is a major risk factor for headache chronification, and it has

previously been shown that medication overuse headache accounts for about half of

the participants who report CDH in the HUNT population (13). A similar proportion

was seen in this study with 17 of 36 cases (47%), but none of the controls, reporting

medication overuse at the time of follow-up. Notably, all 4 participants who reported

medication overuse at baseline had developed chronic headache at follow-up. Given

our design, it is feasible that at least some of the DNA methylation changes result

from medication use. Still, when stratifying the meta-analysis results by cases with

and without medication overuse, the strength and direction of association for the top

10 loci from the meta-analysis were similar between the groups (**Figure 3**). Also,

unsupervised hierarchical clustering of the samples based on methylation levels for

the top 200 CpG sites did not discriminate cases with from cases without medication

overuse (**Figure 4**). The results therefore suggest that medication overuse is unlikely

to explain the identified associations.

The most significant CpG site is at the *SH2D5* gene, which encodes the SH2 Domain

Containing 5 protein. The second strongest associated CpG site is 76 kb downstream

from the nearest gene *NPTX2*, which encodes the Neuronal Pentraxin II protein. Both

proteins are highly expressed in adult human brain (23, 24) (**Supplemental Figure 1**),

and DNA methylation levels at the associated CpG sites are highly correlated between

brain and blood (**Table 3 and Figure 2**). Both proteins have presumed roles in

regulating synaptic plasticity. The SH2 Domain Containing 5 protein regulates

synaptic plasticity through the control of Rac-GTP levels (25). Neuronal Pentraxin II

regulates synaptic plasticity by acting as an inhibitor of excitatory synapses, through

binding and clustering of glutamatergic AMPA receptors. Neuronal Pentraxin II is

induced by neuronal activity, and has an assumed function in rebalancing excitation-

inhibition after periods of increased synaptic activity (23). Of interest, another of the

top associated CpG markers is related to glutamate signaling. *GRID2* encodes the

glutamate receptor δ2, an ionotropic glutamate receptor which is expressed in several

brain regions, with highest expression in the cerebellum (20, 26). Epigenetic

14

alteration of synaptic plasticity has been previously hypothesized as a possible

mechanism of migraine chronification (6). Support for this comes from the

experimental finding that synchronous neuronal activity, which also occurs in cortical

spreading depression of the migraine aura (27), results in DNA methylation changes

in brain-specific genes related to neuronal plasticity (28). One may envisage that as

neuronal activity can cause epigenetic changes altering synaptic plasticity, migraine

attacks in a feed-forward loop may promote subsequent migraine attacks leading to

chronification. Our findings seem to provide at least some support for a role of

synaptic plasticity in headache chronification. Regardless, it is far from clear whether

such changes are a cause or consequence of frequent headache attacks. Furthermore,

our findings may implicate that glutamatergic activity is implicated not only in

migraine pathogenesis (29), but also in the transition from episodic to chronic

headache.

When considering the more extensive list of 200 top-ranking CpG sites, there is a

statistically significant enrichment in genes related to several specific functional

groups and pathways. The strongest enrichment was seen for calcium ion binding.

Calcium ion binding and signaling is important in a wide range of cellular systems,

including the synaptic release of neurotransmitters (30). While this presents a

plausible mechanism for calcium ion imbalance in migraine pathogenesis, no genetic

association to this or other calcium channels have been found for the common forms

of migraine (31). The second strongest enrichment was found for genes involved in

biological cell adhesion. Cell adhesion molecules have multiple roles in tissue

development and function. In the human brain they play a major role at synapses,

where they control synaptic morphology and plasticity, and at the neurovascular unit

15

they form the molecular interface between the vascular and the central nervous

system and are involved in the complex signaling between these systems (32, 33).

Lastly, in this study of female subjects, we found enrichment for genes involved in

estrogen receptor signaling. This corresponds well with lines of evidence suggesting a

role of estrogen in migraine pathogenesis (34). Estrogen receptor signaling has also

been discussed specifically in relation to possible epigenetic mechanisms in migraine,

since the estrogen receptors predominantly assert their action through the epigenetic

programming of target genes (6).

Strengths of this study include the genome-wide assessment of DNA methylation

sites, which allowed us to search for genetic loci associated with headache

chronification without relying on previous hypotheses. Also, the headache

classification used in this study has been validated by clinical interview. Given the

longitudinal design we could determine DNA methylation levels in both cases and

controls at two time points, before and after the development of chronic headache,

which gave the possibility to detect changes in DNA methylation levels occurring

during the natural course of the disease.

Our study has clear limitations. Despite a large baseline population of 51,000

individuals with available DNA who were characterized with respect to headache, and

a follow-up time of 11 years, only 36 cases fulfilled inclusion criteria and could be

used in the final analysis. This resulted in limited statistical power, as indicated by the

failure to replicate signals identified in the discovery stage in the smaller replication

sample. We therefore focused our subsequent analyses on the combined dataset of

both stages in order to maximize the power to detect true signals. The loci identified

16

from the combined analysis aggregate in functionally meaningful pathways,

indicating the presence of true signals. Nevertheless, the reported associations will

require replication in independent samples, but given our complex design this will be

challenging. This clearly demonstrates the difficulty of designing studies of epigenetic

changes during the natural course of incident disease, and suggests that international

collaboration is required to achieve robust sample sizes when using such a study

design. Another limitation was that we determined DNA methylation in leucocytes

from whole blood, whereas DNA methylation patterns are to a considerable extent

tissue-specific, so our findings may not be representative of changes occurring in the

more relevant tissue in headache, in particular the brain (8). However, for the majority

of the top associated markers, DNA methylation levels in blood correlated well with

DNA methylation levels in one or more brain regions, suggesting that our design has

good potential to detect DNA methylation relevant to migraine. This is also in line

with previous studies that have found that DNA methylation patterns in non-affected

tissues such as blood can be of value, even leading to treatment decisions (35).

