

# DNA methylation signatures of aggression and closely related constructs

## A meta-analysis of epigenome-wide studies across the lifespan

Jenny van Dongen, PhD<sup>1</sup>; Fiona A. Hagenbeek, Msc<sup>1</sup>; Matthew Suderman, PhD<sup>2,3</sup>; Peter Roetman, Msc<sup>4</sup>; Karen Sugden, PhD<sup>5,6</sup>; Andreas G. Chiocchetti, PhD<sup>7</sup>; Khadeeja Ismail, Msc<sup>8</sup>; Rosa H. Mulder, Msc<sup>9,10,11</sup>; Jonathan Hafferty, MRCPsych, PhD<sup>12</sup>; Mark J. Adams, PhD<sup>12</sup>; Rosie M. Walker, PhD<sup>13</sup>; Stewart W. Morris, BSc<sup>13</sup>; Jari Lahti, PhD<sup>14,15</sup>; Leanne K. Küpers, PhD<sup>16</sup>; Georgia Escaramis, PhD<sup>17,18,19</sup>; Silvia Alemany, PhD<sup>20,21,22</sup>; Marc Jan Bonder, PhD<sup>23</sup>; Mandy Meijer, Msc<sup>24</sup>; Hill F. Ip, Msc<sup>1</sup>; Rick Jansen, PhD<sup>25</sup>; Bart M. L. Baselmans, PhD<sup>1</sup>; Priyanka Parmar, MPH<sup>26,27</sup>; Estelle Lowry, PhD<sup>26,28</sup>; Fabian Streit, PhD<sup>29</sup>; Lea Sirignano, Msc<sup>29</sup>; Tabea Send, Msc<sup>30</sup>; Josef Frank, PhD<sup>29</sup>; Juulia Jylhävä, PhD<sup>31</sup>; Yunzhang Wang, Msc<sup>31</sup>; Pashupati Prasad Mishra, Msc<sup>32</sup>; Olivier F. Colins, PhD<sup>4,33</sup>; David Corcoran, PhD<sup>6</sup>; Richie Poulton, PhD<sup>34</sup>; Jonathan Mill, PhD<sup>35</sup>; Eilis J. Hannon, PhD<sup>35</sup>; Louise Arseneault, PhD<sup>36</sup>; Tellervo Korhonen, PhD<sup>8</sup>; Eero Vuoksimaa, PhD<sup>8</sup>; Janine Felix, PhD, MD<sup>11,37</sup>; Marian Bakermans-Kranenburg, PhD<sup>38</sup>; Archie Campbell, MA<sup>13</sup>; Darina Czamara, PhD<sup>14</sup>; Elisabeth Binder, PhD<sup>14</sup>; Eva Corpeleijn, PhD<sup>16</sup>; Juan Ramon González, PhD<sup>20,21,22</sup>; Regina Grazuleviciene, PhD, MD<sup>39</sup>; Kristine B. Gutzkow<sup>40</sup>, PhD; Jorunn Evandt<sup>40</sup>, Msc; Marina Vafeiadi<sup>41</sup>, PhD; Marieke Klein<sup>24,42</sup>, PhD; Dennis van der Meer<sup>43,44</sup>, PhD; Lannie Lighthart<sup>1</sup>, PhD; BIOS Consortium\*; Cornelis Klufft<sup>45</sup>, PhD; Gareth E. Davies<sup>46</sup>, PhD; Christian Hakulinen<sup>15</sup>, PhD; Liisa Keltikangas-Järvinen<sup>15</sup>, PhD; Barbara Franke<sup>24,47</sup>, PhD; Christine M. Freitag<sup>7</sup>, PhD; Kerstin Konrad<sup>48,49</sup>, PhD; Amaia Hervas<sup>50</sup>, PhD, MD; Aranzazu Fernández-Rivas<sup>51</sup>, PhD, MD; Agnes Vetro<sup>51</sup>, PhD; Olli Raitakari<sup>53,54,55</sup>, PhD, MD; Terho Lehtimäki<sup>32</sup>, PhD, MD; Robert Vermeiren<sup>4,56</sup>, PhD, MD; Timo Strandberg<sup>57</sup>, PhD; Katri Räikkönen<sup>15</sup>, PhD; Harold Snieder<sup>16</sup>, PhD; Stephanie H. Witt<sup>29</sup>, PhD; Michael Deuschle<sup>30</sup>, MD, Prof; Nancy L. Pedersen<sup>31</sup>, PhD; Sara Hägg<sup>31</sup>, PhD; Jordi Sunyer<sup>20,21,22,58</sup>, PhD, MD; Lude Franke<sup>23</sup>, PhD; Jaakko Kaprio<sup>8</sup>, PhD, MD; Miina Ollikainen<sup>8</sup>, PhD; Terrie E. Moffitt<sup>5,6,36,59</sup>, PhD; Henning Tiemeier<sup>10,60</sup>, PhD, MD; Marinus H. van Ijzendoorn<sup>61,62</sup>, PhD; Caroline Relton<sup>2,3</sup>, PhD; Martine Vrijheid<sup>20,21,22</sup>, PhD; Sylvain Sebert<sup>26,27,63</sup>, PhD; Marjo-Riitta Jarvelin<sup>26,27,64</sup>, PhD, MD; Avshalom Caspi<sup>5,6,36,59</sup>, PhD; Kathryn L. Evans<sup>13</sup>, PhD; Andrew M. McIntosh<sup>12</sup>, MD, FRCPsych; Meike Bartels<sup>1</sup>, PhD; Dorret Boomsma<sup>1</sup>, PhD

1 Department of Biological Psychology, Amsterdam Public Health Research Institute, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands

2 Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK.

3 MRC Medical Research Council Integrative Epidemiology Unit, University of Bristol, Bristol, UK.

4 Curium-LUMC, Department of Child and Adolescent Psychiatry, Leiden University Medical Center in Oegstgeest, The Netherlands

5 Department of Psychology and Neuroscience, Duke University, Durham, NC, USA.

6 Center for Genomic and Computational Biology, Duke University, Durham, NC, USA

7 University Hospital Frankfurt; Department of Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy; Frankfurt am Main, Germany

8 Institute for Molecular Medicine Finland, University of Helsinki, Helsinki, Finland

9 Institute of Education and Child Studies, Leiden University, Leiden, the Netherlands.

10 Department of Child and Adolescent Psychiatry/Psychology, Erasmus MC, University Medical Center Rotterdam, Rotterdam, the Netherlands.

11 Generation R Study Group, Erasmus MC, University Medical Center Rotterdam, Rotterdam, the Netherlands

12 Division of Psychiatry, University of Edinburgh

13 Centre for Genomic and Experimental Medicine, Institute of Genetics and Molecular Medicine, University of Edinburgh

14 Turku Institute for Advances Studies, University of Turku, Turku Finland

15 Department of Psychology and logopedics, Faculty of Medicine, University of Helsinki, Helsinki, Finland

- 1 16 University of Groningen, University Medical Center Groningen, Department of Epidemiology, Groningen, The
- 2 Netherlands
- 3 17 CIBER Epidemiología y Salud Pública (CIBERESP), Barcelona, Spain
- 4 18 Dpt. of Biomedical Science, Faculty of Medicine and Health Science, University of Barcelona, Barcelona, Spain
- 5 19 Research Group on Statistics, Econometrics and Health (GRECS), UdG
- 6 20 ISGlobal, Barcelona, Spain.
- 7 21 Universitat Pompeu Fabra (UPF), Barcelona, Spain.
- 8 22 CIBER Epidemiología y Salud Pública (CIBERESP), Barcelona, Spain
- 9 23 Department of Genetics, University of Groningen, University Medical Centre Groningen, Groningen, The Netherlands
- 10 24 Department of Human Genetics, Donders Institute for Brain, Cognition and Behavior, Radboud University Medical
- 11 Center, Nijmegen, The Netherlands
- 12 25 Department of Psychiatry, Amsterdam UMC, Vrije Universiteit Amsterdam, Amsterdam, the Netherlands
- 13 26 Center for Life Course Health Research, P.O. Box 5000, 90014, University of Oulu, Oulu, Finland.
- 14 27 Biocenter Oulu, University of Oulu, P.O. Box 5000, 90014, Finland
- 15 26 Center for Life Course Health Research, P.O. Box 5000, 90014, University of Oulu, Oulu, Finland.
- 16 28 Queen's University Belfast
- 17 29 Department of Genetic Epidemiology in Psychiatry, Central Institute of Mental Health, Medical Faculty Mannheim,
- 18 University of Heidelberg, Mannheim, Germany
- 19 30 Department of Psychiatry and Psychotherapy, Central Institute of Mental Health, Medical Faculty Mannheim, University
- 20 of Heidelberg, Germany
- 21 31 The Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Nobels väg 12A, 17165 Stockholm,
- 22 Sweden
- 23 32 Department of Clinical Chemistry, Fimlab Laboratories, and Finnish Cardiovascular Research Center - Tampere, Faculty of
- 24 Medicine and Health Technology, Tampere University, Tampere 33520, Finland
- 25 33 Department of Special Needs Education, Ghent University, Ghent, Belgium
- 26 34 Dunedin Multidisciplinary Health and Development Research Unit, Department of Psychology, University of Otago,
- 27 Dunedin, New Zealand
- 28 35 University of Exeter Medical School, University of Exeter, Exeter, UK
- 29 36 Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, Psychology and Neuroscience, King's
- 30 College London, London, UK
- 31 37 Department of Pediatrics, Erasmus MC, University Medical Center Rotterdam, Rotterdam, the Netherlands
- 32 38 Clinical Child & Family Studies, Vrije Universiteit Amsterdam, Amsterdam, the Netherlands
- 33 39 Department of Environmental Sciences, Vytautas Magnus University, K. Donelaičio str. 58, 44248 Kaunas, Lithuania
- 34 40 Norwegian Institute of Public Health
- 35 41 Department of Social Medicine, University of Crete, Greece
- 36 42 University Medical Center Utrecht, UMC Utrecht Brain Center, Department of Psychiatry, Utrecht, The Netherlands
- 37 43 NORMENT, Division of Mental Health and Addiction, Oslo University Hospital & Institute of Clinical Medicine, University
- 38 of Oslo, Oslo, Norway.
- 39 44 School of Mental Health and Neuroscience, Faculty of Health, Medicine and Life Sciences, Maastricht University,
- 40 Maastricht, The Netherlands
- 41 45 Good Biomarker Sciences, Leiden, the Netherlands.
- 42 46 Avera Institute for Human Genetics, 3720 W. 69th Street, Sioux Falls, SD, 57108, USA.
- 43 47 Department of Psychiatry, Donders Institute for Brain, Cognition and Behavior, Radboud University Medical Center,
- 44 Nijmegen, The Netherlands
- 45 48 University Hospital, RWTH Aachen, Child Neuropsychology Section, Department of Child and Adolescent Psychiatry,
- 46 Psychosomatics and Psychotherapy, Aachen, Germany;
- 47 49 JARA-Brain Institute II, Molecular Neuroscience and Neuroimaging (INM-11), RWTH Aachen & Research Centre Juelich,
- 48 Germany
- 49 50 Hospital Universitario Mutua de Terrassa, Child and Adolescent Mental Health Service, Barcelona, Spain
- 50 51 Bilbao University Hospital, Psychiatric Service, Osakidetza, Bilbao, Spain
- 51 52 Szeged University, Department of Pediatrics and Pediatrics health center, Child and Adolescent Psychiatry, Szeged,
- 52 Hungary
- 53 53 Centre for Population Health Research, University of Turku and Turku University Hospital, Turku, Finland
- 54 54 Research Centre of Applied and Preventive Cardiovascular Medicine, University of Turku, Turku, Finland
- 55 55 Department of Clinical Physiology and Nuclear Medicine, Turku University Hospital, Turku, Finland
- 56 56 Youz, Parnassia Group, The Hague, the Netherlands
- 57 57 Helsinki University Central Hospital, Geriatrics, Helsinki, Finland.
- 58 58 IMIM (Hospital del Mar Medical Research Institute), Barcelona, Spain
- 59 59 Department of Psychiatry and Behavioral Sciences, Duke University School of Medicine, Durham, N.C., U.S.A.

