

1 **Running title:**

2 **DNA methylation differences and ART**

3

4 **Title:**

5 **DNA methylation differences at birth after conception through**  
6 **assisted reproductive technologies**

7 Elmar W. Tobi<sup>1,2,\*</sup>, Catarina Almqvist<sup>3,4</sup>, Anna Hedman<sup>5</sup>, Ellika Andolf<sup>5</sup>, Jan Holte<sup>6,7,8</sup>, Jan I. Olofsson<sup>9</sup>,

8 Håkan Wråmsby<sup>10</sup>, Margaretha Wråmsby<sup>11</sup>, Göran Pershagen<sup>12</sup>, Bastiaan T. Heijmans<sup>2</sup>, Anastasia N.

9 Iliadou<sup>3</sup>

- 10 1. Periconceptional Epidemiology, Department of Obstetrics and Gynaecology, Erasmus MC, University  
11 Medical Center Rotterdam, Dr. Molewaterplein 50, 3015GE Rotterdam, The Netherlands  
12 2. Molecular Epidemiology, Department of Biomedical Data Sciences, Leiden University Medical Center,  
13 2300RC Leiden, The Netherlands  
14 3. Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, 171 77 Stockholm, Sweden  
15 4. Pediatric Allergy and Pulmonology Unit, Astrid Lindgren Children's Hospital, 171 76 Stockholm, Sweden  
16 5. Department of Clinical Sciences, Danderyd hospital, Karolinska Institutet 182 88 Stockholm Sweden  
17 6. Carl von Linné Clinic, Uppsala, Sweden  
18 7. Department of Women's and Children's Health, Uppsala University, 751 85 Uppsala, Sweden  
19 8. Center for Reproductive Biology in Uppsala, University of Agricultural Sciences and Uppsala University,  
20 Uppsala, Sweden.  
21 9. Department of Women's and Children's Health, Karolinska Institutet, 171 77 Stockholm, Sweden.  
22 10. Livio Fertilitetscentrum Kungsholmen, S:t Görans Sjukhus, 112 81 Stockholm, Sweden  
23 11. Livio Fertilitetscentrum Gärdet Storängsvägen 10, 115 42 Stockholm  
24 12. Institute of Environmental Medicine, Karolinska Institutet, 171 77 Stockholm, Sweden

25 \*corresponding author: Assoc Prof Anastasia Nyman  
26 Department of medical epidemiology and biostatistics,  
27 Karolinska Institutet  
28 Email: [Anastasia.nyman@ki.se](mailto:Anastasia.nyman@ki.se)  
29 Tel:+ 46 708 267874  
30

31 **Abstract**

32 **Study Question:** What is the relationship with assisted reproductive technology (ART) and DNA

33 methylation (DNAm) patterns in cord-blood and is there any difference between in vitro fertilisation

34 (IVF) and intra-cytoplasmic sperm injection (ICSI)?

35 **Summary Answer:** We identify nineteen CpGs at which DNAm was associated with being conceived via  
36 assisted reproductive technology (ART). No difference was found between IVF and ICSI.

37 **What is known already:** Prior studies on either IVF or ICSI show conflicting outcomes as both wide-  
38 spread effects on DNAm and highly localized associations have been reported. No study on both IVF and  
39 ICSI and genome-wide neonatal DNAm has been performed.

40 **Study design, size, duration:** We measured 87 infants conceived with in vitro fertilisation (IVF) or intra-  
41 cytoplasmic sperm injection (ICSI) and 70 who were not in a cross-sectional study.

42 **Participants/materials, setting, methods:** Participants were from the UppstART study, which were  
43 recruited from fertility and reproductive health clinics, and the Born into Life cohort, which is created  
44 from the larger LifeGene study. We measured DNAm from DNA extracted from cord blood collected at  
45 birth with the Illumina Infinium HumanMethylation450k BeadChip micro-array. Group differences in  
46 DNA methylation at individual CpG dinucleotides (CpGs) were determined using robust linear models  
47 and post-hoc Tukey's tests.

48 **Main results and the role of chance:** We found no association with global methylation levels, imprinted  
49 loci and meta-stable epialleles. In contrast, we identify nineteen CpGs at which DNA methylation was  
50 associated with being conceived via ART (effect estimates: 0.5-4.9%,  $P_{FDR} < 0.05$ ), but no difference was  
51 found between IVF and ICSI. The associated CpGs map to genes related to brain function/development  
52 or genes connected to the plethora of conditions linked to subfertility, but functional annotation did not  
53 point to likely functional consequences.

54 **Limitations, reasons for caution:** We measured DNAm in cord blood and not at later ages and/or other  
55 tissues and findings remain to be replicated in an independent study.

56 **Wider implications of the findings:** We find ART is associated with DNA methylation differences in cord  
57 blood, but these differences are limited in number, effect size and with unknown functional  
58 consequences in adult blood. We did not find indications for a disparity between IVF and ICSI.

59 **Study funding/competing interest(s):** The authors declare no competing interests. The funders  
60 had no role in the study, study design, analysis and decision to publish.

61

## 62 **Keywords**

63 DNA methylation; ART; IVF; ICSI; EWAS; cord-blood; epigenetics

64

65

66

67

68

69

## 70 Introduction

71 The usage of assisted reproductive technology (ART) is increasing world-wide and its possible health  
72 consequences are a topic of intense study. Evidence is building for an association of ART with long-term  
73 health outcomes for the offspring, including autism spectrum disorders (Liu *et al.*, 2017) and  
74 cardiovascular health (higher blood pressure, suboptimal cardiac diastolic function and vessel thickness)  
75 (Guo *et al.*, 2017). A body of evidence links ART to short term health effects, including low birth weight,  
76 placenta associated anomalies (Vermeij *et al.*, 2019) and pregnancy complications (Qin *et al.*, 2016),  
77 congenital malformations and imprinting disorders (Turkgeldi *et al.*, 2016). However, it is still debated if  
78 these associations stem from the application of ART or the underlying infertility leading couples to ART  
79 (Luke *et al.*, 2016). Animal studies provided potential molecular mechanisms for these observations by  
80 highlighting that the ART procedure may induce changes to epigenetic marks (Morgan *et al.*, 2008;  
81 Wang *et al.*, 2010). Epigenetic marks, such as DNA methylation (DNAm), influence the transcription  
82 potential of genomic regions (Jaenisch and Bird, 2003). Human studies likewise point towards a link  
83 between early development, DNAm and (late-life) phenotypes, but the causality of these associations  
84 remains unknown (Tobi *et al.*, 2018).

85 To date, several studies on DNAm and ART have been performed. Multiple studies have focused  
86 on candidate gene regions and/or global methylation levels (Lazaraviciute *et al.*, 2014; Canovas *et al.*,  
87 2017). Genome-wide efforts have focused on samples taken from extra-embryonic lineages (Xu *et al.*,  
88 2017; Choufani *et al.*, 2018) and when focused on material from the infant itself suffered from a small  
89 sample size (Melamed *et al.*, 2015) or batch effects correlating with ART status (Estill *et al.*, 2016). Larger  
90 studies have focused on specific ART techniques, namely intra-cytoplasmic sperm injection (ICSI) (El Hajj  
91 *et al.*, 2017) or in vitro fertilisation (IVF) (Castillo-Fernandez *et al.*, 2017) or IVF and less invasive gamete  
92 intra-fallopian transfer (GIFT) coupled with intra-uterine insemination (IUI) (Novakovic *et al.*, 2019). It  
93 has been hypothesized that the technique used may matter for the possible consequences on DNAm

94 patterns (Loke and Craig, 2016). Overall, these studies report conflicting outcomes, with two reporting  
95 wide-spread associations between ART and (cord-)blood DNAm at birth (El Hajj *et al.*, 2017) which  
96 appear to fade with age (Novakovic *et al.*, 2019), while another study reported a DNA methylation  
97 difference at a single genomic locus only (Castillo-Fernandez *et al.*, 2017).

98           We undertook an epigenome wide association study on ART, comparing DNAm patterns in cord  
99 blood of children conceived via ART (N=87) (Iliadou *et al.*, 2019) with that of children from medically  
100 unassisted conceptions (henceforth “MUC”, N=70) (Almqvist *et al.*, 2011). Since we had detailed  
101 information on the ART technique used for 77 of the children conceived via ART we investigated  
102 possible different outcomes for IVF and ICSI on DNAm which has been hypothesized to be important  
103 (Loke and Craig, 2016), but is yet to be tested. In addition, we explored the possible functional  
104 consequences of these methylation differences using external datasets.

105

## 106 Methods

107

### 108 Study Subjects

#### 109 ART controls: Born into Life

110 The prospective longitudinal birth cohort study Born into Life was originally created from the larger  
111 LifeGene study (Almqvist *et al.*, 2011). Between the years 2010 and 2012, pregnant women who were  
112 already participating in the LifeGene study and living in Stockholm County were recruited to Born into  
113 Life (Smew *et al.*, 2018). The inclusion criteria were that they had responded to baseline questionnaires  
114 from the LifeGene study, were pregnant, and gave informed consent. They were recruited both before  
115 and after 10-14 gestational weeks, but no later than 26-28 weeks. Originally 107 pregnant women were  
116 included in Born into Life. We only included the 77 women for which cord blood had been successfully  
117 collected in this study.