A better understanding of the mechanisms underlying headache chronification will be

important for developing better strategies to treat and prevent headaches in these

severely affected patients. Epigenetic changes represent modifiable factors, and

thereby offer potential targets for novel drug therapies (5). Although candidate

regions identified in this study require independent replication and further

characterization, this is a first step towards understanding epigenetic changes

occurring in headache chronification.

Article highlights

- The two CpG sites with strongest association to headache chronification were found in relation to genes that regulate synaptic plasticity.

- Associated CpG sites also showed enrichment for genes involved in calcium ion binding, cell adhesion and estrogen receptor pathways.

- The results suggest possible mechanisms by which epigenetic DNA methylation may be involved in the chronification of headache.

18

Conflict of interest Statement

None declared.

Figure legends

Figure 1: Diagram of the study design and participants. A total of 36 female headache

patients who transformed from episodic to chronic headache ('cases') between

examinations at baseline and follow-up were matched against 35 female headache

patients who did not transform to chronic headache ('controls').

Figure 2: Correlation of DNA methylation in blood and four brain regions for

cg01010870 (*SH2D5*). Top: Boxplot of the distribution of DNA methylation values

for cg01010870 split by tissue. Bottom: Four scatterplots demonstrating the

relationship between DNA methylation in whole blood and four brain regions:

PFC=pre-frontal cortex; EC=entorhinal cortex; STG=superior temporal gyrus;

CER=cerebellum. Plots are retrieved using the Blood Brain DNA Methylation

Comparison tool (21).

Figure 3. Forest plots showing meta-analysis results separately for cases with and

without medication overuse. Meta-analysis beta-value (Beta) and its 95% confidence

intervals (95% CI) for the top 10 CpG sites (cg01010870 to cg14118171) from the

meta-analysis.

Figure 4: Unsupervised clustering of cases based on the top 200 CpG sites from the

meta-analysis. Unsupervised hierarchical clustering using methylation status at

19

follow-up did not discriminate between cases with (marked in yellow) and without

medication overuse in either Stage 1 (A) or Stage 2 (B).

Supplemental Figure 1. Relative expression of SH2D5 in various human tissues.

Retrieved from the GTEx database (20). RPKM=Reads Per Kilobase of transcript per

Million mapped reads. Box plots are shown as median and 25th and 75th percentiles;

points are displayed as outliers if they are above or below 1.5 times the interquartile

range.

References

[1]   Natoli JL, Manack A, Dean B, Butler Q, Turkel CC, Stovner L, Lipton RB.

Global prevalence of chronic migraine: a systematic review. Cephalalgia.

2010;30:599-609.

[2]   Ashina M. Neurobiology of chronic tension-type headache. Cephalalgia.

2004;24:161-72.

[3]   Lanteri-Minet M, Duru G, Mudge M, Cottrell S. Quality of life impairment,

disability and economic burden associated with chronic daily headache, focusing on

chronic migraine with or without medication overuse: a systematic review.

Cephalalgia. 2011;31:837-50.

[4]   Diener HC, Dodick DW, Goadsby PJ, Lipton RB, Olesen J, Silberstein SD.

Chronic migraine--classification, characteristics and treatment. Nature reviews

Neurology. 2011;8:162-71.

[5]   Yan H, Tian S, Slager SL, Sun Z, Ordog T. Genome-Wide Epigenetic Studies

in Human Disease: A Primer on -Omic Technologies. Am J Epidemiol. 2016;183:96-

109.

[6]   Eising E, N AD, van den Maagdenberg AM, Ferrari MD. Epigenetic

mechanisms in migraine: a promising avenue? BMC medicine. 2013;11:26.

[7]   Feil R, Fraga MF. Epigenetics and the environment: emerging patterns and

implications. Nat Rev Genet. 2011;13:97-109.

[8]   Labruijere S, Stolk L, Verbiest M, de Vries R, Garrelds IM, Eilers PH, et al.

Methylation of migraine-related genes in different tissues of the rat. PloS one.

2014;9:e87616.

[9]   Hagen K, Zwart JA, Aamodt AH, Nilsen KB, Brathen G, Helde G, et al. The

validity of questionnaire-based diagnoses: the third Nord-Trondelag Health Study

2006-2008. J Headache Pain. 2010;11:67-73.

[10]   Hagen K, Zwart JA, Vatten L, Stovner LJ, Bovim G. Head-HUNT: validity

and reliability of a headache questionnaire in a large population-based study in

Norway. Cephalalgia. 2000;20:244-51.

[11]   Linde M, Stovner LJ, Zwart JA, Hagen K. Time trends in the prevalence of

headache disorders. The Nord-Trondelag Health Studies (HUNT 2 and HUNT 3).

Cephalalgia. 2011;31:585-96.

[12]   Zwart JA, Dyb G, Hagen K, Svebak S, Stovner LJ, Holmen J. Analgesic

overuse among subjects with headache, neck, and low-back pain. Neurology.

2004;62:1540-4.

[13]   Hagen K, Linde M, Steiner TJ, Stovner LJ, Zwart JA. Risk factors for

medication-overuse headache: an 11-year follow-up study. The Nord-Trondelag

Health Studies. Pain. 2012;153:56-61.

[14]   Bjelland I, Dahl AA, Haug TT, Neckelmann D. The validity of the Hospital

Anxiety and Depression Scale. An updated literature review. Journal of

psychosomatic research. 2002;52:69-77.

21

[15]   Kvalheim S, Sandven I, Hagen K, Zwart JA. Smoking as a risk factor for

chronic musculoskeletal complaints is influenced by age. The HUNT study. Pain.

2013;154:1073-9.

[16]   Aamodt AH, Stovner LJ, Hagen K, Zwart JA. Comorbidity of headache and

gastrointestinal complaints. The Head-HUNT Study. Cephalalgia. 2008;28:144-51.

[17]   R Core Team (2013). R: A language and environment for statistical

computing. R Foundation for Statistical Computing, Vienna, Austria. URL

http://www.R-project.org/.

[18]   Magi R, Morris AP. GWAMA: software for genome-wide association meta-

analysis. BMC bioinformatics. 2010;11:288.

