

Epigenome-Wide Association Study of Aggressive Behavior

Jenny van Dongen,^{1,2} Michel G. Nivard,¹ Bart M. L. Baselmans,^{1,2} Nuno R. Zilhão,¹ Lannie Ligthart,¹ BIOS Consortium,³ Bastiaan T. Heijmans,⁴ Meike Bartels,^{1,2} and Dorret I. Boomsma^{1,2}

¹Department of Biological Psychology, VU Amsterdam, Amsterdam, The Netherlands

²EMGO Institute for Health and Care Research, VU University Medical Center, Amsterdam, The Netherlands

³The Biobank-Based Integrative Omics Study (BIOS) Consortium (a full list of authors is provided in the Supplementary Material)

⁴Department of Molecular Epidemiology, Leiden University Medical Center, Leiden, The Netherlands

Aggressive behavior is highly heritable, while environmental influences, particularly early in life, are also important. Epigenetic mechanisms, such as DNA methylation, regulate gene expression throughout development and adulthood, and may mediate genetic and environmental effects on complex traits. We performed an epigenome-wide association study (EWAS) to identify regions in the genome where DNA methylation level is associated with aggressive behavior. Subjects took part in longitudinal survey studies from the Netherlands Twin Register (NTR) and participated in the NTR biobank project between 2004 and 2011 ($N = 2,029$, mean age at blood sampling = 36.4 years, $SD = 12.4$, females = 69.2%). Aggressive behavior was rated with the ASEBA Adult Self-Report (ASR). DNA methylation was measured in whole blood by the Illumina HM450k array. The association between aggressive behavior and DNA methylation level at 411,169 autosomal sites was tested. Association analyses in the entire cohort showed top sites at cg01792876 (chr8; 116,684,801, nearest gene = *TRPS1*, $p = 7.6 \times 10^{-7}$, False discovery rate (FDR) = 0.18) and cg06092953 (chr18; 77,905,699, nearest gene = *PAR6G-AS1*, $p = 9.0 \times 10^{-7}$, FDR = 0.18). Next, we compared methylation levels in 20 pairs of monozygotic (MZ) twins highly discordant for aggression. Here the top sites were cg21557159 (chr 11; 107,795,699, nearest gene = *RAB39*, $p = 5.7 \times 10^{-6}$, FDR = 0.99), cg08648367 (chr 19; 51,925,472, nearest gene = *SIGLEC10*, $p = 7.6 \times 10^{-6}$, FDR = 0.99), and cg14212412 (chr 6; 105,918,992, nearest gene = *PREP*, $p = 8.0 \times 10^{-6}$, FDR = 0.99). The two top hits based on the entire cohort showed the same direction of effect in discordant MZ pairs (cg01792876, $P_{\text{discordant twins}} = 0.09$ and cg06092953, $P_{\text{discordant twins}} = 0.24$). The other way around, two of the three most significant sites in discordant MZ pairs showed the same direction of effect in the entire cohort (cg08648367, $P_{\text{entire EWAS}} = 0.59$ and cg14212412, $P_{\text{entire EWAS}} = 3.1 \times 10^{-3}$). Gene ontology analysis highlighted significant enrichment of various central nervous system categories among higher-ranking methylation sites. Higher-ranking methylation sites also showed enrichment for DNase I hypersensitive sites and promoter regions, showing that DNA methylation in peripheral tissues is likely to be associated with aggressive behavior.

■ **Keywords:** aggression, aggressive, DNA methylation, epigenetics, EWAS, 450k, twins

Throughout the lifespan, large individual differences are evident in the expression of aggressive behavior. In children, aggressive behavior is regarded to represent, up to a certain extent, a 'normal' behavior that decreases in frequency when children learn to express themselves in more socially preferable ways (Tremblay, 2008). Longitudinal studies have shown that aggressive behavior emerges in early childhood with a peak between 2 and 4 years of age, but decreases during childhood in most individuals (Alink et al., 2006; Cote et al., 2006, 2007; Tremblay et al., 2004). Abnormal levels of aggression in childhood and adulthood are a character-

istic of several psychiatric disorders (Coccaro 2012; Coie et al., 1992; Nestor, 2002) and may be caused by traumatic brain injury (Rao et al., 2009; Tateno et al., 2003), substance abuse (Boles & Miotto, 2003; Moore et al., 2008), and

RECEIVED 31 August 2015; ACCEPTED 20 September 2015. First published online 28 October 2015.

ADDRESS FOR CORRESPONDENCE: Jenny van Dongen, Department of Biological Psychology, VU Amsterdam, Van der Boechorststraat 1, 1081BT Amsterdam, The Netherlands. E-mail: j.van.dongen@vu.nl

neurodegenerative diseases (Lai et al., 2003; Moechars et al., 1996; Volicer & Hurley, 2003).

Twin studies illustrate that in the general population, a large part of variation in aggressive behavior between individuals is explained by genetic differences. For example, in 3- to 12-year-old children, the heritability ranges between 51% and 72% (Hudziak et al., 2003), and a study of adults reported heritability estimates between 37% and 57% for subtypes of aggressive behavior (Yeh et al., 2010). Besides genetic influences, it is clear that the conditions under which a person grows up are important. In the twin study of 3- to 12-year-old children, 13–27% of the variation in aggressive behavior was explained by the family environment (Hudziak et al., 2003), and a range of adverse early life conditions are known to predict high levels of aggression in childhood, such as low parental income, maternal smoking during pregnancy, separation from a parent, family dysfunction, and coercive-hostile parenting by the mother (Tremblay et al., 2004).

Throughout life, the activity of gene expression is coordinated by epigenetic mechanisms such as DNA methylation at CpG dinucleotides and histone modifications (Bird, 2007). These epigenetic marks are not fixed: they change, for example, to switch different genes on and off during development (Reik, 2007). There is also increasing evidence from human and animal studies that DNA methylation may change in response to external conditions, including early life events (McGowan et al., 2009; Weaver et al., 2004) and prenatal exposures such as maternal smoking and diet (Bakulski et al., 2015; Breton et al., 2009; Richmond et al., 2015; Suter et al., 2011; Tobi et al., 2015). Therefore, the study of epigenetic marks may identify genes that are differentially regulated in individuals with higher, compared with lower levels of aggression. This idea is illustrated by a study of rhesus macaques, which showed that monkeys who were separated from their mother in early life and became highly aggressive exhibit numerous methylation differences in the prefrontal cortex and T-cells compared with monkeys reared by their mother (Provencal et al., 2012). The effects were largely tissue-specific, with only a very small subset of genes and individual methylation sites being significantly differentially methylated in both tissue types (Provencal et al., 2012). One of such overlapping sites was a CpG site upstream of *A2D681*, which is the homolog of the human glucocorticoid receptor gene *NR3C1*. Thus far, epigenetic studies of aggression or related phenotypes in humans have been scarce. A recent candidate gene study reported differences between individuals with antisocial personality disorder from a prisoner population and controls in methylation of the monoamine oxidase A (*MAOA*) gene in blood (Checknita et al., 2015). Another candidate gene study reported higher methylation at several CpG sites in the serotonin transporter gene (*SLC6A4*) promoter in T-cells and monocytes from adult men who exhibited high levels of physical aggression in childhood (Wang et al.,

2012). Third, methylation differences have been reported in candidate genes encoding cytokines and their transcription factors in a study of eight adult men who were classified as being on a chronic physical aggression trajectory between 6 and 15 years of age (Provencal et al., 2013). In the same group of men (Provencal et al., 2014), and in a group of 25 women with childhood aggression (Guillemin et al., 2014), differences in methylation were also identified in an analysis targeting genome-wide promoters in T-cells and monocytes. Hence, DNA methylation differences associated with aggressive behavior are observed in blood cells and may reflect changes that are either the cause or consequence of aggressive behavior. Thus far, a large majority of these previous findings have not been replicated, and the role of DNA methylation in aggressive behavior remains inconclusive. Furthermore, no study of aggression has examined genome-wide methylation beyond gene promoters, or studied DNA methylation in relation to continuous variation in aggressive behavior in a general adult population.

Here we describe the first epigenome-wide association study (EWAS) for aggressive behavior in whole blood performed in a population-based sample from the Netherlands Twin Register (NTR). The aim was to identify genomic locations where DNA methylation level is associated with aggressive behavior.

Methods

Subjects and Samples

The subjects in this EWAS participated in longitudinal survey studies from the NTR and participated in the NTR biobank project. Venous blood samples were drawn in the morning after an overnight fast, and multiple ethylenediaminetetraacetic acid (EDTA) tubes were collected for isolation of DNA and RNA and assessment of hematological profiles. Phenotyping through longitudinal surveys and blood sampling procedures have been described in detail previously (Boomsma et al., 2006; van Beijsterveldt et al., 2013; Willemsen et al., 2010, 2013).

