

## Supplemental Information

### DAG

We used directed acyclic diagrams (DAGs) to investigate suitable sets of covariates for adjustment (using the R *daggity* package (1)). A DAG of only the observed correlations between co-variables and global DNAm (as represented by the first 3 principal components of the DNAm at autosomal CpG dinucleotides, Supplemental Figure 2) indicated that a direct effect of ART on DNAm can be estimated without adjustments. When we also consider in this DAG the small, but here unobserved, influence of ART on gestational age and ART on birthweight(2), and the well reported relation between gender and birthweight, adjustment for cell heterogeneity is necessary to establish the direct effect of ART on DNAm. Force including technical batch effects, gender of the child and cell heterogeneity, all of which have an effect on global methylation patterns, showed that a direct effect is possible to ascertain when including these major influences on global DNAm. Another approach would be the use of Factor Analysis, which may identify (hidden) confounders. We used the R package CATE (3) and found no indication for (hidden) confounders when adjusting cellular heterogeneity for the relationship between DNAm and IVF or ICSI nor when ART is taken as one group (see *Methods*)

Despite no indication for (hidden) confounders that influence DNAm on a large scale, many prenatal influences on DNAm have been reported for individual CpG dinucleotides. Therefore we investigated a DAG (Supplemental Figure 6 and 7) with (direct) effects of SES (4), smoking (5), maternal BMI (6), birthweight (7) and gestational age (8) and also the possible effect of the (in)fertility itself (as proxied by the years to index pregnancy) on DNAm. Maternal age, socio-economic status, maternal BMI, smoking history, maternal age, and years to index pregnancy was found as a minimal adjustment set to ascertain the direct effect of ART on DNAm in this scenario. Alternative, gestational age and parity can

be used in addition to batch and cellular heterogeneity in such a scenario. Therefore both sets of adjustments were used in sensitivity analyses.

Important is direction on the effect of cell heterogeneity on DNAm within these DAGs (9). DNAm may not merely reflect, but also drive cell heterogeneity for specific CpG dinucleotides (10) including in blood (11). When the simple scenario (Supplemental Figure 6) is considered the estimate for the effect of ART on DNAm is robust whatever the causal relationship between cell heterogeneity and DNAm. For the more elaborate DAG (Supplemental Figure 7 ) with more influences on DNAm, adjusting for cell heterogeneity impedes the estimation of a direct when DNAm causally drives cell heterogeneity. Therefore we also ran models without adjustment for cell heterogeneity in our sensitivity analyses.

Supplemental Table 1 Results of EWAS with and without cell type adjustment

CpG	Location hg19	Nearest gene <sup>1</sup>	distance <sup>1</sup>	With cell type adj.		Without cell type adj.		
				Estimate (SE) <sup>2</sup>	P <sup>2</sup>	P <sub>FDR</sub>	Estimate (SE) <sup>3</sup>	P <sup>3</sup>
cg27266479	chr1:9294882	<i>H6PD</i>	0	-1.94(0.27)	3.18e-13	8.28E-08	-1.85(0.31)	1.24e-09
cg04811592	chr3:69834386	<i>MITF</i>	0	0.99(0.21)	1.59e-06	0.039	1.07(0.24)	1.11e-05
cg24959663	chr5:10578618	<i>ANKRD33B</i>	0	3.91(0.54)	3.75e-13	8.28E-08	3.18(0.88)	3.22e-04
cg22916646	chr5:162672583	-	-	2.18(0.36)	1.22e-09	1.08E-04	1.99(0.46)	1.28e-05
cg01500567	chr6:44355777	<i>CDC5L</i>	0	0.48(0.1)	5.46e-07	0.020	0.51(0.11)	2.33e-06
cg00478390	chr7:150703765	<i>NOS3</i>	0	-0.92(0.19)	7.37e-07	0.024	-1.2(0.35)	6.56e-04
cg03207674	chr7:1523569	<i>INTS1</i>	0	0.72(0.14)	1.92e-07	9.43E-03	0.66(0.18)	2.06e-04
cg17123384	chr7:83379152	-	-	2.78(0.51)	6.39e-08	3.53E-03	3.14(0.64)	8.23e-07
cg19347588	chr10:3868336	<i>KLF6</i>	40862	0.81(0.16)	4.49e-07	0.018	0.92(0.19)	7.62e-07
cg07569385	chr13:20766226	<i>GJB2</i>	0	1.53(0.31)	1.24e-06	0.036	1.25(0.38)	1.08e-03
cg06485032	chr13:22615064	<i>AKO54845</i>	0	-3.47(0.67)	2.59e-07	0.011	-3.7(0.93)	6.86e-05
cg13051607	chr15:22956714	<i>CYFIP1</i>	0	1.39(0.24)	1.15e-08	7.24 E-04	1.3(0.32)	5.32e-05
cg01251603	chr15:26874098	<i>GABRB3</i>	0	-4.89(0.8)	1.18e-09	1.08 E-04	-4.7(0.91)	2.39e-07
cg15066197	chr15:26874202	<i>GABRB3</i>	0	-4.71(0.79)	2.48e-09	1.83 E-04	-4.48(0.82)	5.49e-08
cg14859324	chr15:26874363	<i>GABRB3</i>	0	-2.26(0.46)	7.68e-07	0.024	-2.29(0.56)	3.68e-05
cg06450634	chr16:30430044	<i>ZNF771</i>	0	3.26(0.46)	2.15e-12	3.17E-07	3.73(0.54)	4.25e-12
cg08783253	chr17:40996565	<i>AOC2</i>	42	-2.93(0.61)	1.48e-06	0.038	-2.99(0.92)	1.11e-03
cg16771467	chr18:55315872	<i>ATP8B1</i>	0	0.11(0.02)	1.34e-06	0.037	0.13(0.03)	4.09e-06
cg14560133	chr19:51199453	<i>SHANK1</i>	0	-1.42(0.3)	2.07e-06	0.048	-1.08(0.37)	3.51e-03

