Title:

DNA methylation differences at birth after conception through assisted reproductive technologies

Elmar W. Tobi[1,2,*], Catarina Almqvist[3,4], Anna Hedman[5], Ellika Andolf[5], Jan Holte[6,7,8], Jan I. Olofsson[7,9] , Håkan Wramsby[10] , Margaretha Wramsby[11], Göran Pershagen[12], Bastiaan T. Heijmans[2], Anastasia N. Iliadou[3]

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

*corresponding author: Dr. Ir. Elmar W. Tobi, Periconceptional Epidemiology, Department of Obstetrics and Gynaecology, Erasmus MC, University Medical Center Rotterdam, Dr. Molewaterplein 50, 3015 GE Rotterdam, The Netherlands

Email: e.tobi@erasmusmc.nl

Tel. 00 31 (0)10 7038256

Running title:

DNA methylation differences and ART

Summary

The possible health consequences of assisted reproductive technology (ART) are a topic of intense study. Epigenetic mechanisms, such as changes in DNA methylation (DNAm), come forward from animal studies as a potential mechanism linking ART and health. Here we report on an epigenome-wide association study (EWAS) of DNAm in cord blood of 87 infants conceived with in vitro fertilisation (IVF) or intra-cytoplasmic sperm injection (ICSI) and 70 who were not. We found no difference in global methylation levels, but identified nineteen CpG dinucleotides (CpGs) associated with ART (effect estimates: 0.5-4.9%). There was no difference between IVF and ICSI and the CpGs map to genes related to brain function/development or the plethora of conditions linked to subfertility. Functional annotation did not point to likely functional consequences. Concluding, ART is associated with DNA methylation differences in cord blood, but these differences are limited in number, effect size and with unknown functional consequences.

Introduction

The usage of assisted reproductive technology (ART) is increasing world-wide and its possible health consequences are a topic of intense study. Evidence is building for an association of ART with long-term health outcomes for the offspring, including autism spectrum disorders (1) and cardiovascular health (higher blood pressure, suboptimal cardiac diastolic function and vessel thickness) (2). A body of evidence links ART to short term health effects, including to low birth weight, placenta associated anomalies (3) and pregnancy complications (4), congenital malformations and imprinting disorders (5). However, it is still debated if these associations stem from the application of ART or the underlying infertility leading couples to ART (6). Animal studies provided potential molecular mechanisms for these observations by highlighting that the ART procedure may induce changes to epigenetic marks (7,8). Epigenetic marks, such as DNA methylation (DNAm), influence the transcription potential of genomic regions (9). Human studies likewise point towards a link between early development, DNAm and (late-life) phenotypes, but the causality of these associations remains unknown (10).

To date, several studies on DNAm and ART have been performed. Multiple studies have focused on candidate gene regions and/or global methylation levels (11,12). Genome-wide efforts have focused on samples taken from extra-embryonic lineages (13,14) and when focused on material from the infant itself suffered from a small sample size (15) or batch effects correlating with ART status (16). Larger studies have focused on specific ART techniques, namely intra-cytoplasmic sperm injection (ICSI) (17) or in vitro fertilisation (IVF) (18) or IVF and less invasive gamete intra-fallopian transfer (GIFT) coupled with intra-uterine insemination (IUI) (19). These studies resulted in conflicting outcomes, with two reporting wide-spread associations between ART and (cord-)blood DNAm at birth (17) which appear to fade with age (19), while another study reported a DNA methylation difference at a single genomic locus only (18).

We undertook an epigenome wide association study on ART, comparing DNAm patterns in cord blood of children conceived via ART (N=87) (20) with that of children with medically unassisted

conceptions (henceforth “MUC”, N=70) (21). Since we had detailed information on the ART technique used for 77 of the children conceived via ART we investigated possible different outcomes for IVF and ICSI on DNAm which has been hypothesized to be important (22). In addition, we explored the possible functional consequences of these methylation differences using external datasets.

Methods

Study Subjects

ART controls: Born into Life

The prospective longitudinal birth cohort study Born into Life was originally created from the larger LifeGene study (21). Between the years 2010 and 2012, pregnant women who were already participating in the LifeGene study and living in Stockholm County were recruited to Born into Life (23). The inclusion criteria were that they had responded to baseline questionnaires from the LifeGene study, were pregnant, and gave informed consent. They were recruited both before and after 10-14 gestational weeks, but no later than 26-28 weeks. Originally 107 pregnant women were included in Born into Life. We only included the 70 women with cord blood collected in this study. Ethical approval was granted by the Regional Ethics Review Board in Stockholm, Sweden. Informed consent from both parents was obtained for all study participants.

Women in the Born into Life were asked to complete questionnaires regarding pregnancy, lifestyle and health at 10-14 and 26-28 gestational weeks. Data on maternal self-reported smoking during pregnancy and body mass index (BMI), in kg/m2 obtained at the first antenatal care visit in gestational week week 5-12 were collected from birth records . Data regarding highest attained educational level, ranging from mandatory secondary school to high school, university or other, were retrieved from the baseline LifeGene questionnaires. Maternal age at delivery, in years, was calculated from mothers' date of birth. Parity and data regarding the infants' sex, gestational age in weeks and mode of delivery, defined as vaginal delivery or Caesarean section, were also collected from the birth records. Cord blood and placenta samples were obtained from delivery. Cord blood was aspirated 2 minutes after birth by the assisting midwife into a test tube and kept in room temperature awaiting transportation. The birth records regarding both mother and child at delivery were collected from Danderyd Hospital, Stockholm.

ART cases: UppstART

Participants were recruited from three of the four fertility and reproductive health clinics in Stockholm (one public, two private) and one private clinic in Uppsala county, which also serves a large volume of patients from Stockholm (20). The study has been approved by the regional ethics review board at the Karolinska Institutet (Dnr 2011/230-31/1, Dnr 2011/1427-32, Dnr 2012/131-32, Dnr 2012/792-32, 2013/1700-32, 2014/1956-32, Dnr 2015/1604-32). Women and their partners were approached by the clinic nurse and asked to participate in the study. To facilitate the process of informed consent, the couple was provided with information approved by the regional ethical board, both verbally and in written format about the purpose of the study, methods, possible risks, and that participation was voluntary. Additionally, participants were informed that they could withdraw from the study at any time with no impact to their medical care. The requirement for inclusion in the study was an understanding of the Swedish language and exclusion was the use of donor gametes. Recruitment took place from September 2011 to December 2013. IVF treatment(s) of the participants were followed until December 2014 or drop-out/consent withdrawn (n=4), which ever came first.