In total, 3,264 blood samples from 3,221 NTR participants were assessed for genome-wide methylation, of which 3,089 samples from 3,057 subjects passed quality control (QC). For a small group of subjects, methylation was measured in two longitudinal blood samples. The genome-wide methylation dataset is described in Van Dongen et al. (under review). In the current EWAS, we included individuals for whom the following data were available: aggression score, good quality DNA methylation data, and data on white blood cell counts, leaving 2,059 samples from 2,029 subjects. This cohort included 1,247 monozygotic (MZ) twins (479 complete MZ pairs), 663 dizygotic (DZ) twins (199 complete DZ pairs), 48 fathers of twins, 56 mothers of twins, 13 siblings, and 2 spouses of twins. For 30 subjects, longitudinal methylation data (two samples) were included in the EWAS. A separate discordant twin analysis was

TABLE 1**ASR Survey Items Used to Score Aggressive Behavior, Based on the ASEBA Manual**

Item
ASR item 3: I argue a lot.
ASR item 5: I blame others for my problems.
ASR item 16: I am mean to others.
ASR item 28: I get along badly with my family.
ASR item 37: I get into many fights.
ASR item 55: My moods swing between elation and depression.
ASR item 57: I physically attack people.
ASR item 68: I scream or yell a lot.
ASR item 81: My behavior is very changeable.
ASR item 86: I am stubborn, sullen, or irritable.
ASR item 87: My moods or feelings change suddenly.
ASR item 95: I have a quick temper.
ASR item 97: I threaten to hurt people.
ASR item 116: I get upset too easily.
ASR item 118: I am too impatient.

Note: ASR = Adult Self-Report (Achenbach & Rescorla, 2003).

performed on data from a subset of 40 MZ twins (20 pairs) highly discordant for aggressive behavior.

All subjects provided written informed consent, and the study protocols were approved by the Central Ethics Committee on Research Involving Human Subjects of the VU University Medical Centre, Amsterdam, an Institutional Review Board certified by the US Office of Human Research Protections (IRB No. IRB-2991 under Federal-wide Assurance-3703; IRB/Institute codes, NTR 03-180).

Aggression

Data on aggression were collected in multiple NTR surveys. Here data from NTR Survey 8 (data collection in 2009) and Survey 10 (data collection in 2013) were analyzed. Aggressive behavior was rated with the ASEBA Adult Self-Report (ASR; Achenbach & Rescorla, 2003). The aggression score was computed following the ASR guidelines as the sum of 15 items (Table 1). For individuals who completed Surveys 8 and 10, the aggression score that was assessed closest to the moment of blood draw was selected. The aggression data were adjusted for sex and age at the moment of survey completion. Residuals were analyzed in the EWAS.

Aggression Discordance in MZ Twins

To identify MZ twins discordant for aggressive behavior, the within-pair differences in aggression scores at Surveys 8 and 10 were computed. In total, aggression scores were available from at least one survey for 461 MZ pairs (18 pairs were excluded, of which one twin completed only Survey 8 and the other twin completed only Survey 10). For 205 pairs, longitudinal aggression scores were available from two time points (Surveys 8 and 10) and for 256 pairs, one aggression score was available. Discordance was defined as a within-pair difference in aggression score ≥ 7 at the closest available survey to the moment of blood draw. This threshold is equivalent to >2 times the *SD* of aggressive behavior in the entire NTR cohort. For pairs with longitudinal aggression

data, the difference had to be at ≥ 5 (equivalent to >1.5 *SD*) at the other time point. This selection yielded 20 discordant MZ pairs, of which discordance was supported by longitudinal data for six pairs and by one time point for 14 pairs.

Infinium HumanMethylation450 BeadChip Data

DNA methylation was assessed with the Infinium HumanMethylation450 BeadChip Kit (Illumina, Inc.; Bibikova et al., 2011). Genomic DNA, 500 ng, from whole blood was bisulfite-treated using the ZymoResearch EZ DNA Methylation kit (Zymo Research Corp, Irvine, CA, USA) following the standard protocol for Illumina 450K micro-arrays, by the Department of Molecular Epidemiology from the Leiden University Medical Center (LUMC), the Netherlands. Subsequent steps (i.e., sample hybridization, staining, and scanning) were performed by the Erasmus Medical Center micro-array facility, Rotterdam, the Netherlands. QC and processing of the blood methylation dataset have been described in detail previously (Tobi et al., 2015; Van Dongen et al., under review). In short, a number of sample- and probe-level quality checks were performed. Sample-level QC was performed using MethylAid (van Iterson et al., 2014). Probes were set to be missing in a sample if they had an intensity value of exactly zero, or a detection $p > .01$, or a bead count of <3 . After these steps, probes that failed based on the above criteria in $>5\%$ of the samples were excluded from all samples (only probes with a success rate ≥ 0.95 were retained). Probes were also excluded from all samples if they mapped to multiple locations in the genome (Chen et al., 2013), or if they had a single nucleotide polymorphism (SNP) within the CpG site (at the C or G position) irrespective of minor allele frequency in the Dutch population (Genome of the Netherlands Consortium, 2014). Only autosomal methylation sites were analyzed in the EWAS. The methylation data were normalized with functional normalization (Fortin et al., 2014) and normalized intensity values were converted into beta (β) values. The β -value represents the methylation level at a site, ranging from 0 to 1, and is calculated as follows:

$$\beta = \frac{M}{M + U + 100}$$

where M = methylated signal, U = unmethylated signal, and 100 represents a correction term to control the β -value of probes with very low overall signal intensity.

Principal component analysis was conducted on genome-wide methylation sites (after QC and normalization), including those on sex chromosomes.

Covariates

White blood cell percentages were included as covariates in the EWAS to account for variation in cellular composition between whole blood samples. The following subtypes of white blood cells were counted in blood samples: neutrophils, lymphocytes, monocytes, eosinophils, and

basophils (Willemsen et al., 2013). Lymphocyte percentage was not included in models because it was strongly correlated with neutrophil percentage ($r = -0.93$), and basophil percentage was not included because it showed little variation between subjects, with a large number of subjects having 0% of basophils. Inspection of principal components (PCs) from the genome-wide methylation data indicated that PC1 reflected sex, PC2 correlated strongly with lymphocyte percentage ($r = -0.8$), and neutrophil percentage ($r = 0.79$). PC3 correlated moderately with age ($r = -0.41$) and weakly with measured white blood cell percentages (absolute $r \sim 0.1$), but we hypothesize that this PC may potentially be reflective of unmeasured white blood cell subtypes and included this PC in the model. PC4 and PC5 were included to account for technical variability between samples, as these PCs were correlated with several indices related to the lab procedure, such as sample plate and order of processing. PCs 3, 4, and 5 were not correlated with aggression (absolute $r < 0.08$).

Epigenome-Wide Association Analyses

To examine the association between aggressive behavior and DNA methylation level, we performed two analyses: (1) An EWAS on aggressive behavior in the entire NTR cohort ($N = 2,029$ individuals), and (2) an analysis in MZ twin pairs ($N = 20$ MZ pairs) highly discordant for aggression, where we tested for differences in methylation levels between high- and low-scoring twins.

In the entire NTR cohort, we tested for each methylation site whether DNA methylation level was associated with aggressive behavior using generalized estimation equation (GEE) models, with DNA methylation β -value as outcome and the following predictors: aggression residual score, sex, age at blood sampling, monocyte, eosinophil, and neutrophil percentage values, HM450k array row, and PCs 3, 4, and 5. GEE models were fitted with the R package GEE, with the following specifications: Gaussian link function (for continuous data), 100 iterations, and the ‘exchangeable’ option to account for the correlation structure within families and within persons.

In the discordant MZ twin analysis, paired t -tests were employed to test for methylation differences between the twin scoring higher on aggressive behavior and the twin scoring lower on aggressive behavior. Paired t -tests were performed on residual methylation levels, which were obtained by adjusting the DNA methylation β -values for the same covariates as used in the EWAS on all subjects (sex, age at blood sampling, monocyte, eosinophil, and neutrophil percentage values, HM450k array row, and PCs 3, 4, and 5).

False discovery rate (FDR) q -values were computed with the R package q -value with default settings. The genomic inflation factor (λ) was calculated with the default regression method from the R package GenABEL. In all analyses, an FDR q -value < 0.05 was considered statistically significant. Follow-up analyses, including a test for enrichment

of genomic locations and gene ontologies (GOs), were performed based on the output from the EWAS on the entire NTR cohort.