[19]   Reimand J, Arak T, Vilo J. g:Profiler--a web server for functional

interpretation of gene lists (2011 update). Nucleic Acids Res. 2011;39:W307-15.

[20]   The Broad Institute of MIT and Harvard. GTEx Portal, database version 6.

Retrieved February 26, 2016, from http://www.gtexportal.org/ [WWW document].

[21]   Hannon E, Lunnon K, Schalkwyk L, Mill J. Interindividual methylomic

variation across blood, cortex, and cerebellum: implications for epigenetic studies of

neurological and neuropsychiatric phenotypes. Epigenetics. 2015;10:1024-32.

http://epigenetics.iop.kcl.ac.uk/bloodbrain/.

[22]   Viechtbauer W. Conducting Meta-Analyses in R with the metafor Package. J

Stat Softw. 2010;36:1-48.

[23]   Chang MC, Park JM, Pelkey KA, Grabenstatter HL, Xu D, Linden DJ, et al.

Narp regulates homeostatic scaling of excitatory synapses on parvalbumin-expressing

interneurons. Nat Neurosci. 2010;13:1090-7.

[24] Pelkey KA, Barksdale E, Craig MT, Yuan X, Sukumaran M, Vargish GA, et al. Pentraxins coordinate excitatory synapse maturation and circuit integration of parvalbumin interneurons. Neuron. 2015;85:1257-72.

[25] Gray EJ, Petsalaki E, James DA, Bagshaw RD, Stacey MM, Rocks O, et al. Src homology 2 domain containing protein 5 (SH2D5) binds the breakpoint cluster region protein, BCR, and regulates levels of Rac1-GTP. J Biol Chem. 2014;289:35397-408.

[26] Coutelier M, Burglen L, Mundwiller E, Abada-Bendib M, Rodriguez D, Chantot-Bastaraud S, et al. GRID2 mutations span from congenital to mild adult-onset cerebellar ataxia. Neurology. 2015;84:1751-9.

[27] Lauritzen M. Pathophysiology of the migraine aura. The spreading depression theory. Brain. 1994;117 (Pt 1):199-210.

[28] Guo JU, Ma DK, Mo H, Ball MP, Jang MH, Bonaguidi MA, et al. Neuronal activity modifies the DNA methylation landscape in the adult brain. Nat Neurosci. 2011;14:1345-51.

[29] Vecchia D, Pietrobon D. Migraine: a disorder of brain excitatory-inhibitory balance? Trends Neurosci. 2012;35:507-20.

[30] Clapham DE. Calcium signaling. Cell. 2007;131:1047-58.

[31] Nyholt DR, LaForge KS, Kallela M, Alakurtti K, Anttila V, Farkkila M, et al. A high-density association screen of 155 ion transport genes for involvement with common migraine. Hum Mol Genet. 2008;17:3318-31.

[32] McCarty JH. Cell adhesion and signaling networks in brain neurovascular units. Curr Opin Hematol. 2009;16:209-14.

[33] Sakisaka T, Takai Y. Cell adhesion molecules in the CNS. J Cell Sci. 2005;118:5407-10.

23

[34]   Sacco S, Ricci S, Degan D, Carolei A. Migraine in women: the role of

hormones and their impact on vascular diseases. J Headache Pain. 2012;13:177-89.

[35]   Schubeler D. Function and information content of DNA methylation. Nature.

2015;517:321-6.

Table 1. Demographic and clinical characteristics at baseline for cases and controls

	Controls (n = 35)	Cases (n = 36)	p-value
Age (mean [SE])	26.1 [0.71]	26.3 [0.76]	0.85[a]
Higher education (n, [%])	6 [17.1]	10 [27.8]	0.40[b]
Acute medication overuse (n, [%])	0 [0]	4 [11.1]	0.12[b]
Daily use of tranquilizers (n, [%])	0 [0]	0 [0]	n.a.
Daily smoking (n, [%])	9 [25.7]	10 [27.8]	1.0[b]
Low physical activity (n, [%])	22 [66.7]	20 [57.1]	0.46[b]
Body-mass index (mean [SE])	25.9 [0.70]	25.2 [0.68]	0.44[a]
HADS Anxiety score (mean [SE])	4.20 [0.52]	5.28 [0.64]	0.20[a]
HADS Depression score (mean [SE])	2.46 [0.37]	2.51 [0.51]	0.94[a]
Chronic musculoskeletal complaints (n, [%])	9 [25.7]	16 [45.7]	0.13[b]
Gastrointestinal complaints (n, [%])	22 [62.9]	25 [69.4]	0.62[b]
Diabetes mellitus (n, [%])	0 [0]	1 [2.8]	1.0[b]
Arthritis (n, [%])	0 [0]	1 [2.8]	1.0[b]

n.a.=not applicable.
[a]Independent samples *t*-test. [b]Fisher's exact test.

Table 2. Summary statistics of the 20 CpG sites that showed strongest association with headache chronification in Stage 1