At the time when informed consent was given, the clinic nurse withdrew blood samples and provided the participant with a kit for at home saliva collection. The participants were asked to answer a web-based baseline questionnaire within a few days of their clinic visit and prior to their IVF treatment start, which included an extensive list of questions on sociodemographic, anthropometric and life-style factors. Once the IVF treatment began and the participants reached the stage of oocyte retrieval, they were asked to respond to a second online follow-up questionnaire, with a shorter version of the baseline questionnaire, to identify any changes in lifestyle factors since the initiation of their treatment. Staff at seven delivery units in Stockholm and Uppsala were recruited to assist in collecting samples from UppStART participants during delivery. Delivery clinics were provided with a sample collection kit

including tubes for collection of maternal blood (EDTA, PAXgene), cord blood (EDTA, PAXgene), amniotic fluid, and placenta samples. Samples were stored in -20°C freezers at the delivery units until they were collected by UppStART study staff and deposited into the KI Biobank.

DNA methylation measurements

Genome-wide DNA methylation data were generated using the Illumina Infinium Human Methylation 450K BeadChip (450k array). A total of 500ng of genomic DNA isolated from cord blood was bisulfite treated using the EZ-96 DNA methylation kit (Zymo Research, Orange County, USA). We used the D-optimum criterion to assign samples over two 96-well plates and individual 450k arrays, ensuring even distributions of ART cases and controls, ART methods (IVF or ICSI), gender, gestational age and birth month across the two 96-well bisulfite plates and each 450k array. The 450K arrays were measured at ServiceXS (Leiden, The Netherlands). The quality of the generated 450K array data was assessed using both sample dependent and sample independent quality metrics using the *Bioconductor* package *MethylAid* (24) with default settings. We used the *Bioconductor* package *omicsPrint* (25) to check for sample duplications and admixtures and sample sexes were checked using the X-chromosomal CpGs. One sample was not the correct gender and was deleted. We used principal component analysis (PCA) and hierarchical clustering on the raw autosomal beta-values to search for outliers and suspect patterns in the dataset, finding none. The cell proportions of cord blood were imputed using IDOL (26) on the "FlowSorted.CordBloodCombined.450k" reference set using the estimateCellCounts2() function in the *FlowSorted.Blood.EPIC* R package(27,28). Normalization of the dataset was performed by NOOB background and color correction in combination with Functional Normalization (29) using four principal components (eigenvalue > 1). All measurements with <3 beads (0.05% of probes), < 1 intensity value (0.012% of probes) and a detection p-value >0.01 (0.11% of CpGs) were set as missing. The measurement success rate per sample was >99%. CpGs with probes that did not map to unique genomic locations (30) or with a < 95% measurement success rate (0.23% of CpGs) were then removed. We used

custom scripts that add on the functions from the *minfi* package (31) and implements parallelization where possible (https://github.com/molepi/DNAmArray).

Outlier detection was performed by PCA and hierarchical clustering on the autosomal beta-values. Hierarchical clustering identified one possible outlier. This individual had a very low gestational age (30 weeks), the lowest in the dataset, and this clustering is therefore likely based on biological reasons. High-quality DNA methylation data were obtained for 157 individuals for 441,836 autosomal CpGs.

Measures of global methylation

Genome-wide average DNA methylation (GWAM) (32) was calculated by averaging all beta-value measurements across the autosomes for each individual. The *R* package *REMP* (33) was used to infer DNA methylation at either *ALU* or *LINES*-1 sequences and then the average across all *ALU* or *LINES*-1 elements was calculated for each neonate.

TFBS enrichment

We calculated transcription factor binding site enrichments using the R package *PWMEnrich* (34) using binding motifs from motifDB. We calculated position weight matrices using the DNA base background frequencies calculated for the CpGs tested with 25bp flanking sequences. Enrichment was tested relative to this background.

Statistics

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

methylation levels were compared via t.test (unassisted vs ART) or ANOVA (medically unassisted fertilization versus IVF versus ICSI).

We used the R package *cate* (35) to identify latent variables (e.g. "hidden variables") influencing the relation between DNA methylation and ART (yes/no) or ART method (IVF/ICSI/unassisted). Two latent variables were identified (P<0.001) for ART and four latent variables for ART method (P<0.001). In both cases the first two latent variables correlated with the proportion of nucleated red blood cells (nRBC) in blood (rho>0.94, P<0.001) and the other two latent variables correlated with CD4T or B cell proportions in blood and several batch effects (rho>|0.3|, P<0.001) arguing together with directed acyclic graphs (DAG) (*Supplemental text*), for a basic model adjusting for batch effects and imputed cell proportions.

We performed epigenome-wide association analyses (EWAS) on being conceived by ART or not using robust linear regression (rlm) from the R *MASS* package with White's estimator for robust standard errors as implemented in the R package *sandwich* (36) (which leads to a model robust for outlying beta-values and heteroscedasticity) To test for differences in neonates conceived via IVF or ICSI or without medical assistance we used type II ANOVA with White's estimator for robust standard errors to account for any heteroscedasticity and unequal variances between groups using the R *car* package. We employed heteroscedasticity robust Tukey tests to test for DNA methylation differences between IVF, ICSI and unassisted conceptions post-hoc. In all instances we used the beta-value of an individual CpG dinucleotide as the outcome and ART status as the dependent. We adjusted for the sex of the individual, height of the sample on the 450k array glass slide (continues variable from 1-6), scan batch and bisulfite plate and the imputed cell proportions (CD4T, CD8T, Mono, NK, Bcell, Gran, nRBC). We ran additional sensitivity analyses by omitting the cell proportions or expanding the adjustment with minimal adjustment sets coming from DAG analysis with the following models, i: years till the index pregnancy, maternal age, pre-pregnancy BMI, educational attainment as proxy for socio-economic

status, smoking history ii) gestational length and parity. In addition, we performed look-ups in other published meta-EWAS of prenatal exposures.

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

All analyses were performed in R (version 3.6.1) (37). All p-values reported are two-sided and multiple testing correction is done using FDR (false discovery rate).

Results

Study subjects

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

DNAm is a key determinant of cell identity (9) and cord blood consists of multiple cell types that influence DNAm (38). Seven major cell types, including nucleated red blood cells, were imputed from the genome-wide DNAm data of the newborns (28). In concordance with a prior report on unassisted and ICSI newborns (17), there were no differences in these seven cell proportions between ART and MUC cord blood samples ($P_{nominal}$>0.13, Supplemental Figure S1).

Global DNAm comparison between ART and unassisted conceptions

First, we investigated genome-wide average DNA methylation (GWAM) (32) in the cord blood of neonates conceived via IVF or ICSI or unassisted. There was no difference between these 3 groups (P=0.89, Supplemental Figure S2). Next, we investigated the genome-wide average DNA methylation of ALU and LINES-1 elements, an often used proxy for global DNA methylation (33), finding no difference between groups (P>0.2, Figure S3).