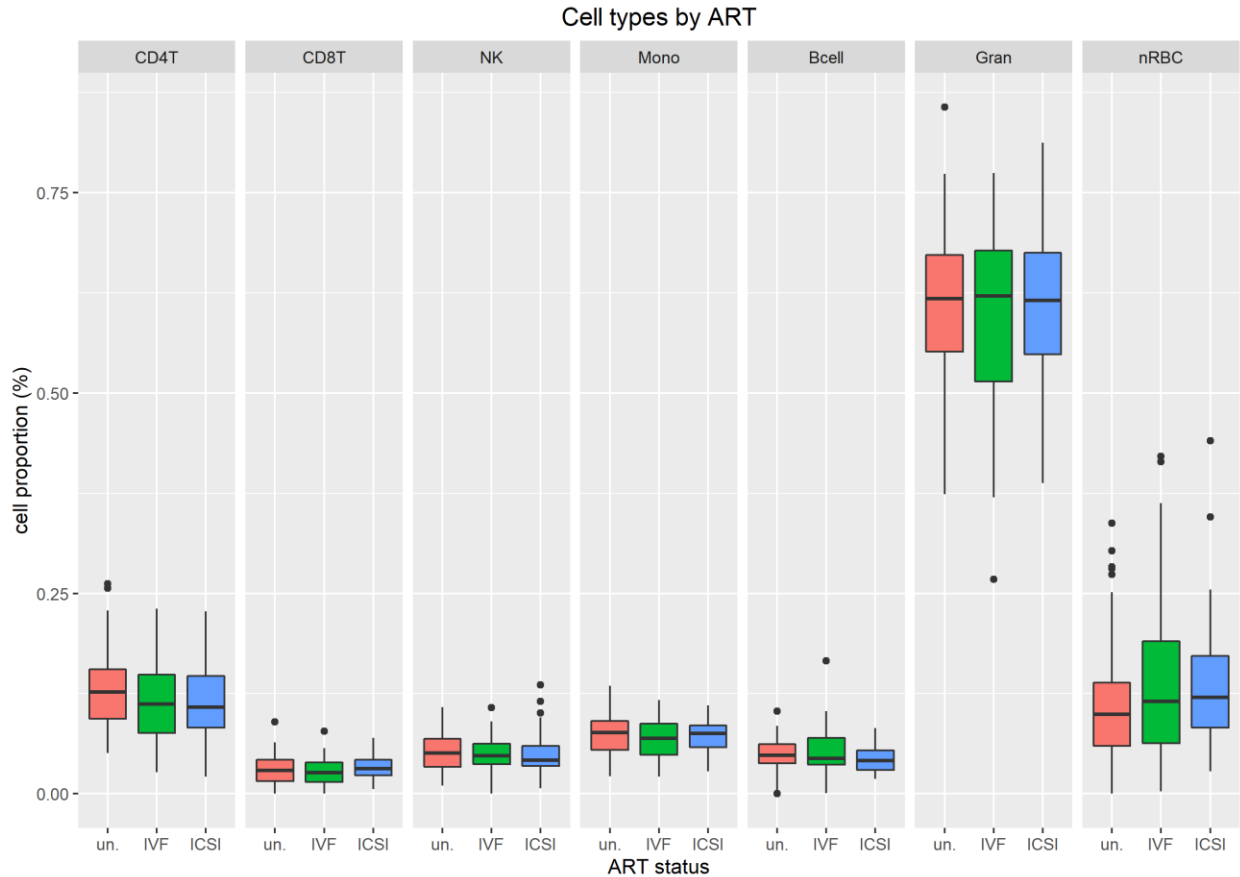
1. Nearest gene within 100kb and the distance to the TSS.
2. Estimate and standard error (SE) and P value, with and without false discovery rate (FDR) correction of the estimate of ART vs unassisted pregnancies in the model: Beta ~ ART (yes/no) + height on micro-array slide + scan batch + bisulfite plate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