Annotation

Methylation sites were mapped to genomic features and DNase I hypersensitive sites (DHS) as described by Sliker et al. (2013). Genomic features included five gene-centric regions: intergenic region (> 10 kb from the nearest transcription start site [TSS]), distal promoter (-10 to -1.5 kb from the nearest TSS), proximal promoter (-1.5 kb to $+500$ bp from the nearest TSS), gene body ($+500$ bp to $3'$ end of the gene), and downstream region ($3'$ end to $+5$ kb from $3'$ end). Next, CpGs were mapped to CpG islands (CGIs) (CG content $> 50\%$, length > 200 bp, and observed/expected ratio of CpGs > 0.6 ; locations were obtained from the UCSC genome browser [44]), CGI shore (2-kb region flanking CGI), CGI shelf (2-kb region flanking CGI shore), or non-CGI regions. The locations of DHS, mapped by the ENCODE project (ENCODE Project Consortium, 2012), were downloaded from the UCSC genome browser (Kent et al., 2002). Finally, we annotated methylation sites to genes of which the methylation level was previously reported to be associated with aggression in humans: *MAOA* (Checknita et al., 2015), *SLC6A4* (Wang et al., 2012), cytokine genes and their transcription factors (*IL6*, *IL1A*, *IL8*, *IL4*, *IL10*, *NFKB1*, *NFAT5*, *STAT6*; Provencal et al., 2013), and all genes significant in men and women from Guillemin et al. (2014). For each annotation category, a variable with two levels (0 and 1) was created to indicate whether a methylation site mapped to the category (1) or not (0).

Enrichment of Genomic Locations

We developed a custom enrichment analysis method to examine whether certain genomic locations on average showed a stronger association between DNA methylation and aggressive behavior, based on the EWAS test statistics from all genome-wide methylation sites. The following nine categories were tested for being enriched among sites more strongly associated with aggression: gene body, proximal promoter, distal promoter, downstream region, CGI, CGI shore, CGI shelf, DHS, and loci previously associated with aggression (all annotation categories described above, minus two to allow for identification of the model):

$$\left(\frac{\beta_{EWAS}}{se_{EWAS}} \right)_{CpGi}^2 = \text{Intercept} + \beta_{\text{category } 1} X \text{Category } 1_{CpGi} \dots + \beta_{\text{category } x} X \text{Category } x_{CpGi}, \quad (1)$$

where $\left(\frac{\beta_{EWAS}}{se_{EWAS}} \right)_{CpGi}^2$ represents the squared standardized effect size from the EWAS for methylation site i ; β_{EWAS} is the estimate for the association between aggression and methylation level at site i , and se_{EWAS} is its standard error. $\beta_{\text{category } x}$ represents the estimate for category x , that is, the change in

the EWAS test statistic associated with a one unit change in category x (e.g., being located within a proximal promoter area), and Category x_{CpGi} is a dichotomous variable indicating whether methylation site i maps to category x . To account for differences in variability between methylation sites, we also included the mean and *SD* of DNA methylation level in the model as covariates. As an additional check, we repeated the analysis restricting to the most variable methylation sites only (*SD* of the methylation proportion >0.03), which resulted in the same conclusions (results not shown).

The regression is weighted by the reciprocal of the fitted values of the variance function (i.e., the predicted values of regressing the squared residuals of regression Equation 1 on the predictors in Equation 1). To account for the fact that the errors in this regression are not normally distributed, Jackknife standard errors were computed. Since the results based on Jackknife standard errors (see Supplementary material) led to the same conclusions as the normal linear regression standard errors, only the standard linear regression estimates are presented in the main text.

In contrast to commonly used enrichment strategies (e.g., chi-square and Fisher's exact test), in which it is tested whether certain genomic annotation groups are enriched within a selected group of methylation sites (e.g., sites significantly associated with a phenotype), our approach uses information from all genome-wide methylation sites to test whether there is a stronger overall signal in certain genomic annotation groups (thus including information from all methylation sites, rather than limiting the analysis to sites that are significantly associated with the phenotype).

Enrichment of Gene Ontology Terms

To test for enrichment of GO terms among methylation sites with a stronger association with aggression, all methylation sites tested were ranked by EWAS p -value, and the resulting ranked gene list (keeping the highest rank for each gene for which multiple CpG sites were measured) was supplied to the online software tool GOrilla (Eden et al., 2009). This tool performs GO enrichment analysis based on gene rank, thus not requiring a p -value cut-off for defining the input list of genes. The background in this analysis consists of all genes for which methylation sites were analyzed in the EWAS. In all analyses, we accounted for multiple testing by controlling FDR. An FDR q -value < 0.05 was considered statistically significant.

Results

Characteristics of NTR Cohort

Data on aggressive behavior scored with the ASR were available for 14,173 individuals who filled out Survey 8 and for 15,535 individuals who filled out Survey 10. The EWAS was performed on 2,059 blood samples from 2,029 subjects (mean age at blood sampling = 36.4 years, *SD* = 12.4,

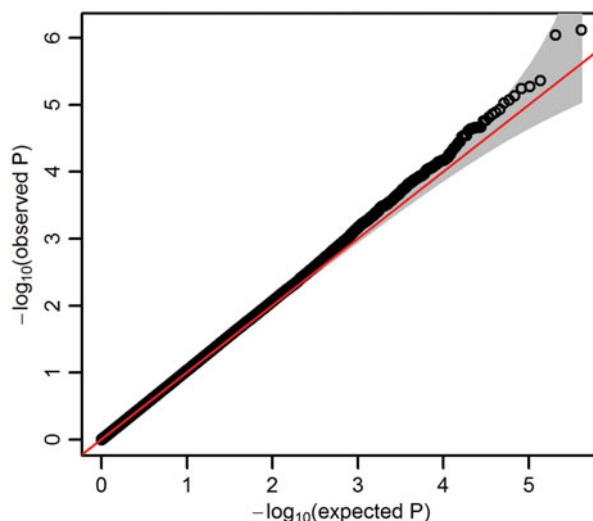


FIGURE 1

(Colour online) Quantile-quantile (QQ) plot from the EWAS of aggressive behavior in the entire NTR cohort. The observed p -values (y-axis) are plotted against the p -values expected under the null hypothesis (x-axis). The straight diagonal line denotes the pattern expected under the null hypothesis, with 95% confidence intervals indicated by the shaded grey area.

females = 69.2%), for whom the aggression score closest to the moment of blood draw (from which DNA was extracted) was selected: For 1,906 samples, aggression was assessed after blood draw (range: 0–10 years, mean value = 3.6) and for 153 samples, aggression was assessed before blood draw (range: 0–8 years, mean value = 2.1). The characteristics of the aggression data and EWAS cohort are described in Table 2. The average aggression score of the EWAS cohort was similar to the averages in the entire NTR cohort (mean EWAS cohort = 2.8, mean NTR cohort in survey 8 = 2.9, mean NTR cohort in survey 10 = 2.8). On average, women scored 0.15 *SD* higher on aggressive behavior compared with men ($p_{\text{survey 8}} = 9.6 \times 10^{-19}$; $p_{\text{survey 10}} = 2.3 \times 10^{-20}$) and with each one-year increase in age, aggression scores decreased on average 0.01 *SD* (Supplementary Table S1, $p_{\text{survey 8}} = 4.2 \times 10^{-81}$ and $p_{\text{survey 10}} = 2.4 \times 10^{-72}$). The EWAS was performed on the residuals derived after correcting aggression scores for sex and age at completion of survey (for a distribution of aggression data, see Supplementary Figure S1).

EWAS in the Entire NTR Cohort

For 411,169 autosomal sites in the genome, we tested the association between DNA methylation and (age- and sex-adjusted) aggression score, while correcting for white blood cell counts, age at blood sampling, sex, array row, and three PCs from the methylation data. The quantile–quantile (QQ) plot is shown in Figure 1; λ was 1.047.

None of the p -values was below the genome-wide significance threshold of $p < 1 \times 10^{-7}$ (Bonferroni) or

TABLE 2
Characteristics of the NTR Aggression Data and the EWAS Group

Data/analysis group	ASR aggression score						Age at survey			Age at blood sampling			Sex
	N	Mean	Median	SD	Min	Max	Mean	Median	SD	Mean	Median	SD	Female (%)
Survey 8	14,173	2.9	2	3.2	0	26	40.6	40.0	15.4				64.3
Survey 10	15,535	2.8	2	3.1	0	24	39.7	41.0	14.7				64.0
EWAS ^a	2,059	2.8	2	3.1	0	23	38.9	36.0	12.4	36.4	33	12.0	69.2
Discordant twins	High aggression	20	11.5	1	3.9	7	36.7	35.5	13.5	34.5	33	12.0	85.0
	Low aggression	20	2.0	1	2.4	0	10	36.6	35.5	13.5	34.5	33	12.0

Note: ^aIncludes all individuals with 450k methylation data and data on white blood cell counts who were included in the EWAS of the entire NTR cohort. The descriptives are given for samples. For each blood sample, the aggression data closest to blood sampling was selected. ASR = Adult Self-Report. EWAS = Epigenome-Wide Association Study.

