Duration of breastfeeding is associated with leptin (*LEP*) DNA methylation profiles and

BMI in 10-year-old children

William B Sherwood[1], Victoria Bion[1], Gabrielle A Lockett[1], Ali H Ziyab[2], Nelís Soto-Ramírez[3], Nandini Mukherjee[4], Ramesh J Kurukulaaratchy[5,6], Susan Ewart[7], Hongmei Zhang[4], S Hasan Arshad[5,6], Wilfried Karmaus[4], John W Holloway[1,5], Faisal I Rezwan[1,*]

*Corresponding author

[1] Human Development and Health, Faculty of Medicine, University of Southampton, Southampton General Hospital, Southampton, SO16 6YD, UK

[2] Department of Community Medicine and Behavioral Sciences, Faculty of Medicine, Kuwait University, Kuwait City, Kuwait

[3] College of Social Work, University of South Carolina, Columbia, SC, USA

[4] Division of Epidemiology, Biostatistics and Environmental Health, School of Public Health, University of Memphis, 236A Robison Hall, Memphis, TN 38152

[5] Clinical and Experimental Sciences, Faculty of Medicine, University of Southampton, UK

[6] The David Hide Asthma and Allergy Research Centre, Isle of Wight, UK

[7] College of Veterinary Medicine, Michigan State University, East Lansing, MI, USA

Correspondence to:

Faisal I Rezwan:

Postal address: Duthie Building, MP808, Southampton General Hospital, Tremona Road, Southampton, SO16 6YD, Hampshire, UK.

Tel: +44 (0)238 120 8783

Email: F.Rezwan@soton.ac.uk

Abstract

Background: Breastfeeding is protective against many long-term diseases, yet the mechanisms involved are unknown. Leptin gene (*LEP*) is reported to be associated with body mass index (BMI). On the other hand, breastfeeding duration has been found to be associated with DNA methylation (DNAm) of the *LEP* gene. Therefore, epigenetic regulation of *LEP* may represent the mechanism underlying the protective effect of breastfeeding duration against obesity.

Methods: In the Isle of Wight Birth Cohort, peripheral blood DNAm at 23 Cytosine-phosphate-Guanine sites (CpGs) in the *LEP* locus in 10-year-old (n=297) samples and 16 CpGs in 18-year-old (n=305) samples, were generated using the Illumina Infinium MethylationEPIC and HumanMethylation450 Beadchips respectively and tested for association with breastfeeding duration (total and exclusive) using linear regression. To explore the association between breastfeeding durations and genome-wide DNAm, epigenome-wide association studies (EWASs) and differential methylation region (DMR) analyses were performed. BMI trajectories spanning the first 18 years of life were used as the outcome to test the association with breastfeeding duration (exposure) using multi-nominal logistic regression. Mediation analysis was performed for significant CpG sites.

Results: Both total and exclusive breastfeeding duration were associated with DNAm at four *LEP* CpG sites at 10 years (P-value <0.05), and not at 18 years. Though no association was observed between breastfeeding duration and genome-wide DNAm, DMR analyses identified five significant differentially methylated regions (Sidak adjusted P-value <0.05). Breastfeeding duration was also associated with the early transient overweight trajectory. Furthermore, DNAm of *LEP* was associated with this trajectory at one CpG site and early persistent obesity at another, though mediation analysis was not significant.

Conclusions: Breastfeeding duration is associated with *LEP* methylation at age 10 years and BMI trajectory. *LEP* DNAm is also significantly associated with BMI trajectories throughout childhood, though sample sizes were small.. However, mediation analysis did not demonstrate that DNAm of *LEP* explained the protective effect of breastfeeding against childhood obesity.

Keywords

BMI trajectories, Breastfeeding, DNA methylation, exclusive breastfeeding, epigenetics, obesity, *Leptin*.

Introduction

The multiple health benefits of breastfeeding are well recognised, with the World Health Organisation (WHO) recommending that infants should be exclusively breastfed up until six months and then breastfed alongside complementary food until two years of age (1), though globally, a high proportion of infants are not breastfed according to international guidelines (2,3).