1 60 Department of Social and Behavioral Science, Harvard TH Chan School of Public Health, Boston USA  
2 61 Department of Psychology, Education and Child Studies, Erasmus University Rotterdam, Rotterdam, the Netherlands.  
3 62 School of Clinical Medicine, University of Cambridge, Cambridge, UK  
4 63 Section of Genomics of Common Disease, Department of Medicine, Imperial College. Paddington, W2 1PG, London,  
5 United Kingdom London, Hammersmith Hospital Campus, Burlington Danes Building, Du Cane Road, London, W12 0NN, UK  
6 64 MRC-PHE Centre for Environment and Health, Imperial College London, Praed Street Wing, St Mary's Campus,  
7 Paddington, W2 1PG, London, United Kingdom London, Hammersmith Hospital Campus, Burlington Danes Building, Du Cane  
8 Road, London, W12 0NN, UK  
9 \*Biobank-based Integrative Omics Study Consortium. For a complete list of authors, see the  
10 acknowledgements.

11 Corresponding author: Jenny van Dongen, Ph.D., Department of Biological Psychology, Vrije  
12 Universiteit Amsterdam, Van der Boerhorststraat 7-9, 2081BT, Amsterdam, Netherlands, tel:  
13 +31 (20) 5983570, [j.van.dongen@vu.nl](mailto:j.van.dongen@vu.nl) .

14

15

16

17

18

19

20

21

22

23

24

25

26

27

28

29

30

31

32

1 **Abstract**

2 DNA methylation profiles of aggressive behavior may capture lifetime cumulative effects of  
3 genetic, stochastic, and environmental influences associated with aggression. Here, we  
4 report the first large meta-analysis of epigenome-wide association studies (EWAS) of  
5 aggressive behavior (N=15,324 participants). In peripheral blood samples of 14,434  
6 participants from 18 cohorts with mean ages ranging from 7 to 68 years, 13 methylation sites  
7 were significantly associated with aggression ( $\alpha=1.2 \times 10^{-7}$ ; Bonferroni correction). In cord  
8 blood samples of 2,425 children from five cohorts with aggression assessed at mean ages  
9 ranging from 4 to 7 years, 83% of these sites showed the same direction of association with  
10 childhood aggression ( $r=0.74$ ,  $p=0.006$ ) but no epigenome-wide significant sites were found.  
11 Top-sites (48 at a false discovery rate of 5% in the peripheral blood meta-analysis or in a  
12 combined meta-analysis of peripheral blood and cord blood) have been associated with  
13 chemical exposures, smoking, cognition, metabolic traits, and genetic variation (mQTLs).  
14 Three genes whose expression levels were associated with top-sites were previously linked  
15 to schizophrenia and general risk tolerance. At six CpGs, DNA methylation variation in blood  
16 mirrors variation in the brain. On average 44% (range=3-82%) of the aggression–methylation  
17 association was explained by current and former smoking and BMI. These findings point at  
18 loci that are sensitive to chemical exposures with potential implications for neuronal  
19 functions. We hope these results to be a starting point for studies leading to applications as  
20 peripheral biomarkers and to reveal causal relationships with aggression and related traits.

21

22

23

24

25

26

27

28

29

30

31

## 1 Introduction

2 Aggression encompasses a range of behaviors, such as bullying, verbal abuse, fighting, and  
3 destroying objects. Early life social conditions, including low parental income, separation  
4 from a parent, family dysfunction, and maternal smoking during pregnancy are risk factors for  
5 childhood aggression<sup>1,2,3</sup>. High levels of aggression are a characteristic of several psychiatric  
6 disorders and may also be caused by traumatic brain injury<sup>3</sup>, neurodegenerative diseases<sup>4</sup>  
7 and alcohol and substance abuse<sup>5,6</sup>.

8 DNA methylation mediates effects of genetic variants in regulatory regions on gene  
9 expression<sup>7</sup> and is modifiable by early life social environment, as demonstrated by animal  
10 studies<sup>8,9</sup>, and by chemical exposures including (prenatal) exposure to cigarette smoke, as  
11 illustrated by numerous human studies<sup>10</sup>. Despite the large tissue-specificity of DNA  
12 methylation, effects of genetic variants on nearby DNA methylation (*cis* mQTLs) correlate  
13 strongly between blood and brain cells<sup>11</sup>. DNA methylation signatures of chemical  
14 exposures<sup>12</sup> and maternal rearing<sup>9</sup> show a certain (but less understood) degree of  
15 conservation across tissues.

16 Large-scale epigenome-wide association studies (EWASs) have become feasible  
17 through DNA methylation microarrays applied to blood samples from large cohorts,  
18 identifying thousands of loci where methylation in cord blood is associated with maternal  
19 smoking<sup>13</sup>. Methylation in blood is associated with depressive symptoms<sup>14</sup> and brain  
20 morphology<sup>15</sup>, with some evidence for blood DNA methylation signatures being a marker for  
21 methylation levels<sup>15</sup> or gene expression<sup>14</sup> in the brain. For several traits, DNA methylation  
22 scores based on multiple CpGs from EWAS show better predictive value than currently  
23 available polygenic scores<sup>16,17</sup>.

24 Small-scale studies (maximum sample size=260) have provided some evidence that  
25 DNA methylation differences in blood, cord blood, and buccal cells are associated with  
26 severe forms of aggressive behavior and related problems in children and adults, including  
27 (chronic) physical aggression and early onset conduct problems<sup>18-20</sup>, but studies on violent  
28 aggression in schizophrenia patients (N=134)<sup>21</sup> and a population-based study of continuous  
29 aggression symptoms in adults (N=2,029)<sup>22</sup> did not detect epigenome-wide significant sites.

30 We performed an EWAS meta-analysis of aggressive behavior and closely related  
31 constructs. We chose to meta-analyze multiple measures of aggression across ages and sex  
32 to maximize sample size. The contribution of genetic influences to aggression is largely  
33 stable, at least throughout childhood<sup>23</sup>, whereas epigenetic signatures may be dynamic and  
34 may differ across cell types and age. Therefore, we performed separate meta-analyses of  
35 peripheral blood collected after birth (N=14,434) and cord blood (N=2,425), followed by a

1 combined meta-analysis (N=15,324) including an examination of heterogeneity of effects.  
2 Next, we tested the relationship between aggressive behavior and epigenetic clocks, as  
3 associations of lifetime stress<sup>24</sup>, exposure to violence<sup>25</sup>, and psychiatric disorders<sup>26,27</sup> with  
4 accelerated epigenetic ageing have been reported. We performed extensive functional  
5 follow-up by integrating our findings with data on gene expression, mQTLs and DNA  
6 methylation in brain samples.

7

8

9

10

11

12

13

14

15

16

17

18

19

20

21

22

23

24

25

26

27

## 1 **Methods**

### 2 **Cohorts**

3 Demographic information for the cohorts is provided in **Table 1**. Detailed cohort information  
4 is provided in **eAppendix 1**. Informed consent was obtained from all participants. The  
5 protocol for each study was approved by the ethical review board of each institution.

### 6 **Aggressive behavior**

7 Aggressive behavior was assessed by self-report or reported by parents and teachers.  
8 Multiple instruments were used (**eTable 1**): ASEBA Child Behavior Check List (CBCL)<sup>28</sup>,  
9 Strengths and Difficulties Questionnaire (SDQ) conduct problem scale<sup>29</sup>, multidimensional  
10 Peer Nomination Inventory (MNPI) aggression scale<sup>30</sup>, ASEBA adult self-report (ASR)  
11 aggression scale<sup>31</sup>, DSM-IV Conduct Disorder Symptom Scale<sup>32</sup>, Multidimensional  
12 Personality Questionnaire (MPQ) aggression scale<sup>33</sup>, and the Hunter-Wolf aggressive  
13 behavior scale<sup>34,35</sup>. In four cohorts, a single aggression-related item from personality  
14 questionnaires was used. Distributions of aggression scores are provided in **eFigure 1**.

### 15 **DNA Methylation BeadChips**

16 DNA methylation was assessed with Illumina BeadChips: the Illumina Infinium  
17 HumanMethylation450 BeadChip (450k array; majority of cohorts), or the Illumina  
18 MethylationEPIC BeadChip (EPIC array). Most cohorts analyzed DNA methylation  $\beta$ -values,  
19 which range from 0 to 1, indicating the proportion of DNA that is methylated at a CpG in a  
20 sample. Cohort-specific details about DNA methylation profiling, quality control, and  
21 normalization are described in **eAppendix 1** and summarized in **eTable 2**.

### 22 **Epigenome-wide Association Analysis**

23 EWAS analyses were performed according to a standard operating procedure  
24 (<http://www.action-euproject.eu/content/data-protocols>). In each cohort, the association  
25 between DNA methylation level and aggressive behavior was specified under a linear model  
26 with DNA methylation as outcome, and correction for relatedness of individuals where  
27 applicable. Two models were tested. Model 1 included aggressive behavior, sex, age at  
28 blood sampling (not in cohorts with invariable age), white blood cell percentages (measured  
29 or imputed), and technical covariates. Model 2 included the same predictors plus body-mass-  
30 index (BMI) and smoking status in adolescents and adults (current smoker, former smoker or  
31 never smoked). Cohort-specific details and R-code are provided in **eAppendix 1** and **eTable**  
32 **3**, respectively. The relationship between aggressive behavior and covariates is provided in  
33 **eTable 4** based on data from the Netherlands Twin Register (N=2059).