118 Women in the Born into Life were asked to complete questionnaires regarding pregnancy,  
119 lifestyle and health at 10-14 and 26-28 gestational weeks. Data on maternal self-reported smoking  
120 during pregnancy and body mass index (BMI), in kg/m<sup>2</sup> obtained at the first antenatal care visit in  
121 gestational week week 5-12 were collected from birth records. Data regarding highest attained  
122 educational level, ranging from mandatory secondary school to high school, university or other, were  
123 retrieved from the baseline LifeGene questionnaires. Maternal age at delivery, in years, was calculated  
124 from mothers' date of birth. Parity and data regarding the infants' sex, gestational age in weeks and  
125 mode of delivery, defined as vaginal delivery or Caesarean section, were also collected from the birth  
126 records. Cord blood and placenta samples were obtained from delivery. Cord blood was aspirated 2  
127 minutes after birth by the assisting midwife into a test tube and kept in room temperature awaiting  
128 transportation. The birth records regarding both mother and child at delivery were collected from  
129 Danderyd Hospital, Stockholm.

## 130 ART cases: UppstART

131 The UppstART study has been described in detail elsewhere (Iliadou *et al.*, 2019). Participants of the  
132 UppstART study were recruited from three of the four fertility and reproductive health clinics in  
133 Stockholm (one public, two private) and one private clinic in Uppsala county, which also serves a large  
134 volume of patients from Stockholm. Recruitment took place from September 2011 to December 2013.  
135 IVF treatment(s) of the participants were followed until December 2014 or drop-out/consent withdrawn  
136 (n=4), which ever came first. The participants were asked to answer a web-based baseline questionnaire  
137 within a few days of their clinic visit and prior to their IVF treatment start, which included an extensive  
138 list of questions on sociodemographic, anthropometric and life-style factors. Once the IVF treatment  
139 began and the participants reached the stage of oocyte retrieval, they were asked to respond to a  
140 second online follow-up questionnaire, with a shorter version of the baseline questionnaire, to identify  
141 any changes in lifestyle factors since the initiation of their treatment. Staff at seven delivery units in  
142 Stockholm and Uppsala were recruited to assist in collecting samples from UppStART participants during  
143 delivery. Delivery clinics were provided with a sample collection kit including tubes for collection of  
144 cord-blood for DNA extraction (EDTA). Samples were stored in -20°C freezers at the delivery units until  
145 they were collected by UppStART study staff and deposited into the KI Biobank.

## 146 DNA methylation measurements

147 Genome-wide DNA methylation data were generated using the Illumina Infinium Human Methylation  
148 450K BeadChip (450k array). A total of 500ng of genomic DNA isolated from cord blood was bisulfite  
149 treated using the EZ-96 DNA methylation kit (Zymo Research, Orange County, USA). We used the D-  
150 optimum criterion to assign samples over two 96-well plates and individual 450k arrays, ensuring even  
151 distributions of ART cases and controls, ART methods (IVF or ICSI), gender, gestational age and birth  
152 month across the two 96-well bisulfite plates and each 450k array. The 450K arrays were measured at  
153 ServiceXS (Leiden, The Netherlands). The quality of the generated 450K array data was assessed using  
154 both sample dependent and sample independent quality metrics using the *Bioconductor* package

155 *MethylAid* (van Iterson *et al.*, 2014) with default settings. We used the *Bioconductor* package *omicsPrint*  
156 (van Iterson *et al.*, 2018) to check for sample duplications and admixtures and sample sexes were  
157 checked using the X-chromosomal CpGs. One sample was not the correct gender and was deleted. We  
158 used principal component analysis (PCA) and hierarchical clustering on the raw autosomal beta-values  
159 to search for outliers and suspect patterns in the dataset, finding none. The cell proportions of cord  
160 blood were imputed using IDOL (Koestler *et al.*, 2016) on the “FlowSorted.CordBloodCombined.450k”  
161 reference set using the `estimateCellCounts2()` function in the *FlowSorted.Blood.EPIC* R package (Salas *et*  
162 *al.*, 2018; Gervin *et al.*, 2019). Normalization of the dataset was performed by NOOB background and  
163 color correction in combination with Functional Normalization (Fortin *et al.*, 2014) using four principal  
164 components (eigenvalue > 1). All measurements with <3 beads (0.05% of probes), < 1 intensity value  
165 (0.012% of probes) and a detection p-value >0.01 (0.11% of CpGs) were set as missing. The  
166 measurement success rate per sample was >99%. CpGs with probes that did not map to unique genomic  
167 locations (Chen *et al.*, 2013) or with a < 95% measurement success rate (0.23% of CpGs) were then  
168 removed. We used custom scripts that add on the functions from the *minfi* package (Aryee *et al.*, 2014)  
169 and implements parallelization where possible (<https://github.com/molepi/DNAarray>).  
170 Outlier detection was performed by PCA and hierarchical clustering on the autosomal beta-values.  
171 Hierarchical clustering identified one possible outlier. This individual had a very low gestational age (30  
172 weeks), the lowest in the dataset, and this clustering is therefore likely based on biological reasons.  
173 High-quality DNA methylation data were obtained for 157 individuals, 70 MUC from Born into Life, 6 ART  
174 from Born into Life (specific type of ART unknown), 81 ART from UppStART (specific type of ART  
175 unknown for N=4) for a total of 441,836 autosomal CpGs. The datasets generated and analyzed during  
176 the current study are not publicly available due to Swedish privacy and data safety laws but are available  
177 from ANI and CA on reasonable request and after meeting legal requirements.

178

## 179 Measures of global methylation

180 Genome-wide average DNA methylation (GWAM) (Li *et al.*, 2018) was calculated by averaging all beta-  
181 value measurements across the autosomes for each individual. The R package *REMP* (Zheng *et al.*, 2017)  
182 was used to infer DNA methylation at either *ALU* or *LINES-1* sequences and then the average across all  
183 *ALU* or *LINES-1* elements was calculated for each neonate.

## 184 TFBS enrichment

185 We calculated transcription factor binding site enrichments using the R package *PWME* (Stojnic and  
186 Diez, 2018) using binding motifs from motifDB. We calculated position weight matrices using the DNA  
187 base background frequencies calculated for the CpGs tested with 25bp flanking sequences. Enrichment  
188 was tested relative to this background.

## 189 Statistics

190 We compared unassisted and ART cohort descriptives via Kruskal Wallis (maternal BMI, maternal SES,  
191 years till index pregnancy, CD4T, NK and nucleated red blood cells), t-test (maternal and gestational age,  
192 birthweight, CD8T, Monocytes, granulocytes and B cell proportions) assuming equal or unequal variance  
193 where appropriate and chi-square tests (child sex, parity and maternal smoking). DNA methylation was  
194 always analyzed as beta-value which is reported as % (e.g. x100) in Tables. GWAM, *ALU* and *LINES-1*  
195 methylation levels were compared via t.test (unassisted vs ART) or ANOVA (medically unassisted  
196 fertilization versus IVF versus ICSI).

197 We used the R package *cate* (Wang and Zhao, 2015) to identify latent variables (e.g. “hidden  
198 variables”) influencing the relation between DNA methylation and ART (yes/no) or ART method  
199 (IVF/ICSI/unassisted). Two latent variables were identified ( $P < 0.001$ ) for ART and four latent variables for  
200 ART method ( $P < 0.001$ ). In both cases the first two latent variables correlated with the proportion of  
201 nucleated red blood cells (nRBC) in blood ( $\rho > 0.94$ ,  $P < 0.001$ ) and the other two latent variables  
202 correlated with CD4T or B cell proportions in blood and several batch effects ( $\rho > |0.3|$ ,  $P < 0.001$ )

203 arguing together with directed acyclic graphs (DAG) (*Supplemental text*), for a basic model adjusting for  
204 batch effects and imputed cell proportions.

205 We performed epigenome-wide association analyses (EWAS) on being conceived by ART or not  
206 using robust linear regression (rlm) from the R *MASS* package with White's estimator for robust  
207 standard errors as implemented in the R package *sandwich* (Zeileis, 2006) (which leads to a model  
208 robust for outlying beta-values and heteroscedasticity). To test for differences in neonates conceived via  
209 IVF or ICSI or without medical assistance we used type II ANOVA with White's estimator for robust  
210 standard errors to account for any heteroscedasticity and unequal variances between groups using the R  
211 *car* package. We employed heteroscedasticity robust Tukey tests to test for DNA methylation  
212 differences between IVF, ICSI and unassisted conceptions post-hoc. In all instances we used the beta-  
213 value of an individual CpG dinucleotide as the outcome and ART status as the dependent. We adjusted  
214 for the sex of the individual, height of the sample on the 450k array glass slide (continuous variable from  
215 1-6), scan batch and bisulfite plate and the imputed cell proportions (CD4T, CD8T, Mono, NK, Bcell,  
216 Gran, nRBC). We ran additional sensitivity analyses by omitting the cell proportions or expanding the  
217 adjustment with minimal adjustment sets coming from DAG analysis with the following models, i: years  
218 till the index pregnancy, maternal age, pre-pregnancy BMI, educational attainment as proxy for socio-  
219 economic status, smoking history ii) gestational length and parity. In addition, we performed look-ups in  
220 other published meta-EWAS of prenatal exposures.

221 Regions were tested with the same adjustments as defined above in a linear mixed model with a  
222 random effect for individual and a factorial co-variate denoting the specific CpG dinucleotide of each  
223 DNAm measurement in the R *lme4* package using compound symmetry as correlation structure and the  
224 R *lmerTest* package for the Satterthwaite's approximation of the degree of freedom of the fixed effects  
225 within each linear mixed model.