TABLE 3
Top-Ranking CpG Sites From the EWAS of Aggression in the Entire NTR Cohort

CpG site	Chr	Position (hg19)	Nearest gene name	Mean methylation level ^a	SD methylation level ^a	Estimate ^b	Robust SE ^c	p-value	FDR q-value
cg01792876	8	116684801	TRPS1	0.91	0.013	0.00038	0.00008	7.6×10^{-7}	0.18
cg06092953	18	77905699	PARD6G-AS1	0.35	0.096	0.00299	0.00061	9.0×10^{-7}	0.18
cg04168050	1	220263278	RNU5F-1	0.11	0.018	-0.00053	0.00011	4.3×10^{-6}	0.46
cg18089145	14	39572493	SEC23A	0.11	0.013	-0.00028	0.00006	5.3×10^{-6}	0.46
cg13535736	9	111863775	C9orf5	0.60	0.058	0.00144	0.00032	5.7×10^{-6}	0.46
cg19946699	21	40817645	C21orf13	0.21	0.048	0.00138	0.00031	7.3×10^{-6}	0.46
cg04690395	2	19558507	OSR1	0.12	0.015	-0.00035	0.00008	8.5×10^{-6}	0.46
cg05874450	6	105627246	POPDC3	0.12	0.013	-0.00033	0.00008	9.3×10^{-6}	0.46

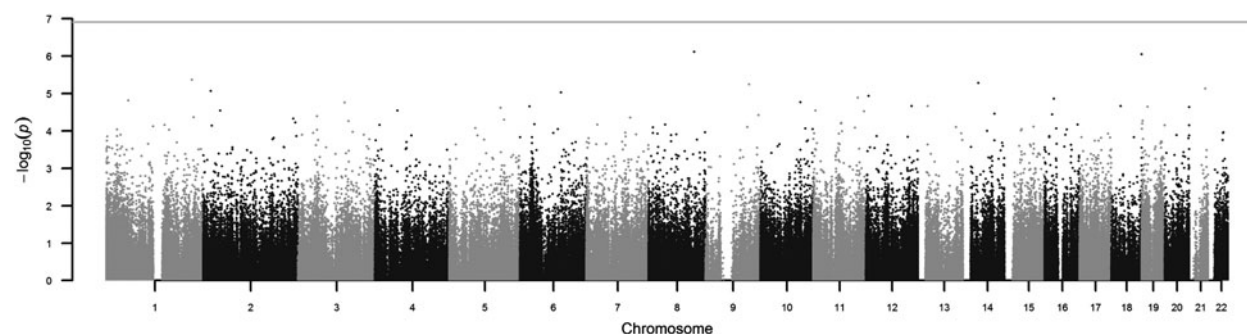
Note: Top hits from the EWAS of aggression (p -value $< 1 \times 10^{-5}$ for the association between methylation level and aggression).

^aMean and SD of the methylation proportion (β -value) in the entire NTR cohort.

^bEstimate from the regression of methylation proportion on aggression score.

^cRobust standard error of the estimate (accounting for the clustering of observations within families).

FDR = False Discovery Rate.

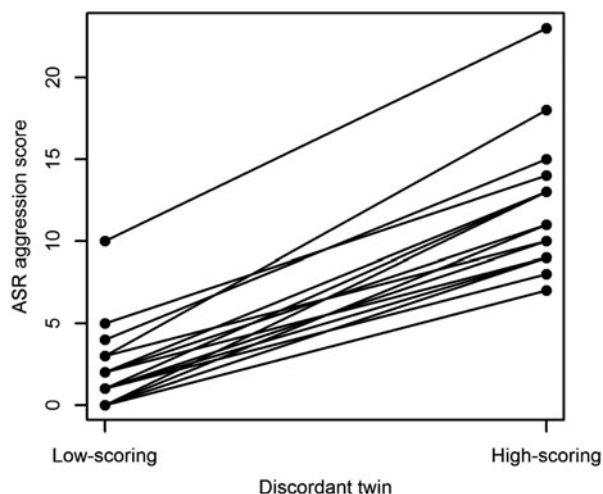
**FIGURE 2**
Manhattan plot showing p-values for the association between aggression score and DNA methylation level in the entire NTR cohort at genome-wide autosomal sites. The horizontal gray line represents the Bonferroni-adjusted p-value threshold.

FDR q-value < 0.05 (Figure 2). The two highest ranking methylation sites, cg01792876 ($p = 7.6 \times 10^{-7}$, FDR = 0.18) and cg06092953 ($p = 9.0 \times 10^{-7}$, FDR = 0.18), located on chromosomes 8 and 18, respectively, showed a positive relationship between DNA methylation level and aggressive behavior (Supplementary Figure S2). At these CpG sites, an increase of one SD in aggressive behavior (3.1 points) was associated with a methylation change of 0.1% (cg01792876) and 0.9% (cg06092953), respectively, corresponding to 9.1% of the SD of methylation level at cg01792876 and 9.7% of the SD of methylation level at

cg06092953. Characteristics of all methylation sites with an EWAS p -value $< 1 \times 10^{-5}$ are provided in Table 3.

Discordant MZ Twin Analysis

We next performed an alternative EWAS approach, in which we tested for DNA methylation differences at 411,169 sites within 20 MZ pairs discordant for aggressive behavior (Table 2, mean age at blood draw = 34.5 years, SD = 12.0, N [females] = 17). At the survey closest to blood draw, aggressive behavior was on average 9.5 points higher (range: 7–15)

**FIGURE 3**

Aggression scores of discordant MZ twins. The ASR aggression score at the survey closest to the moment of blood draw is plotted for low-scoring and high-scoring twins of 20 discordant MZ twin pairs. The scores of co-twins are connected by lines. ASR = Adult Self-Report (Achenbach & Rescorla, 2003).

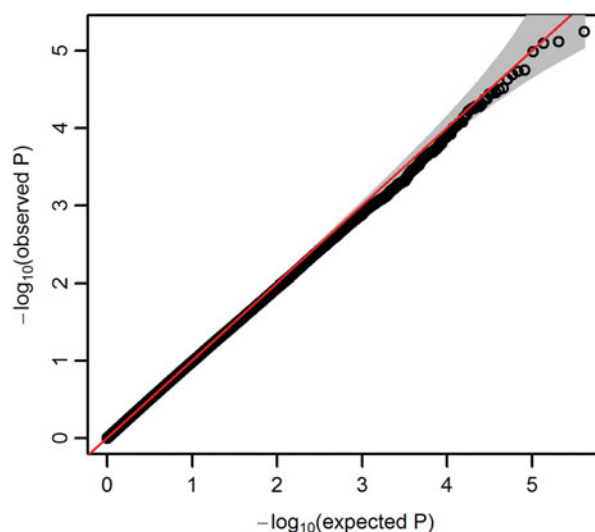
in the high-scoring twins compared with the low-scoring twins (Figure 3).

The within-pair difference in DNA methylation across the genome is plotted for each aggression-discordant pair in Supplementary Figure S3. Screening for large DNA methylation differences, we observed that each aggression-discordant MZ pair had on average 24 methylation sites (range: 0–311), where the within-pair methylation difference was larger than 30% (methylation β -value difference > 0.30), and 291 sites with a difference larger than 15% (range: 39–793; highlighted in Supplementary Figure S3); however, these highly discordant methylation sites were generally twin pair-specific, as has been previously observed in autism-discordant MZ twins (Wong et al., 2014).

We tested for consistent differences in DNA methylation levels between aggression-discordant MZ twins by applying paired t -tests to genome-wide autosomal methylation sites. The QQ plot from this analysis is shown in Figure 4 ($\lambda = 0.98$). No genome-wide significant methylation differences were identified between discordant twins (Figure 5). Characteristics of the top methylation sites from the discordant twin analysis (p -values < 1×10^{-5}) are provided in Table 4, and their methylation levels in the high- and low-scoring twins from discordant pairs are plotted in Figure 6.

Enrichment of Signal in Proximal Promoter Regions and DHS

We performed further follow-up analyses based on the results from the EWAS in the entire NTR cohort. Regression

**FIGURE 4**

(Colour online) Quantile-Quantile (QQ) plot from the comparison of aggression-discordant MZ twins. The observed p -values (y-axis) are plotted against the p -values expected under the null hypothesis (x-axis). The straight diagonal line denotes the pattern expected under the null hypothesis, with 95% confidence intervals indicated by the shaded gray area.

of the EWAS test statistics on annotation categories across all genome-wide sites (Table 5 and Supplementary Table S2) revealed a small but significant enrichment of signal in proximal promoter regions ($p = 1.5 \times 10^{-7}$) and DNase hypersensitivity sites (DHS, $p = 1.0 \times 10^{-5}$). Based on their test-statistic for the association with aggressive behavior, methylation sites in proximal promoters and DHS on average ranked higher than expected (median rank proximal promoter = 203,439, median rank DHS = 203,382, expected median rank under the null hypothesis = 205,585). Sites with a lower mean methylation level also on average showed stronger evidence for being associated with aggressive behavior ($p = 9.0 \times 10^{-5}$). These findings indicate that methylation sites associated with aggressive behavior are enriched in promoter areas and other regions of regulatory active DNA.