1           Quality control and filtering of cohort-level EWAS summary statistics is described in  
2 **eAppendix 2**. The following probes were removed: on sex chromosomes, methylation sites  
3 with more than 5% missing data in a cohort, probes overlapping SNPs affecting the CpG or  
4 single base extension site with a minor allele frequency (MAF) > 0.01 in the 1000G EU or  
5 GONL population<sup>7</sup>, and ambiguous mapping probes reported with an overlap of at least 47  
6 bases per probe<sup>36</sup>. The R package Bacon was used to compute the Bayesian inflation factor  
7 and to obtain bias- and inflation-corrected test statistics (**eFigure 2**) prior to meta-analysis<sup>37</sup>.

## 8 **Meta-analysis**

9 Fixed-effects meta-analyses were performed in METAL<sup>38</sup>. We used the p-value-based  
10 (sample size-weighted) method because the measurement scale of aggressive behavior  
11 differs across studies. First, results based on peripheral blood and cord blood data were  
12 meta-analyzed separately. Second, a combined meta-analysis was performed of all data.  
13 The following cohorts had data available for both cord blood and peripheral blood (from the  
14 same children): INMA (which is part of HELIX) and ALSPAC. In the combined meta-analysis,  
15 the cord blood data from ALSPAC and INMA were excluded to avoid sample overlap.  
16 Statistical significance was assessed considering Bonferroni correction for the number of  
17 sites tested ( $\alpha=1.2 \times 10^{-7}$ ). Methylation sites that were associated with aggression at the  
18 less conservative false discovery rate (FDR) threshold (5%) were included in follow-up  
19 analyses. The  $I^2$  statistic from METAL was used to describe heterogeneity.

## 20 **Follow-up Analyses**

21 DNA methylation score analyses and epigenetic clock analyses are described in **eAppendix**  
22 **3** and **eAppendix 4**. Follow-up analyses (**eAppendix 5- eAppendix 10**) were performed on  
23 meta-analysis top-sites (FDR<0.05), including a comparison of top-sites with all previously  
24 reported associations in the EWAS atlas<sup>39</sup>, follow-up analysis of top-sites in two clinical  
25 cohorts with blood methylation data (**Table 2**), a cross-tissue analysis (blood, buccal, brain),  
26 and association with gene expression level and mQTLs. Analyses of differentially methylated  
27 regions (DMRs) are described in **eAppendix 8**. Finally, we performed replication analysis of  
28 a previously reported DMR associated with aggression<sup>20</sup> (**eAppendix 9**).

29



## 1 **Results**

### 2 ***Peripheral blood meta-analysis***

3 We performed a meta-analysis of 13 studies with peripheral blood DNA methylation data  
4 (N=14,434). The meta-analysis test statistics showed no inflation (**eTable 5, eFigure 3**). In  
5 model 1, methylation at 13 CpGs was associated with aggression (Bonferroni correction;  
6  $\alpha=1.2 \times 10^{-7}$ ), and 35 passed a less conservative threshold (FDR 5%; **Figure 1a**). At 28  
7 out of the 35 sites (80%), higher levels of aggression were associated with lower methylation  
8 levels. Top-sites showed varying degrees of between-study heterogeneity (mean  $I^2=50\%$ ;  
9 range=0- 86%, **eTable 6**). Five sites showed significant heterogeneity ( $\alpha=1.2 \times 10^{-7}$ ).

10

### 11 ***Cord blood meta-analysis***

12 The meta-analysis of cord blood (five cohorts; N=2,425) detected no significant CpGs  
13 (**eTable 7**). Examining top-sites from the peripheral blood meta-analysis, 12 of the  
14 significant, and 33 of the FDR top-sites were assessed in cord blood; 10 (83%), and 25  
15 (71%), respectively, showed the same direction of association (**Figure 1b**). Effect sizes in  
16 cord blood correlated significantly with effect sizes in peripheral blood ( $r=0.74$ ,  $p=0.006$  for  
17 epigenome-wide significant and  $r=0.51$ ,  $p=0.003$  for FDR top-sites).

18

### 19 ***Combined meta-analysis***

20 In the combined meta-analysis of peripheral and cord blood data (total sample size=15,324,  
21 **eTable 6**), methylation at 13 CpGs was associated with aggression after Bonferroni  
22 correction, including ten CpGs from the peripheral blood meta-analysis, and 43 passed a  
23 less conservative threshold (FDR 5%, **Table 3**). Among FDR top-sites from both analyses,  
24 13 CpGs were only found in the combined meta-analysis but not in the peripheral blood  
25 meta-analysis, while five CpGs from the peripheral blood meta-analysis were no longer  
26 significant in the combined meta-analysis (**Figure 1c**).

27

### 28 ***CBCL meta-analysis***

29 We compared our meta-analysis results to a meta-analysis of cohorts that applied the same  
30 aggression instrument; i.e. CBCL (four studies; N=2,286; **Table 1**). No epigenome-wide  
31 significant sites were detected (**eFigure 4a**). Examining top-sites from the overall meta-  
32 analysis (Model 1), 38 (79%) showed the same direction of association for CBCL aggression  
33 in children, and effect sizes correlated strongly ( $r=0.75$ ,  $p=6.8 \times 10^{-10}$ , **eFigure 4b**).

34

### 35 ***Overlap with CpGs detected in previous EWASs***

36 We performed enrichment analyses against all previously reported associations with  
37 diseases and environmental exposures recorded in the EWAS Atlas<sup>39</sup>. The top ten most

1 strongly enriched traits are shown in **Figure 1e**. CpGs associated with aggressive behavior  
2 showed large overlap with CpGs previously associated with smoking (37 CpGs;  
3 corresponding to 77% of aggression-associated CpGs and 0.3% of CpGs that have been  
4 previously associated with smoking), and smaller overlap with other smoking traits (e.g.  
5 maternal smoking), other chemical exposures (e.g. perinatal exposure to polychlorinated  
6 biphenyls (PCBs) and polychlorinated dibenzofurans (PCDFs)). Further overlap includes  
7 CpGs associated with alcohol consumption, cognitive function, educational attainment,  
8 ageing, and metabolic traits (**eTable 8**).

### 9 10 ***Controlling for smoking and BMI***

11 Model 2 was fitted to test whether the association between DNA methylation and aggressive  
12 behavior attenuated after adjusting for the most important postnatal lifestyle factors that  
13 influence DNA methylation (smoking and BMI). Examining the 35 CpGs associated with  
14 aggression at FDR 5% in peripheral blood, all CpGs showed the same direction of  
15 association after adjusting for smoking and BMI (**eTable 6, Figure 1d**). Effect sizes were  
16 attenuated to varying degrees (mean reduction=44%, range=3-83%). Changes in effect sizes  
17 are likely primarily driven by the correction for smoking, since only one top-site has been  
18 associated previously with BMI. Some CpGs showed little attenuation, in particular CpGs that  
19 have not been previously associated with smoking (e.g.; cg02895948; *PLXN2A*,  
20 cg00891184; *KIF1B*, cg1215892; intergenic, and cg05432213; *ACT1*). In model 2, between-  
21 study heterogeneity at top-sites was greatly reduced (adjusted: mean  $I^2$ =28%, range=0-  
22 77%). No CpGs were epigenome-wide significant or FDR-significant in the adjusted meta-  
23 analyses.

### 24 25 ***DNA methylation scores***

26 We computed weighted sumscores in NTR (peripheral blood, mean age=36.4, SD=12,  
27 N=2,059) based on summary statistics from the peripheral blood meta-analysis without NTR  
28 (**Figure 2**). The best score, based on CpGs with  $p < 1 \times 10^{-3}$  in model 2 (745 CpGs), explained  
29 0.29% of the variance in aggression ( $p=0.02$ , not significant after multiple testing correction).  
30 This effect was attenuated when age and sex were added to the prediction equation.

### 31 ***Epigenetic clocks***

32 Horvath and Hannum epigenetic age acceleration were not associated with aggression  
33 (**eTable 9**) in a meta-analysis of 12 studies with peripheral blood DNA methylation data  
34 (N=9,554), five studies with cord blood DNA methylation (N=2,225), or in a combined meta-

1 analysis of 15 studies (N=9,740). There was no significant heterogeneity between cohorts  
2 (mean  $I^2=16\%$ , range=0-60%).

3

#### 4 ***Follow-up in clinical cohorts***

5 To assess the translation of our observations to aggression-related problem behavior in  
6 psychiatric disorders that show comorbidity with aggression, we performed follow-up  
7 analyses of top-sites in two clinical cohorts (**Table 2**): the NeuroIMAGE<sup>40</sup> cohort of ADHD  
8 cases and controls (N<sub>total</sub>=71) and the FemNAT-CD<sup>41</sup> cohort of female conduct disorder  
9 cases and controls (N<sub>total</sub>=100). Results did not replicate (**eAppendix 6, eTable 10, eTable**  
10 **11, eFigure 5, eFigure 6**).

11

#### 12 ***Cross-tissue analysis***

13 To assess the generalizability of our observations in blood to other tissues, we examined the  
14 association with CBCL aggression in buccal DNA methylation data (EPIC array), available for  
15 38 top-sites, in children (N=1237) and a child clinical cohort (N=172;**Table 2, eTable 12**)<sup>42</sup>.  
16 We also tested associations with maternal smoking and with child nervous system  
17 medication (as indexed by the Anatomical Therapeutic Chemical classification system (ATC  
18 N-class))

19 Correlations between DNA methylation levels in blood and buccal cells, based on  
20 450k data from matched samples (N=22, age=18 years)<sup>43</sup> were available for 36 of these  
21 CpGs. The average correlation was weak ( $r=0.25$ , range=-0.40-0.76). Five CpGs showed a  
22 strong correlation between blood and buccal cells ( $r>0.5$ , **eTable 13**), of which three have  
23 been previously associated with (maternal) smoking.

24 In line with the weak correlation between blood and buccal cell methylation for most  
25 top-sites, none of the top-sites was associated with aggression in buccal samples  
26 ( $\alpha=0.001$ , **eTable 14**). Regression coefficients based on analyses in buccal cells and  
27 blood overall showed no directional consistency (twin cohort:  $r=0.03$ ,  $p=0.86$ ; concordant  
28 direction: 47%,  $p=0.87$ , binomial test, clinical cohort:  $r=0.27$ ,  $p=0.10$ ; concordant direction:  
29 61%,  $p=0.26$ ). Exclusion of ancestry outliers did not change these results (**eTable 14**). Of the  
30 five CpGs with a large blood-buccal correlation, three showed the same direction of  
31 association with aggression in buccal cells from twins, four in clinical cases, and one CpG  
32 was nominally associated with aggression in buccal samples from twins; cg11554391  
33 (**AHRR**),  $r_{\text{blood-buccal}}=0.69$ ,  $\beta_{\text{aggression}}=-0.0002$ ,  $p=0.007$ .