226 All analyses were performed in R (version 3.6.1) (R core team, 2019). All p-values reported are  
227 two-sided and multiple testing correction is done using FDR (false discovery rate).

## 228 Ethical approval

229 Ethical approval for Born into Life was granted by the Regional Ethics Review Board in Stockholm,  
230 Sweden. Informed consent from both parents was obtained for all study participants. The UppstART  
231 study has been approved by the regional ethics review board at the Karolinska Institutet (Dnr 2011/230-  
232 31/1, Dnr 2011/1427-32, Dnr 2012/131-32, Dnr 2012/792-32, 2013/1700-32, 2014/1956-32, Dnr  
233 2015/1604-32). Women and their partners were approached by the clinic nurse and asked to participate  
234 in the study. To facilitate the process of informed consent, the couple was provided with information  
235 approved by the regional ethical board, both verbally and in written format about the purpose of the  
236 study, methods, possible risks, and that participation was voluntary. Additionally, participants were  
237 informed that they could withdraw from the study at any time with no impact to their medical care. The  
238 requirement for inclusion in the study was an understanding of the Swedish language and exclusion was  
239 the use of donor gametes.

240

## 241 Results

### 242 Study subjects

243 Pre-processing and normalization resulted in DNAm data of 157 infants from cord blood, 87 of which  
244 were conceived with the help of ART (Table 1). We had detailed information on the ART technique used  
245 for 77 of the 87 newborns conceived by ART and with DNAm data, of which 44 were conceived via IVF  
246 and 33 via ICSI. The mothers that conceived with the help of ART were not different in body mass index  
247 (BMI,  $P=0.25$ ) and age ( $P=0.25$ ), but had a lower socio-economic status ( $P<0.001$ ) and more were former  
248 smokers ( $P<0.001$ ) than mothers with a MUC. It took on average two years longer for the mothers who  
249 conceived through ART to become pregnant than for mothers with MUC ( $P<0.001$ ). There was no  
250 difference between ART and MUC in length of gestation ( $P=0.16$ ) and parity ( $P=0.20$ ). The percentage of  
251 male newborns in the MUC group was higher (61.4% vs 50.6%), although this difference was not  
252 significant ( $P=0.23$ ).

253 DNAm is a key determinant of cell identity (Jaenisch and Bird, 2003) and cord blood consists of  
254 multiple cell types that influence DNAm (Jones *et al.*, 2007). Seven major cell types, including nucleated  
255 red blood cells, were imputed from the genome-wide DNAm data of the newborns (Gervin *et al.*, 2019).  
256 In concordance with a prior report on unassisted and ICSI newborns (El Hajj *et al.*, 2017), there were no  
257 differences in these seven cell proportions between ART and MUC cord blood samples ( $P_{\text{nominal}}>0.13$ ,  
258 Supplemental Figure S1).

### 259 Global DNAm comparison between ART and unassisted conceptions

260 First, we investigated genome-wide average DNA methylation (GWAM) (Li *et al.*, 2018) in the cord blood  
261 of neonates conceived via IVF or ICSI or MUC. There was no difference between these 3 groups ( $P=0.89$ ,  
262 Supplemental Figure S2). Next, we investigated the genome-wide average DNA methylation of ALU and

263 LINES-1 elements, an often used proxy for global DNA methylation (Zheng *et al.*, 2017), finding no  
264 difference between groups ( $P > 0.2$ , Figure S3).

## 265 DNAm comparison between ART and medically unassisted conceptions

266 We compared DNA methylation at 441,836 autosomal CpGs between 87 ART and 70 MUC neonates  
267 (Supplemental Figure S4). Nineteen CpGs were associated with being conceived via ART in a model  
268 adjusting for sex, batches and cell heterogeneity ( $P_{\text{FDR}} < 0.05$ , Figure S5). The associations extended to  
269 neighboring probes (Figure 1) for the proximal promoter of *AK054845*, a gene of unknown function (five  
270 CpGs  $P_{\text{nominal}} < 0.05$ ), and the proximal promoter of a smaller variant of *GABRB3* (NM\_001191321.2) which  
271 is expressed in the brain (eight CpGs  $P_{\text{nominal}} < 0.05$ ) and has a neuro-developmental role (Tanaka *et al.*,  
272 2012). Indeed, several CpGs can be linked to genes with roles in brain function or (early) brain  
273 development. *ZNF771* (cg06450634), as a transcription factor (TF), is a driving factor behind a large gene  
274 network in the brain (Maulik *et al.*, 2018) while *CYFIP1* (cg13051607) (Hsiao *et al.*, 2016), *GABRB3*  
275 (cg01251603, cg15066197, cg14859324) (Tanaka *et al.*, 2012) and *SHANK1* (cg14560133) (Sungur *et al.*,  
276 2018) are all autism spectrum disorder candidate genes, a late life phenotype tentatively associated  
277 with ART (Liu *et al.*, 2017).

278 Omitting the adjustment for cell heterogeneity had little to no effect on the effect estimates for  
279 all nineteen CpGs (Supplemental Table S1) just like additional adjustment for maternal and pregnancy  
280 characteristics (Supplemental Table S2). The nineteen CpGs were not among those previously reported  
281 in other EWASs (FDR corrected P-value of  $< 0.05$ ) with prenatal smoking (Joubert *et al.*, 2016b), folate  
282 use (Joubert *et al.*, 2016a), maternal hypertension and pre-eclampsia (Kazmi *et al.*, 2019), maternal BMI  
283 (Sharp *et al.*, 2017) and birthweight (Küpers *et al.*, 2019). Moreover, a look-up of these previously  
284 reported CpGs in our study did not yield more nominally associated CpGs with ART than may be  
285 expected by chance ( $P > 0.07$ ).

## 286 DNA methylation comparison between IVF, ICSI and unassisted conceptions

287 To investigate any differences between IVF and ICSI we performed post-hoc Tukey's tests for DNA  
288 methylation differences between neonates conceived via IVF and ICSI for these nineteen CpGs. No  
289 differences were found ( $P_{\text{FDR-19 tests}} > 0.15$ ). Next, we extended our analysis to all CpGs by performing an  
290 ANOVA analysis to test for differences in DNA methylation across the 441,836 autosomal CpGs between  
291 neonates conceived with IVF (N=44) or ICSI (N=33) or MUC (N=70). This yielded five CpGs already  
292 identified in the EWAS for ART status and in all cases the difference was between the MUC and both IVF  
293 and ICSI, thus finding no evidence for IVF or ICSI specific associations.

## 294 Functional annotation

295 We performed KEGG and Gene Ontology tests (Phipson *et al.*, 2015), finding no enrichment for the  
296 genes linked to nineteen CpGs associated with ART, nor when we relaxed the significance threshold to  
297 suggestive associations ( $P < 10^{-5}$ , N=37). We did not find significant enrichments of single transcription  
298 factor (TF) binding motifs (Stojnic and Diez, 2018), although multiple CpGs overlapped TF binding sites of  
299 TFs with a role in early development like *TFAP2C* (Sharma *et al.*, 2016) overlapping cg27266479 mapping  
300 to *H6PD* and *RAX* (Bennett *et al.*, 2008) overlapping cg16771467 to *ATP8B1*. Look-up of the nineteen  
301 CpGs in reference data (Bonder *et al.*, 2017) uncovered no link between DNAm and gene-expression in  
302 adult blood. We measured DNAm in cord blood and average DNAm levels may vary between tissues, but  
303 still reflect variation in other tissues (Slieker *et al.*, 2013). There was little to no correlation of  
304 methylation at the nineteen CpGs between adult whole blood and tissues from sixteen cadavers (Slieker  
305 *et al.*, 2013; Relton *et al.*, 2015), with the notable exception of cg14560133 (*SHANK1*,  $\beta = -1.42\%$   
306 (SE=0.3%),  $P = 2.1 \times 10^{-6}$ ) which showed moderate correlation with 5/8 tissues ( $\rho > 0.53$ ,  $P < 0.04$ ), Figure  
307 2). The three CpG dinucleotides at *GABRB3* (cg01251603, cg15066197, cg14859324), a gene active in the  
308 brain, showed weak to moderate correlation between adult blood and four brain regions ( $\rho = 0.39$ -  
309 0.62,  $P < 0.001$ , N=71-74) in another reference dataset (Hannon *et al.*, 2015).

## 310 Imprinted regions and metastable epialleles

311 Multiple candidate gene studies have looked at DNA methylation or gene-expression differences of  
312 imprinted genes between ART and MUC with mixed results. Therefore, we looked at 374 CpGs on the  
313 450k array known to overlap imprinted differentially methylated regions (DMRs) (Yuen *et al.*, 2011).  
314 There was no overlap with the CpGs showing a (suggestive) association with ART in our study ( $P < 10^{-5}$ ,  
315  $N = 22$ ). These 374 CpGs were distributed over 59 regions. There was no difference in DNAm between  
316 MUC and ART infants when we tested each of these 59 regions ( $P_{FDR} > 0.86$ ). Similarly, so-called meta-  
317 stable epialleles (MEs) are hypothesized to be especially sensitive to the early prenatal environment  
318 (Kessler *et al.*, 2018). There were 187 CpGs overlapping MEs, none of which were associated with ART  
319 ( $P_{FDR} > 0.06$ ).