DNAm comparison between ART and medically unassisted conceptions

We compared DNA methylation at 441,836 autosomal CpGs between 87 ART and 70 MUC neonates (Supplemental Figure S4). Nineteen CpGs were associated with being conceived via ART in a model adjusting for sex, batches and cell heterogeneity (P_{FDR}<0.05, Figure S5). Associations were confined to one CpG dinucleotide and did not extend across a larger region (Figure 1), with the exception of probes covering CpGs in the proximal promoter of *AK054845*, a gene of unknown function (five CpGs $P_{nominal}$<0.05), and the proximal promoter of a smaller variant of *GABRB3* (NM_001191321.2) which is expressed in the brain (eight CpGs $P_{nominal}$<0.05). *GABRB3* is a gene with a neuro-developmental role (39). Indeed, several CpGs can be linked to genes with roles in brain function or (early) brain development. *ZNF771* (cg06450634), as a transcription factor (TF), is a driving factor behind a large gene network in the brain (40) while *CYFIP1* (cg13051607) (41), *GABRB3* (cg01251603, cg15066197,cg14859324) (39) and *SHANK1* (cg14560133) (42) are all autism spectrum disorder candidate genes, a late life phenotype associated with ART (1).

Omitting the adjustment for cell heterogeneity had little to no effect on the effect estimates for all nineteen CpGs (Supplemental Table S1) just like additional adjustment for maternal and pregnancy characteristics (Supplemental Table S2). The nineteen CpGs were not among those previously reported in other EWASs (FDR corrected P-value of <0.05) with prenatal smoking (43), folate use (44), maternal hypertension and pre-eclampsia (45), maternal BMI (46) and birthweight (47). Moreover, a look-up of

these previously reported CpGs in our study did not yield more nominally associated CpGs with ART than may be expected by chance (P>0.07).

DNA methylation comparison between IVF, ICSI and unassisted conceptions

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

Functional annotation

We performed KEGG and Gene Ontology tests (48), finding no enrichment for the genes linked to nineteen CpGs associated with ART, nor when we relaxed the significance threshold to suggestive associations ($P<10^{-5}$, N=37). We did not find significant enrichments of single transcription factor (TF) binding motifs (34), although multiple CpGs overlapped TF binding sites of TFs with a role in early development like *TFAP2C* (49) overlapping cg27266479 mapping to *H6PD* and *RAX* (50) overlapping cg16771467 to *ATP8B1*. Look-up of the nineteen CpGs in reference data (51) uncovered no link between DNAm and gene-expression in adult blood. We measured DNAm in cord blood and average DNAm levels may vary between tissues, but still reflect variation in other tissues (52). There was little to no correlation of methylation at the nineteen CpGs between adult whole blood and tissues from sixteen cadavers (52,53), with the notable exception of cg14560133 (*SHANK1*, β = -1.42%(SE=0.3%) $P=2.1x10^{-6}$) which showed moderate correlation with 5/8 tissues (rho>0.53, P<0.04), Figure 2). The three CpG dinucleotides at *GABRB3* (cg01251603, cg15066197,cg14859324), a gene active in the brain, showed

weak to moderate correlation between adult blood and four brain regions (rho=0.39-0.62, P<0.001, N=71-74) in another reference dataset (54).

Imprinted regions and metastable epialleles

Multiple candidate gene studies have looked at DNA methylation or gene-expression differences of imprinted genes between ART and MUC with mixed results. Therefore, we looked at 374 CpGs on the 450k array known to overlap imprinted differentially methylated regions (DMRs) (55). There was no overlap with the CpGs showing a (suggestive) association with ART in our study ($P<10^{-5}$, N=22). These 374 CpGs where distributed over 59 regions. There was no difference in DNAm between unassisted and ART infants when we tested each of these 59 regions ($P_{FDR}>0.86$). Similarly, so-called meta-stable epialleles (MEs) are hypothesized to be especially sensitive to the early prenatal environment (56). There were 187 CpGs overlapping MEs, none of which were associated with ART ($P_{FDR}>0.06$).

Prior studies

We performed a look-up of CpGs and regions associated in prior studies with one form of ART. The sole region associated with IVF in the meDIP sequencing study from Castillo-Fernandez *et al*. was not covered by our genome-scale screen with the 450k array (18). Next, El Haij *et al*. found 4730 CpG dinucleotides associated with ICSI after FDR adjustment (which are unavailable as Supplemental information) of which two regions, out of five regions selected for validation with pyrosequencing, were validated in an additional sample (17). The seven CpGs in these two regions, the promoters of *ATG4C* and *SNORD114-9*, were not associated with ART in our study (P>0.23) nor with ICSI only (post hoc Tukey's test: P=0.40).

Using the 850k array, Novakovic *et al*. identified 2340 CpG dinucleotides at which DNA methylation in Guthrie card blood was associated with IVF and GIFT coupled with IUI (19), 1228 of which were covered in our 450k study. Of these, 38 were nominally associated (P<0.05) and with the same direction of effect in our study. This decreased to 12 loci when we tested for DNA methylation

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

Discussion

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

The nineteen CpG dinucleotides could be mapped to fifteen genes, several of which can be linked genes with a role in brain function and/or development. *ZNF771* (cg06450634) has been identified as a driving factor behind a large gene network in the brain (40) and *CYFIP1* (cg13051607) (41), *GABRB3* (cg01251603, cg15066197,cg14859324) (39) and *SHANK1* (cg14560133) (42) are all autism spectrum disorder candidate genes, a late life phenotype tentatively associated with ART (1). In addition, several CpG dinucleotides can be linked to genes connected to the plethora of conditions that can be linked to subfertility. *H6PD* (cg27266479) is a candidate gene for polycystic ovary syndrome (PCOS) (57). *INTS1* (cg03207674) has a crucial role in the developing blastocyst, as inhibition of *INTS1* function causes growth arrest (58). *NOS3* (cg00478390) knockout mice are used as an *in vivo* model of (recurrent) embryo loss, as nitric oxide metabolism plays an important role in implantation (59).

The latter may hint that the various medical reasons for ART may underlie the associations, rather than the ART process itself, which is a key question in the study of ART (6). We employed directed acyclic graphs (DAG) (60) and DAG indicated that and ascertainment of a direct effect is possible with a small minimal adjustment set. Indeed, there was little to no effect on the effect estimates from adjustment for various maternal characteristics, gestational age and the number of years until pregnancy (as a proxy for in-/subfertility). In addition, we did not find any evidence for different or stronger effects of ICSI on DNAm, as has been hypothesized (22), and the fertility clinics from which we

recruited all used the same culturing media for ICSI and IVF. Both might argue that not the (reasons for) infertility, but the ART process itself may explain the associations.