Gene Ontology Analysis

Gene ontology enrichment analysis based on EWAS p -value rank (from the analysis in the entire NTR cohort) identified a large number of GO terms that were significantly enriched among higher ranking methylation sites (Supplementary Table S3). The strongest enriched GO term was single-organism developmental process (GO:0044767, $p = 4.5 \times 10^{-23}$), and many other GO terms related to development and metabolism were enriched. The list of significantly enriched GO processes also included many brain or central nervous system processes, such as

TABLE 4
Top-Ranking CpG Sites From the Comparison of Aggression-Discordant MZ Twins

CpG site	Chr	Position (hg19)	Nearest gene name	Mean methylation level ^a	SD methylation level ^a	Estimate ^b	p-value	FDR q-value
cg21557159	11	107795699	RAB39	0.89	0.02	-0.0144	5.7×10^{-6}	0.99
cg08648367	19	51925472	SIGLEC10	0.04	0.01	0.0078	7.6×10^{-6}	0.99
cg14212412	6	105918992	PREP	0.91	0.02	0.0184	8.0×10^{-6}	0.99

Note: Top hits from the comparison of aggression-discordant MZ twins (p -value $< 1 \times 10^{-5}$ from a paired t-test comparing methylation level in the higher-scoring versus the lower-scoring twins).

^aMean and SD of the methylation proportion (β -value) in the entire NTR cohort.

^bMean difference in residual methylation level between aggression-discordant twins (aggression high-scoring twin minus aggression low-scoring twin). FDR = False Discovery Rate.

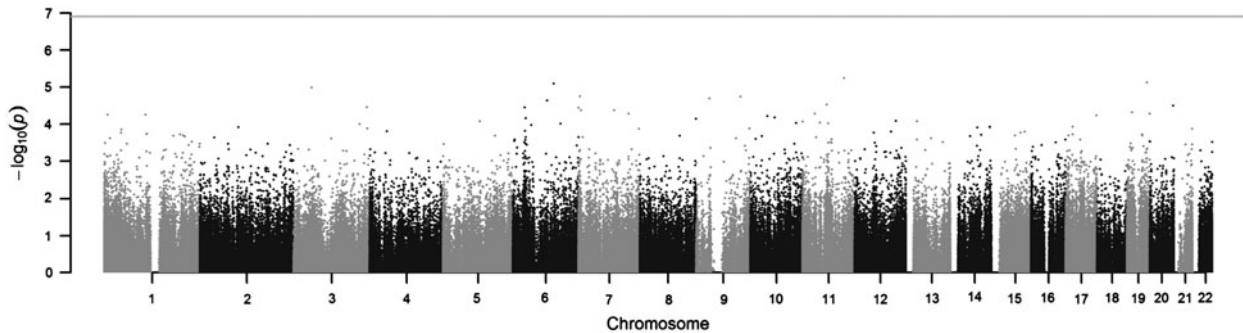


FIGURE 5

Manhattan plot showing the p -values from paired t -tests comparing DNA methylation level at genome-wide autosomal sites between aggression-discordant twins. The horizontal gray line represents the Bonferroni-adjusted p -value threshold.

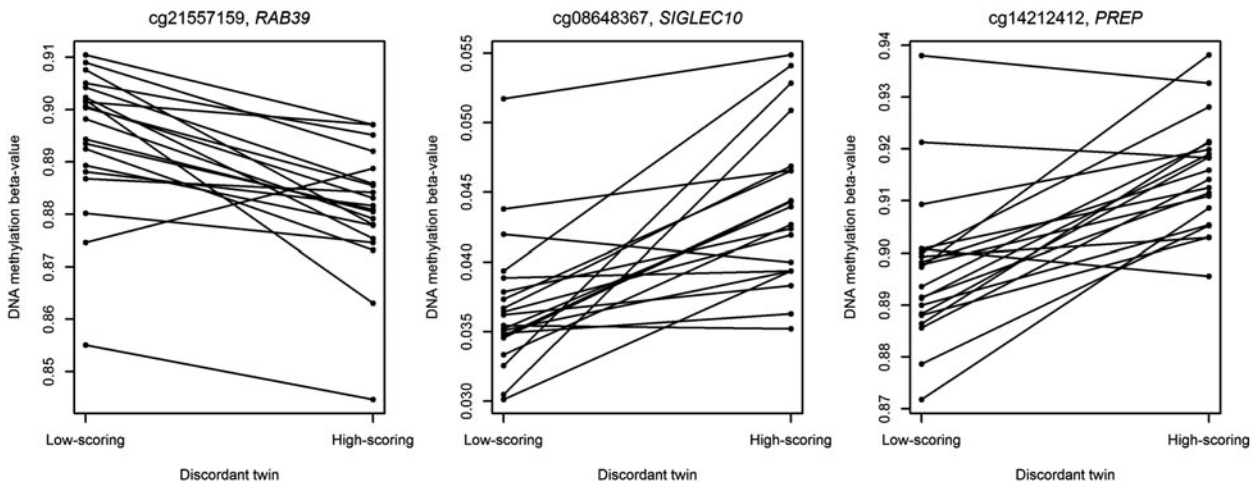


FIGURE 6

Methylation levels at the three most significant differentially methylated sites in aggression-discordant MZ twins. The residual DNA methylation β -values (adjusted for covariates) are plotted for low- and high-scoring twins of 20 discordant MZ twin pairs. The DNA methylation levels of co-twins are connected by lines.

regulation of neurogenesis (GO:0050767, $p = 3.9 \times 10^{-13}$), axon guidance (GO:0007411, $p = 8.0 \times 10^{-10}$), excitatory post-synaptic potential (GO:0060079, $p = 6.9 \times 10^{-6}$), cognition (GO:0050890, $p = 6.0 \times 10^{-4}$), and behavior (GO:0007610, $p = 6.0 \times 10^{-4}$); for a complete list of significant GO terms, see Supplementary Table S3. In line with the enrichment of genes connected to brain

processes, significantly enriched GO components included many neuron-specific components; for example, axon part (GO:0033267, $p = 2.3 \times 10^{-6}$), dendrite (GO:0030425, $p = 1.6 \times 10^{-5}$), post-synaptic density (GO:0014069, $p = 9.5 \times 10^{-5}$), and pre-synaptic active zone (GO:0048786, $p = 5.5 \times 10^{-4}$). The most strongly enriched GO component was cell junction (GO:0030054, $p = 7.4 \times 10^{-13}$).

TABLE 5
Results From the Regression of EWAS Test Statistics on Genomic Annotation Categories

Regression parameter	Estimate	Standard error	t-value	p-value
Intercept	1.0207	0.0110	92.76	0
Mean methylation level ^a	-0.0373	0.0095	-3.92	8.96×10^{-5}
SD methylation level ^b	0.5456	0.1457	3.74	1.81×10^{-4}
Downstream region	0.0095	0.0175	0.54	0.59
Gene body	0.0114	0.0070	1.63	0.10
Proximal promoter	0.0401	0.0076	5.25	1.49×10^{-7}
Distal promoter	0.0068	0.0124	0.55	0.58
CpG island (CGI)	0.0046	0.0071	0.66	0.51
CGI shelf	-0.0026	0.0087	-0.30	0.77
CGI shore	-0.0047	0.0065	-0.72	0.47
DNase I hypersensitive site (DHS)	0.0239	0.0054	4.41	1.01×10^{-5}
Aggression loci ^c	0.0366	0.0594	0.62	0.54

Note: ^aMean methylation level of the site.

^bStandard deviation of methylation level.

^cLoci where DNA methylation level was previously reported to be associated with aggression. Results based on Jackknife are presented in Supplementary Table S2.

Discussion

We performed an EWAS to identify genomic locations where DNA methylation level is associated with aggressive behavior in a population-based cohort of adults. No genome-wide significant methylation hits were identified after correction for multiple testing. GO analysis, in which categories of genes rather than single methylation sites are tested, highlighted that genes involved in developmental and central nervous system processes are enriched among higher-ranking genes from our EWAS. Here we describe our two top-ranked CpGs based on the entire NTR cohort, which have an FDR of 0.18. Our top CpG, cg01792876 (genomic location: chr8; 116,684,801) is located near the trichorhinophalangeal syndrome I (*TRPS1*) gene. This gene encodes a zinc finger transcription factor that represses GATA-regulated genes, and mutations in this gene are linked to tricho-rhino-phalangeal syndromes I–III, which are characterized by distinctive craniofacial and skeletal malformations (ONIM 190350, ONIM 150230, and ONIM 190351). Interestingly, the region also harbors a suggestive SNP association for major depressive disorder (rs2721937, chr8: 116,702,874) based on the genome-wide association study meta-analysis from the PGC consortium published in 2013 (Ripke et al., 2013). *TRPS1* is also linked to a whole range of other complex traits based on genome-wide association studies (UCSC genome browser, accessed on August 17, 2015). ENCODE transcription factor binding site (TFBS) ChIP-seq data (Mar 2012 Freeze, accessed through the UCSC genome browser) indicate that our top CpG, cg01792876, is located within the binding sites of transcription factors FOXA1, GATA3, and EP300.