## 320 Prior studies

321 We performed a look-up of CpGs and regions associated in prior studies with one form of ART. The sole  
322 region associated with IVF in the meDIP sequencing study from Castillo-Fernandez *et al.* was not covered  
323 by our genome-scale screen with the 450k array (Castillo-Fernandez *et al.*, 2017). Next, El Hajj *et al.*  
324 found 4730 CpG dinucleotides associated with ICSI after FDR adjustment (which are unavailable as  
325 Supplemental information) of which two regions, out of five regions selected for validation with  
326 pyrosequencing, were validated in an additional sample (El Hajj *et al.*, 2017). The seven CpGs in these  
327 two regions which cover the proximal promoters of *ATG4C* and *SNORD114-9* were not associated with  
328 ART in our study ( $P > 0.23$ ) nor with ICSI only (post hoc Tukey's test:  $P > 0.40$ ).

329 Using the 850k array, Novakovic *et al.* identified 2340 CpG dinucleotides at which DNA  
330 methylation in Guthrie card blood was associated with IVF and GIFT coupled with IUI (Novakovic *et al.*,  
331 2019), 1228 of which were covered in our 450k study. Of these, 38 were nominally associated ( $P < 0.05$ )  
332 and with the same direction of effect in our study. This decreased to 12 loci when we tested for DNA  
333 methylation differences between MUC and IVF only. Novakovic *et al.* also identified three differentially

334 methylated regions where the difference in methylation associated with IVF remained in an adult  
335 sample. CpG dinucleotides at *CHRNE* (chr17:4803506–4805392), *PRSS16* (chr6:27185896–27186199)  
336 and *TMEM18* (chr2:731073–732037) showed a similar direction of effect in our study as that of  
337 Novakovic *et al.* and the association was almost nominally significant for *PRSS16* (cg10279314:  
338  $\beta=2.3\%[SE=1.2\%]$ ,  $P=0.057$ , cg09395805:  $\beta=3.1[SE=1.6]$ ,  $P=0.049$ , cg07555084:  $\beta=2.2[SE=1.2]$ ,  $P=0.074$ ).

339

340

## 341 Discussion

342 We identified nineteen CpGs at which DNAm was associated with ART and the association extended to  
343 neighboring CpG dinucleotides at *AK054845* and *GABRB3*. We found no indication for a disparity  
344 between IVF or ICSI on DNAm patterns. The associations were robust to adjustment for cellular  
345 heterogeneity and maternal characteristics and did not overlap with those loci associated with other  
346 common prenatal exposures. We found no evidence that imprinted loci and meta-stable epialleles are  
347 especially sensitive to ART. The nineteen CpGs can be annotated to relevant genes, but the functional  
348 implications of variation in DNA methylation at these CpGs remains to be elucidated.

349 The nineteen CpG dinucleotides could be mapped to fifteen genes. Four of the nineteen were  
350 located in a proximal promoter and here the neighboring CpG dinucleotides were also nominally  
351 associated with ART. In other cases the associations were either limited to one CpG dinucleotide, which  
352 may stem from the sparse coverage of the 450K array across the methylome and in other cases is  
353 consistent with the fact that DNAm acts through the altering of the binding potential of a specific TFBS  
354 (Bonder *et al.*, 2017). Indeed, none of the CpGs were located in a CpG island, which is consistent with  
355 data from other exposures showing that associations between DNAm and prenatal and/or  
356 environmental conditions are enriched at regulatory regions like enhancers and CpG poor promoters  
357 that show intermediate levels of methylation (Tobi *et al.*, 2014).

358 Several of the CpGs can be linked to genes with a role in brain function and/or development.  
359 *ZNF771* (cg06450634) has been identified as a driving factor behind a large gene network in the brain  
360 (Maulik *et al.*, 2018) and *CYFIP1* (cg13051607) (Hsiao *et al.*, 2016), *GABRB3* (cg01251603,  
361 cg15066197,cg14859324) (Tanaka *et al.*, 2012) and *SHANK1* (cg14560133) (Sungur *et al.*, 2018) are all  
362 autism spectrum disorder candidate genes, a late life phenotype tentatively associated with ART (Liu *et*  
363 *al.*, 2017). In addition, several CpG dinucleotides can be linked to genes connected to the plethora of  
364 conditions that can be linked to subfertility. *H6PD* (cg27266479) is a candidate gene for polycystic ovary

365 syndrome (PCOS) (Martínez-García *et al.*, 2012). *INTS1* (cg03207674) has a crucial role in the developing  
366 blastocyst, as inhibition of *INTS1* function causes growth arrest (Hata and Nakayama, 2007). *NOS3*  
367 (cg00478390) knockout mice are used as an *in vivo* model of (recurrent) embryo loss, as nitric oxide  
368 metabolism plays an important role in implantation (Pallares and Gonzalez-Bulnes, 2010).

369         The latter may hint that the various medical reasons for ART may underlie the associations,  
370 rather than the ART process itself, which is a key question in the study of ART (Luke *et al.*, 2016). We  
371 employed directed acyclic graphs (DAG) (Krieger and Davey Smith, 2016) and DAG indicated that and  
372 ascertainment of a direct effect is possible with a small minimal adjustment set. There was little to no  
373 effect on the effect estimates from adjustment for various maternal characteristics, gestational age and  
374 the number of years until pregnancy (as a proxy for in-/subfertility). In addition, we did not find any  
375 evidence for different or stronger effects of ICSI on DNAm, as has been hypothesized (Loke and Craig,  
376 2016), and the fertility clinics from which we recruited all used the same culturing media for ICSI and  
377 IVF. The last two items highlighted might argue that not the (reasons for) infertility, but the ART process  
378 itself may explain the associations.

379         Our analysis did not find wide-spread genome-wide differences like other genome-scale studies  
380 to date (El Hajj *et al.*, 2017; Novakovic *et al.*, 2019), but is consistent with a genome-scale study on only  
381 IVF using immunoprecipitation of methylated DNA with next generation sequencing (Castillo-Fernandez  
382 *et al.*, 2017). We extensively compared individual results, as far as possible, and found little overlap. This  
383 may be due to study power, the different scope of the measurement techniques used (Castillo-  
384 Fernandez *et al.*, 2017) and the different sources of genomic DNA (Guthrie card versus whole cord  
385 blood) (Novakovic *et al.*, 2019). In addition, it is possible that different media was used between  
386 countries or that the fact that only a small proportion of our IVF and ICSI group consisted of embryos  
387 that were frozen and/or cultured extensively in comparison to other studies. The influence of this latter

388 aspect we were underpowered to test as most ART procedures entailed fresh embryos that were not  
389 cultured extensively *in vitro*.

390 Our study is one of the larger genome-scale DNAm studies on ART to date and we are the first to  
391 study both IVF and ICSI on this scale, nonetheless our study has important limitations to consider. First,  
392 DNAm is one of the drivers of cell identity (Jaenisch and Bird, 2003) and we measured cord blood, which  
393 may be a tissue less relevant in relation to study outcomes although it may still mark processes in  
394 relevant tissues due to mitotic inheritance (Heijmans and Mill, 2012). We used the latest methods to  
395 impute the seven major cell types in cord-blood (Gervin *et al.*, 2019) and in conjuncture with an earlier  
396 report (El Hajj *et al.*, 2017) these imputed cell types did not differ between infants conceived with or  
397 without ART and adjustment did not alter the effect estimates. However, most of these major cell types  
398 consist of subsets of more specialized cells (subsequently showing unique DNAm profiles at increasingly  
399 select loci) and their influence could not be investigated. However, none of the nineteen CpG  
400 dinucleotides is linked to a gene with a role in (auto)immune function or hematopoiesis. Another major  
401 driver of DNAm variation is genetic variation (Bonder *et al.*, 2017) and we did not control for genetic  
402 variation in our analyses. Nine out of nineteen CpG dinucleotides did not have a single nucleotide  
403 polymorphism (SNP) influencing the DNA methylation levels, as determined in a large meta-analysis  
404 (Bonder *et al.*, 2017) and none of these 450k probes had an overlap with a single SNP arguing that  
405 genetic heterogeneity is unlikely to underlie our findings. In addition, the influence of the prenatal  
406 environment on DNAm variation may be independent and additive of genetic variation (Tobi *et al.*,  
407 2012).

408 Also important is to consider the fact that an ART population is inherently different from the  
409 general population (Luke *et al.*, 2016). Despite controlling for possible confounding factors, and cross-  
410 checking our results with other EWAS on prenatal complications/exposures, unresolved confounding  
411 remains a possibility. Our study differs with some of the earlier described in ART studies, which is useful

412 for triangulation between studies (Lawlor *et al.*, 2017), as our control group has a higher, rather than a  
413 lower, SES as is normal in most of studies on ART.

414 The data presented showcases modest and specific DNAm differences associated with ART, of  
415 which the functional relevance in adult tissues is unknown. We did not find any difference in DNAm  
416 patterns between IVF and ICSI. Our study found little evidence for the hypothesis that ART, be it IVF or  
417 ICSI, leads to wide-spread disturbances of DNAm patterns nor for the hypothesis that ICSI has a  
418 different/larger relation with DNAm patterns. Our findings warrant cautious interpretation and awaits  
419 confirmation, given the sample size, tissue studied and the unknown functional consequences of the  
420 identified DNAm differences.