1 One CpG was significantly associated with maternal smoking in both cohorts:  
2 cg04180046 (*MYO1G*), NTR:  $\beta_{\text{maternal smoking}}=0.041$ ,  $p=6.0 \times 10^{-6}$ , Curium:  $\beta_{\text{maternal smoking}}=0.048$ ,  
3  $p=7.9 \times 10^{-5}$  (**eTable 14**). None of the CpGs was associated with medication use (**eTable 14**).

4 We examined the correlation between DNA methylation levels in blood and brain  
5 (N=122)<sup>44</sup> in published DNA methylation data from matched blood samples and four brain  
6 regions. Six aggression top-sites (13%) showed significantly correlated DNA methylation  
7 levels between blood and one or multiple brain regions: mean  $r=0.52$ ; range=0.45-0.63,  
8  $\alpha=2.6 \times 10^{-4}$ , **eTable 15, eFigure 7**), two of which have not been previously associated  
9 with smoking or BMI: cg14560430 (*TRIM71*), and cg20673321 (*ZNF541*).

10

### 11 **DMRs**

12 DMR analysis showed that 14 DMPs from our combined meta-analysis reside in regions  
13 where multiple correlated methylation sites showed evidence for association with aggressive  
14 behavior. DMR analysis also detected additional regions that were not significant in DMP  
15 analysis (**eTable 16- eTable 21**). These analyses are described in detail in **eAppendix 8**.

16

### 17 **Replication analysis**

18 A previous EWAS based on Illumina array data detected a significant DMR in *DRD4* in  
19 buccal cells associated with engagement in physical fights<sup>20</sup>. This locus did not replicate in  
20 our meta-analyses or in the two cohorts with buccal methylation data (**eTable 22, eAppendix**  
21 **9**).

22

### 23 **Gene Expression**

24 Based on peripheral blood RNA-seq and DNA methylation data (N=2,101)<sup>7</sup>, 17 significant  
25 DNA methylation-gene expression associations were identified among 15 CpGs and ten  
26 transcripts (**Table 3, eTable 23**). For most transcripts, a higher methylation level at a CpG  
27 site in *cis* correlated with lower expression (82.4%): cg03935116 and *FAM60A*, cg00310412  
28 and *SEMA7A*, cg03707168 and *PPP1R15A*, cg03636183 and *F2RL3*, two intergenic CpGs  
29 on chromosome 6, where methylation level correlated negatively with expression levels of  
30 *FLOT1*, *TUBB*, *LINC00243*, and six CpGs annotated to *AHRR* were negatively associated  
31 with *EXOC3* expression level. Positive correlations were observed between methylation  
32 levels at 2 CpGs on chromosome 7 and levels of *RP4-647J21.1* (novel transcript,  
33 overlapping *MYO1G*) and between cg02895948 and *PLXNA2*.

34

35

36

1 ***mQTLs***

2 To gain insight into genetic causes of variation underlying top-sites, we obtained whole-blood  
3 mQTL data (N=3,841)<sup>7</sup>. In total, 75 mQTL associations were identified among 34 aggression  
4 top-sites (70.8%) and 66 SNPs at the experiment-wide threshold applied by the mQTL study  
5 FDR<0.05): 80% were *cis* mQTLs and 20% were *trans* mQTLs (**eTable 24**).

6

7

8

9

10

11

12

13

14

15

16

17

18

19

20

21

22

23

24

25

26

27

28

## 1 Discussion

2 We identified 13 epigenome-wide significant sites (Bonferroni corrected) in the meta-analysis  
3 of blood and 13 in the combined meta-analysis of blood and cord blood (16 unique sites). We  
4 prioritized 48 top-sites (FDR 5%) for follow-up analyses. Methylation level at three top-sites  
5 was associated with expression levels of genes that have been previously linked to  
6 psychiatric or behavioral traits in GWASs: *FLOT1* (schizophrenia<sup>45</sup>), *TUBB* (schizophrenia)<sup>45</sup>,  
7 and *PLXNA2* (general risk tolerance)<sup>46</sup>. Several other loci have functions in the brain and six  
8 CpGs showed correlated methylation levels between blood and brain.

9 The majority of top-sites (77%) were associated with smoking, 46% were associated  
10 with maternal smoking, 25% were associated with alcohol consumption, and 15% were  
11 associated with perinatal PCB and PCDF exposure. This overlap of aggression top-sites with  
12 smoking and other chemical exposures is noteworthy. Methylation levels of top-sites in the  
13 Aryl-Hydrocarbon Receptor Repressor gene *AHRR* and several other genes are known to be  
14 strongly associated with exposure to cigarette smoke<sup>13,47</sup> and persistent organic pollutants<sup>48</sup>.  
15 The best characterized exogenous ligands of the widely expressed Aryl-Hydrocarbon  
16 Receptor are environmental contaminants such as benzo[a]pyrene (B[a]P), and TCDD  
17 (dioxin), whose neurotoxic and neuro-endocrine effects, including disruption of neuronal  
18 proliferation, differentiation, and survival, have been well-characterized<sup>49</sup>. Human prenatal  
19 exposure to B[a]P is associated with delayed mental development, lower IQ, anxiety and  
20 attention problems<sup>50</sup>. Research on B[a]P neurotoxicity in adults is scarce but a study on coke  
21 oven workers found that occupational B[a]P exposure correlates with reduced monoamine,  
22 amino acid and choline neurotransmitter levels and with impaired learning and memory<sup>51</sup>.

23 On average 44% (range=3-82%) of the aggression–methylation association was  
24 explained by current and former smoking and BMI. Our findings do not merely reflect effects  
25 of own smoking: 71% of the top-sites showed the same direction for the prospective  
26 association of cord blood methylation at birth and aggression in childhood, and 46% have  
27 been associated with maternal prenatal smoking. There is a weak observational association  
28 between maternal smoking and child aggression<sup>52</sup>. Our analyses did not adjust for prenatal  
29 and postnatal second-hand smoking. Future studies can examine if the link between prenatal  
30 maternal smoking and aggression is mediated by DNA methylation.

31 We found that DNA methylation scores for aggression explained less variation  
32 compared to DNA methylation scores for traits such as BMI, smoking and educational  
33 attainment. For these traits, EWASs tended to identify more epigenome-wide significant  
34 hits<sup>16,17</sup>. The variance in aggression explained by DNA methylation scores was in the same  
35 order of magnitude as the variance in height explained by DNA methylation scores (based on

1 EWASs of height in smaller samples), i.e. less than 1%<sup>16</sup>. More research is needed in  
2 particular to delineate if there is a causal link between these methylation sites and  
3 aggressive behaviour, since our results may also reflect (residual) confounding by (exposure  
4 to second-hand) smoking. One approach to address this could be Mendelian Randomization,  
5 in which genetic information (SNPs) is used for causal inference of the effect of an exposure  
6 (e.g. DNA methylation) on an outcome (e.g. aggression). This approach previously supported  
7 a causal effect of maternal smoking-associated methylation sites in blood on various traits  
8 and diseases for which well-powered GWASs have been performed, including  
9 schizophrenia<sup>53,54</sup>. For aggressive behavior, the currently available<sup>55</sup> largest GWASs of  
10 aggressive behavior included ~16,000<sup>56</sup> and ~75,000 participants (Ip et al, Multivariate GWA  
11 meta-analysis in over 500K observations on aggressive behavior and ADHD symptoms,  
12 *submitted*), respectively. The GWAS by Ip et al detected 3 significant genes in gene-based  
13 analysis, but both GWASs did not detect genome-wide significant SNPs and are likely still  
14 underpowered. In the future, larger GWASs of aggressive behavior and larger mQTL  
15 analyses will allow for powerful Mendelian Randomization for aggression-associated  
16 methylation sites.

17

### 18 **Strengths and limitations**

19 This is the largest EWAS of aggressive behavior to date. The large sample size was  
20 achieved by applying a broad phenotype definition, including participants from multiple  
21 countries and all ages in a meta-analysis, and analyzing DNA methylation data from blood. A  
22 limitation of this approach is that it reduces power to detect age-, sex-, and symptom-specific  
23 effects, and that genetic and environmental backgrounds of different populations, as well as  
24 non-identical processing methods of methylation data play a role. A limitation of population-  
25 based cohorts and even clinical populations is that individuals with extreme levels of  
26 aggressive behavior who cause most societal problems are likely underrepresented.  
27 Moreover, some studies used measures that tapped features that overlap with but are not  
28 necessarily indicative of aggression (e.g. personality traits, anger, oppositional defiant  
29 disorder). Future EWASs that specifically focus on more homogeneous aggression  
30 measures are therefore warranted. Our meta-analysis approach may identify a common  
31 epigenomic signature of aggression-related problems.

32 Follow-up analysis in independent datasets indicated that these findings do not  
33 generalize strongly to buccal cells, and results did not replicate in two clinical cohorts. These  
34 were small, used different aggression measures, and one used a different technology  
35 (sequencing) in females only.

36

## 1 **Conclusions**

2 We identified associations between aggressive behavior and DNA methylation in blood at  
3 CpGs whose methylation level is also associated with exposure to smoking, alcohol  
4 consumption, other chemical exposures, and genetic variation. Methylation levels at three  
5 top-sites were associated with expression levels of genes that have been previously linked to  
6 psychiatric or behavioral traits in GWAS. Our study illustrates both the merit of EWASs  
7 based on peripheral tissues to identify environmentally-driven molecular variation associated  
8 with behavioral traits and their challenges to tease-out confounders and mediators of the  
9 association, and causality. Pursuing full control of potential confounders in behavioral EWAS  
10 meta-analyses (including smoking-exposure and other substance-use across the life course,  
11 socioeconomic conditions and other, perhaps less obvious, ones) might be unrealistic, and  
12 has the potential disadvantage of over-correction. Future studies, including those that  
13 integrate EWAS results for multiple traits and exposures, DNA methylation in multiple  
14 tissues, and GWASs of multiple traits are warranted to unravel the utility of our results as  
15 peripheral biomarkers for pathological mechanisms in other tissues (such as neurotoxicity)  
16 and to unravel possible causal relationships with aggression and related traits. We consider  
17 this study to be the starting point for such follow-up studies.

## 18 **Acknowledgements**

19 This work was supported by ACTION. ACTION receives funding from the European Union  
20 Seventh Framework Program (FP7/2007-2013) under grant agreement no 602768. Cohort-  
21 specific acknowledgements are provided in **eAppendix 1**. The following authors declare a  
22 conflict of interest: Barbara Franke received educational speaking fees from Medice. Andrew  
23 M. McIntosh has received research support from Eli Lilly, Janssen and The Sackler Trust and  
24 speaker fees from Illumina and Janssen. Christine M. Freitag has received funding by the  
25 DFG, BMBF, State of Hessen, and the EU. She receives royalties for books on ASD, ADHD,  
26 and MDD.