Supplemental Table 2. Sensitivity analyses EWAS findings

probeID	Beta (SD) <sup>1</sup>	P <sup>1</sup>	Beta (SD) <sup>2</sup>	P <sup>2</sup>	Beta (SD) <sup>3</sup>	P <sup>3</sup>
cg27266479	-2.02(0.42)	1.48E-06	-1.87(0.3)	3.97E-10	-2.01(0.45)	6.58E-06
cg04811592	1.23(0.36)	0.000559	0.82(0.23)	0.000361	1.25(0.4)	0.0016
cg24959663	3.48(0.98)	0.000399	3.72(0.61)	9.15E-10	2.64(1.05)	0.011
cg22916646	2.58(0.54)	2.05E-06	2.03(0.38)	9.46E-08	2.22(0.61)	2.8E-04
cg01500567	0.54(0.17)	0.00187	0.5(0.1)	4.15E-07	0.43(0.17)	0.015
cg00478390	-1(0.3)	0.000831	-0.9(0.22)	3.24E-05	-0.98(0.33)	2.7E-03
cg03207674	0.72(0.2)	0.0033	0.74(0.17)	8.48E-06	0.66(0.21)	2.2E-04
cg17123384	2.47(0.71)	0.000506	2.59(0.54)	1.29E-06	2.2(0.92)	0.016
cg19347588	0.68(0.32)	0.033	0.68(0.17)	1.06E-05	0.45(0.33)	0.18
cg07569385	2.16(0.51)	2.39E-05	1.57(0.34)	3.92E-06	2.16(0.5)	1.49E-05
cg06485032	-5.11(1.19)	1.77E-05	-3.51(0.79)	8.06E-06	-4.65(1.31)	3.9E-04
cg13051607	1.59(0.34)	2.39E-06	1.35(0.28)	1.25E-06	1.79(0.42)	1.96E-05
cg01251603	-4.58(1.47)	0.00184	-5.1(0.81)	2.94E-10	-3.62(1.53)	0.017
cg15066197	-4.2(1.41)	0.00286	-4.93(0.8)	9.08E-10	-3.63(1.48)	0.014
cg14859324	-1.99(0.84)	0.017	-2.47(0.54)	3.93E-06	-1.61(0.96)	0.092
cg06450634	2.9(0.78)	0.000184	3.04(0.55)	3.26E-08	2.62(0.78)	0.00076
cg08783253	-2.6(1.12)	0.02	-2.59(0.67)	0.000102	-2.52(1.21)	0.038
cg16771467	0.14(0.03)	1.07E-05	0.11(0.02)	2.03E-05	0.14(0.04)	2.64E-4
cg14560133	-1.44(0.75)	0.054	-1.41(0.35)	5.16E-05	-0.99(0.62)	0.11