421

422

## 423 Author contributions

424 Conceptualization: ANI, CA. Methodology: EWT, BTH. Investigation: EWT, BTH, ANI, CA. Formal Analysis:  
425 EWT. Validation: EWT. Resources: ANI, AH, CA, GP, EA, JIO, MW, HW. Data curation: ANI, AH. Writing –  
426 original draft: EWT. Writing – Review & Editing: BTH, ANI, AH, CA, GP, EA, JIO, MW, HW. Visualization:  
427 EWT. Supervision: BTH, ANI. Project administration: ANI, AH. Funding Acquisition: ANI, CA, GP, EA

## 428 Acknowledgements

429 We wish to thank the Biobank at the Karolinska Institutet for professional biobank service.

## 430 Competing interests

431 The authors declare no conflicts of interest

432

433

434

## 435 Abbreviations

436 ART: assisted reproductive technology

437 IVF: *in vitro* fertilisation

438 ICSI: intra-cytoplasmic sperm injection

439 DNAm: DNA methylation

440 EWAS: epigenome-wide association analysis

441 CpGs: CpG dinucleotides

442 IUI: intra-uterine insemination

443 GIFT: gamete intra-fallopian transfer

444 MUC: medically unassisted conceptions

445 450K array: Illumina Infinium Human Methylation 450K BeadChip

446 PCA: principal component analysis

447 TFBS: transcription factor binding site

448 TF: transcription factor

449 SNP: single nucleotide polymorphism

450 SES: socio-economic status

451 BMI: body mass index

## 452 Funding

453 EWT was supported by a VENI grant from the Netherlands Organization for Scientific Research  
454 (91617128) and JPI-H2020 Joint Programming Initiative a Healthy Diet for a Healthy Life (JPI HDHL)  
455 under proposal number 655 (PREcisE Project) through ZonMw (529051023). Financial support was  
456 provided from the European Union's Seventh Framework Program IDEAL (259679), the Swedish  
457 Research Council (K2011-69X-21871-01-6, 2011-3060 and 2018-02640), and the Strategic Research  
458 Program in Epidemiology Young Scholar Awards, Karolinska Institute (to ANI) and through the Swedish  
459 Initiative for Research on Microdata in the Social And Medical Sciences (SIMSAM) framework grant no  
460 340-2013-5867, grants provided by the Stockholm County Council (ALF-projects), the Strategic Research  
461 Program in Epidemiology at Karolinska Institutet, the Swedish Heart-Lung Foundation and Danderyd  
462 University Hospital (Stockholm, Sweden). The funders had no role in study design, data collection,  
463 analysis, decision to publish or preparation of the manuscript.

464

## 465 References

466

- 467 Almqvist C, Adami HO, Franks PW, Groop L, Ingelsson E, Kere J, Lissner L, Litton JE, Maeurer M,  
468 Michaëlsson K, *et al.* LifeGene - A large prospective population-based study of global relevance. *Eur*  
469 *J Epidemiol* 2011;**26**:67–77.
- 470 Aryee MJ, Jaffe AE, Corrada-Bravo H, Ladd-Acosta C, Feinberg AP, Hansen KD, Irizarry RA. Minfi: a  
471 flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation  
472 microarrays. *Bioinformatics* 2014;**30**:1363–1369. England.
- 473 Bennett RL, Blalock WL, Choi E-J, Lee YJ, Zhang Y, Zhou L, Paul Oh S, Stratford May W. RAX is required for  
474 fly neuronal development and mouse embryogenesis. *Mech Dev* [Internet] 2008;**125**:777–785.
- 475 Bonder MJ, Luijk R, Zhernakova D V, Moed M, Deelen P, Vermaat M, Iterson M van, Dijk F van, Galen M  
476 van, Bot J, *et al.* Disease variants alter transcription factor levels and methylation of their binding  
477 sites. *Nat Genet* 2017;**49**:131–138. University of Groningen, University Medical Center Groningen,  
478 Department of Genetics, Groningen, the Netherlands Molecular Epidemiology Section, Department  
479 of Medical Statistics and Bioinformatics, Leiden University Medical Center, Leiden, the  
480 Netherlands, United States.
- 481 Canovas S, Ross PJ, Kelsey G, Coy P. DNA Methylation in Embryo Development: Epigenetic Impact of ART  
482 (Assisted Reproductive Technologies). *BioEssays* 2017;**39**:1–11.
- 483 Castillo-Fernandez JE, Loke YJ, Bass-Stringer S, Gao F, Xia Y, Wu H, Lu H, Liu Y, Wang J, Spector TD, *et al.*  
484 DNA methylation changes at infertility genes in newborn twins conceived by in vitro fertilisation.  
485 *Genome Med* [Internet] 2017;**9**:28. Genome Medicine.
- 486 Chen YA, Lemire M, Choufani S, Butcher DT, Grafodatskaya D, Zanke BW, Gallinger S, Hudson TJ,

487           Weksberg R. Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium  
488           HumanMethylation450 microarray. *Epigenetics* 2013;**8**:203–209. Genetics and Genome Biology  
489           Program, Hospital for Sick Children, Toronto, ON, Canada.

490   Choufani S, Turinsky AL, Melamed N, Greenblatt E, Brudno M, Bérard A, Fraser WD, Weksberg R, Trasler  
491           J, Monnier P, *et al.* Impact of Assisted Reproduction, Infertility, Sex, and Paternal Factors on the  
492           Placental DNA Methylome. *Hum Mol Genet* [Internet] 2018;Available from:  
493           [http://fdslive.oup.com/www.oup.com/pdf/production\\_in\\_progress.pdf](http://fdslive.oup.com/www.oup.com/pdf/production_in_progress.pdf).

494   Estill MS, Bolnick JM, Waterland RA, Bolnick AD, Diamond MP, Krawetz SA. Assisted reproductive  
495           technology alters deoxyribonucleic acid methylation profiles in bloodspots of newborn infants.  
496           *Fertil Steril* [Internet] 2016;**106**:629-639.e10. Elsevier Inc.

497   Fortin J-P, Labbe A, Lemire M, Zanke BW, Hudson TJ, Fertig EJ, Greenwood CM, Hansen KD. Functional  
498           normalization of 450k methylation array data improves replication in large cancer studies. *Genome*  
499           *Biol* [Internet] 2014;**15**:503.

500   Gervin K, Salas LA, Bakulski KM, Zelm MC Van, Koestler DC, Wiencke JK, Duijts L, Moll HA, Kelsey KT,  
501           Kobor MS, *et al.* Systematic evaluation and validation of reference and library selection methods  
502           for deconvolution of cord blood DNA methylation data. *Clin Epigenetics* [Internet] 2019;**11**:125.

503   Guo XY, Liu XM, Jin L, Wang TT, Ullah K, Sheng JZ, Huang HF. Cardiovascular and metabolic profiles of  
504           offspring conceived by assisted reproductive technologies: a systematic review and meta-analysis.  
505           *Fertil Steril* [Internet] 2017;**107**:622-631.e5. Elsevier.

506   Hajj N El, Haertle L, Dittrich M, Denk S, Lehnen H, Hahn T, Schorsch M, Haaf T. DNA methylation  
507           signatures in cord blood of ICSI children. *Hum Reprod* [Internet] 2017;**32**:1761–1769. Oxford  
508           University Press.

509 Hannon E, Lunnon K, Schalkwyk L, Mill J. Interindividual methylomic variation across blood, cortex, and  
510 cerebellum: implications for epigenetic studies of neurological and neuropsychiatric phenotypes.  
511 *Epigenetics* [Internet] 2015;**10**:1024–1032.

512 Hata T, Nakayama M. Targeted disruption of the murine large nuclear KIAA1440/Ints1 protein causes  
513 growth arrest in early blastocyst stage embryos and eventual apoptotic cell death. *Biochim Biophys*  
514 *Acta* [Internet] 2007;**1773**:1039–1051.

515 Heijmans BT, Mill J. Commentary: The seven plagues of epigenetic epidemiology. *IntJEpidemiol*  
516 2012;**41**:74–78. Molecular Epidemiology, Leiden University Medical Center, Leiden, The  
517 Netherlands, Netherlands. bas.heijmans@lumc.nl.

518 Hsiao K, Harony-Nicolas H, Buxbaum JD, Bozdagi-Gunal O, Benson DL. Cyfip1 Regulates Presynaptic  
519 Activity during Development. *J Neurosci* [Internet] 2016;**36**:1564–1576.

520 Iliadou AN, Öberg AS, Pege J, Rodriguez-Wallberg KA, Olofsson JI, Holte J, Wramsby H, Wramsby M,  
521 Cnattingius S, Cesta CE. The Uppsala-Stockholm Assisted Reproductive Techniques (UppStART)  
522 study. *BMJ Open* 2019;**9**:. BMJ Publishing Group.

523 Iterson M van, Cats D, Hop P, Heijmans BT. omicsPrint: detection of data linkage errors in multiple omics  
524 studies. In Stegle O, editor. *Bioinformatics* [Internet] 2018;**34**:2142–2143.

525 Iterson M van, Tobi EW, Sliker RC, Hollander W den, Luijk R, Slagboom PE, Heijmans BT. MethyIAid:  
526 visual and interactive quality control of large Illumina 450k datasets. *Bioinformatics* [Internet]  
527 2014;**30**:3435–3437. Department of Molecular Epidemiology, Leiden University Medical Center,  
528 2333 ZC Leiden, the Netherlands Department of Molecular Epidemiology, Leiden University  
529 Medical Center, 2333 ZC Leiden, the Netherlands Department of Molecular Epidemiology, Leiden  
530 Un.