27  
28  
29  
30  
31  
32  
33  
34  
35  
36



1 **BIOS Consortium (Biobank-based Integrative Omics Study)**

2 **Management Team** Bastiaan T. Heijmans (chair)<sup>1</sup>, Peter A.C. 't Hoen<sup>2</sup>, Joyce van Meurs<sup>3</sup>, Aaron Isaacs<sup>4</sup>, Rick  
3 Jansen<sup>5</sup>, Lude Franke<sup>6</sup>.

4 **Cohort collection** Dorret I. Boomsma<sup>7</sup>, René Pool<sup>7</sup>, Jenny van Dongen<sup>7</sup>, Jouke J. Hottenga<sup>7</sup> (Netherlands Twin  
5 Register); Marleen MJ van Greevenbroek<sup>8</sup>, Coen D.A. Stehouwer<sup>8</sup>, Carla J.H. van der Kallen<sup>8</sup>, Casper G.  
6 Schalkwijk<sup>8</sup> (Cohort study on Diabetes and Atherosclerosis Maastricht); Cisca Wijmenga<sup>6</sup>, Lude Franke<sup>6</sup>, Sasha  
7 Zhernakova<sup>6</sup>, Ettje F. Tigchelaar<sup>6</sup> (LifeLines Deep); P. Eline Slagboom<sup>1</sup>, Marian Beekman<sup>1</sup>, Joris Deelen<sup>1</sup>, Diana  
8 van Heemst<sup>9</sup> (Leiden Longevity Study); Jan H. Veldink<sup>10</sup>, Leonard H. van den Berg<sup>10</sup> (Prospective ALS Study  
9 Netherlands); Cornelia M. van Duijn<sup>4</sup>, Bert A. Hofman<sup>11</sup>, Aaron Isaacs<sup>4</sup>, André G. Uitterlinden<sup>3</sup> (Rotterdam  
10 Study).

11 **Data Generation** Joyce van Meurs (Chair)<sup>3</sup>, P. Mila Jhama<sup>3</sup>, Michael Verbiest<sup>3</sup>, H. Eka D. Suchiman<sup>1</sup>, Marijn  
12 Verkerk<sup>3</sup>, Ruud van der Breggen<sup>1</sup>, Jeroen van Rooij<sup>3</sup>, Nico Lakenberg<sup>1</sup>.

13 **Data management and computational infrastructure** Hailiang Mei (Chair)<sup>12</sup>, Maarten van Iterson<sup>1</sup>, Michiel van  
14 Galen<sup>2</sup>, Jan Bot<sup>13</sup>, Dasha V. Zhernakova<sup>6</sup>, Rick Jansen<sup>5</sup>, Peter van 't Hof<sup>12</sup>, Patrick Deelen<sup>6</sup>, Irene Nooren<sup>13</sup>, Peter  
15 A.C. 't Hoen<sup>2</sup>, Bastiaan T. Heijmans<sup>1</sup>, Matthijs Moed<sup>1</sup>.

16 **Data Analysis Group** Lude Franke (Co-Chair)<sup>6</sup>, Martijn Vermaat<sup>2</sup>, Dasha V. Zhernakova<sup>6</sup>, René Luijk<sup>1</sup>, Marc Jan  
17 Bonder<sup>6</sup>, Maarten van Iterson<sup>1</sup>, Patrick Deelen<sup>6</sup>, Freerk van Dijk<sup>14</sup>, Michiel van Galen<sup>2</sup>, Wibowo Arindrarto<sup>12</sup>,  
18 Szymon M. Kielbasa<sup>15</sup>, Morris A. Swertz<sup>14</sup>, Erik. W van Zwet<sup>15</sup>, Rick Jansen<sup>5</sup>, Peter-Bram 't Hoen (Co-Chair)<sup>2</sup>,  
19 Bastiaan T. Heijmans (Co-Chair)<sup>1</sup>.

20 1. Molecular Epidemiology Section, Department of Medical Statistics and Bioinformatics, Leiden University Medical Center,  
21 Leiden, The Netherlands

22 2. Department of Human Genetics, Leiden University Medical Center, Leiden, The Netherlands

23 3. Department of Internal Medicine, ErasmusMC, Rotterdam, The Netherlands

24 4. Department of Genetic Epidemiology, ErasmusMC, Rotterdam, The Netherlands

25 5. Department of Psychiatry, VU University Medical Center, Neuroscience Campus Amsterdam, Amsterdam, The  
26 Netherlands

27 6. Department of Genetics, University of Groningen, University Medical Centre Groningen, Groningen, The Netherlands

28 7. Department of Biological Psychology, VU University Amsterdam, Neuroscience Campus Amsterdam, Amsterdam, The  
29 Netherlands

30 8. Department of Internal Medicine and School for Cardiovascular Diseases (CARIM), Maastricht University Medical Center,  
31 Maastricht, The Netherlands

32 9. Department of Gerontology and Geriatrics, Leiden University Medical Center, Leiden, The Netherlands

33 10. Department of Neurology, Brain Center Rudolf Magnus, University Medical Center Utrecht, Utrecht, The Netherlands

34 11. Department of Epidemiology, ErasmusMC, Rotterdam, The Netherlands

35 12. Sequence Analysis Support Core, Leiden University Medical Center, Leiden, The Netherlands

36 13. SURFsara, Amsterdam, the Netherlands

37 14. Genomics Coordination Center, University Medical Center Groningen, University of Groningen, Groningen, the  
38 Netherlands

39 15. Medical Statistics Section, Department of Medical Statistics and Bioinformatics, Leiden University Medical Center,  
40 Leiden, The Netherlands

41

42

1 **Figure 1** DNA methylation associated with aggressive behavior in a large blood-based meta-  
2 analysis  
3 a) Manhattan plot showing the fixed effects meta-analysis p-values for the association  
4 between aggressive behavior and DNA methylation level based on the meta-analysis  
5 of peripheral blood. The blue horizontal line denotes the FDR-threshold (5%) and the red  
6 line indicates the Bonferroni threshold. b) Effects sizes of top-sites from the meta-analysis  
7 of aggression in peripheral blood (x-axis) versus effects sizes from the meta-analysis  
8 of aggression in cord blood (y-axis). c) Venn diagram showing the numbers and overlap of  
9 CpGs detected at FDR 5% in the meta-analysis of peripheral blood and the combined meta-  
10 analysis and cord blood and peripheral blood. d) Effects sizes of top-sites from the meta-  
11 analysis of aggression in peripheral blood model 1 (x-axis) versus effects sizes from the  
12 meta-analysis of aggression in peripheral blood model 2; adjusted for smoking and BMI (y-  
13 axis). e) Top enriched traits based on enrichment analysis with all 48 top-sites. The third  
14 column shows how many of the 48 CpGs have been previously associated with the trait in  
15 the first column. The last column shows the overlap as a percentage of the total number of  
16 CpGs previously associated with the trait in column 1 (e.g. 0.3% of all CpGs previously  
17 associated with smoking are also associated with aggression in the current meta-analysis. d)  
18 In panel b and d, CpGs that have not been previously associated with smoking in the meta-  
19 analysis by Joehanes et al<sup>47</sup> are plotted in red.

20

21 **Figure 2** Prediction of aggression by DNA methylation scores

22 The bars indicate how much of the variance in ASEBA adult self-report (ASR) aggression  
23 scores were explained by DNA methylation scores in NTR (N=2059, peripheral blood, 450k  
24 array). Scores were created based on weights from the peripheral blood meta-analysis with  
25 NTR excluded (N=12,375). The y-axis shows percentage of variance explained. Different  
26 colors denote DNA methylation scores created with different numbers of CpGs that were  
27 selected on their p-value in the meta-analysis (see legend). From left to right, the first three  
28 plots show DNA methylation scores created based on weights obtained from the meta-  
29 analysis of EWAS model 1, and plots 4 till 6 show DNA methylation scores created based on  
30 weights obtained from the meta-analysis of EWAS model 2. Each DNA methylation score  
31 was tested for association with aggression in three model: the simplest model (first plot)  
32 included aggression as outcome variable, and DNA methylation score as predicted plus  
33 technical covariates and cell counts. The second model additionally included sex and age as  
34 predictors. The third model additionally included sex, age, and smoking as predictors. Stars  
35 denote nominal p-values < 0.05 (not corrected for multiple testing).

36

37

1 **Table 1** Discovery cohorts

Cohort	N, M1	N, M2	% female	% current smoker	% former smoker	DNA age, Mean (SD), y <sup>e</sup>	Aggression survey	Array	Aggression, Mean (SD)	Time between survey and DNA, Mean (min, max), y <sup>f</sup>
Peripheral blood										
ALSPAC <sup>57</sup>	865	865	49.4	0	0	7.5 (0.2)	SDQ <sup>29</sup>	450k	1.5 (1.4)	0.7 (0.0, 2.1)
Dunedin <sup>58</sup>	767	764	46.3	33.8	13.7	26.0 (0)	MPQ <sup>33</sup>	450k	23.3 (19.3)	0
E-Risk <sup>59</sup>	1629	1601	49.8	22.7	0	18.0 (0)	DSM-IV Conduct Disorder <sup>32</sup>	450k	2.2 (2.3)	0
FinnTwin12 <sup>60</sup>	757	757	59.2	46.0*	NA	22.4 (0.7)	MNPI <sup>30</sup>	450k	0.6 (0.7)	10.4 (9.0,13.0)
GS:SFHS <sup>61</sup>	4609	4421	67.9	18.9	29.5	46.6 (14.0)	1 item, from GHQ 28 <sup>62a</sup>	EPIC	0.1(0.3)	0
GLAKU <sup>63</sup>	192	177	56.3	1.7	0	12.3 (0.5)	CBCL <sup>28</sup>	EPIC	3.9 (3.8)	0
HELIX <sup>64</sup>	1058	1058	44.9	NA	NA	8.0 (1.6)	CBCL <sup>28</sup>	450k	5.2 (5.0)	0
LLD <sup>65</sup>	683	683	59.4	19.0	33.1	43.9 (11.6)	1 item, personality questionnaire <sup>b</sup>	450k	1.9 (0.9)	0.1, (0.0,0.3)
NFBC1966 <sup>66</sup>	740	740	56.9	29.9	23.8	31.0 (0)	1 item, from TCI-NS4 <sup>c</sup>	450k	0.8 (0.4)	0.6 (0.0,10)
NFBC1986 <sup>66</sup>	517	517	53.8	36.7	41.9	16.0 (0)	ASR <sup>31</sup>	450k	4.3 (2.6)	0.6 (0.0,10)
NTR <sup>67</sup>	2059	2049	69.2	18.3	22.5	36.4 (12.0)	ASR <sup>31</sup>	450k	2.8 (3.1)	-2.6 (-10.0, 8.0)
SATSA <sup>68</sup>	377	377	60.2	17.0	4.0	70.2 (9.7)	1 item, from EAS <sup>69,70d</sup>	450k	2.0 (1.07)	-2.0 (-10.0,5.0)
YFS <sup>71</sup>	181	181	63.0	30.9	27.5	19.2 (3.3)	Hunter-Wolf <sup>34,35</sup>	450k	3.5 (0.9)	0
Cord blood										
ALSPAC <sup>57</sup>	808	808	50.4	0	0	0 (0)	SDQ <sup>29</sup>	450k	1.5 (1.4)	-6.8 (-6.8,-6.8)
GECKO <sup>72</sup>	196	186	51.5	0	0	0 (0)	SDQ <sup>29</sup>	450k	1.1 (1.4)	-5.9(-5.1,- 6.9)
Generation R <sup>73</sup>	806	718	49.4	0	0	0 (0)	CBCL <sup>28</sup>	450k	5.2 (5.1)	-5.9 (-5.2, -8.3)
INMA <sup>74</sup>	385	385	48.8	0	0	0 (0)	SDQ <sup>29</sup>	450k	1.8 (1.7)	-6,9 (-8,3,-6,2)
Poseidon <sup>75</sup>	230	230	54.3	0	0	0 (0)	CBCL <sup>28</sup>	450k	9.4 (5.9)	-3.8 (-3.6, -4)