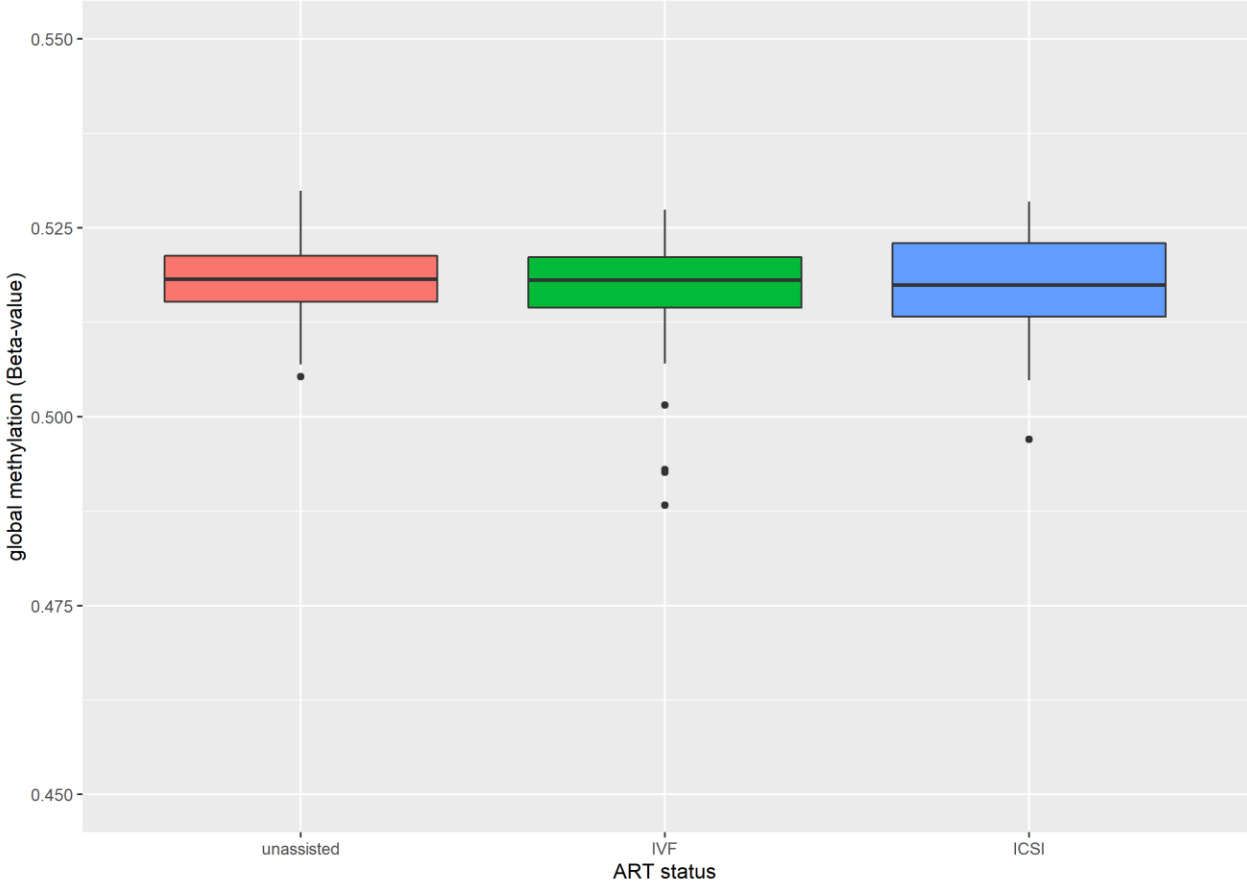
1. Estimate (SE) and P value for ART yes/no analysis with adjustment of the DAG minimal adjustment set:  $B \sim \text{ART (yes/no)} + \text{position} + \text{scan batch} + \text{Bisulfite plate} + \text{sex} + \text{cell heterogeneity} + \text{maternal BMI} + \text{maternal smoking history} + \text{years to index pregnancy}$ .
2. Estimate (SE) and P value for ART yes/no analysis with adjustment of :  $B \sim \text{ART (yes/no)} + \text{position} + \text{scan batch} + \text{Bisulfite plate} + \text{sex} + \text{cell heterogeneity} + \text{gestational age} + \text{parity (ordinal factor)}$ .
3. Estimate (SE) and P value for ART yes/no analysis with adjustment of maternal characteristics:  $B \sim \text{ART (yes/no)} + \text{position} + \text{scan batch} + \text{Bisulfite plate} + \text{sex} + \text{cell heterogeneity} + \text{maternal BMI} + \text{maternal smoking history} + \text{years to index pregnancy} + \text{maternal age} + \text{socio-economic status (ordinal factor)}$ .