Our analysis did not find wide-spread genome-wide differences like other genome-scale studies to date (17,19), but is consistent with a genome-scale study on only IVF using immunoprecipitation of methylated DNA with next generation sequencing (18). We extensively compared individual results as far as possible and found little overlap. This may be due to study power, the different scope of the measurement techniques used (18) and the different sources of genomic DNA (Guthrie card versus whole cord blood) (19). In addition, it is possible that different media was used between countries or that the fact that only a small proportion of our IVF and ICSI group consisted of *concepti* that were frozen and/or cultured extensively in comparison to other studies. This latter aspect we were underpowered to test as most ART procedures entailed fresh *concepti* that were not cultured extensively *in vitro*.

Our study is one of the larger genome-scale DNAm studies on ART to date and we are the first to study both IVF and ICSI on this scale, nonetheless our study has important limitations to consider. First, DNAm is one of the drivers of cell identity (9) and we measured cord blood, which may be a tissue less relevant in relation to study outcomes although it may still mark processes in relevant tissues due to mitotic inheritance (61). We used the latest methods to impute the seven major cell types in cord-blood (28) and in conjuncture with an earlier report (17) these imputed cell types did not differ between infants conceived with or without ART and adjustment did not alter the effect estimates. However, most of these major cell types consist of subsets of more specialized cells (subsequently showing unique DNAm profiles at increasingly select loci) and their influence could not be investigated. However, none of the nineteen CpG dinucleotides is linked to a gene with a role in (auto)immune function or hematopoiesis. Another major driver of DNAm variation is genetic variation (51) and we did not control for genetic variation in our analyses. Nine out of nineteen CpG dinucleotides did not have a meQTL in a

large reference set (51) and none of these 450k probes had an overlap with a single nucleotide polymorphism (SNP) arguing that genetic heterogeneity is unlikely to underlie our findings. In addition, the influence of the prenatal environment on DNAm variation may be independent and additive of genetic variation (62).

Also important is to consider the fact that an ART population is inherently different from the general population (6) and despite controlling for possible confounding factors and cross-checking our results with other EWAS on prenatal complications/exposures unresolved confounding remains a possibility. Our study differs with some of the earlier described in ART studies, which is useful for triangulation between studies (63), as our control group has a higher, rather than a lower, SES as is normal in most of studies on ART.

In conclusion, the data presented showcases modest and specific DNAm differences associated with ART which are unlikely to be functionally relevant in adult tissues. We did not find any difference in DNAm patterns between IVF and ICSI. Our study found little evidence for the hypothesis that ART, be it IVF or ICSI, leads to wide-spread disturbances of DNAm patterns. Our findings warrant cautious interpretation given the sample size, tissue studied and the unknown functional consequences of the identified DNAm differences.

Acknowledgements

EWT was supported by a VENI grant from the Netherlands Organization for Scientific Research (91617128) and JPI-HDHL project PREcisE funded through ZonMw (529051023). Financial support was provided from the European Union's Seventh Framework Program IDEAL (259679), the Swedish Research Council (K2011-69X-21871-01-6, 2011-3060 and 2018-02640), and the Strategic Research Program in Epidemiology Young Scholar Awards, Karolinska Institute (to ANI) and through the Swedish Initiative for Research on Microdata in the Social And Medical Sciences (SIMSAM) framework grant no 340-2013-5867, grants provided by the Stockholm County Council (ALF-projects), the Strategic Research Program in Epidemiology at Karolinska Institutet, the Swedish Heart-Lung Foundation and Danderyd University Hospital (Stockholm, Sweden). The funders had no role in study design, data collection, analysis, decision to publish or preparation of the manuscript. We wish to thank the Biobank at the Karolinska Institutet for professional biobank service.

Conflicts of interest

The authors declare no conflicts of interest

References

1. Liu L, Gao J, He X, Cai Y, Wang L, Fan X. Association between assisted reproductive technology and the risk of autism spectrum disorders in the offspring: a meta-analysis. Sci Rep [Internet] 2017 [cited 2018 Jun 13];7(1):46207. Available from: http://www.nature.com/articles/srep46207

2. Guo XY, Liu XM, Jin L, Wang TT, Ullah K, Sheng JZ, et al. Cardiovascular and metabolic profiles of offspring conceived by assisted reproductive technologies: a systematic review and meta-analysis. Fertil Steril [Internet] 2017 [cited 2018 Nov 2];107(3):622-631.e5. Available from: http://www.ncbi.nlm.nih.gov/pubmed/28104241

3. Vermey BG, Buchanan A, Chambers GM, Kolibianakis EM, Bosdou J, Chapman MG, et al. Are singleton pregnancies after assisted reproduction technology (ART) associated with a higher risk of placental anomalies compared with non-ART singleton pregnancies? A systematic review and meta-analysis [Internet]. BJOG An Int. J. Obstet. Gynaecol. 2019 [cited 2020 Mar 10];126(2):209–18. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1111/1471-0528.15227

4. Qin J, Liu X, Sheng X, Wang H, Gao S. Assisted reproductive technology and the risk of pregnancy-related complications and adverse pregnancy outcomes in singleton pregnancies: A meta-analysis of cohort studies. Fertil Steril 2016;105(1):73-85.e6.

5. Turkgeldi E, Yagmur H, Seyhan A, Urman B, Ata B. Short and long term outcomes of children conceived with assisted reproductive technology. Eur J Obstet Gynecol Reprod Biol [Internet] 2016 [cited 2018 Sep 19];207:129–36. Available from: https://www.sciencedirect.com/science/article/pii/S0301211516309587?via%3Dihub

6. Luke B, Stern JE, Hornstein MD, Kotelchuck M, Diop H, Cabral H, et al. Is the wrong question being asked in infertility research? J Assist Reprod Genet [Internet] 2016 [cited 2020 Jan 8];33(1):3–8. Available from: http://link.springer.com/10.1007/s10815-015-0610-3

7. Morgan HD, Jin XL, Li A, Whitelaw E, O'Neill C. The culture of zygotes to the blastocyst stage changes the postnatal expression of an epigentically labile allele, agouti viable yellow, in mice. BiolReprod 2008;79(0006-3363 (Print)):618–23.

8. Wang Z, Xu L, He F. Embryo vitrification affects the methylation of the H19/Igf2 differentially methylated domain and the expression of H19 and Igf2. Fertil Steril [Internet] 2010 [cited 2018 Nov 1];93(8):2729–33. Available from: https://www.sciencedirect.com/science/article/pii/S0015028210004620

9. Jaenisch R, Bird A. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. Nat Genet 2003;33 Suppl:245–54.

10. Tobi EW, Slieker RC, Luijk R, Dekkers KF, Stein AD, Xu KM, et al. DNA methylation as a mediator of the association between prenatal adversity and risk factors for metabolic disease in adulthood. Sci Adv [Internet] 2018 [cited 2018 Feb 1];4(1):eaao4364. Available from: http://advances.sciencemag.org/lookup/doi/10.1126/sciadv.aao4364

11. Canovas S, Ross PJ, Kelsey G, Coy P. DNA Methylation in Embryo Development: Epigenetic Impact of ART (Assisted Reproductive Technologies). BioEssays 2017;39(11):1–11.