2 ALSPAC=Avon Longitudinal Study of Parents and Children, Dunedin= Dunedin Multidisciplinary Health and Development Study, E-Risk= E-Risk Twin Study,  
 3 FinnTwin12=Finnish Twin Cohort, GS:SFHS= Generation Scotland: Scottish Family Health Study, GLAKU= Glycyrrhizin in Licorice cohort,HELIX=The Human Early-Life  
 4 Exposome, LLD= LifeLines-DEEP, NFBC1966=Northern Finland Birth Cohort 1966, NFBC1986= Northern Finland Birth Cohort 1986, NTR= Netherlands Twin  
 5 Register, SATSA= Swedish Adoption/Twin Study of Aging, YFS= Young Finns Study, GECKO= Groningen Expert Center for Kids with Obesity, Generation  
 6 R=Generation R Study, INMA= The INMA-INfancia y Medio Ambiente (Environment and Childhood) Project. Poseidon= Pre-, peri- and postnatal Stress in  
 7 human and non-human offspring: A translational approach to study Epigenetic Impact on DepressiON. SDQ= Strengths and Difficulties Questionnaire (SDQ),  
 8 conduct problems. MPQ= Multidimensional Personality Questionnaire aggression. DSM-IV Conduct Disorder =DSM-IV Conduct Disorder Symptom Scale.  
 9 MNPI= Multidimensional Peer Nomination Inventory, aggression. CBCL= Child Behavior Checklist, Aggressive Behavior scale. GHQ= General Health  
 10 Questionnaire. TCI-NS4=Temperament and Character Inventory- Novelty Seeking. ASR=Adult self-report, aggression scale. EAS= Emotionality, activity,  
 11 sociability scale. Hunter-Wolf= Hunter-Wolf aggressive behavior scale. <sup>a</sup>Have you recently been getting edgy and bad-tempered? <sup>b</sup>Could you indicate to what

1 extent the following statement applies to you? I am known for being short-tempered and irritable? <sup>c</sup>I lose my temper more quickly than most people. <sup>d</sup>People  
2 think I am hot-tempered an temperamental. <sup>e</sup>Age at DNA sample collection. <sup>f</sup>Time between DNA sample collection and phenotype measure: DNA minus  
3 phenotype. M1=model1. M2=model2. Model 1 included the following predictors: aggressive behavior, sex, age at blood sampling (if applicable), white blood  
4 cell percentages (measured or imputed), and technical covariates. Model 2 included the same predictors as model 1 plus BMI and smoking status (smoking  
5 status was not included in model 2 in cohorts that assessed DNA methylation in children). \*The percentage shows current and former smokers combined.  
6 NA=not assessed.

7

8

9

10

11

1 **Table 2** Follow-up cohorts

<b>Cohort</b>	<b>Type</b>	<b>DNA methylation</b>	<b>Phenotype</b>	<b>N</b>	<b>% female</b>	<b>Mean age (SD)</b>	<b>Aggression mean (SD)</b>
NeuroIMAGE <sup>40</sup>	Clinical cohort; ADHD	Illumina EPIC	Callous Traits	71	28.2	21 (2.9)	9.3 (4.4)
FemNAT-CD <sup>41</sup>	Clinical cohort; Conduct disorder	HpaII methylation Sequencing	Case-control status	Total: 100 Cases: 50 Controls:50	100	Cases: 16.1(1.6) Controls: 15.8(1.5)	NA
ACTION – NTR <sup>42</sup>	Twin cohort, selected on aggression (high-low)	Illumina EPIC	CBCL aggression	1237	47.4	9.6 (1.9)	5.0 (5.4)
ACTION-Curium-LUMC <sup>42</sup>	Clinical cohort; children with severe and complex mental health problems	Illumina EPIC	CBCL aggression	172	25.6	9.6 (1.7)	13.1 (7.6)

2 NeuroIMAGE=The NeuroIMAGE study is a follow-up of the Dutch part of the International Multicenter ADHD  
3 Genetics (IMAGE) project. FemNAT-CD= Neurobiology and Treatment of Adolescent Female Conduct Disorder.  
4 ACTION= Aggression in Children: Unraveling gene-environment interplay to inform Treatment and InterventiON  
5 strategies. NTR= Netherlands Twin Register.

6

7

8

9

10

11

12

13

14

15

16

17

18

1 **Table 3** Top-sites associated with aggressive behavior from the combined EWAMA of cord  
 2 blood and peripheral blood (FDR 5%)

CpG ID	CHR	Position*	Gene	Gene Expression Associated With CpGs	N M1	Zscore M1	P value M1	Zscore M2	P value M2
cg05575921	5	373378	AHRR	EXOC3	15666	-8.995	2.36E-19	-4.159	3.20E-05
cg21161138	5	399360	AHRR	EXOC3	15661	-7.573	3.66E-14	-3.155	1.61E-03
cg26703534	5	377358	AHRR	EXOC3	15665	-6.695	2.16E-11	-2.058	3.96E-02
cg14753356	6	30720108		FLOT1	15666	-6.672	2.52E-11	-3.342	8.33E-04
cg22132788	7	45002486	MYO1G		10847	6.313	2.74E-10	3.637	2.76E-04
cg06126421	6	30720080		FLOT1, TUBB, LINC00243	10864	-6.196	5.78E-10	-2.154	3.13E-02
cg07826859	7	45020086	MYO1G		10863	-6.017	1.77E-09	-3.665	2.48E-04
cg09935388	1	92947588	GFI1		15661	-5.906	3.51E-09	-3.222	1.27E-03
cg25648203	5	395444	AHRR	EXOC3	15657	-5.583	2.37E-08	-2.233	2.55E-02
cg12062133	8	142548839			14482	5.462	4.71E-08	4.881	1.06E-06
cg05951221	2	233284402			10864	-5.443	5.25E-08	-1.679	9.32E-02
cg14817490	5	392920	AHRR	EXOC3	10863	-5.407	6.43E-08	-2.152	3.14E-02
cg14179389	1	92947961	GFI11		15666	-5.35	8.80E-08	-3.888	1.01E-04
cg05432213	15	35086985	ACTC1		15666	5.144	2.68E-07	4.87	1.12E-06
cg03636183	19	17000585	F2RL3	F2RL3	15666	-5.124	3.00E-07	-0.909	3.63E-01
cg09022230	7	5457225	TNRC18		15666	-5.071	3.95E-07	-3.024	2.49E-03
cg12803068	7	45002919	MYO1G	RP4-647J21.1	15666	4.93	8.22E-07	2.493	1.27E-02
cg23916896	5	368804	AHRR		15652	-4.915	8.86E-07	-2.332	1.97E-02
cg04180046	7	45002736	MYO1G	RP4-647J21.1	15665	4.884	1.04E-06	2.989	2.80E-03
cg02228160	5	143192067	HMHB1		10852	4.867	1.13E-06	3.451	5.58E-04
cg03519879	14	74227499	C14orf43		15663	-4.859	1.18E-06	-3.609	3.08E-04
cg00310412	15	74724918	SEMA7A	SEMA7A	15666	-4.854	1.21E-06	-2.608	9.11E-03
cg13165240	17	3715743	C17orf85		15664	4.838	1.31E-06	4.436	9.16E-06
cg02895948	1	208204062	PLXNA2	PLXNA2	10865	-4.811	1.51E-06	-4.448	8.68E-06
cg12147622	10	74021432			15662	-4.796	1.62E-06	-3.312	9.26E-04
cg26883434	5	111091560	C5orf13		14540	4.773	1.81E-06	4.739	2.15E-06
cg03991871	5	368447	AHRR	EXOC3	10857	-4.753	2.01E-06	-2.374	1.76E-02
cg06946797	16	11422409			15666	-4.75	2.03E-06	-3.317	9.08E-04
cg00891184	1	10272185	KIF1B		15662	4.746	2.07E-06	4.421	9.82E-06
cg09243533	1	19281949	IFFO2		15666	-4.74	2.14E-06	-4.003	6.26E-05
cg03935116	12	31476565	FAM60A	FAM60A	15665	-4.735	2.19E-06	-3.664	2.48E-04
cg11554391	5	321320	AHRR		15666	-4.717	2.39E-06	-2.731	6.32E-03
cg19825437	3	169383292			15664	-4.663	3.12E-06	-3.094	1.98E-03
cg00624037	12	89315201			15663	4.633	3.61E-06	4.081	4.49E-05
cg01940273	2	233284934			15666	-4.621	3.82E-06	-0.305	7.61E-01
cg25949550	7	145814306	CNTNAP2		15666	-4.615	3.94E-06	-2.333	1.96E-02
cg23067299	5	323907	AHRR		10865	4.615	3.94E-06	3.21	1.33E-03
cg04387347	16	88537187	ZFPM1		9563	4.603	4.17E-06	2.678	7.42E-03

cg02325250	5	131409289	<i>CSF2</i>		15664	-4.597	4.28E-06	-3.635	2.78E-04
cg14560430	3	32863175	<i>TRIM71</i>		15665	-4.569	4.90E-06	-3.924	8.70E-05
cg03844894	15	35086967	<i>ACTC1</i>		15666	4.567	4.94E-06	4.176	2.97E-05
cg21611682	11	68138269	<i>LRP5</i>		14859	-4.561	5.08E-06	-1.721	8.53E-02
cg20673321	19	48049233	<i>ZNF541</i>		15666	4.538	5.67E-06	4.672	2.98E-06