Supplemental Figure S1. Imputed cell type proportions in cord blood



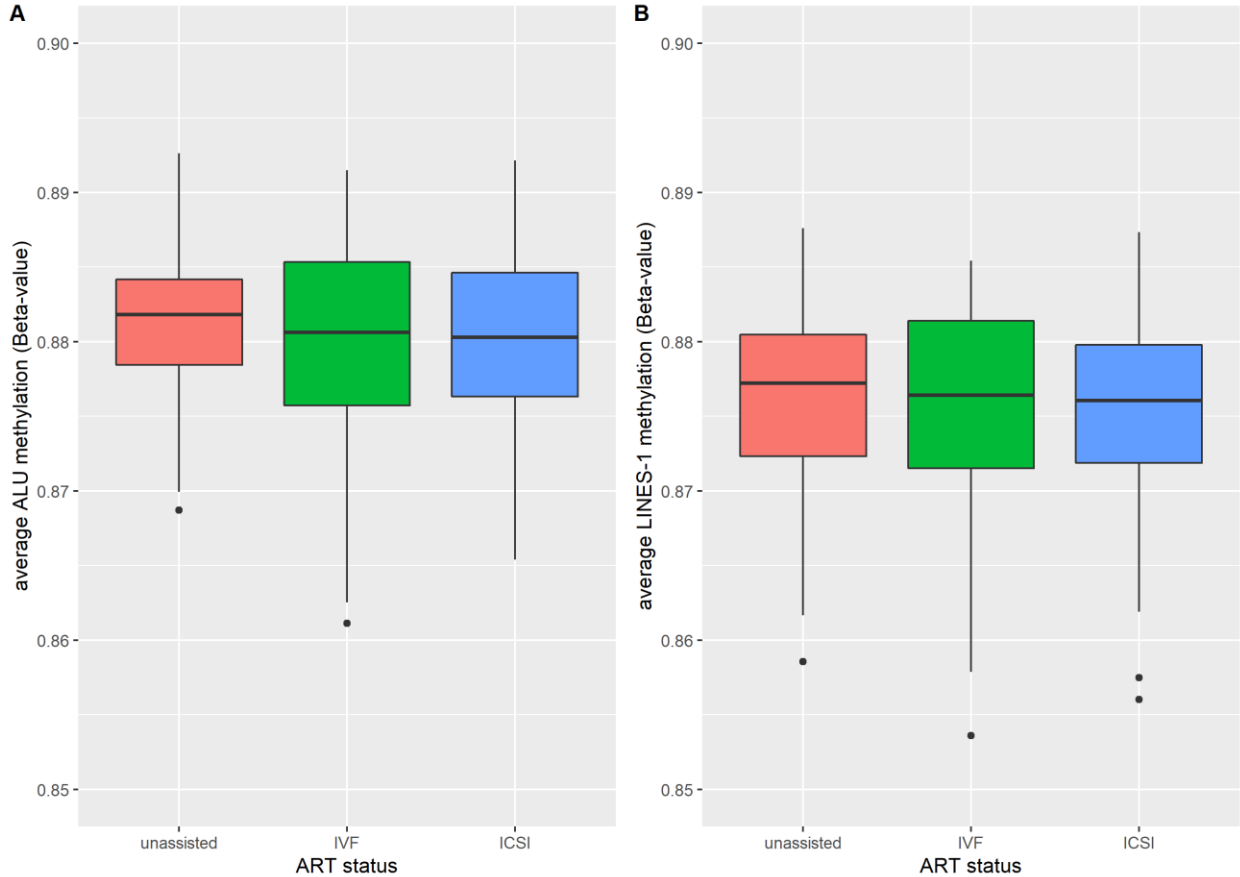
Overview of cell imputed cell proportions in unassisted (un.), IVF and ICSI newborn cord blood as imputed via the DNA methylation data.

Supplemental Figure S2. Genome-wide average methylation (GWAM)



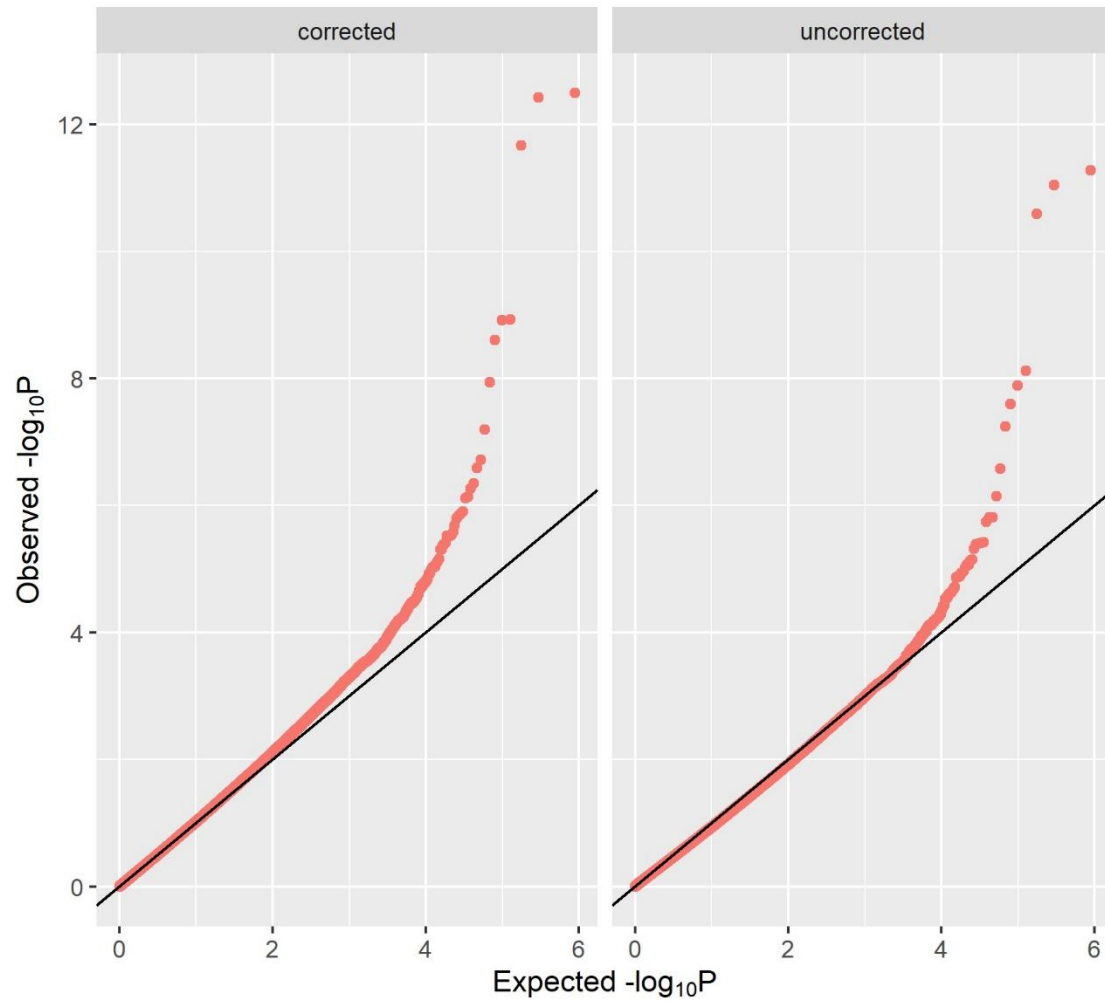
The genome-wide average methylation as denoted in beta-value (ranging from 0-1), as measured by the 450k array, in the cord blood of children conceived unassisted (control) or via IVF or ICSI.