Marker	Chr	Position	Gene(s)*	Discovery (Stage 1)			Replication (Stage 2)		
				β	SE	p	β	SE	p
cg00113321	7	157737485	PTPRN2	0.031	0.0054	1.17E-06	0.003	0.0112	0.79
cg01010870	1	20732775	SH2D5	0.022	0.0041	3.93E-06	0.009	0.0077	0.27
cg20968963	2	2767914	MYT1L; LINC01250	0.021	0.0040	4.41E-06	0.000	0.0077	0.99
cg07195851	5	3517408	LINC01019	-0.039	0.0075	5.17E-06	0.003	0.0123	0.78
cg18321073	5	12795593	LINC01194	-0.033	0.0063	7.17E-06	-0.012	0.0110	0.29
cg18221862	2	192194504	TMEFF2	0.019	0.0038	8.58E-06	-0.009	0.0058	0.13
cg12614711	7	1496305	INTS1	0.011	0.0022	8.98E-06	-0.012	0.0042	0.007
cg13432241	4	153468133	KIAA0922	-0.016	0.0032	1.25E-05	-0.005	0.0043	0.21
cg03091590	19	42196421	POU2F2	0.023	0.0047	2.14E-05	0.005	0.0087	0.56
cg24993607	13	112602844	RP11-88E10.4	0.023	0.0049	2.25E-05	0.001	0.0072	0.92
cg14006774	19	44676311	CEACAM19	0.027	0.0059	3.81E-05	0.003	0.0102	0.79
cg15127702	7	101436336	COL26A1	0.048	0.0104	3.97E-05	0.008	0.0176	0.65
cg17509967	19	13506280	CACNA1A	0.011	0.0025	3.99E-05	-0.004	0.0063	0.53
cg26835520	16	32253720	TP53TG3D	-0.022	0.0049	4.90E-05	0.002	0.0050	0.69
cg24281422	8	67064238	COPS5; CSPP1	-0.014	0.0030	5.03E-05	0.010	0.0069	0.16
cg08231664	14	102050295	DYNC1H1	0.023	0.0050	5.38E-05	-0.012	0.0065	0.08
cg20137441	12	132553440	FBRSL1	0.023	0.0051	5.52E-05	-0.003	0.0061	0.58
cg26764019	19	2462520	LOC101928602; GADD45B	-0.016	0.0036	6.08E-05	0.003	0.0061	0.67
cg04388053	X	153904893	AVPR2	0.021	0.0048	6.49E-05	-0.013	0.0062	0.04
cg14209854	16	11276493	RMI2	0.022	0.0050	6.52E-05	0.004	0.0078	0.63

β, SE (standard error) and p (p-value) for association between headache chronification and change in CpG site methylation (linear regression model), adjusted for age.
Chr=chromosome. Position=genome position (build GRCh38). *For probes not located in direct relation to a gene, the two nearest genes are listed, separated by semicolon.

Table 3. CpG loci demonstrating strongest association to headache chronification in combined meta-analysis

Marker	Chr	Position	Gene**	β	SE	p	β (Stage 1)	β (Stage 2)
cg01010870*	1	20732775	*SH2D5*	0.018	0.0041	6.17E-06	0.022	0.009
cg11086695*	7	98705889	*NPTX2; TMEM130*	-0.039	0.0086	6.95E-06	-0.038	-0.041
cg00113321*	7	157737485	*PTPRN2*	0.024	0.0055	1.23E-05	0.031	0.003
cg18321073	5	12795593	*LINC01194*	-0.026	0.0061	1.41E-05	-0.033	-0.012
cg02969243*	X	11427642	*ARHGAP6*	-0.027	0.0062	1.50E-05	-0.028	-0.024
cg05678960*	Y	6265797	*TSPY2; LINC00280*	0.056	0.0130	1.89E-05	0.054	0.078
cg13952745	4	92304315	*GRID2*	-0.016	0.0038	1.91E-05	-0.016	-0.016
cg20458019*	4	2275622	*ZFYVE28*	0.015	0.0036	2.30E-05	0.015	0.015
cg23174760	1	63150382	*RP11-335E6.2*	0.022	0.0052	2.51E-05	0.024	0.019
cg14118171	9	131475765	*PRRC2B*	0.014	0.0033	2.70E-05	0.012	0.016

* Methylation levels in blood showing high correlation with methylation levels in one or more regions of the human brain at p-value < 0.05.

** For CpG markers not located in direct relation to a gene, the two nearest genes are listed.

β, SE (standard error) and p (p-value) for association between headache chronification and change in CpG site methylation in meta-analysis of results from Stages 1 and 2. Chr=chromosome. Position=genome position (build GRCh38).

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Table 4. Functional profiling of top associated CpG sites and corresponding genes

p-value*	Term ID	Term name
1.21E-05	GO:0005509	**calcium ion binding**
4.23E-05	GO:0022610	**biological adhesion**
3.87E-05	GO:0007155	cell adhesion
7.21E-06	GO:0098609	cell-cell adhesion
6.52E-10	GO:0098742	cell-cell adhesion via plasma-membrane adhesion molecules
6.68E-12	GO:0007156	homophilic cell adhesion via plasma membrane adhesion molecules
0.024	GO:0030520	**intracellular estrogen receptor signaling pathway**
0.037	GO:0034111	**negative regulation of homotypic cell-cell adhesion**

* Functional profiling using *g:Profiler*. The p-values are corrected for multiple testing. Highest-level hierarchical groups and pathways ('main terms') that were significant are marked in bold. Increased indent indicates deeper level in the local hierarchy under the group or pathway.
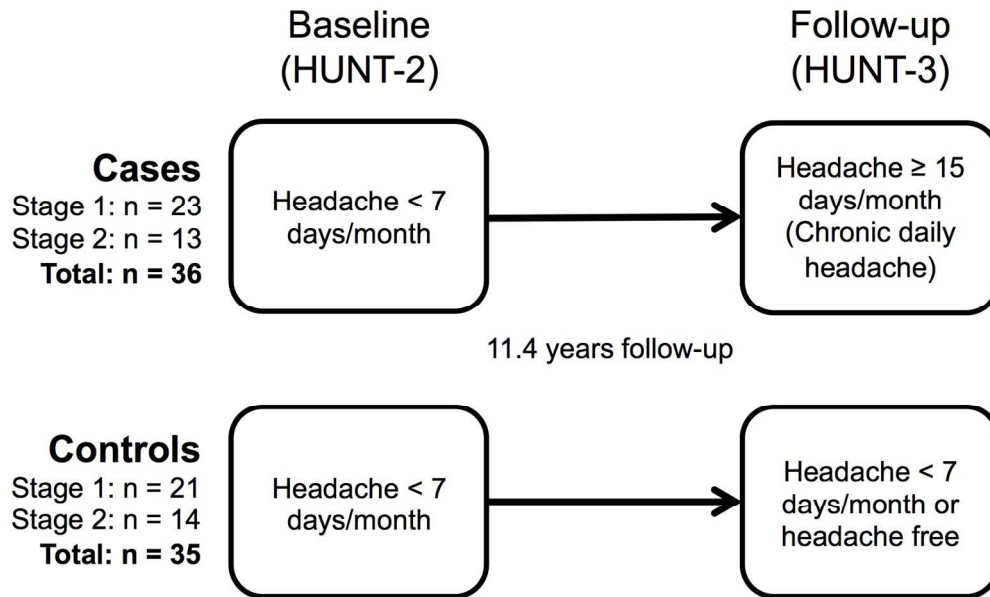
Figure 1: Diagram of the study design and participants. A total of 36 female headache patients who transformed from episodic to chronic headache ('cases') between examinations at baseline and follow-up were matched against 35 female headache patients who did not transform to chronic headache ('controls')
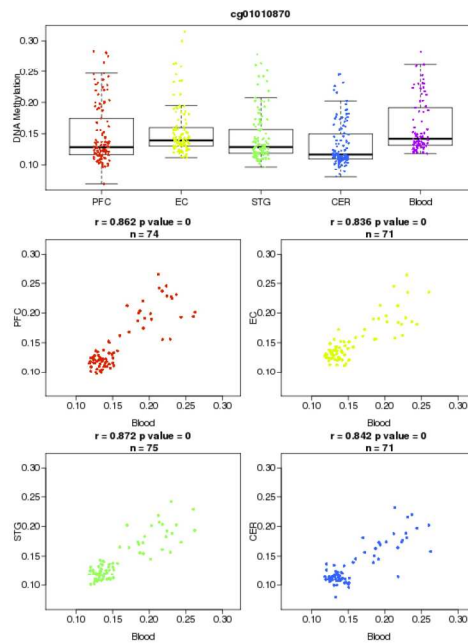
151x91mm (300 x 300 DPI)