12. Lazaraviciute G, Kauser M, Bhattacharya S, Haggarty P, Bhattacharya S. A systematic review and meta-analysis of DNA methylation levels and imprinting disorders in children conceived by IVF/ICSI compared with children conceived spontaneously. Hum Reprod Update [Internet] 2014 [cited 2018 Nov 2];20(6):840–52. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24961233

13. Choufani S, Turinsky AL, Melamed N, Greenblatt E, Brudno M, Bérard A, et al. Impact of Assisted Reproduction, Infertility, Sex, and Paternal Factors on the Placental DNA Methylome. Hum Mol

Genet [Internet] 2018 [cited 2018 Nov 2];Available from: http://fdslive.oup.com/www.oup.com/pdf/production_in_progress.pdf

14. Xu N, Barlow GM, Cui J, Wang ET, Lee B, Akhlaghpour M, et al. Comparison of Genome-Wide and Gene-Specific DNA Methylation Profiling in First-Trimester Chorionic Villi From Pregnancies Conceived With Infertility Treatments. Reprod Sci [Internet] 2017 [cited 2018 May 15];24(7):996–1004. Available from: http://journals.sagepub.com/doi/pdf/10.1177/1933719116675056

15. Melamed N, Choufani S, Wilkins-Haug LE, Koren G, Weksberg R. Comparison of genome-wide and gene-specific DNA methylation between ART and naturally conceived pregnancies. Epigenetics [Internet] 2015 [cited 2018 May 15];10(6):474–83. Available from: http://www.tandfonline.com/doi/full/10.4161/15592294.2014.988041

16. Estill MS, Bolnick JM, Waterland RA, Bolnick AD, Diamond MP, Krawetz SA. Assisted reproductive technology alters deoxyribonucleic acid methylation profiles in bloodspots of newborn infants. Fertil Steril [Internet] 2016;106(3):629-639.e10. Available from: http://dx.doi.org/10.1016/j.fertnstert.2016.05.006

17. El Hajj N, Haertle L, Dittrich M, Denk S, Lehnen H, Hahn T, et al. DNA methylation signatures in cord blood of ICSI children. Hum Reprod [Internet] 2017 [cited 2018 May 15];32(8):1761–9. Available from: https://academic.oup.com/humrep/article/32/8/1761/3859168

18. Castillo-Fernandez JE, Loke YJ, Bass-Stringer S, Gao F, Xia Y, Wu H, et al. DNA methylation changes at infertility genes in newborn twins conceived by in vitro fertilisation. Genome Med [Internet] 2017;9(1):28. Available from: http://genomemedicine.biomedcentral.com/articles/10.1186/s13073-017-0413-5%5Cnhttp://www.ncbi.nlm.nih.gov/pubmed/28340599%5Cnhttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC5364659

19. Novakovic B, Lewis S, Halliday J, Kennedy J, Burgner DP, Czajko A, et al. Assisted reproductive technologies are associated with limited epigenetic variation at birth that largely resolves by adulthood. Nat Commun [Internet] 2019;10(1):3922. Available from: http://dx.doi.org/10.1038/s41467-019-11929-9

20. Iliadou AN, Öberg AS, Pege J, Rodriguez-Wallberg KA, Olofsson JI, Holte J, et al. The Uppsala-Stockholm Assisted Reproductive Techniques (UppStART) study. BMJ Open 2019;9(8).

21. Almqvist C, Adami HO, Franks PW, Groop L, Ingelsson E, Kere J, et al. LifeGene - A large prospective population-based study of global relevance. Eur J Epidemiol 2011;26(1):67–77.

22. Loke YJ, Craig JM. Are the effects of IVF on DNA methylation driven by intracytoplasmic sperm injection and male infertility? Epigenomics [Internet] 2016 [cited 2018 May 15];8(7):881–4. Available from: http://www.futuremedicine.com/doi/10.2217/epi-2016-0041

23. Smew AI, Hedman AM, Chiesa F, Ullemar V, Andolf E, Pershagen G, et al. Limited association between markers of stress during pregnancy and fetal growth in 'Born into Life', a new prospective birth cohort. Acta Paediatr Int J Paediatr 2018;107(6):1003–10.

24. van Iterson M, Tobi EW, Slieker RC, den Hollander W, Luijk R, Slagboom PE, et al. MethylAid: visual and interactive quality control of large Illumina 450k datasets. Bioinformatics [Internet] 2014;30(23):3435–7. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25147358

25. van Iterson M, Cats D, Hop P, Heijmans BT. omicsPrint: detection of data linkage errors in multiple omics studies. Bioinformatics [Internet] 2018 [cited 2020 Feb 3];34(12):2142–3. Available from: https://academic.oup.com/bioinformatics/article/34/12/2142/4840579

26. Koestler DC, Jones MJ, Usset J, Christensen BC, Butler RA, Kobor MS, et al. Improving cell mixture deconvolution by identifying optimal DNA methylation libraries (IDOL). BMC Bioinformatics

[Internet] 2016 [cited 2019 Nov 11];17(1):120. Available from: http://www.biomedcentral.com/1471-2105/17/120

27. Salas LA, Koestler DC, Butler RA, Hansen HM, Wiencke JK, Kelsey KT, et al. An optimized library for reference-based deconvolution of whole-blood biospecimens assayed using the Illumina HumanMethylationEPIC BeadArray. Genome Biol [Internet] 2018 [cited 2019 Nov 11];19(1):64. Available from: https://genomebiology.biomedcentral.com/articles/10.1186/s13059-018-1448-7

28. Gervin K, Salas LA, Bakulski KM, Van Zelm MC, Koestler DC, Wiencke JK, et al. Systematic evaluation and validation of reference and library selection methods for deconvolution of cord blood DNA methylation data. Clin Epigenetics [Internet] 2019 [cited 2019 Nov 11];11(1):125. Available from: https://clinicalepigeneticsjournal.biomedcentral.com/articles/10.1186/s13148-019-0717-y

29. Fortin J-P, Labbe A, Lemire M, Zanke BW, Hudson TJ, Fertig EJ, et al. Functional normalization of 450k methylation array data improves replication in large cancer studies. Genome Biol [Internet] 2014;15(12):503. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25599564

30. Chen YA, Lemire M, Choufani S, Butcher DT, Grafodatskaya D, Zanke BW, et al. Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray. Epigenetics 2013;8(1559-2308 (Electronic)):203–9.