- 531 Jaenisch R, Bird A. Epigenetic regulation of gene expression: how the genome integrates intrinsic and  
532 environmental signals. *Nat Genet* 2003;**33 Suppl**:245–254. United States.
- 533 Jones SC, Jordan WJ, Meiners SJ, Miller AN, Methven AS. Fungal spore dispersal by the Eastern box turtle  
534 (*Terrapene carolina carolina*). *Am Midl Nat* [Internet] 2007;**157**:121–126.
- 535 Joubert BR, Dekker HT den, Felix JF, Bohlin J, Ligthart S, Beckett E, Tiemeier H, Meurs JB van, Uitterlinden  
536 AG, Hofman A, *et al*. Maternal plasma folate impacts differential DNA methylation in an  
537 epigenome-wide meta-analysis of newborns. *Nat Commun* [Internet] 2016a;**7**:10577.
- 538 Joubert BR, Felix JF, Yousefi P, Bakulski KM, Just AC, Breton C, Reese SE, Markunas CA, Richmond RC, Xu  
539 CJ, *et al*. DNA Methylation in Newborns and Maternal Smoking in Pregnancy: Genome-wide  
540 Consortium Meta-analysis. *Am J Hum Genet* [Internet] 2016b;**98**:680–696.
- 541 Kazmi N, Sharp GC, Reese SE, Vehmeijer FO, Lahti J, Page CM, Zhang W, Rifas-Shiman SL, Rezwan FI,  
542 Simpkin AJ, *et al*. Hypertensive Disorders of Pregnancy and DNA Methylation in Newborns.  
543 *Hypertens (Dallas, Tex 1979)* [Internet] 2019;**74**:375–383.
- 544 Kessler NJ, Waterland RA, Prentice AM, Silver MJ. Establishment of environmentally sensitive DNA  
545 methylation states in the very early human embryo. *Sci Adv* [Internet] 2018;**4**:eaat2624.
- 546 Koestler DC, Jones MJ, Usset J, Christensen BC, Butler RA, Kobor MS, Wiencke JK, Kelsey KT. Improving  
547 cell mixture deconvolution by identifying optimal DNA methylation libraries (IDOL). *BMC*  
548 *Bioinformatics* [Internet] 2016;**17**:120.
- 549 Krieger N, Davey Smith G. The tale wagged by the DAG: broadening the scope of causal inference and  
550 explanation for epidemiology. *Int J Epidemiol* [Internet] 2016;**45**:1787–1808. Oxford University  
551 Press.
- 552 Küpers LK, Monnereau C, Sharp GC, Yousefi P, Salas LA, Ghantous A, Page CM, Reese SE, Wilcox AJ,

- 553 Czamara D, *et al.* Meta-analysis of epigenome-wide association studies in neonates reveals  
554 widespread differential DNA methylation associated with birthweight. *Nat Commun* 2019;**10**.  
555 Nature Publishing Group.
- 556 Lawlor DA, Tilling K, Davey Smith G. Triangulation in aetiological epidemiology. *Int J Epidemiol* [Internet]  
557 2017;dyw314 Available from: [https://academic.oup.com/ije/article-](https://academic.oup.com/ije/article-lookup/doi/10.1093/ije/dyw314)  
558 [lookup/doi/10.1093/ije/dyw314](https://academic.oup.com/ije/article-lookup/doi/10.1093/ije/dyw314).
- 559 Lazaraviciute G, Kauser M, Bhattacharya S, Haggarty P, Bhattacharya S. A systematic review and meta-  
560 analysis of DNA methylation levels and imprinting disorders in children conceived by IVF/ICSI  
561 compared with children conceived spontaneously. *Hum Reprod Update* [Internet] 2014;**20**:840–  
562 852.
- 563 Li S, Wong EM, Dugué P-A, McRae AF, Kim E, Joo J-HE, Nguyen TL, Stone J, Dite GS, Armstrong NJ, *et al.*  
564 Genome-wide average DNA methylation is determined in utero. *Int J Epidemiol* [Internet]  
565 2018;**47**:908–916.
- 566 Liu L, Gao J, He X, Cai Y, Wang L, Fan X. Association between assisted reproductive technology and the  
567 risk of autism spectrum disorders in the offspring: a meta-analysis. *Sci Rep* [Internet] 2017;**7**:46207.  
568 Nature Publishing Group.
- 569 Loke YJ, Craig JM. Are the effects of IVF on DNA methylation driven by intracytoplasmic sperm injection  
570 and male infertility? *Epigenomics* [Internet] 2016;**8**:881–884. Future Medicine Ltd London, UK.
- 571 Luke B, Stern JE, Hornstein MD, Kotelchuck M, Diop H, Cabral H, Declercq ER. Is the wrong question  
572 being asked in infertility research? *J Assist Reprod Genet* [Internet] 2016;**33**:3–8. Springer New York  
573 LLC.
- 574 Martínez-García MA, San-Millán JL, Escobar-Morreale HF. The R453Q and D151A polymorphisms of

- 575 Hexose-6-Phosphate Dehydrogenase Gene (H6PD) influence the polycystic ovary syndrome (PCOS)  
576 and obesity. *Gene* [Internet] 2012;**497**:38–44.
- 577 Maulik U, Sen S, Mallik S, Bandyopadhyay S. Detecting TF-miRNA-gene network based modules for  
578 5hmC and 5mC brain samples: a intra- and inter-species case-study between human and rhesus.  
579 *BMC Genet* [Internet] 2018;**19**:9.
- 580 Melamed N, Choufani S, Wilkins-Haug LE, Koren G, Weksberg R. Comparison of genome-wide and gene-  
581 specific DNA methylation between ART and naturally conceived pregnancies. *Epigenetics* [Internet]  
582 2015;**10**:474–483. Taylor & Francis.
- 583 Morgan HD, Jin XL, Li A, Whitelaw E, O’Neill C. The culture of zygotes to the blastocyst stage changes the  
584 postnatal expression of an epigenetically labile allele, agouti viable yellow, in mice. *BiolReprod*  
585 2008;**79**:618–623. Royal North Shore Hospital, Disciplines of Medicine and Physiology, University of  
586 Sydney, Sydney, 2065 New South Wales, Australia.
- 587 Novakovic B, Lewis S, Halliday J, Kennedy J, Burgner DP, Czajko A, Kim B, Sexton-Oates A, Juonala M,  
588 Hammarberg K, *et al.* Assisted reproductive technologies are associated with limited epigenetic  
589 variation at birth that largely resolves by adulthood. *Nat Commun* [Internet] 2019;**10**:3922.  
590 Springer US.
- 591 Pallares P, Gonzalez-Bulnes A. The effect of embryo and maternal genotypes on prolificacy, intrauterine  
592 growth retardation and postnatal development of Nos3-knockout mice. *Reprod Biol* 2010;**10**:241–  
593 248. Elsevier B.V.
- 594 Phipson B, Maksimovic J, Oshlack A. missMethyl: an R package for analyzing data from Illumina’s  
595 HumanMethylation450 platform. *Bioinformatics* [Internet] 2015;**32**:btv560.
- 596 Qin J, Liu X, Sheng X, Wang H, Gao S. Assisted reproductive technology and the risk of pregnancy-related

597 complications and adverse pregnancy outcomes in singleton pregnancies: A meta-analysis of  
598 cohort studies. *Fertil Steril* 2016;**105**:73-85.e6. Elsevier Inc.

599 R core team. R: A language and environment for statistical computing. In Team RDC, editor. *R Found Stat*  
600 *Comput* [Internet] 2019; R Foundation for Statistical Computing Available from: [http://www.r-](http://www.r-project.org/)  
601 [project.org/](http://www.r-project.org/).

602 Relton CL, Gaunt T, McArdle W, Ho K, Duggirala A, Shihab H, Woodward G, Lyttleton O, Evans DM, Reik  
603 W, *et al.* Data Resource Profile: Accessible Resource for Integrated Epigenomic Studies (ARIES).  
604 *IntJEpidemiol* 2015;**44**:1181–1190. MRC Integrative Epidemiology Unit, and School of Social and  
605 Community Medicine, University of Bristol, Bristol, UK, Institute of Genetic Medicine, Newcastle  
606 University, Newcastle upon Tyne, UK, [caroline.relton@bristol.ac.uk](mailto:caroline.relton@bristol.ac.uk) MRC Integrative Epidemiology  
607 Un.

608 Salas LA, Koestler DC, Butler RA, Hansen HM, Wiencke JK, Kelsey KT, Christensen BC. An optimized  
609 library for reference-based deconvolution of whole-blood biospecimens assayed using the Illumina  
610 HumanMethylationEPIC BeadArray. *Genome Biol* [Internet] 2018;**19**:64.

611 Sharma N, Kubaczka C, Kaiser S, Nettersheim D, Mughal SS, Riesenberger S, Hölzel M, Winterhager E,  
612 Schorle H. *Tpbpa-Cre* -mediated deletion of TFAP2C leads to deregulation of *Cdkn1a* , *Akt1* and the  
613 ERK pathway, causing placental growth arrest. *Development* [Internet] 2016;**143**:787–798.