Supplemental Figure S3. Genome-wide average methylation of ALU and LINES-1 elements



The genome-wide average methylation as denoted in beta-value (ranging from 0-1), as measured by the 450k array for **A.** CpG dinucleotides within *ALU* elements **B.** CpG dinucleotides within *ALU* elements *LINES-1* elements.

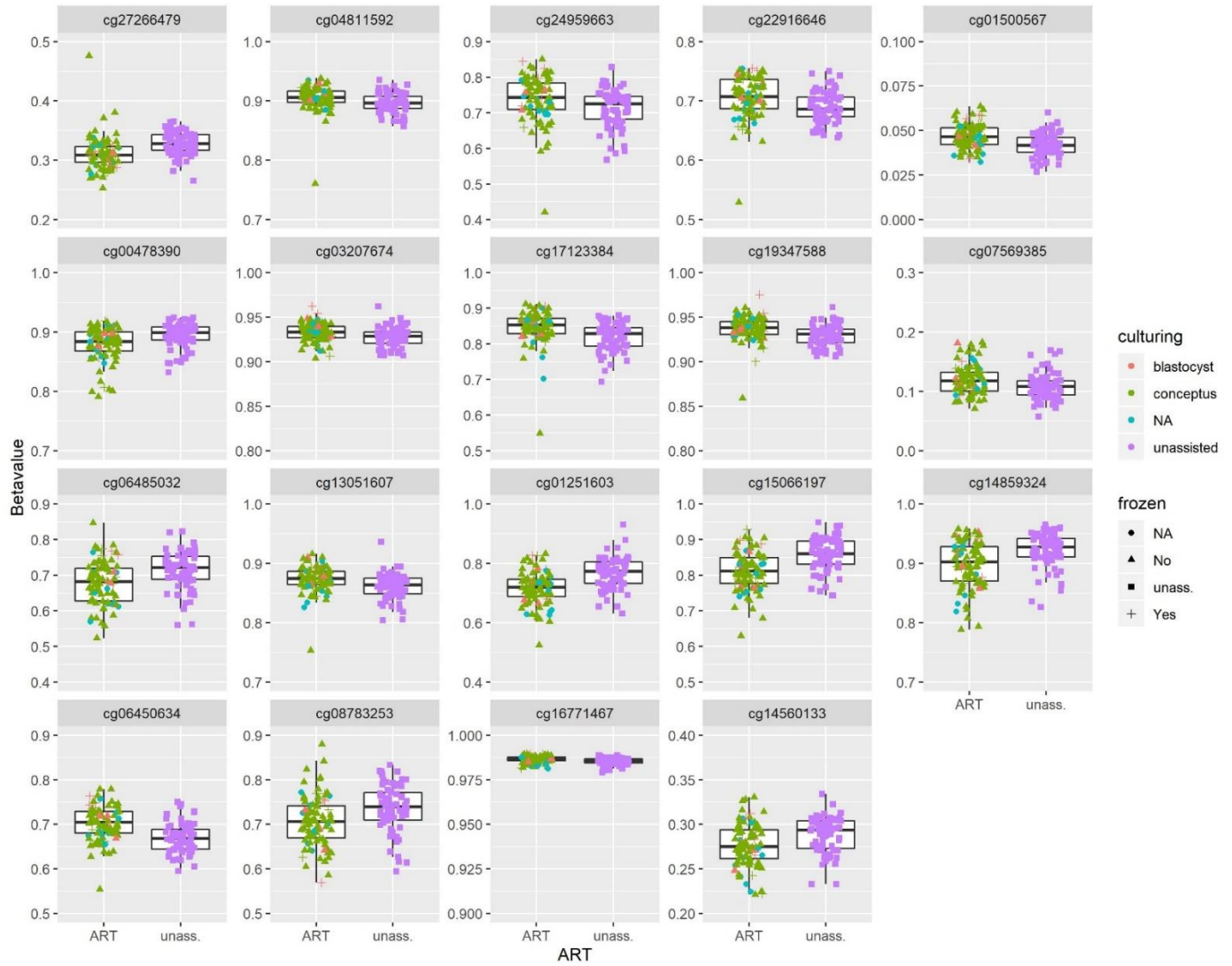
Supplemental Figure S4. QQ-plot of the EWAS for ART status