The second-ranked CpG site from our EWAS (cg06092953, genomic location: chr18; 77,905,699) is located between the non-coding RNA PARD6G antisense RNA 1 (*PARD6G-AS1*) and the activity-dependent neuroprotective protein 2 gene (*ADNP2*). *ADNP2* is ubiquitously expressed in mice tissues, but especially highly expressed in the brain (particularly in the cerebral cortex),

and is thought to play a role in cellular survival pathways (Kushnir et al., 2008). *ADNP2* expression has been linked to schizophrenia (Dresner et al., 2011) and to the action of the mood-stabilizing drug lithium that is prescribed in bipolar disorder (McEachin et al., 2010). Our second-ranked site cg06092953 overlaps with the binding sites of a number of transcription factors (ENCODE TFBS ChIP-seq data Mar 2012 Freeze).

In the comparison of DNA methylation levels within aggression-discordant MZ twin pairs, we also found no genome-wide significant methylation differences, but high-ranking sites may nevertheless be of potential interest. Here we describe the sites with a *p*-value $< 10^{-5}$ (FDR = 0.99, $N = 3$). The effect sizes of these sites range from an average within-pair difference in DNA methylation percentage of 0.8% to 1.8% (Table 4). The top site (cg21557159, chr 11: 107,795,699, $p = 5.7 \times 10^{-6}$) is located upstream of the RAS oncogene family 39 gene (*RAB39/RAB39A*). *RAB39* is involved in cellular endocytosis, and is considered to be ubiquitously expressed in human tissues (Chen et al., 2003). In a neuronal cell line from mice, this gene was shown to play a role in cell differentiation, with knock-down of *Rab39A* resulting in a decrease in the number of neurites formed (Mori et al., 2013). The second-ranked CpG site, cg08648367 (chr 19; 51,925,472, $p = 7.6 \times 10^{-6}$) overlaps the binding sites of a number of transcription factors near the sialic acid-binding Ig-like lectin 10 gene (*SIGLEC10*). The encoded Siglec-10 protein is involved in regulating immune responses induced by tissue damage (Chen et al., 2009). The third-ranked CpG site, cg14212412 (chr6; 105,918,992, $p = 8.0 \times 10^{-6}$) is located near the prolyl endopeptidase (*PREP*) gene. Prolyl endopeptidase is a proteinase that cleaves small neuropeptides and peptide hormones, such as angiotensin, thyrotropin-releasing hormone, gonadotropin-releasing hormone, neurotensin, vasopressin, and oxytocin (Garcia-Horsman et al., 2007). A recent study reported significant associations between serum prolyl endopeptidase (protein) level and problem

behavior scored by the Child Behavior Check List (CBCL) delinquent, aggressive, externalizing, and internalizing behavior subscales (Frenssen et al., 2015). Abnormal *PREP* activity has been also linked to several neurological and psychiatric conditions, including depression, mania, and schizophrenia (Maes et al., 1995), autism spectrum disorders (Momeni et al., 2005), neurodegenerative diseases such as Alzheimer's disease (Mantle et al., 1996), and to the aggregation of α -synuclein in Parkinson's disease (Myohanen et al., 2012).

A cross-comparison of results from different analyses showed that the two top hits based on the entire study showed the same direction of effect in discordant MZ pairs: cg01792876, near *TRPS1* showed a mean difference of 0.4% between aggression-discordant MZ twins ($p = .09$) and cg06092953 (nearest gene *PARD6G-AS1*), a mean difference of 1.9% ($p = .24$). The other way around, two of the three most significant sites in discordant MZ pairs showed the same direction of effect in the entire NTR cohort: cg08648367, near *SIGLEC10*, $\beta_{EWAS} = 2.6 \times 10^{-5}$, $p = .59$, and cg14212412, near *PREP*, $\beta_{EWAS} = 3.3 \times 10^{-4}$, $p = 3.1 \times 10^{-3}$. Finally, we also looked up the heritability of the top methylation sites in our dataset to gain insight into the extent to which they are influenced by genetic and environmental influences (van Dongen et al., under review). Each of the three most significant sites from the comparison of discordant MZ twins had a low heritability: cg21557159, heritability = 10%; cg08648367, heritability = 7%; and cg14212412, heritability = 0%; indicating that most of the variation at these methylation sites is caused by individual-specific environmental and/or stochastic effects. The top hit based on the overall NTR cohort, cg01792876, also had a heritability of 0% in our data, while the second hit cg06092953 was strongly influenced by genetic effects, with an MZ twin correlation of 0.67, and a DZ twin correlation of 0.15.

We limited this study to aggressive behavior only and did not consider other aspects of behavior or psychiatric symptoms. Since aggressive behavior is connected to a range of comorbidities in psychiatric and neurodegenerative disease domains, DNA methylation levels associated with aggressive behavior may to some extent be related to other comorbid conditions of aggression (e.g., depression and substance abuse). Importantly, aggressive behavior as scored on a continuous scale by the ASR by definition overlaps to a certain extent with other psychiatric dimensions. For instance, the ASR aggression scale includes items that score mood-related symptoms (e.g., 'My moods swing between elation and depression'), which are considered to be part of the aggression phenotype in adults. Yet, for this reason, it is possible that our EWAS results to some extent reflect methylation sites related to mood. To distinguish between methylation sites primarily associated with aggression and methylation sites related to other aspects of psychiatric symptomatology, the ideal psychiatric EWAS approach would involve an elabo-

rate analysis of multiple symptom dimensions of aggressive behavior and comorbidities (e.g., substance abuse, depression). Such an approach would give insight into which methylation sites are common to multiple psychiatric symptoms, which are specific to certain symptoms, and which are mainly reflective of effects of comorbid exposures.

While we did not identify genome-wide significant methylation hits for aggressive behavior in whole blood, a previous much smaller study reported numerous methylation hits for aggressive behavior based on a study of individual cell populations (monocytes and T-cells) extracted from blood samples of adults who had shown a trajectory of chronic physical aggression in childhood (Guillemin et al., 2014; Provencal et al., 2013). Three studies of candidate genes, including *MAOA* (Checknita et al., 2015), *SLC6A4* (Wang et al., 2012), and genes encoding cytokines and their transcription factors (Provencal et al., 2013), also reported associations between DNA methylation level and aggression or antisocial personality disorder. We had data on 694 CpGs in or near 35 genes identified in these previous studies (Checknita et al., 2015; Guillemin et al., 2014; Provencal et al., 2013; Wang et al., 2012). After correcting for the number of loci (35), one site was significantly replicated in our EWAS (cg14260773, chr2: 33,226,136, in *LTBP1*, $p = 8.01 \times 10^{-4}$) and eight additional sites, in or near the genes *PCDH8*, *UCK2*, *IL4*, *TGIF1*, *PCDHB2*, *TGIF1*, *AGTR1*, and *NDUFS2*, had a nominal p -value in our EWAS ($p < .01$, Supplementary Table S4). Possible explanations for differences between our findings and those from previous studies may include differences in the technique used to measure methylation, differences in the examined cell populations, but perhaps most importantly, differences in the study populations. While we examined the relationship between DNA methylation and a continuously scored measure of aggression in an unselected population-based adult cohort of adults, the previous studies focused on much more extreme cohorts. The *MAOA* gene was studied in a prisoner population (Checknita et al., 2015) and the other previous studies included individuals who had shown a chronic physical aggression trajectory in childhood, while the relationship between DNA methylation and aggression in adulthood was not examined (Guillemin et al., 2014; Provencal et al., 2013, 2014; Wang et al., 2012). Aggressive behavior scored with the ASR in the NTR cohort is highly left-skewed, with very few people scoring high on aggression, and very few people responding positively to the physical aggression items (e.g., 'I physically attack people'). Future EWASs on aggression in the general population may benefit from focusing on childhood or adolescent ages, where variation in aggressive behavior is more pronounced. Moreover, longitudinal studies with repeated measures of DNA methylation and aggression, analysis strategies such as Mendelian randomization, and analyses of DNA methylation in relevant brain regions (taking into account genetic variation) are required to unravel whether aggression-associated

methylation changes in a peripheral tissue such as whole blood are related to the cause or consequence of aggression (Mill & Heijmans, 2013). Once genome-wide association studies of aggressive behavior have identified genetic risk variants for aggression, future studies integrating genetic information (e.g., SNP data), and epigenetic information (e.g., methylation data) are warranted to unravel how the interplay of genetic risk variants and epigenetic mechanisms gives rise to aggressive behavior.