Figure 2: Correlation of DNA methylation in blood and four brain regions for cg01010870 (SH2D5). Top: Boxplot of the distribution of DNA methylation values for cg01010870 split by tissue. Bottom: Four scatterplots demonstrating the relationship between DNA methylation in whole blood and four brain regions: PFC=pre-frontal cortex; EC=entorhinal cortex; STG=superior temporal gyrus; CER=cerebellum. Plots are retrieved using the Blood Brain DNA Methylation Comparison tool (21)
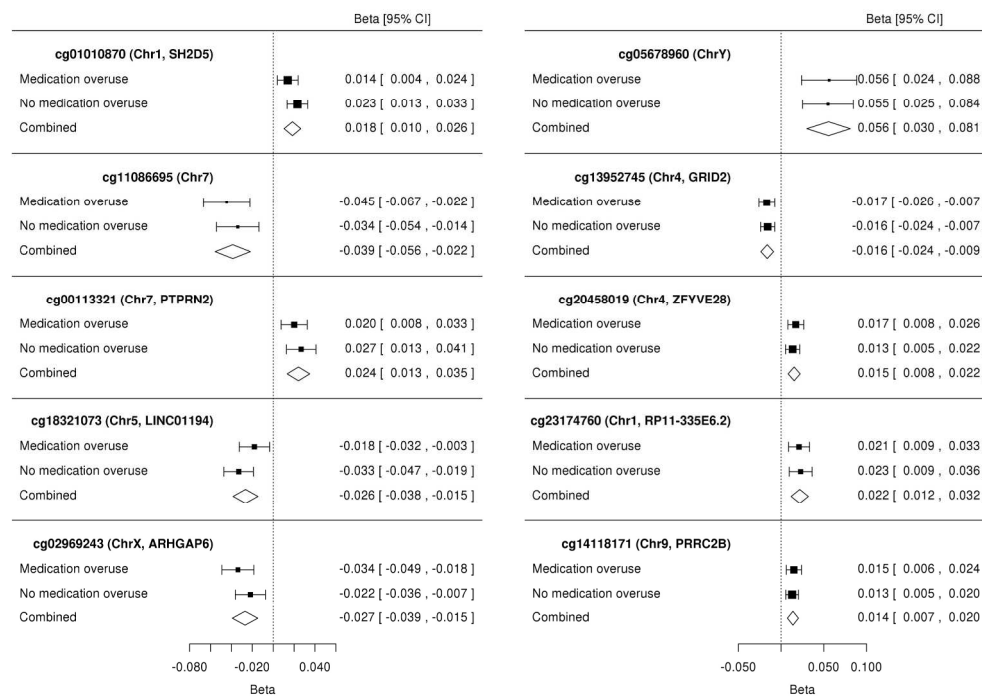
297x209mm (300 x 300 DPI)

Figure 3. Forest plots showing meta-analysis results separately for cases with and without medication overuse. Meta-analysis beta-value (Beta) and its 95% confidence intervals (95% CI) for the top 10 CpG sites (cg01010870 to cg14118171) from the meta-analysis
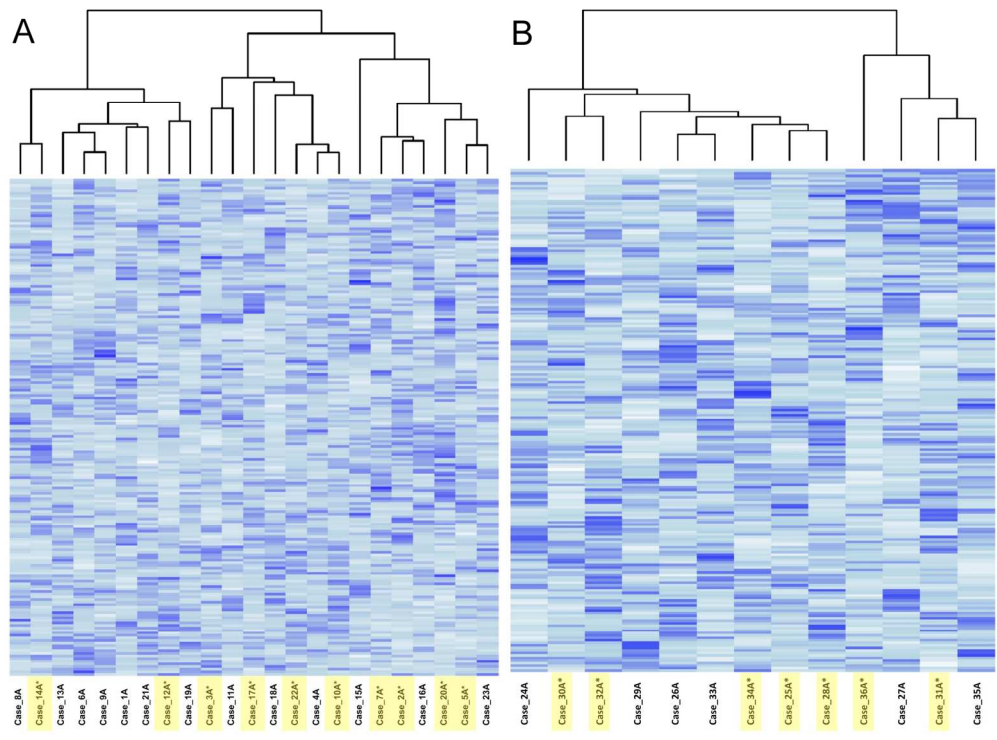
254x179mm (300 x 300 DPI)