1 \*Genome build 37. M1=Model 1. M2=Model 2

2

3

4

5

6

7

8

9

10

11

12

13

14

15

16

17

18

19

20

21

22

23

24

## 1    **References**

- 2    1.    Tremblay, R. E. Physical Aggression During Early Childhood: Trajectories and  
3       Predictors. *Pediatrics* (2004). doi:10.1542/peds.114.1.e43
- 4    2.    Tremblay, R. E., Vitaro, F. & Côté, S. M. Developmental Origins of Chronic Physical  
5       Aggression: A Bio-Psycho-Social Model for the Next Generation of Preventive  
6       Interventions. *Annu. Rev. Psychol.* (2017). doi:10.1146/annurev-psych-010416-  
7       044030
- 8    3.    Tateno, A., Jorge, R. E. & Robinson, R. G. Clinical Correlates of Aggressive Behavior  
9       After Traumatic Brain Injury. *J. Neuropsychiatry Clin. Neurosci.* (2014).  
10      doi:10.1176/jnp.15.2.155
- 11   4.    Volicer, L. & Hurley, A. C. Review Article: Management of Behavioral Symptoms in  
12      Progressive Degenerative Dementias. *Journals Gerontol. Ser. A Biol. Sci. Med. Sci.*  
13      (2003). doi:10.1093/gerona/58.9.m837
- 14   5.    Moore, T. M. *et al.* Drug abuse and aggression between intimate partners: A meta-  
15      analytic review. *Clinical Psychology Review* (2008). doi:10.1016/j.cpr.2007.05.003
- 16   6.    Boles, S. M. & Miotto, K. Substance abuse and violence: A review of the literature.  
17      *Aggression and Violent Behavior* (2003). doi:10.1016/S1359-1789(01)00057-X
- 18   7.    Bonder, M. J. *et al.* Disease variants alter transcription factor levels and methylation of  
19      their binding sites. *Nat. Genet.* **49**, 131–138 (2017).
- 20   8.    Weaver, I. C. G. *et al.* Epigenetic programming by maternal behavior. *Nat. Neurosci.*  
21      **7**, 847–54 (2004).
- 22   9.    Provencal, N. *et al.* The Signature of Maternal Rearing in the Methylome in Rhesus  
23      Macaque Prefrontal Cortex and T Cells. *J. Neurosci.* (2012).  
24      doi:10.1523/jneurosci.1470-12.2012
- 25   10.   Martin, E. M. & Fry, R. C. Environmental Influences on the Epigenome: Exposure-  
26      Associated DNA Methylation in Human Populations. *Annu. Rev. Public Health* (2018).  
27      doi:10.1146/annurev-publhealth-040617-014629
- 28   11.   Qi, T. *et al.* Identifying gene targets for brain-related traits using transcriptomic and  
29      methylomic data from blood. *Nat. Commun.* (2018). doi:10.1038/s41467-018-04558-1
- 30   12.   Tsai, P. C. *et al.* Smoking induces coordinated DNA methylation and gene expression  
31      changes in adipose tissue with consequences for metabolic health. *Clin. Epigenetics*  
32      (2018). doi:10.1186/s13148-018-0558-0
- 33   13.   Joubert, B. R. *et al.* DNA Methylation in Newborns and Maternal Smoking in  
34      Pregnancy: Genome-wide Consortium Meta-analysis. *Am. J. Hum. Genet.* **98**, 680–  
35      696 (2016).
- 36   14.   Jovanova, O. S. *et al.* DNA methylation signatures of depressive symptoms in middle-



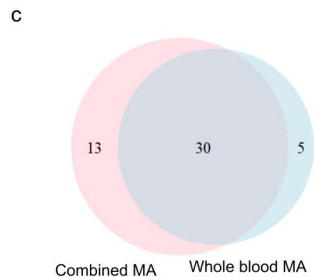
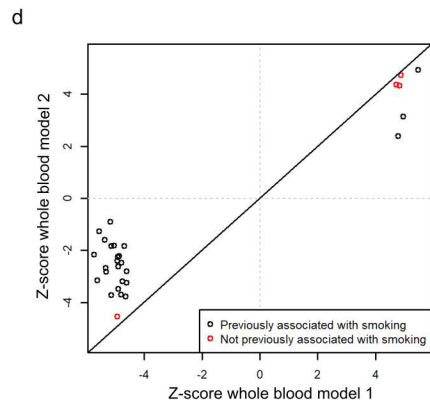
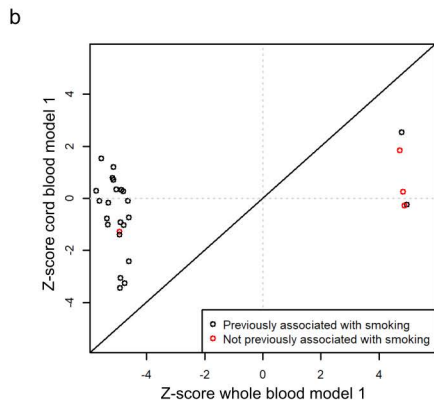
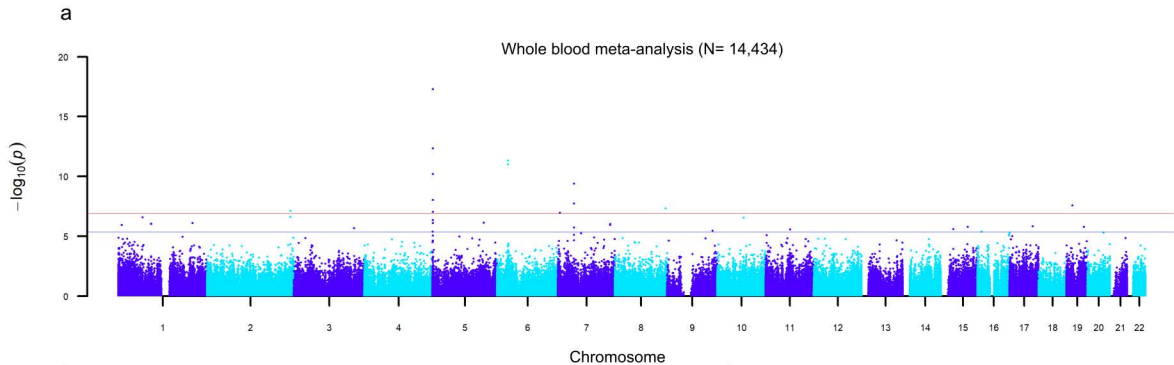
- 1 aged and elderly persons: Meta-analysis of multiethnic epigenome-wide studies.
- 2 *JAMA Psychiatry* (2018). doi:10.1001/jamapsychiatry.2018.1725
- 3 15. Jia, T. *et al.* Epigenome-wide meta-analysis of blood DNA methylation and its
- 4 association with subcortical volumes: findings from the ENIGMA Epigenetics
- 5 Working. *bioRxiv* (2018). doi:<https://doi.org/10.1101/460444>
- 6 16. Shah, S. *et al.* Improving Phenotypic Prediction by Combining Genetic and Epigenetic
- 7 Associations. *Am. J. Hum. Genet.* (2015). doi:10.1016/j.ajhg.2015.05.014
- 8 17. McCartney, D. L. *et al.* Epigenetic prediction of complex traits and death. *Genome*
- 9 *Biol.* (2018). doi:10.1186/s13059-018-1514-1
- 10 18. Guillemin, C. *et al.* DNA methylation signature of childhood chronic physical
- 11 aggression in T cells of both men and women. *PLoS One* (2014).
- 12 doi:10.1371/journal.pone.0086822
- 13 19. Cecil, C. A. M. *et al.* Neonatal DNA methylation and early-onset conduct problems: A
- 14 genome-wide, prospective study. *Dev. Psychopathol.* (2018).
- 15 doi:10.1017/S095457941700092X
- 16 20. Cecil, C. A. M. *et al.* DRD4 methylation as a potential biomarker for physical
- 17 aggression: An epigenome-wide, cross-tissue investigation. *Am. J. Med. Genet. Part B*
- 18 *Neuropsychiatr. Genet.* (2018). doi:10.1002/ajmg.b.32689
- 19 21. Mitjans, M. *et al.* Violent aggression predicted by multiple pre-adult environmental hits.
- 20 *Mol. Psychiatry* (2018). doi:10.1038/s41380-018-0043-3
- 21 22. van Dongen, J. *et al.* Epigenome-wide association study of aggressive behavior. *Twin*
- 22 *Res. Hum. Genet.* **18**, 686–698 (2015).
- 23 23. Lubke, G. H., McArtor, D. B., Boomsma, D. I. & Bartels, M. Genetic and environmental
- 24 contributions to the development of childhood aggression. *Dev. Psychol.* **54**, 39–50
- 25 (2018).
- 26 24. Zannas, A. S. *et al.* Lifetime stress accelerates epigenetic aging in an urban, African
- 27 American cohort: relevance of glucocorticoid signaling. *Genome Biol.* **16**, 266 (2015).
- 28 25. Jovanovic, T. *et al.* Exposure to Violence Accelerates Epigenetic Aging in Children.
- 29 *Sci. Rep.* **7**, (2017).
- 30 26. Han, L. K. M. *et al.* Epigenetic Aging in Major Depressive Disorder. *Am. J. Psychiatry*
- 31 *appi.ajp.2018.1* (2018). doi:10.1176/appi.ajp.2018.17060595
- 32 27. Ori, A. P. S. *et al.* Schizophrenia is characterized by age- and sex-specific effects on
- 33 epigenetic aging. *bioRxiv* 727859 (2019). doi:10.1101/727859
- 34 28. Thomas M Achenbach, C. E. Manual for the Child Behavior Checklist. *Burlingt.* **7**,
- 35 (1991).
- 36 29. Goodman, R. The Strengths and Difficulties Questionnaire: a research note. *J. Child*
- 37 *Psychol. Psychiatry* **38**, 581–6 (1997).