Acknowledgments

We would like to thank the twins and their family members who participated in the NTR studies. We thank Hill Fung Ip for writing the Jackknife software that we used to test for the enrichment of genomic annotations. This work was supported by ACTION. ACTION receives funding from the European Union Seventh Framework Program (FP7/2007–2013) under grant agreement No. 602768. This study was also supported by the BBRMI-NL-financed BIOS Consortium (NWO 184.021.007) and by the European Research Council (ERC): ERC-230374: Genetics of mental illness, a lifespan approach to the genetics of childhood and adult neuropsychiatric disorders and comorbid conditions, and ERC-284167: Beyond the genetics of addiction. MV is supported by Royal Netherlands Academy of Science Professor Award (PAH/6635) to DIB.

Supplementary Material

To view supplementary material for this article, please visit <http://dx.doi.org/10.1017/thg.2015.74>.

References

- Achenbach, T. M., & Rescorla, L. A. (2003). *Manual for the ASEBA adult forms and profiles*. Burlington, VT: Research Center for Children, Youth, & Families, University of Vermont.
- Alink, L. R., Mesman, J., van Zeijl, J., Stolk, M. N., Juffer, F., Koot, H. M., Bakermans-Kranenburg, M. J., & van Ijzendoorn, M. H. (2006). The early childhood aggression curve: Development of physical aggression in 10- to 50-month-old children. *Child Development, 77*, 954–66.
- Bakulski, K. M., Lee, H., Feinberg, J. I., Wells, E. M., Brown, S., Herbstman, J. B., . . . Fallin, M. D. (2015). Prenatal mercury concentration is associated with changes in DNA methylation at TCEANC2 in newborns. *International Journal of Epidemiology, 44*, 1249–62, doi: 10.1093/ije/dyv032.
- Bibikova, M., Barnes, B., Tsan, C., Ho, V., Klotzle, B., Le, J. M., . . . Shen, R. (2011). High density DNA methylation array with single CpG site resolution. *Genomics, 98*, 288–295.
- Bird, A. (2007). Perceptions of epigenetics. *Nature, 447*, 396–398.
- Boles, S. M., & Miotto, K. (2003). Substance abuse and violence: A review of the literature. *Aggression and Violent Behavior, 8*, 155–174.
- Boomsma, D. I., de Geus, E. J., Vink, J. M., Stubbe, J. H., Distel, M. A., Hottenga, J. J., . . . Willemsen, G. (2006). Netherlands twin register: From twins to twin families. *Twin Research and Human Genetics, 9*, 849–857.
- Breton, C. V., Byun, H. M., Wenten, M., Pan, F., Yang, A., & Gilliland, F. D. (2009). Prenatal tobacco smoke exposure affects global and gene-specific DNA methylation. *American Journal of Respiratory and Critical Care Medicine, 180*, 462–467.
- Checknita, D., Maussion, G., Labonte, B., Comai, S., Tremblay, R. E., Vitaro, F., . . . Turecki, G. (2015). Monoamine oxidase a gene promoter methylation and transcriptional downregulation in an offender population with antisocial personality disorder. *British Journal of Psychiatry, 206*, 216–222.
- Chen, G. Y., Tang, J., Zheng, P., & Liu, Y. (2009). CD24 and Siglec-10 selectively repress tissue damage-induced immune responses. *Science, 323*, 1722–1725.
- Chen, T., Han, Y., Yang, M., Zhang, W., Li, N., Wan, T., . . . Cao, X. (2003). Rab39, a novel Golgi-associated Rab GTPase from human dendritic cells involved in cellular endocytosis. *Biochemical and Biophysical Research Communications, 303*, 1114–1120.
- Chen, Y. A., Lemire, M., Choufani, S., Butcher, D. T., Grafodatskaya, D., Zanke, B. W., . . . Weksberg, R. (2013). Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray. *Epigenetics, 8*, 203–209.
- Coccaro, E. F. (2012). Intermittent explosive disorder as a disorder of impulsive aggression for DSM-5. *American Journal of Psychiatry, 169*, 577–588.
- Coie, J. D., Lochman, J. E., Terry, R., & Hyman, C. (1992). Predicting early adolescent disorder from childhood aggression and peer rejection. *Journal of Consulting and Clinical Psychology, 60*, 783–792.
- Cote, S. M., Boivin, M., Nagin, D. S., Japel, C., Xu, Q., Zoccolillo, M., . . . Tremblay, R. E. (2007). The role of maternal education and nonmaternal care services in the prevention of children's physical aggression problems. *Archives of General Psychiatry, 64*, 1305–1312.
- Cote, S. M., Vaillancourt, T., LeBlanc, J. C., Nagin, D. S., & Tremblay, R. E. (2006). The development of physical aggression from toddlerhood to pre-adolescence: A nationwide longitudinal study of Canadian children. *Journal of Abnormal Child Psychology, 34*, 71–85.
- Dresner, E., Agam, G., & Gozes, I. (2011). Activity-dependent neuroprotective protein (ADNP) expression level is correlated with the expression of the sister protein ADNP2: Deregulation in schizophrenia. *European Neuropsychopharmacology, 21*, 355–361.
- Eden, E., Navon, R., Steinfeld, I., Lipson, D., & Yakhini, Z. (2009). GOrilla: A tool for discovery and visualization of enriched GO terms in ranked gene lists. *BMC Bioinformatics, 10*, 48.
- ENCODE Project Consortium. (2012). An integrated encyclopedia of DNA elements in the human genome. *Nature, 489*, 57–74.
- Fortin, J. P., Labbe, A., Lemire, M., Zanke, B. W., Hudson, T. J., Fertig, E. J., . . . Hansen, K. D. (2014). Functional