31. Aryee MJ, Jaffe AE, Corrada-Bravo H, Ladd-Acosta C, Feinberg AP, Hansen KD, et al. Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. Bioinformatics 2014;30(10):1363–9.

32. Li S, Wong EM, Dugué P-A, McRae AF, Kim E, Joo J-HE, et al. Genome-wide average DNA methylation is determined in utero. Int J Epidemiol [Internet] 2018 [cited 2020 Jan 9];47(3):908–

16. Available from: https://academic.oup.com/ije/article/47/3/908/4922690

33. Zheng Y, Joyce BT, Liu L, Zhang Z, Kibbe WA, Zhang W, et al. Prediction of genome-wide DNA methylation in repetitive elements. Nucleic Acids Res [Internet] 2017 [cited 2019 Nov 20];45(15):8697–711. Available from: https://academic.oup.com/nar/article/45/15/8697/3933583

34. Stojnic R, Diez D. PWMEnrich: PWM enrichment analysis R package version [Internet]. 2018;Available from: https://bioconductor.org/packages/release/bioc/html/PWMEnrich.html

35. Wang J, Zhao Q. The CATE Package for High Dimensional Factor Analysis and Confounder Adjusted Multiple Testing [Internet]. 2015 [cited 2019 Nov 11];(1.0.4). Available from: https://cran.r-project.org/package=cate

36. Zeileis A. Object-oriented computation of sandwich estimators. J Stat Softw 2006;16(9):1–16.

37. R core team. R: A language and environment for statistical computing [Internet]. R Found. Stat. Comput. 2019;(3.6.1). Available from: http://www.r-project.org/

38. Jones SC, Jordan WJ, Meiners SJ, Miller AN, Methven AS. Fungal spore dispersal by the Eastern box turtle (Terrapene carolina carolina). Am Midl Nat [Internet] 2007 [cited 2018 Sep 5];157(1):121–6. Available from: http://www.ncbi.nlm.nih.gov/pubmed/27019159

39. Tanaka M, DeLorey TM, Delgado-Escueta A, Olsen RW. GABRB3, Epilepsy, and Neurodevelopment [Internet]. 2012 [cited 2019 Nov 27]. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22787634

40. Maulik U, Sen S, Mallik S, Bandyopadhyay S. Detecting TF-miRNA-gene network based modules for 5hmC and 5mC brain samples: a intra- and inter-species case-study between human and rhesus. BMC Genet [Internet] 2018 [cited 2018 Nov 2];19(1):9. Available from:

https://bmcgenet.biomedcentral.com/articles/10.1186/s12863-017-0574-7

41. Hsiao K, Harony-Nicolas H, Buxbaum JD, Bozdagi-Gunal O, Benson DL. Cyfip1 Regulates Presynaptic Activity during Development. J Neurosci [Internet] 2016 [cited 2018 Nov 2];36(5):1564–76. Available from: http://www.jneurosci.org/lookup/doi/10.1523/JNEUROSCI.0511-15.2016

42. Sungur AÖ, Schwarting RKW, Wöhr M. Behavioral phenotypes and neurobiological mechanisms in the Shank1 mouse model for autism spectrum disorder: A translational perspective. Behav Brain Res [Internet] 2018 [cited 2020 Jan 8];352:46–61. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0166432817308926

43. Joubert BR, Felix JF, Yousefi P, Bakulski KM, Just AC, Breton C, et al. DNA Methylation in Newborns and Maternal Smoking in Pregnancy: Genome-wide Consortium Meta-analysis. Am J Hum Genet [Internet] 2016;98(4):680–96. Available from: http://www.ncbi.nlm.nih.gov/pubmed/27040690

44. Joubert BR, den Dekker HT, Felix JF, Bohlin J, Ligthart S, Beckett E, et al. Maternal plasma folate impacts differential DNA methylation in an epigenome-wide meta-analysis of newborns. Nat Commun [Internet] 2016;7(May 2015):10577. Available from: http://www.nature.com/doifinder/10.1038/ncomms10577

45. Kazmi N, Sharp GC, Reese SE, Vehmeijer FO, Lahti J, Page CM, et al. Hypertensive Disorders of Pregnancy and DNA Methylation in Newborns. Hypertens (Dallas, Tex 1979) [Internet] 2019 [cited 2020 Jan 9];74(2):375–83. Available from: http://www.ncbi.nlm.nih.gov/pubmed/31230546

46. Sharp GC, Salas LA, Monnereau C, Allard C, Yousefi P, Everson TM, et al. Maternal BMI at the

start of pregnancy and offspring epigenome-wide DNA methylation: findings from the pregnancy and childhood epigenetics (PACE) consortium. Hum Mol Genet [Internet] 2017 [cited 2017 Nov 28];26(20):4067–85. Available from: http://academic.oup.com/hmg/article/26/20/4067/3988352/Maternal-BMI-at-the-start-of-pregnancy-and

47. Küpers LK, Monnereau C, Sharp GC, Yousefi P, Salas LA, Ghantous A, et al. Meta-analysis of epigenome-wide association studies in neonates reveals widespread differential DNA methylation associated with birthweight. Nat Commun 2019;10(1).

48. Phipson B, Maksimovic J, Oshlack A. missMethyl: an R package for analyzing data from Illumina's HumanMethylation450 platform. Bioinformatics [Internet] 2015 [cited 2018 Nov 9];32(2):btv560. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26424855

49. Sharma N, Kubaczka C, Kaiser S, Nettersheim D, Mughal SS, Riesenberg S, et al. *Tpbpa-Cre* -mediated deletion of TFAP2C leads to deregulation of *Cdkn1a* , *Akt1* and the ERK pathway, causing placental growth arrest. Development [Internet] 2016 [cited 2018 Oct 16];143(5):787–98. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26811378

50. Bennett RL, Blalock WL, Choi E-J, Lee YJ, Zhang Y, Zhou L, et al. RAX is required for fly neuronal development and mouse embryogenesis. Mech Dev [Internet] 2008 [cited 2018 Oct 16];125(9–10):777–85. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18634873

51. Bonder MJ, Luijk R, Zhernakova D V, Moed M, Deelen P, Vermaat M, et al. Disease variants alter transcription factor levels and methylation of their binding sites. Nat Genet 2017;49(1):131–8.

52. Slieker RC, Bos SD, Goeman JJ, Bovée JV, Talens RP, van der Breggen R, et al. Identification and systematic annotation of tissue-specific differentially methylated regions using the Illumina 450k

array. Epigenetics Chromatin [Internet] 2013;6(1):26. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23919675

53. Relton CL, Gaunt T, McArdle W, Ho K, Duggirala A, Shihab H, et al. Data Resource Profile: Accessible Resource for Integrated Epigenomic Studies (ARIES). IntJEpidemiol 2015;44(1464-3685 (Electronic)):1181–90.