614 Sharp GC, Salas LA, Monnereau C, Allard C, Yousefi P, Everson TM, Bohlin J, Xu Z, Huang R-C, Reese SE, *et*  
615 *al.* Maternal BMI at the start of pregnancy and offspring epigenome-wide DNA methylation:  
616 findings from the pregnancy and childhood epigenetics (PACE) consortium. *Hum Mol Genet*  
617 [Internet] 2017;**26**:4067–4085. Oxford University Press.

618 Sliker RC, Bos SD, Goeman JJ, Bovée JV, Talens RP, Breggen R van der, Suchiman HED, Lameijer E-W,

- 619 Putter H, Akker EB van den, *et al.* Identification and systematic annotation of tissue-specific  
620 differentially methylated regions using the Illumina 450k array. *Epigenetics Chromatin* [Internet]  
621 2013;**6**:26. Molecular Epidemiology, Leiden University Medical Center, Leiden, The Netherlands.  
622 B.T.Heijmans@lumc.nl.
- 623 Smew AI, Hedman AM, Chiesa F, Ullemar V, Andolf E, Pershagen G, Almqvist C. Limited association  
624 between markers of stress during pregnancy and fetal growth in 'Born into Life', a new prospective  
625 birth cohort. *Acta Paediatr Int J Paediatr* 2018;**107**:1003–1010.
- 626 Stojnic R, Diez D. PWMEnrich: PWM enrichment analysis R package version. 2018;Available from:  
627 <https://bioconductor.org/packages/release/bioc/html/PWMEnrich.html>.
- 628 Sungur AÖ, Schwarting RKW, Wöhr M. Behavioral phenotypes and neurobiological mechanisms in the  
629 Shank1 mouse model for autism spectrum disorder: A translational perspective. *Behav Brain Res*  
630 [Internet] 2018;**352**:46–61. Elsevier B.V.
- 631 Tanaka M, DeLorey TM, Delgado-Escueta A, Olsen RW. *GABRB3, Epilepsy, and Neurodevelopment*  
632 [Internet]. *Jasper's Basic Mech Epilepsies* [Internet] 2012;**51**..
- 633 Tobi EW, Goeman JJ, Monajemi R, Gu H, Putter H, Zhang Y, Slieker RC, Stok AP, Thijssen PE, Müller F, *et*  
634 *al.* DNA methylation signatures link prenatal famine exposure to growth and metabolism. *Nat*  
635 *Commun* [Internet] 2014;**5**:5592.
- 636 Tobi EW, Slagboom PE, Dongen J van, Kremer D, Stein AD, Putter H, Heijmans BT, Lumey LH. Prenatal  
637 famine and genetic variation are independently and additively associated with dna methylation at  
638 regulatory loci within IGF2/H19. *PLoS One* 2012;**7**..
- 639 Tobi EW, Slieker RC, Luijk R, Dekkers KF, Stein AD, Xu KM, Slagboom PE, Zwet EW van, Lumey LH,  
640 Heijmans BT. DNA methylation as a mediator of the association between prenatal adversity and

641 risk factors for metabolic disease in adulthood. *Sci Adv* [Internet] 2018;**4**:eaao4364. American  
642 Association for the Advancement of Science.

643 Turkgeldi E, Yagmur H, Seyhan A, Urman B, Ata B. Short and long term outcomes of children conceived  
644 with assisted reproductive technology. *Eur J Obstet Gynecol Reprod Biol* [Internet] 2016;**207**:129–  
645 136. Elsevier.

646 Vermey BG, Buchanan A, Chambers GM, Kolibianakis EM, Bosdou J, Chapman MG, Venetis CA. Are  
647 singleton pregnancies after assisted reproduction technology (ART) associated with a higher risk of  
648 placental anomalies compared with non-ART singleton pregnancies? A systematic review and  
649 meta-analysis. *BJOG An Int J Obstet Gynaecol* [Internet] 2019;**126**:209–218. Blackwell Publishing  
650 Ltd.

651 Wang J, Zhao Q. The CATE Package for High Dimensional Factor Analysis and Confounder Adjusted  
652 Multiple Testing. 2015; Comprehensive R Archive Network (CRAN) Available from: [https://cran.r-](https://cran.r-project.org/package=cate)  
653 [project.org/package=cate](https://cran.r-project.org/package=cate).

654 Wang Z, Xu L, He F. Embryo vitrification affects the methylation of the H19/Igf2 differentially methylated  
655 domain and the expression of H19 and Igf2. *Fertil Steril* [Internet] 2010;**93**:2729–2733. Elsevier.

656 Xu N, Barlow GM, Cui J, Wang ET, Lee B, Akhlaghpour M, Kroener L, Williams J, Rotter JI, Chen YI, *et al.*  
657 Comparison of Genome-Wide and Gene-Specific DNA Methylation Profiling in First-Trimester  
658 Chorionic Villi From Pregnancies Conceived With Infertility Treatments. *Reprod Sci* [Internet]  
659 2017;**24**:996–1004.

660 Yuen RK, Jiang R, Peñaherrera MS, McFadden DE, Robinson WP. Genome-wide mapping of imprinted  
661 differentially methylated regions by DNA methylation profiling of human placentas from triploidies.  
662 *Epigenetics Chromatin* [Internet] 2011;**4**:10. BioMed Central.

663 Zeileis A. Object-oriented computation of sandwich estimators. *J Stat Softw* 2006;**16**:1–16. American

664 Statistical Association.

665 Zheng Y, Joyce BT, Liu L, Zhang Z, Kibbe WA, Zhang W, Hou L. Prediction of genome-wide DNA

666 methylation in repetitive elements. *Nucleic Acids Res* [Internet] 2017;**45**:8697–8711.

667

668 **Tables**

669

670 **Table 1. Cohort characteristics**

Variable	MUC <sup>#</sup>	ART	IVF	ICSI
N	70	87	44	33
Frozen (N) <sup>1</sup>	-	10	7	3
Blastocyst (N) <sup>2</sup>	-	8	5 (3 frozen)	3 (2 frozen)
BMI in kg/m <sup>2</sup> (SD)	22.8 (3.5)	23.4 (3.4)	23.5 (3.3)	23.4 (3.0)
Age in y(SD)	32.4 (3.6)	33.1 (3.6)	33.0 (3.7)	33.0 (3.7)
SES <sup>3</sup> (SD)	3.0 (0.3)	2.6 (0.8) <sup>*</sup>	2.7 (0.7)	2.3 (0.9)
Smoking				
Never	65	56	27	23
Ever	4	26	17	9
NA	1	5	0	1
Years to index pregnancy				
Mean (SD)	0.2 (0.7)	2.6 (1.3) <sup>*</sup>	2.7 (1.2)	2.4 (1.5)
Min - max	0- 4	0 – 7	1- 7	0-7
Parity				
First	47	62	34	22
Second	21	17	10	7
Third	2	1	0	1
NA	0	7	0	3
N Male children (%)	43 (61.4)	44 (50.6)	22 (50.0)	16 (48.5)
Gestational age in weeks	39.1 (1.6)	39.5 (2.0)	39.2 (2.3)	39.7 (1.6)

- 671 1. The number of embryos that were frozen and thawed before placement *in utero*.
- 672 2. The number of embryos that were grown *ex utero* to the blastocyst stage before placement *in utero*. Off
- 673 note, five of these blastocysts were frozen and thawed.
- 674 3. Socio-economic status (SES) graded on a 4 level scale based on the highest attained educational level.
- 675 <sup>\*</sup>Significantly different P<0.001, <sup>#</sup> medically unassisted conception

676

677

678 Table 2. Results of ART EWAS

CpG ID	Location hg19	Nearest gene <sup>1</sup>	distance <sup>1</sup>	Estimate (SE) <sup>2</sup>	P <sup>2</sup>	P <sub>FDR</sub> <sup>679</sup>
cg27266479	chr1:9294882	<i>H6PD</i>	0	-1.94(0.27)	3.18e-13	8.28E-08
cg04811592	chr3:69834386	<i>MITF</i>	0	0.99(0.21)	1.59e-06	0.039
cg24959663	chr5:10578618	<i>ANKRD33B</i>	0	3.91(0.54)	3.75e-13	8.28E-08
cg22916646	chr5:162672583	-	-	2.18(0.36)	1.22e-09	1.08E-04
cg01500567	chr6:44355777	<i>CDC5L</i>	0	0.48(0.1)	5.46e-07	0.020
cg00478390	chr7:150703765	<i>NOS3</i>	0	-0.92(0.19)	7.37e-07	0.024
cg03207674	chr7:1523569	<i>INTS1</i>	0	0.72(0.14)	1.92e-07	9.43E-03
cg17123384	chr7:83379152	-	-	2.78(0.51)	6.39e-08	3.53E-03
cg19347588	chr10:3868336	<i>KLF6</i>	40862	0.81(0.16)	4.49e-07	0.018
cg07569385	chr13:20766226	<i>GJB2</i>	0	1.53(0.31)	1.24e-06	0.036
cg06485032	chr13:22615064	<i>AKO54845</i>	0	-3.47(0.67)	2.59e-07	0.011
cg13051607	chr15:22956714	<i>CYFIP1</i>	0	1.39(0.24)	1.15e-08	7.24E-04
cg01251603	chr15:26874098	<i>GABRB3</i>	0	-4.89(0.8)	1.18e-09	1.08E-04
cg15066197	chr15:26874202	<i>GABRB3</i>	0	-4.71(0.79)	2.48e-09	1.83E-04
cg14859324	chr15:26874363	<i>GABRB3</i>	0	-2.26(0.46)	7.68e-07	0.024
cg06450634	chr16:30430044	<i>ZNF771</i>	0	3.26(0.46)	2.15e-12	3.17E-07
cg08783253	chr17:40996565	<i>AOC2</i>	42	-2.93(0.61)	1.48e-06	0.038
cg16771467	chr18:55315872	<i>ATP8B1</i>	0	0.11(0.02)	1.34e-06	0.037
cg14560133	chr19:51199453	<i>SHANK1</i>	0	-1.42(0.3)	2.07e-06	0.048

689 The results of the ART vs MUC EWAS are ordered on genomic location (hg19).

- 690 1. Nearest gene within 100kb and the distance to the TSS.
- 691 2. Estimate and standard error (SE) and P value, with and without false discovery rate (FDR) correction of the
- 692 estimate of ART vs unassisted pregnancies in the model: Beta ~ ART (yes/no) + height on micro-array slide
- 693 + scan batch + bisulfite plate + sex + CD4T + CD8T + Mono + Bcell + Nk + Gran + nRBC.