Figure 4: Unsupervised clustering of cases based on the top 200 CpG sites from the meta-analysis. Unsupervised hierarchical clustering using methylation status at follow-up did not discriminate between cases with (marked in yellow) and without medication overuse in either Stage 1 (A) or Stage 2 (B)

178x132mm (300 x 300 DPI)

1

Epigenetic DNA methylation changes associated with headache chronification: A retrospective case-control study

SUPPLEMENTAL METHODS

Blood sampling and DNA isolation

Blood samples were donated from subjects in the HUNT2 or HUNT3 studies at the time of inclusion. Serum was separated from whole blood within 2 hours by centrifugation, placed in a refrigerator together with a blood clot or EDTA whole blood, and transported in a cooler within 1 - 3 days to the HUNT Biobank, where it was stored at -70°C. DNA was isolated from leukocytes from thawed blood clots or EDTA whole blood by using Gentra Puregene Blood Kit (Qiagen Diagnostics GmbH, Hamburg, Germany) or MasterPure DNA purification Kit for blood, Buffy Coat DNA Purification Protocol (Epicentre) and stored at -20°C. DNA concentrations were quantified using PicoGreen reagents, and every eight normalized sample was spectrophotometrically tested a second time using NanoDrop (ND1000 or ND8000) (Thermo Fisher Scientific Inc., Waltham, MA, USA). Purity was assessed based on the A260/A280 ratios.

Genome-wide quantification of DNA methylation

Bisulfite conversion and array quantification of DNA methylation was performed in two separate stages. In Stage 1, 96 samples from 24 cases and 24 controls were analyzed at the Institute for Molecular Medicine Finland (FIMM), Helsinki, Finland; in Stage 2, 64 samples from 16 cases and 16 controls were analyzed at the Wellcome Trust Sanger Institute, Cambridge, UK. Bisulfite conversion of 1 µg of the DNA sample was performed using the EZ-96 DNA Methylation Kit (Zymo Research, Irvine, CA, USA) according to the manufacturer's protocol. 5 µl of the bisulfite converted DNA were used for genome-wide methylation profiling using the Illumina Infinium HumanMethylation450 BeadChip (Illumina, San Diego, CA, USA) according to manufacturer's protocol. After hybridization and washing, the BeadChips were scanned using the Illumina HiScan SQ scanner. The efficiency of the bisulfite conversion was checked according to the manufacturer's protocol.

Data normalization and quality control

Raw image data were imported into the GenomeStudio (Illumina version 2011.1) Methylation module (version 1.9), which was used to extract the fluorescent signal intensities and pre-normalize the data. Briefly, background signal intensity was subtracted from the raw signal intensity values and data were normalized with respect to the Illumina internal control probes. The methylation level was calculated as a β-value (*AVG_Beta*), scaling from 0 (not methylated) to 1 (fully methylated, representing the fraction of methylated DNA molecules at the corresponding CpG site in the assayed sample. These data were exported from the GenomeStudio as standard Illumina report files and further normalized and analyzed with custom scripts. Data for stages 1 and 2 were jointly quantile-normalized (1) with R software (version 3.2.3) (2) as described by Grundberg et al. (3), followed by the PCA analysis of β-values to detect any potential outliers and other quality control steps.

   In sample-wise quality control steps, altogether 9 samples were excluded. Of the 40 cases and 40 controls selected for methylation typing (24 cases and 24 controls

2

in Stage 1; 16 cases and 16 controls in Stage 2), samples were excluded because of low DNA concentration (n = 2); because they were red staining outliers (n = 2); had gender mismatch between methylation typing and phenotype annotation (n = 1); or were sample outliers as identified by PCA analysis (n = 4). Samples were also excluded if after above-mentioned exclusions they did not have data for both time points available, resulting in 36 cases (23 in Stage 1 and 13 in Stage 2) and 35 controls (21 in Stage 1 and 14 in Stage 2) contributing to the final analysis.

Marker (CpG site) wise quality control was carried out similarly to Grundberg and colleagues (3). Briefly, out of 485,775 CpG markers initially interrogated by the HumanMethylation450 chip, all markers were excluded that had a call rate ≤ 0.9; overlapped with single nucleotide variants with minor allele frequency ≥ 5% in the CEU panel of the 1000Genomes Project (4); or overlapped with copy number variants in 1000Genomes Project. After quality control 431,282 and 436,517 markers were available for analysis in Stage 1 and Stage 2, respectively.

Genomic annotation of interrogated CpG sites

To assign each assayed CpG marker to their corresponding gene, genomic and gene-level annotations provided by Illumina (HumanMethylation450_15017482_v.1.1.csv, based on the human reference sequence assembly version GRCh37/hg19) were used. In order to get the latest genomic annotation of the meta-analysis top-ranking CpG sites for functional profiling, chromosomal coordinates of the top 1000 CpG sites were converted to the current version of the human reference sequence assembly (GRCh38/hg38) with the liftOver web-based batch conversion tool (http://genome.ucsc.edu/cgi-bin/hgLiftOver, by using default parameters).