Plotted on the x-axis are the expected p values by chance versus the observed p-values in the analysis of ART vs unassisted EWAS corrected for technical batches and cell heterogeneity. The genome-scale analysis across probes showed signs of deflation ( $\lambda=0.9428$ ) and deviations from a perfect normal distribution (e.g. "bias",  $\sigma=0.0485$ ). Correction for deflation and bias was performed with the R package *bacon*.

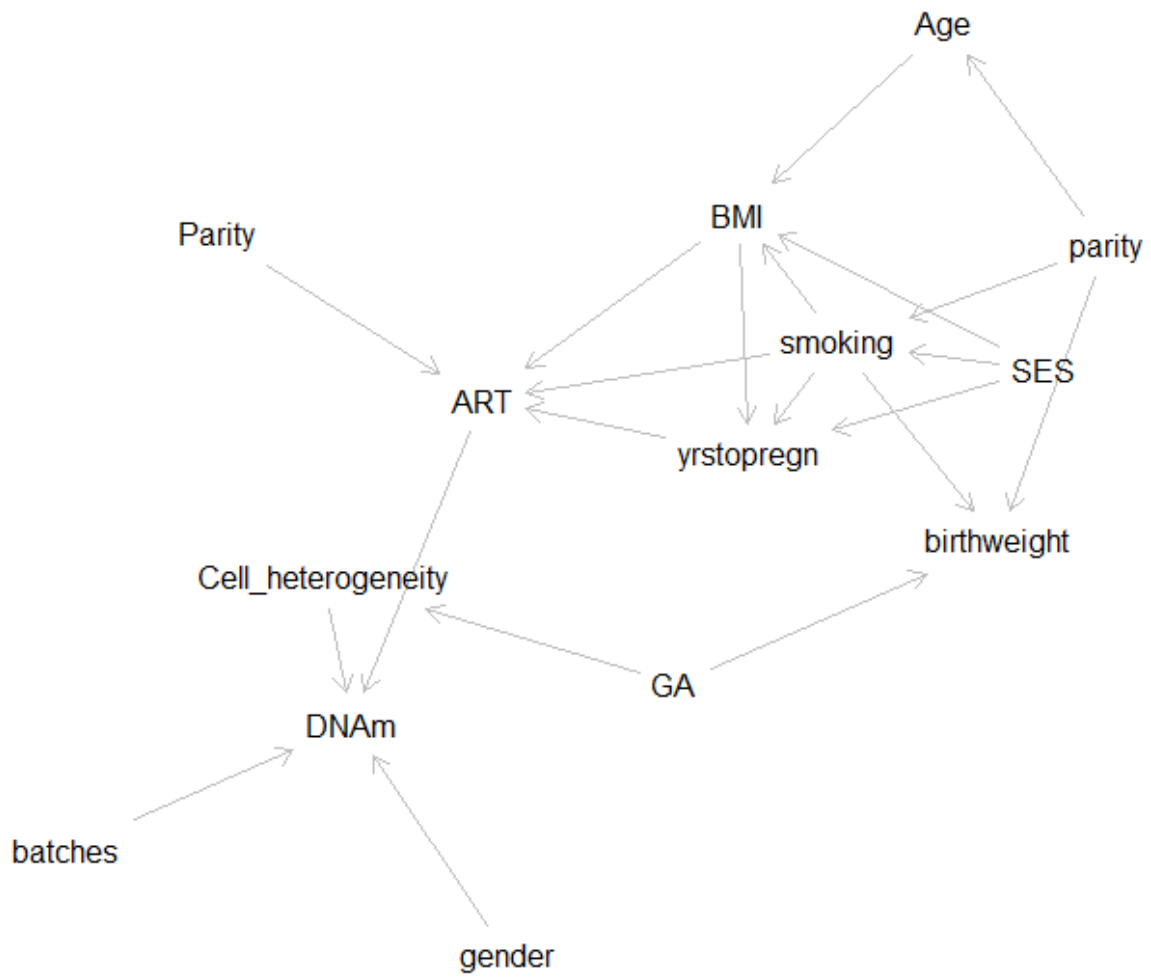


Supplemental Figure S5. Boxplot for the difference in methylation between neonates conceived via ART or unassisted.