54. Hannon E, Lunnon K, Schalkwyk L, Mill J. Interindividual methylomic variation across blood, cortex, and cerebellum: implications for epigenetic studies of neurological and neuropsychiatric phenotypes. Epigenetics [Internet] 2015 [cited 2018 Nov 2];10(11):1024–32. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26457534

55. Yuen RK, Jiang R, Peñaherrera MS, McFadden DE, Robinson WP. Genome-wide mapping of imprinted differentially methylated regions by DNA methylation profiling of human placentas from triploidies. Epigenetics Chromatin [Internet] 2011 [cited 2018 Nov 9];4(1):10. Available from: http://epigeneticsandchromatin.biomedcentral.com/articles/10.1186/1756-8935-4-10

56. Kessler NJ, Waterland RA, Prentice AM, Silver MJ. Establishment of environmentally sensitive DNA methylation states in the very early human embryo. Sci Adv [Internet] 2018;4(7):eaat2624. Available from: http://www.ncbi.nlm.nih.gov/pubmed/30009262

57. Martínez-García MA, San-Millán JL, Escobar-Morreale HF. The R453Q and D151A polymorphisms of Hexose-6-Phosphate Dehydrogenase Gene (H6PD) influence the polycystic ovary syndrome (PCOS) and obesity. Gene [Internet] 2012 [cited 2018 Nov 2];497(1):38–44. Available from: http://linkinghub.elsevier.com/retrieve/pii/S037811191200090X

58. Hata T, Nakayama M. Targeted disruption of the murine large nuclear KIAA1440/Ints1 protein causes growth arrest in early blastocyst stage embryos and eventual apoptotic cell death.

Biochim Biophys Acta [Internet] 2007 [cited 2018 Nov 2];1773(7):1039–51. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0167488907000900

59. Pallares P, Gonzalez-Bulnes A. The effect of embryo and maternal genotypes on prolificacy, intrauterine growth retardation and postnatal development of Nos3-knockout mice. Reprod Biol 2010;10(3):241–8.

60. Krieger N, Davey Smith G. The tale wagged by the DAG: broadening the scope of causal inference and explanation for epidemiology. Int J Epidemiol [Internet] 2016 [cited 2018 Feb 13];45(6):1787–808. Available from: https://academic.oup.com/ije/article-lookup/doi/10.1093/ije/dyw114

61. Heijmans BT, Mill J. Commentary: The seven plagues of epigenetic epidemiology. IntJEpidemiol 2012;41(1464-3685 (Electronic)):74–8.

62. Tobi EW, Slagboom PE, van Dongen J, Kremer D, Stein AD, Putter H, et al. Prenatal famine and genetic variation are independently and additively associated with dna methylation at regulatory loci within IGF2/H19. PLoS One 2012;7(5).

63. Lawlor DA, Tilling K, Davey Smith G. Triangulation in aetiological epidemiology. Int J Epidemiol [Internet] 2017 [cited 2020 Jan 9];dyw314. Available from: https://academic.oup.com/ije/article-lookup/doi/10.1093/ije/dyw314

Tables

Table 1. Cohort characteristics

Variable	MUC[#]	ART	IVF	ICSI
N	70	87	44	33
Frozen (N)[1]	-	10	7	3
Blastocyst (N)[2]	-	8	5 (*3 frozen*)	3 (*2 frozen*)
BMI in kg/m^2 (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[3] (SD)	3.0 (0.3)	2.6 (0.8)[*]	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)[*]	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)

1. The number of *concepti* that were frozen and thawed before placement *in utero*.
2. The number of *concepti* that were grown *ex utero* to the blastocyst stage before placement *in utero.* Off note, five of these blastocysts were frozen and thawed.
3. Socio-economic status (SES) graded on a 4 level scale based on the highest attained educational level.

[*]Significantly different P<0.001, [#] medically unassisted conceptions

Table 2. Results of ART EWAS

CpG ID	Location hg19	Nearest gene[1]	distance[1]	Estimate (SE)[2]	P[2]	P_{FDR}
cg27266479	chr1:9294882	*H6PD*	0	-1.94(0.27)	3.18e-13	8.28E-08
cg04811592	chr3:69834386	*MITF*	0	0.99(0.21)	1.59e-06	0.039
cg24959663	chr5:10578618	*ANKRD33B*	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	*CDC5L*	0	0.48(0.1)	5.46e-07	0.020
cg00478390	chr7:150703765	*NOS3*	0	-0.92(0.19)	7.37e-07	0.024
cg03207674	chr7:1523569	*INTS1*	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	*KLF6*	40862	0.81(0.16)	4.49e-07	0.018
cg07569385	chr13:20766226	*GJB2*	0	1.53(0.31)	1.24e-06	0.036
cg06485032	chr13:22615064	*AK054845*	0	-3.47(0.67)	2.59e-07	0.011
cg13051607	chr15:22956714	*CYFIP1*	0	1.39(0.24)	1.15e-08	7.24 E-04
cg01251603	chr15:26874098	*GABRB3*	0	-4.89(0.8)	1.18e-09	1.08 E-04
cg15066197	chr15:26874202	*GABRB3*	0	-4.71(0.79)	2.48e-09	1.83 E-04
cg14859324	chr15:26874363	*GABRB3*	0	-2.26(0.46)	7.68e-07	0.024
cg06450634	chr16:30430044	*ZNF771*	0	3.26(0.46)	2.15e-12	3.17E-07
cg08783253	chr17:40996565	*AOC2*	42	-2.93(0.61)	1.48e-06	0.038
cg16771467	chr18:55315872	*ATP8B1*	0	0.11(0.02)	1.34e-06	0.037
cg14560133	chr19:51199453	*SHANK1*	0	-1.42(0.3)	2.07e-06	0.048