Blood cell composition estimates

White blood cell (WBC) composition is regarded a main source of variation in DNA methylation analyses of whole blood. To check for possible WBC composition biases in our samples, the WBC composition was estimated from the DNA methylation profiles using the R-package minfi (version 1.8.9) developed by Houseman et al. (5). No significant differences were found in white blood cell composition between cases and controls in either Stage 1 or Stage 2.

Functional profiling of top-ranking CpG sites using g:profiler

Functional profiling of the meta-analysis top-ranking CpG sites was carried out by utilizing g:Profiler web-based software (version r1536, http://biit.cs.ut.ee/gprofiler) using Ensembl 83 genomic annotations, Gene Ontology release 2016-01-21 and KEGG biological pathways release 2016-01-25 (6). Considering all genes in user-provided gene lists or chromosomal regions, g:Profiler performs statistical gene set enrichment analysis to find which functional gene groups or biological pathways are significantly over-represented in user-provided input. As our input CpG site coordinates were ordered by their meta-analysis association p-values, the 'Ordered query' option was used, and we also required at least 3 CpG site-related genes in any tested functional group or pathway (option 'Size of query / term intersection'). Only CpG markers that were successfully called and showed the same effect direction in both stages were included, in order to enrich the input for true positive signals. The optimal size of the input set of CpG sites for the gene set enrichment analysis depends on a balance between including a sufficient number of sites to enable the detection of enrichment, while limiting the input set so that the highest possible proportion of true signals is obtained. To this end, the analysis was run using incrementally larger input

3

sets, including 10, 20, 50, 100 and 200 meta-analysis top-ranking CpG markers
resepectively. The sets including 10, 20, 50 and 100 markers did not return significant
enriched functional groups or pathways and are therefore not presented. The results
were considered significant if the multiple testing corrected p-value for enriched
group or pathway was ≤ 0.05.

References

[1]   Bolstad BM, Irizarry RA, Astrand M, Speed TP. A comparison of normalization
methods for high density oligonucleotide array data based on variance and bias.
Bioinformatics. 2003;19:185-93.

[2]   R Core Team (2013). R: A language and environment for statistical computing. R
Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

[3]   Grundberg E, Meduri E, Sandling JK, Hedman AK, Keildson S, Buil A, et al. Global
analysis of DNA methylation variation in adipose tissue from twins reveals links to disease-
associated variants in distal regulatory elements. Am J Hum Genet. 2013;93:876-90.

[4]   Genomes Project C, Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, et al. A
global reference for human genetic variation. Nature. 2015;526:68-74.

[5]   Houseman EA, Accomando WP, Koestler DC, Christensen BC, Marsit CJ, Nelson HH,
et al. DNA methylation arrays as surrogate measures of cell mixture distribution. BMC
bioinformatics. 2012;13:86.

[6]   Reimand J, Arak T, Vilo J. g:Profiler--a web server for functional interpretation of
gene lists (2011 update). Nucleic Acids Res. 2011;39:W307-15.

Supplemental Table 1. Significantly over-represented functional groups, and their corresponding genes

Term name	Term ID	Term type	p-value	Term group	Depth in group	Gene list*
Negative regulation of homotypic cell-cell adhesion	GO:0034111	Biological process	3.67E-02	4	1	*IL4R* (ENSG00000077238), *KIAA0922* (ENSG00000121210), *C1QTNF1* (ENSG00000173918), *ABAT* (ENSG00000183044)
Intracellular estrogen receptor signaling pathway	GO:0030520	Biological process	2.39E-02	3	1	*ESR1* (ENSG00000091831), *CNOT1* (ENSG00000125107), *DYX1C1* (ENSG00000256061)
Biological adhesion	GO:0022610	Biological process	4.23E-05	2	1	*FYN* (ENSG00000010810), *ARHGAP6* (ENSG00000047648), *PVR* (ENSG00000073008), *IL4R* (ENSG00000077238), *PCDHGA2* (ENSG00000081853), *PCDHB15* (ENSG00000113248), *NRP2* (ENSG00000118257), *TGFBI* (ENSG00000120708), *KIAA0922* (ENSG00000121210), *NLGN4X* (ENSG00000146938), *HSD17B12* (ENSG00000149084), *GRID2* (ENSG00000152208), *COL26A1* (ENSG00000160963), *C1QTNF1* (ENSG00000173918), *CNTNAP2* (ENSG00000174469), *ABAT* (ENSG00000183044), *DLL1* (ENSG00000198719), *PCDHGA1* (ENSG00000204956), *PCDHGA5* (ENSG00000253485), *PCDHGA7* (ENSG00000253537), *PCDHGA6* (ENSG00000253731), *PCDHGA8* (ENSG00000253767), *PCDHGB2* (ENSG00000253910), *PCDHGB4* (ENSG00000253953), *PCDHGB1* (ENSG00000254221), *PCDHGA3* (ENSG00000254245), *PCDHGB3* (ENSG00000262209), *PCDHGA4* (ENSG00000262576)
Cell adhesion	GO:0007155	Biological process	3.87E-05	2	2	*FYN* (ENSG00000010810), *ARHGAP6* (ENSG00000047648), *PVR* (ENSG00000073008), *IL4R* (ENSG00000077238), *PCDHGA2* (ENSG00000081853), *PCDHB15* (ENSG00000113248), *NRP2* (ENSG00000118257), *TGFBI* (ENSG00000120708), *KIAA0922* (ENSG00000121210), *NLGN4X* (ENSG00000146938), *HSD17B12* (ENSG00000149084), *GRID2* (ENSG00000152208), *COL26A1* (ENSG00000160963), *C1QTNF1* (ENSG00000173918), *CNTNAP2* (ENSG00000174469), *ABAT* (ENSG00000183044), *DLL1* (ENSG00000198719), *PCDHGA1* (ENSG00000204956), *PCDHGA5* (ENSG00000253485), *PCDHGA7*