- normalization of 450k methylation array data improves replication in large cancer studies. *Genome Biology*, 15, 503.
- Frenssen, F., Croonenberghs, J., Van den Steene, H., & Maes, M. (2015). Prolyl endopeptidase and dipeptidyl peptidase IV are associated with externalizing and aggressive behaviors in normal and autistic adolescents. *Life Sciences*, 136, 157–162.
- Garcia-Horsman, J. A., Mannisto, P. T., & Venalainen, J. I. (2007). On the role of prolyl oligopeptidase in health and disease. *Neuropeptides*, 41, 1–24.
- Genome of the Netherlands Consortium. (2014). Whole-genome sequence variation, population structure and demographic history of the Dutch population. *Nature Genetics*, 46, 818–825.
- Guillemin, C., Provencal, N., Suderman, M., Cote, S. M., Vitaro, F., Hallett, M., . . . Szyf, M. (2014). DNA methylation signature of childhood chronic physical aggression in T cells of both men and women. *PLoS One*, 9, e86822.
- Hudziak, J. J., van Beijsterveldt, C. E., Bartels, M., Rietveld, M. J., Rettew, D. C., Derks, E. M., . . . Boomsma, D. I. (2003). Individual differences in aggression: Genetic analyses by age, gender, and informant in 3-, 7-, and 10-year-old Dutch twins. *Behavior Genetics*, 33, 575–589.
- Kent, W. J., Sugnet, C. W., Furey, T. S., Roskin, K. M., Pringle, T. H., Zahler, A. M., & Haussler, D. (2002). The human genome browser at UCSC. *Genome Research*, 12, 996–1006.
- Kushnir, M., Dresner, E., Mandel, S., & Gozes, I. (2008). Silencing of the ADNP-family member, ADNP2, results in changes in cellular viability under oxidative stress. *Journal of Neurochemistry*, 105, 537–545.
- Lai, M. K., Tsang, S. W., Francis, P. T., Esiri, M. M., Keene, J., Hope, T., & Chen, C. P. (2003). Reduced serotonin 5-HT_{1A} receptor binding in the temporal cortex correlates with aggressive behavior in Alzheimer's disease. *Brain Research*, 974, 82–87.
- Maes, M., Goossens, F., Scharpe, S., Calabrese, J., Desnyder, R., & Meltzer, H. Y. (1995). Alterations in plasma prolyl endopeptidase activity in depression, mania, and schizophrenia: Effects of antidepressants, mood stabilizers, and antipsychotic drugs. *Psychiatric Research*, 58, 217–225.
- Mantle, D., Falkous, G., Ishiura, S., Blanchard, P. J., & Perry, E. K. (1996). Comparison of proline endopeptidase activity in brain tissue from normal cases and cases with Alzheimer's disease, Lewy body dementia, Parkinson's disease and Huntington's disease. *Clinica Chimica Acta*, 249, 129–139.
- McEachin, R. C., Chen, H., Sartor, M. A., Saccone, S. F., Keller, B. J., Prossin, A. R., . . . McInnis, M. G. (2010). A genetic network model of cellular responses to lithium treatment and cocaine abuse in bipolar disorder. *BMC Systems Biology*, 4, 158.
- McGowan, P. O., Sasaki, A., D'Alessio, A. C., Dymov, S., Labonte, B., Szyf, M., . . . Meaney, M. J. (2009). Epigenetic regulation of the glucocorticoid receptor in human brain associates with childhood abuse. *Nature Neuroscience*, 12, 342–348.
- Mill, J., & Heijmans, B. T. (2013). From promises to practical strategies in epigenetic epidemiology. *Nature Reviews Genetics*, 14, 585–594.
- Moechars, D., Lorent, K., De Strooper, B., Dewachter, I., & Van Leuven, F. (1996). Expression in brain of amyloid precursor protein mutated in the alpha-secretase site causes disturbed behavior, neuronal degeneration and premature death in transgenic mice. *EMBO Journal*, 15, 1265–1274.
- Momeni, N., Nordstrom, B. M., Horstmann, V., Avarseji, H., & Sivberg, B. V. (2005). Alterations of prolyl endopeptidase activity in the plasma of children with autistic spectrum disorders. *BMC Psychiatry*, 5, 27.
- Moore, T. M., Stuart, G. L., Meehan, J. C., Rhatigan, D. L., Hellmuth, J. C., & Keen, S. M. (2008). Drug abuse and aggression between intimate partners: A meta-analytic review. *Clinical Psychology Review*, 28, 247–274.
- Mori, Y., Matsui, T., Omote, D., & Fukuda, M. (2013). Small GTPase Rab39A interacts with UACA and regulates the retinoic acid-induced neurite morphology of Neuro2A cells. *Biochemical and Biophysical Research Communications*, 435, 113–119.
- Myohanen, T. T., Hannula, M. J., Van, E. R., Gerard, M., Van, D. V., Garcia-Horsman, J. A., . . . Lambear, A. M. (2012). A prolyl oligopeptidase inhibitor, KYP-2047, reduces alpha-synuclein protein levels and aggregates in cellular and animal models of Parkinson's disease. *British Journal of Pharmacology*, 166, 1097–1113.
- Nestor, P. G. (2002). Mental disorder and violence: Personality dimensions and clinical features. *American Journal of Psychiatry*, 159, 1973–1978.
- Provencal, N., Suderman, M. J., Caramaschi, D., Wang, D., Hallett, M., Vitaro, F., . . . Szyf, M. (2013). Differential DNA methylation regions in cytokine and transcription factor genomic loci associate with childhood physical aggression. *PLoS One*, 8, e71691.
- Provencal, N., Suderman, M. J., Guillemin, C., Massart, R., Ruggiero, A., Wang, D., . . . Szyf, M. (2012). The signature of maternal rearing in the methylome in rhesus macaque prefrontal cortex and T cells. *Journal of Neuroscience*, 32, 15626–15642.
- Provencal, N., Suderman, M. J., Guillemin, C., Vitaro, F., Cote, S. M., Hallett, M., . . . Szyf, M. (2014). Association of childhood chronic physical aggression with a DNA methylation signature in adult human T cells. *PLoS One*, 9, e89839.
- Rao, V., Rosenberg, P., Bertrand, M., Salehinia, S., Spiro, J., Vaishnavi, S., . . . Miles, Q. S. (2009). Aggression after traumatic brain injury: Prevalence and correlates. *Journal of Neuropsychiatry & Clinical Neurosciences*, 21, 20–29.
- Reik, W. (2007). Stability and flexibility of epigenetic gene regulation in mammalian development. *Nature*, 447, 425–432.
- Richmond, R. C., Simpkin, A. J., Woodward, G., Gaunt, T. R., Lytton, O., McArdle, W. L., . . . Relton, C. L. (2015). Prenatal exposure to maternal smoking and offspring DNA methylation across the lifecourse: Findings from the Avon Longitudinal Study of Parents and Children (ALSPAC). *Human Molecular Genetics*, 24, 2201–2217.
- Ripke, S., Wray, N. R., Lewis, C. M., Hamilton, S. P., Weissman, M. M., Breen, G., . . . Sullivan, P. F. (2013). A mega-analysis of genome-wide association studies for major depressive disorder. *Molecular Psychiatry*, 18, 497–511.

- Slieker, R. C., Bos, S. D., Goeman, J. J., Bovee, J. V., Talens, R. P., van der Breggen, R., . . . Heijmans, B. T. (2013). Identification and systematic annotation of tissue-specific differentially methylated regions using the Illumina 450k array. *Epigenetics & Chromatin*, *6*, 26.
- Suter, M., Ma, J., Harris, A., Patterson, L., Brown, K. A., Shope, C., . . . Aagaard-Tillery, K. M. (2011). Maternal tobacco use modestly alters correlated epigenome-wide placental DNA methylation and gene expression. *Epigenetics*, *6*, 1284–1294.
- Tateno, A., Jorge, R. E., & Robinson, R. G. (2003). Clinical correlates of aggressive behavior after traumatic brain injury. *Journal of Neuropsychiatry & Clinical Neurosciences*, *15*, 155–160.
- Tobi, E. W., Slieker, R. C., Stein, A. D., Suchiman, H. E., Slagboom, P. E., van Zwet, E. W., . . . Lumey, L. H. (2015). Early gestation as the critical time-window for changes in the prenatal environment to affect the adult human blood methylome. *International Journal of Epidemiology*, *44*, 1211–23. doi: 10.1093/ije/dyv043.
- Tremblay, R. E. (2008). Understanding development and prevention of chronic physical aggression: Towards experimental epigenetic studies. *Philosophical Transactions of the Royal Society of London. Series B: Biological Science*, *363*, 2613–2622.
- Tremblay, R. E., Nagin, D. S., Seguin, J. R., Zoccolillo, M., Zelazo, P. D., Boivin, M., . . . Japel, C. (2004). Physical aggression during early childhood: Trajectories and predictors. *Pediatrics*, *114*, e43–e50.
- van Beijsterveldt, C. E., Groen-Blokhuis, M., Hottenga, J. J., Franic, S., Hudziak, J. J., Lamb, D., . . . Boomsma, D. I. (2013). The Young Netherlands Twin Register (YNTR): Longitudinal twin and family studies in over 70,000 children. *Twin Research and Human Genetics*, *16*, 252–267.
- Van Dongen, J., Heijmans, B. T., Nivard, M. G., Willemsen, G., Hottenga, J.-J., Helmer, Q., . . . Boomsma, D. I. (under review). Genetic and environmental influences interact with age and sex in shaping the human methylome.
- van Iterson, M., Tobi, E. W., Slieker, R. C., den Hollander, W., Luijk, R., Slagboom, P. E., & Heijmans, B. T. (2014). MethylAid: Visual and interactive quality control of large Illumina 450k datasets. *Bioinformatics*, *30*, 3435–3437.
- Volicer, L., & Hurley, A. C. (2003). Management of behavioral symptoms in progressive degenerative dementias. *Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, *58*, M837–M845.
- Wang, D., Szyf, M., Benkelfat, C., Provencal, N., Turecki, G., Caramaschi, D., . . . Booij, L. (2012). Peripheral SLC6A4 DNA methylation is associated with in vivo measures of human brain serotonin synthesis and childhood physical aggression. *PLoS One*, *7*, e39501.
- Weaver, I. C., Cervoni, N., Champagne, F. A., D'Alessio, A. C., Sharma, S., Seckl, J. R., . . . Meaney, M. J. (2004). Epigenetic programming by maternal behavior. *Nature Neuroscience*, *7*, 847–854.
- Willemsen, G., de Geus, E. J. C., Bartels, M., van Beijsterveldt, C. E. M., Brooks, A. I., Estourgie-van Burk, G. F., . . . Boomsma, D. I. (2010). The Netherlands twin register biobank: A resource for genetic epidemiological studies. *Twin Research and Human Genetics*, *13*, 231–245.
- Willemsen, G., Vink, J. M., Abdellaoui, A., den Braber, A., van Beek, J. H., Draisma, H. H., . . . Boomsma, D. I. (2013). The Adult Netherlands Twin Register: Twenty-five years of survey and biological data collection. *Twin Research and Human Genetics*, *16*, 271–281.
- Wong, C. C., Meaburn, E. L., Ronald, A., Price, T. S., Jeffries, A. R., Schalkwyk, L. C., . . . Mill, J. (2014). Methylomic analysis of monozygotic twins discordant for autism spectrum disorder and related behavioural traits. *Molecular Psychiatry*, *19*, 495–503.
- Yeh, M. T., Coccaro, E. F., & Jacobson, K. C. (2010). Multivariate behavior genetic analyses of aggressive behavior subtypes. *Behavior Genetics*, *40*, 603–617.