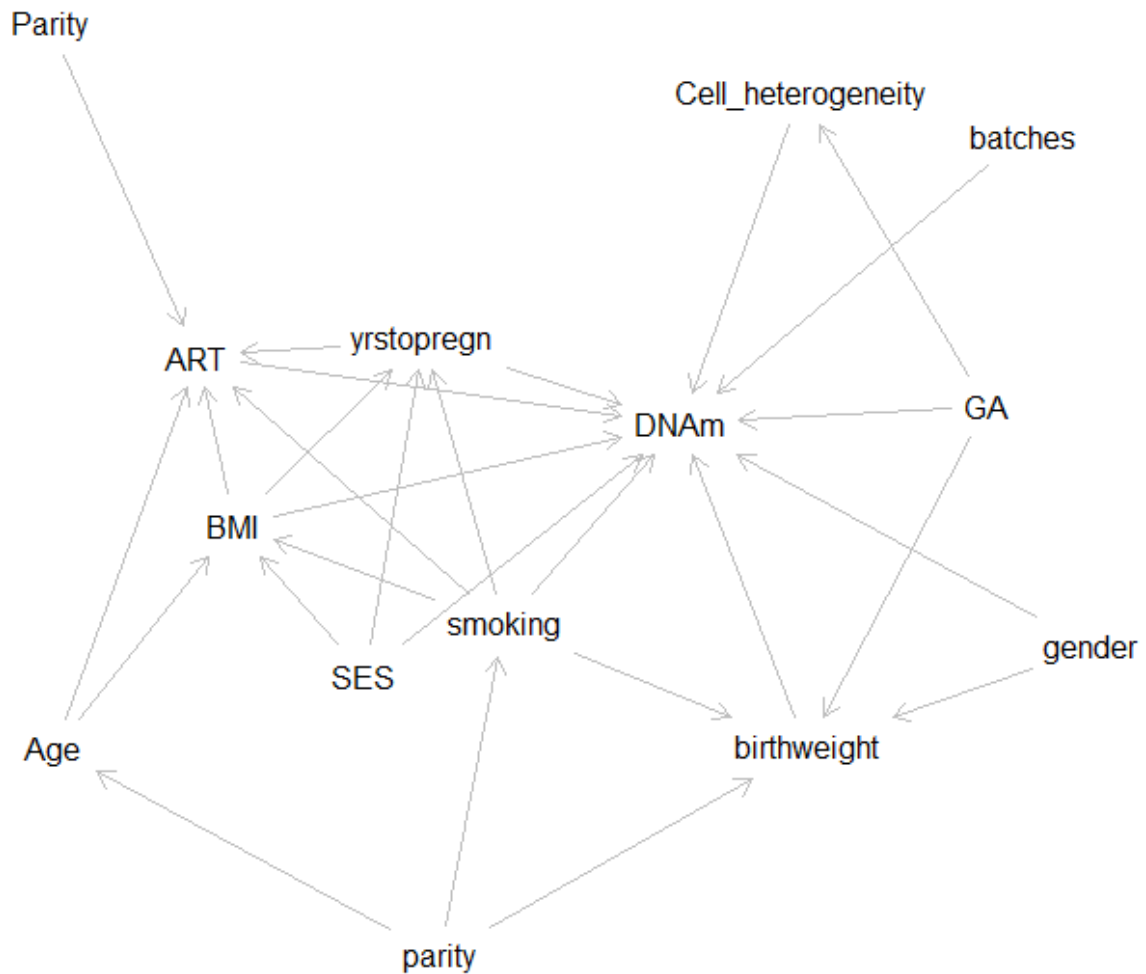
Boxplot with overlaying “jitter” for the 19 CpG dinucleotides associated with ART. With differing colors the degree of in vitro culturing is shown (orange= culturing till the blastocyst stage, dark green = culturing stopped

Supplemental figure S6. DAG of observed correlations



The directed acyclic diagram of the observed correlations in the studied individuals. GA = gestational age, BMI = body mass index, SES = socio-economic status as assessed by maternal education, smoking = is a variable denoting if the mother has smoked during her lifetime (yes/no), Age = maternal age at birth child, gender = sex of child.

Supplemental figure S7. DAG of observed correlations and additional possible links with DNAm



The directed acyclic diagram of the observed correlations in the studied individuals and additional possible associations between these variables and DNAm. GA = gestational age, BMI = body mass index, SES = socio-economic status as assessed by maternal education, smoking = is a variable if denoting if the mother has smoked during her lifetime (yes/no), Age = maternal age at birth child, gender = sex of child.