						(ENSG00000253537), *PCDHGA6* (ENSG00000253731), *PCDHGA8* (ENSG00000253767), *PCDHGB2* (ENSG00000253910), *PCDHGB4* (ENSG00000253953), *PCDHGB1* (ENSG00000254221), *PCDHGA3* (ENSG00000254245), *PCDHGB3* (ENSG00000262209), *PCDHGA4* (ENSG00000262576)
Cell-cell adhesion	GO:0098609	Biological process	7.21E-06	2	3	*IL4R* (ENSG00000077238), *PCDHGA2* (ENSG00000081853), *PCDHB15* (ENSG00000113248), *KIAA0922* (ENSG00000121210), *C1QTNF1* (ENSG00000173918), *ABAT* (ENSG00000183044), *PCDHGA1* (ENSG00000204956), *PCDHGA5* (ENSG00000253485), *PCDHGA7* (ENSG00000253537), *PCDHGA6* (ENSG00000253731), *PCDHGA8* (ENSG00000253767), *PCDHGB2* (ENSG00000253910), *PCDHGB4* (ENSG00000253953), *PCDHGB1* (ENSG00000254221), *PCDHGA3* (ENSG00000254245), *PCDHGB3* (ENSG00000262209), *PCDHGA4* (ENSG00000262576)
Cell-cell adhesion via plasma-membrane adhesion molecules	GO:0098742	Biological process	6.52E-10	2	4	*PCDHGA2* (ENSG00000081853), *PCDHB15* (ENSG00000113248), *PCDHGA1* (ENSG00000204956), *PCDHGA5* (ENSG00000253485), *PCDHGA7* (ENSG00000253537), *PCDHGA6* (ENSG00000253731), *PCDHGA8* (ENSG00000253767), *PCDHGB2* (ENSG00000253910), *PCDHGB4* (ENSG00000253953), *PCDHGB1* (ENSG00000254221), *PCDHGA3* (ENSG00000254245), *PCDHGB3* (ENSG00000262209), *PCDHGA4* (ENSG00000262576)
Homophilic cell adhesion via plasma membrane adhesion molecules	GO:0007156	Biological process	6.68E-12	2	5	*PCDHGA2* (ENSG00000081853), *PCDHB15* (ENSG00000113248), *PCDHGA1* (ENSG00000204956), *PCDHGA5* (ENSG00000253485), *PCDHGA7* (ENSG00000253537), *PCDHGA6* (ENSG00000253731), *PCDHGA8* (ENSG00000253767), *PCDHGB2* (ENSG00000253910), *PCDHGB4* (ENSG00000253953), *PCDHGB1* (ENSG00000254221), *PCDHGA3* (ENSG00000254245), *PCDHGB3* (ENSG00000262209), *PCDHGA4* (ENSG00000262576)

Calcium ion binding	GO:0005509	Molecular function	1.21E-05	1	1	*PCDHGA2* (ENSG00000081853), *PCDHB15* (ENSG00000113248), *SLC25A12* (ENSG00000115840), *DLL1* (ENSG00000198719), *PCDHGA1* (ENSG00000204956), *PCDHGA5* (ENSG00000253485), *PCDHGA7* (ENSG00000253537), *PCDHGA6* (ENSG00000253731), *PCDHGA8* (ENSG00000253767), *PCDHGB2* (ENSG00000253910), *PCDHGB4* (ENSG00000253953), *PCDHGB1* (ENSG00000254221), *PCDHGA3* (ENSG00000254245), *PCDHGB3* (ENSG00000262209), *PCDHGA4* (ENSG00000262576)

Gene set enrichment analysis using *g:Profiler*. The p-values are corrected for multiple testing.
*Genes corresponding to the top 200 CpG sites (See Supplementary Methods, Genomic annotation of interrogated CpG sites), which are annotated to the given term name.
HUGO Gene Nomenclature Committee name followed (in parentheses) by Ensembl ID.
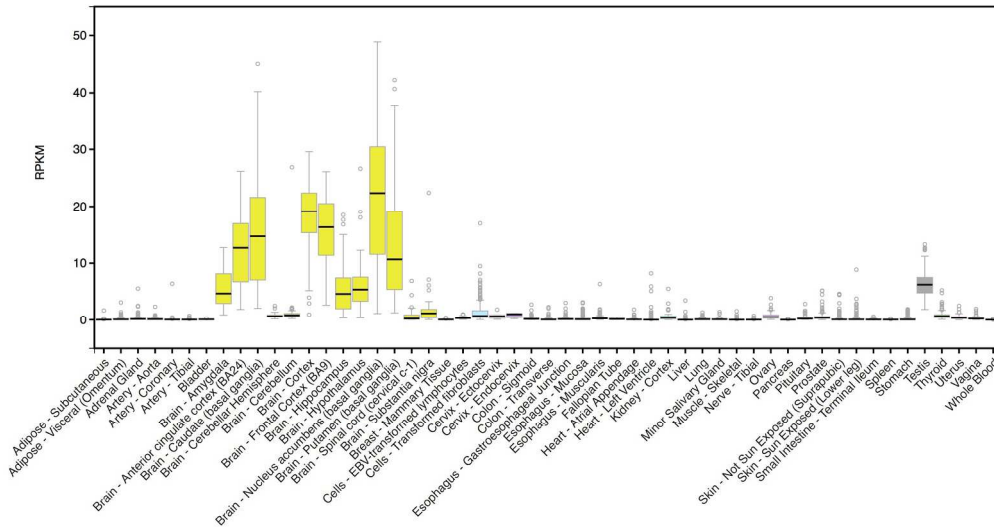
Supplemental Figure 1. Relative expression of SH2D5 in various human tissues. Retrieved from the GTEx database (20). RPKM=Reads Per Kilobase of transcript per Million mapped reads. Box plots are shown as median and 25th and 75th percentiles; points are displayed as outliers if they are above or below 1.5 times the interquartile range

204x108mm (300 x 300 DPI)