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

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

Figures

Figure 1. Manhattan plot of EWAS for ART status

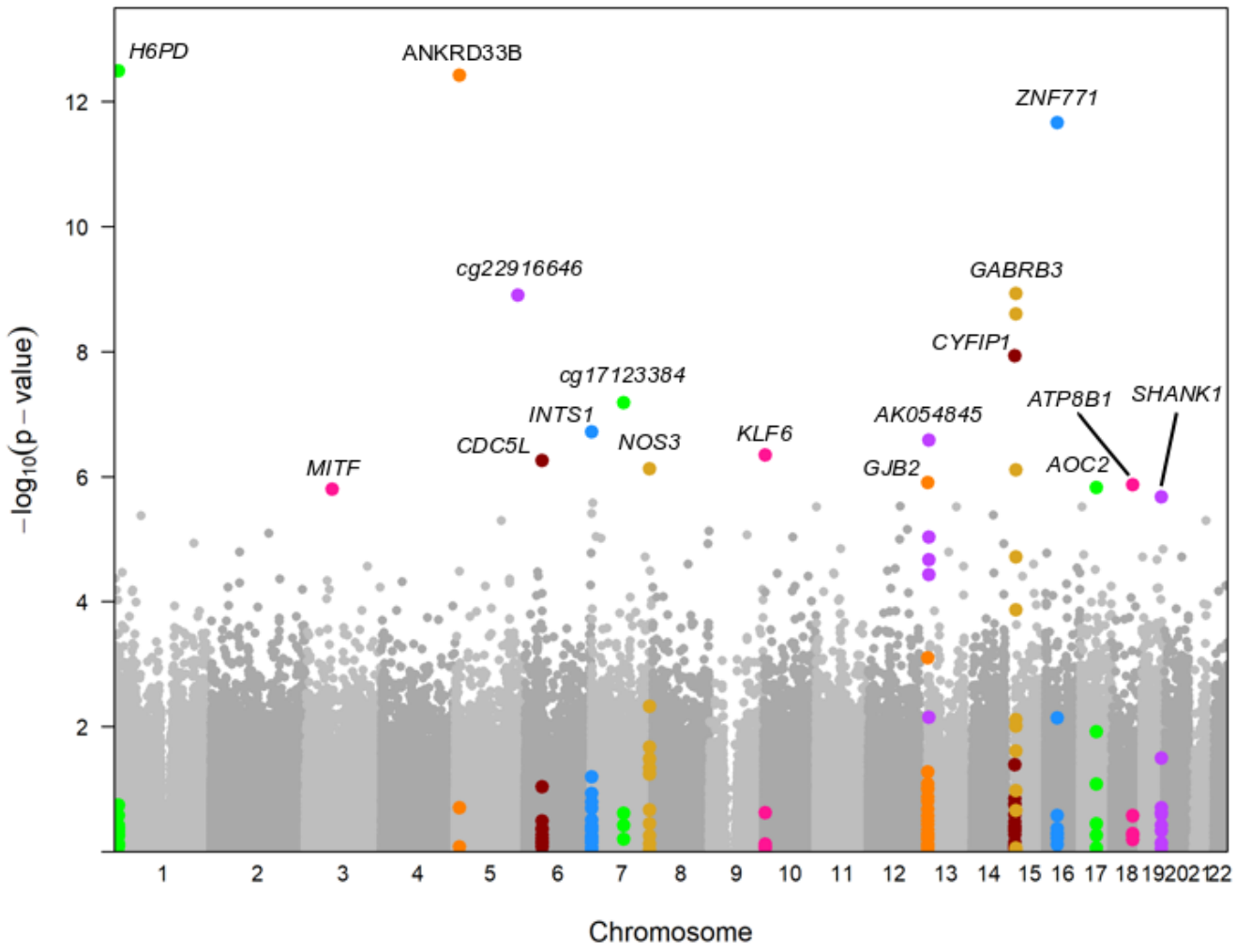

Figure 1. A manhattan plot showing the –log10 P-values (y-axis) for the association of DNA methylation at individual CpG dinucleotides across the 22 autosomal chromosomes (x-axis). CpG dinucleotides 5kb up- and downstream of the lead association have been colored in the same color as the lead association.

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

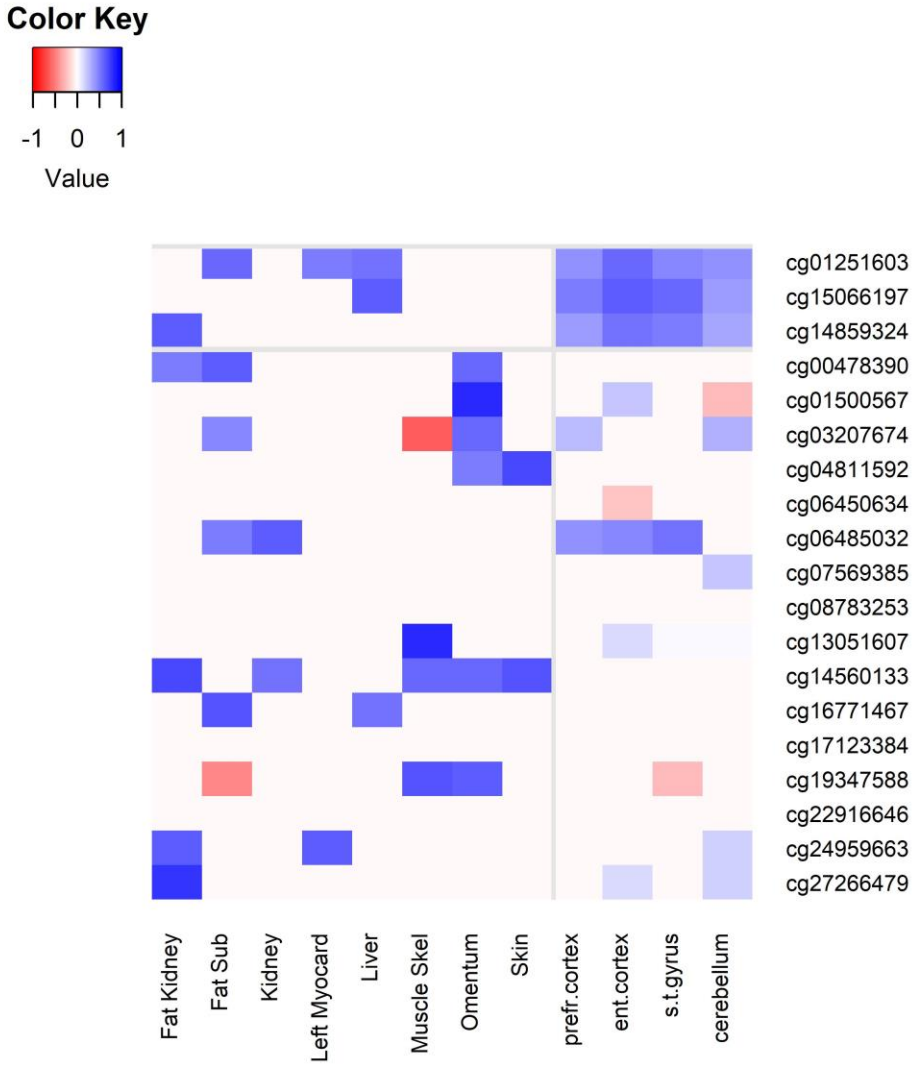

Figure 2. The spearman correlation plotted for 8 tissues from the body (N=16, P<0.1) and 4 from the brain (N>71, P<0.01) in colors ranging from bright red (*rho*= -1.0) to dark blue (*rho*=1.0). The light gray lines along the rows demark three CpG dinucleotides from the *GABRB3* region associated with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