- 1 30. Pulkkinen, L., Kaprio, J. & Rose, R. J. Peers, teachers and parents as assessors of  
2 the behavioural and emotional problems of twins and their adjustment: the  
3 Multidimensional Peer Nomination Inventory. *Twin Res.* **2**, 274–285 (1999).
- 4 31. Achenbach, T. M. & Rescorla, L. a. Manual for the ASEBA Adult Forms & Profiles.  
5 *English* University of Vermont, Research Center for Childre (2003).  
6 doi:10.1017/CBO9781107415324.004
- 7 32. American Psychiatric Association. Diagnostic and Statistical Manual of Mental  
8 Disorders, (DSM IV). *Washingt. DC, APA Fourth Ed.*, 915 (1994).
- 9 33. Tellegen, A. *et al.* Personality Similarity in Twins Reared Apart and Together. *J. Pers.*  
10 *Soc. Psychol.* **54**, 1031–1039 (1988).
- 11 34. Wolf, T. M., Sklov, M. C., Wenzl, P. A., Hunter, S. M. & Berenson, G. S. Validation of a  
12 measure of type A behavior pattern in children: Bogalusa heart study. *Child Dev.* **53**,  
13 126–135 (1982).
- 14 35. Ravaja, N., Keltikangas-Järvinen, L. & Keskivaara, P. Type A factors as predictors of  
15 changes in the metabolic syndrome precursors in adolescents and young adults--a 3-  
16 year follow-up study. *Health Psychol.* **15**, 18–29 (1996).
- 17 36. Chen, Y. *et al.* Discovery of cross-reactive probes and polymorphic CpGs in the  
18 Illumina Infinium HumanMethylation450 microarray. *Epigenetics* **2294**, (2013).
- 19 37. Iterson, M. Van, Zwet, E. W. Van & Heijmans, B. T. Controlling bias and inflation in  
20 association studies using the empirical null distribution. *Genome Biol.* 1–13 (2017).  
21 doi:10.1186/s13059-016-1131-9
- 22 38. Willer, C. J., Li, Y. & Abecasis, G. R. METAL: Fast and efficient meta-analysis of  
23 genomewide association scans. *Bioinformatics* (2010).  
24 doi:10.1093/bioinformatics/btq340
- 25 39. Li, M. *et al.* EWAS Atlas: A curated knowledgebase of epigenome-wide association  
26 studies. *Nucleic Acids Res.* (2019). doi:10.1093/nar/gky1027
- 27 40. von Rhein, D. *et al.* The NeuroIMAGE study: a prospective phenotypic, cognitive,  
28 genetic and MRI study in children with attention-deficit/hyperactivity disorder. Design  
29 and descriptives. *Eur. Child Adolesc. Psychiatry* (2015). doi:10.1007/s00787-014-  
30 0573-4
- 31 41. Freitag, C. M. *et al.* Conduct disorder in adolescent females: current state of research  
32 and study design of the FemNAT-CD consortium. *European Child and Adolescent*  
33 *Psychiatry* (2018). doi:10.1007/s00787-018-1172-6
- 34 42. Hagenbeek, F. A. *et al.* Urinary Amine and Organic Acid Metabolites Evaluated as  
35 Markers for Childhood Aggression: The ACTION Biomarker Study. *Front. Psychiatry*  
36 **11**, (2020).
- 37 43. van Dongen, J. *et al.* Genetic and environmental influences interact with age and sex

- 1 in shaping the human methylome. *Nat. Commun.* **7**, 11115 (2016).
- 2 44. Hannon, E., Lunnon, K., Schalkwyk, L. & Mill, J. Interindividual methylomic variation  
3 across blood, cortex, and cerebellum: Implications for epigenetic studies of  
4 neurological and neuropsychiatric phenotypes. *Epigenetics* **10**, 1024–1032 (2015).
- 5 45. Ripke, S. *et al.* Biological insights from 108 schizophrenia-associated genetic loci.  
6 *Nature* **511**, 421–427 (2014).
- 7 46. Karlsson Linnér, R. *et al.* Genome-wide association analyses of risk tolerance and  
8 risky behaviors in over 1 million individuals identify hundreds of loci and shared  
9 genetic influences. *Nat. Genet.* (2019). doi:10.1038/s41588-018-0309-3
- 10 47. Joehanes, R. *et al.* Epigenetic Signatures of Cigarette Smoking. *Circ. Cardiovasc.*  
11 *Genet.* **9**, 436–447 (2016).
- 12 48. Su, K. Y. *et al.* Perinatal polychlorinated biphenyls and polychlorinated dibenzofurans  
13 exposure are associated with DNA methylation changes lasting to early adulthood:  
14 Findings from Yucheng second generation. *Environ. Res.* (2019).  
15 doi:10.1016/j.envres.2019.01.001
- 16 49. Juricek, L. & Coumoul, X. The aryl hydrocarbon receptor and the nervous system.  
17 *International Journal of Molecular Sciences* (2018). doi:10.3390/ijms19092504
- 18 50. Chepelev, N. L., Moffat, I. D., Bowers, W. J. & Yauk, C. L. Neurotoxicity may be an  
19 overlooked consequence of benzo[a]pyrene exposure that is relevant to human health  
20 risk assessment. *Mutation Research - Reviews in Mutation Research* **764**, 64–89  
21 (2015).
- 22 51. Niu, Q., Zhang, H., Li, X. & Li, M. Benzo[a]pyrene-induced neurobehavioral function  
23 and neurotransmitter alterations in coke oven workers. *Occup. Environ. Med.* **67**, 444–  
24 448 (2010).
- 25 52. Malanchini, M. *et al.* Aggressive behaviour in childhood and adolescence: The role of  
26 smoking during pregnancy, evidence from four twin cohorts in the EU-ACTION  
27 consortium. *Psychol. Med.* (2019). doi:10.1017/S0033291718001344
- 28 53. Richardson, T. G. *et al.* An integrative approach to detect epigenetic mechanisms that  
29 putatively mediate the influence of lifestyle exposures on disease susceptibility. *Int. J.*  
30 *Epidemiol.* **48**, 887–898 (2019).
- 31 54. Wiklund, P. *et al.* DNA methylation links prenatal smoking exposure to later life health  
32 outcomes in offspring. *Clin. Epigenetics* **11**, (2019).
- 33 55. Odintsova, V. V. *et al.* GENOMICS OF HUMAN AGGRESSION: current state of  
34 genome-wide studies and an automated systematic review tool. *Psychiatr. Genet.* **29**,  
35 170–190 (2019).
- 36 56. Tielbeek, J. J. *et al.* Genome-Wide Association Studies of a Broad Spectrum of  
37 Antisocial Behavior. *JAMA Psychiatry* **74**, 1242 (2017).

- 1 57. Relton, C. L. *et al.* Data Resource Profile: Accessible Resource for Integrated  
2 Epigenomic Studies (ARIES). *Int. J. Epidemiol.* (2015). doi:10.1093/ije/dyv072
- 3 58. Poulton, R., Moffitt, T. E. & Silva, P. A. The Dunedin Multidisciplinary Health and  
4 Development Study: overview of the first 40 years, with an eye to the future. *Social*  
5 *Psychiatry and Psychiatric Epidemiology* **50**, 679–693 (2015).
- 6 59. Moffitt, T. E. *et al.* Teen-aged mothers in contemporary Britain. *J. Child Psychol.*  
7 *Psychiatry Allied Discip.* **43**, 727–742 (2002).
- 8 60. Kaprio, J. The Finnish Twin Cohort Study: An update. *Twin Res. Hum. Genet.* **16**,  
9 157–162 (2013).
- 10 61. Smith, B. H. *et al.* Cohort profile: Generation scotland: Scottish family health study  
11 (GS: SFHS). The study, its participants and their potential for genetic research on  
12 health and illness. *Int. J. Epidemiol.* (2013). doi:10.1093/ije/dys084
- 13 62. Goldberg, D. P. & Hillier, V. F. A scaled version of the General Health Questionnaire.  
14 *Psychol. Med.* (1979). doi:10.1017/S0033291700021644
- 15 63. Strandberg, T. E., Järvenpää, A. L., Vanhanen, H. & McKeigue, P. M. Birth outcome in  
16 relation to licorice consumption during pregnancy. *Am. J. Epidemiol.* (2001).  
17 doi:10.1093/aje/153.11.1085
- 18 64. Vrijheid, M. *et al.* The human early-life exposome (HELIX): Project rationale and  
19 design. *Environmental Health Perspectives* **122**, 535–544 (2014).
- 20 65. Tigchelaar, E. F. *et al.* Cohort profile: LifeLines DEEP, a prospective, general  
21 population cohort study in the northern Netherlands: study design and baseline  
22 characteristics. *BMJ Open* **5**, e006772 (2015).
- 23 66. Rantakallio, P. The longitudinal study of the northern Finland birth cohort of 1966.  
24 *Paediatr. Perinat. Epidemiol.* **2**, 59–88 (1988).
- 25 67. Boomsma, D. I. *et al.* Netherlands Twin Register: From Twins to Twin Families. *Twin*  
26 *Res. Hum. Genet.* **9**, 849–857 (2006).
- 27 68. Pedersen, N. L. *et al.* The Swedish Adoption Twin Study of Aging: An update. *Acta*  
28 *Genet. Med. Gemellol. (Roma)*. **40**, 7–20 (1991).
- 29 69. Buss, A. H. & Plomin, R. *Temperament Early developing personality traits*.  
30 (Lawrence Erlbaum Associates Inc, 1984).
- 31 70. Plomin, R., Pedersen, N. L., McClearn, G. E., Nesselroade, J. R. & Bergeman, C. S.  
32 EAS temperaments during the last half of the life span: twins reared apart and twins  
33 reared together. *Psychol. Aging* **3**, 43–50 (1988).
- 34 71. Raitakari, O. T. *et al.* Cohort profile: The cardiovascular risk in young Finns study. *Int.*  
35 *J. Epidemiol.* **37**, 1220–1226 (2008).
- 36 72. L'Abée, C. *et al.* Cohort Profile: The GECKO Drenthe study, overweight programming  
37 during early childhood. *Int. J. Epidemiol.* **37**, 486–489 (2008).

- 1 73. Kruithof, C. J. *et al.* The Generation R Study: Biobank update 2015. *Eur. J. Epidemiol.*
- 2 **29**, 911–927 (2014).
- 3 74. Guxens, M. *et al.* Cohort profile: The INMA-Infancia y Medio Ambiente-(environment
- 4 and childhood) project. *Int. J. Epidemiol.* **41**, 930–940 (2012).
- 5 75. Witt, S. H. *et al.* Impact on birth weight of maternal smoking throughout pregnancy
- 6 mediated by DNA methylation. *BMC Genomics* (2018). doi:10.1186/s12864-018-4652-
- 7 7
- 8



**e**

Trait(s)	$-\log_{10}(p)$	CpGs	%
smoking	>309	37	0.3
maternal smoking	197.0	22	0.7
educational attainment	172.5	21	29.2
HIV frailty	101.0	18	7.9
smoking cessation	89.8	18	2.8
lung function	88.2	14	13.3
cognitive function	74.8	15	3.8
alcohol consumption	55.0	12	0.7
perinatal PCB and PCDF exposure	48.7	7	35.0
lgG glycosylation	46.9	6	85.7

# Prediction of ASR aggression by DNA methylation scores