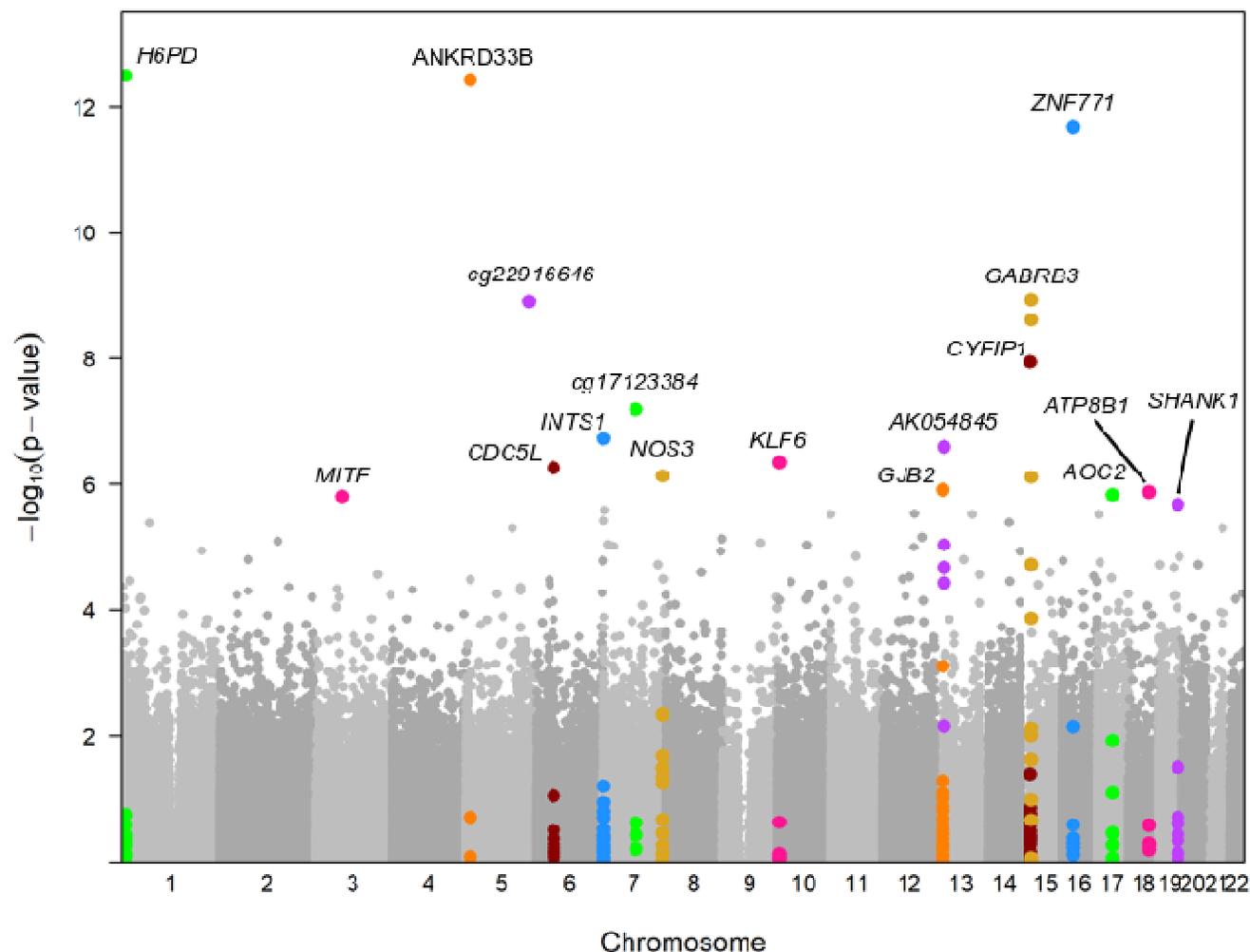
694

695 **Figures**

696

697 **Figure 1. Manhattan plot of EWAS for ART status**

698



699

700 **Figure 1.** A Manhattan plot showing the  $-\log_{10}$  P-values (y-axis) for the association of DNA methylation  
701 at individual CpG dinucleotides across the 22 autosomal chromosomes (x-axis). CpG dinucleotides 5kb  
702 up- and downstream of the lead association have been colored in the same color as the lead association.

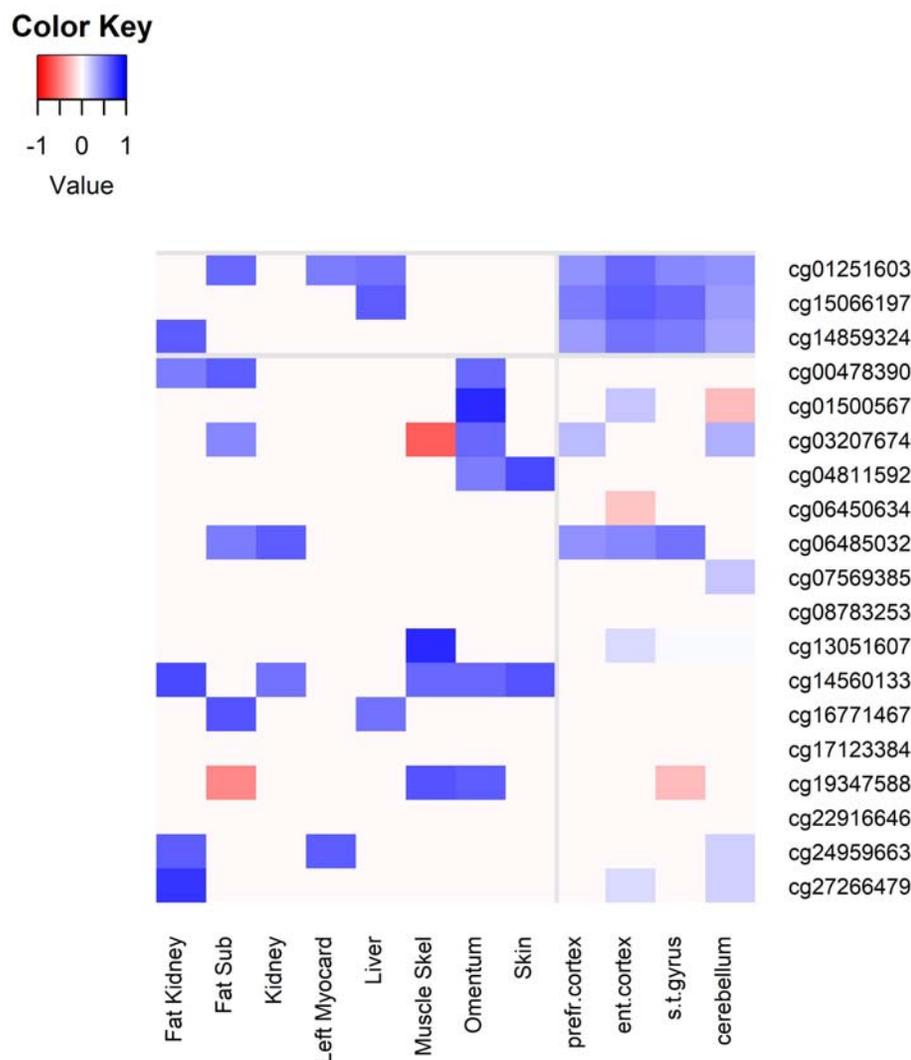
703

704

705

706

707 Figure 2. correlation between adult blood DNA methylation and internal tissues



708

709

710

711 **Figure 2.** The spearman correlation plotted for 8 tissues from the body (N=16, P<0.1) and 4 from the  
 712 brain (N>71, P<0.01) in colors ranging from bright red ( $\rho = -1.0$ ) to dark blue ( $\rho = 1.0$ ). The light gray  
 713 lines along the rows demark three CpG dinucleotides from the *GABRB3* region associated with ART. The  
 714 light gray line running along one column denotes the separation of the two tissue reference datasets  
 715 used. “Fat sub”: subcutaneous fat, “Left myocard”: left myocardium, “Muscle Skel”: skeleton muscle,  
 716 “prefr. Cortex”: prefrontal cortex, “ent.cortex”: entorhinal cortex, “s.t.gyrus”: superior temporal gyrus.

717