Infants who are not breastfed are at increased risk of all-cause and infection-related mortality (4). For infants with higher intensity of breastfeeding, there is a reduction in morbidities such as respiratory infection and diarrhoea in their first five years of life (5). Long-term, adequate breastfeeding has been associated with a range of phenotypes including lowered risk of Type 2 diabetes (6), improved performance in intelligence tests (7), and reduced risk of childhood leukaemia (8). Conversely, associations between breastfeeding and later disease risk are not always as clear cut, such as in the context of allergic disease (9–11). Methodical disparities including defining exposure and adjustment for confounding factors make the comparison between studies difficult (12). Similarly, conflicting results on the relationship between breastfeeding and obesity have been attributed to heterogeneity between studies, residual confounding and insufficient statistical power (13). However several studies and meta-analyses have recently reported a significant association between breastfeeding duration and childhood obesity (6,14,15). Based on current evidence, the WHO reports that breastfeeding appears to reduce the risk of being overweight and obese in childhood (16).

Early life nutrition, and breastfeeding in particular, up to two years of life, are considered to be critical for programming of long-term health. The composition of human breastmilk, as a source of nutrition and bioactive factors, is crucial in understanding the mechanism behind its beneficial effect (17). For example, breastmilk is high in long-chain polyunsaturated fatty

acids (LCPUFAs), which in adults are associated with a lower risk of metabolic syndrome and cardiovascular disease (18), though evidence of their benefit in children is less clear (19,20). Moreover, maternal antibodies in human milk are taken up by an infant's intestinal dendritic cells to support their immature acquired immunity (17). Exactly how these metabolites and antibodies convey the long-term protective effects of breastmilk needs further investigation (21), though recent evidence suggests that some may be directly related to epigenetic changes (22) and infants' gut microbiota (23).

A variety of nutritional sources are thought to induce epigenetic changes, including changes in DNA methylation (DNAm) (24). In particular, early life nutritional exposures, such as maternal protein restriction in pregnancy and methyl donor nutrients including folate, methionine and some B-vitamins, may leave long-lasting changes to DNAm in humans (25,26). Furthermore, components of human breastmilk are thought to induce epigenetic changes such as to LCPUFAs (27), and lactoferrin, which by suppressing the NF-κB signalling pathway may explain why breastmilk is protective against neonatal necrotizing enterocolitis (22).

However to date, few studies have investigated the association between breastfeeding duration and DNAm levels (28,29). In one of these studies, Obermann-Borst and colleagues demonstrated that breastfeeding duration was significantly associated with a reduction in methylation in whole blood DNA of the leptin gene *(LEP)* in young children (mean age 17 months) (28). *LEP* is expressed in adipocytes and encodes the hormone leptin, important in the regulation of energy intake by inducing early satiety. However, defective production of *LEP* is associated with obesity (30), and *LEPR/LEPROT* DNAm has been demonstrated to interact with genotype to influence both leptin levels and BMI (31). This suggests that breastfeeding duration, associated with postnatal DNAm, may contribute to childhood obesity. The aim of this study was to replicate and extend the study by Obermann-Borst *et*

al., in the Isle of Wight Birth Cohort (IOWBC) to test the association between breastfeeding

and DNAm of Cytosine-phosphate-Guanines (CpGs) in *LEP* gene in peripheral blood at later

ages (10 and 18 years) and further explore the association between breastfeeding duration and

body mass index (BMI), where DNAm may act as a mediator.

Methods

Isle of Wight (IoW) birth cohort (IOWBC)

The IOWBC (2nd generation also known as IoW F_1,) was recruited between 1989 and 1990

(n=1,536) (32,33). The parents (1st generation, IoW F_0) of all infants born over this period

were contacted at birth, and subsequently 95% of infants (n=1456) were enrolled following

informed consent and exclusion. Follow ups were conducted at 1, 2, 4, 10, and 18 years,

where information was gathered about the participants, such as infant nutrition and

breastfeeding practice (34). Peripheral blood samples were collected from participants at age

10 and 18 years and DNAm measured for 297 and 305 samples, at ages 10 and 18

respectively, of whom 162 participants were matched between the two groups. Descriptive

statistics of the sub-cohorts for exposure and covariates are presented in Table 1.

Categorisation of breastfeeding duration

Two categories, exclusive and total breastfeeding durations, were used to define the duration

of breastfeeding for each participant. The exclusive breastfeeding duration was defined as the

length of time (in weeks) a child was breastfed, until formula feed and/or solid foods were

144 introduced. The total breastfeeding duration, equates to the total number of weeks a mother

145 breastfed her child irrespective of the introduction of formula feed and/or solid foods.

146

DNA extraction and arraying

148 DNA was extracted from peripheral blood samples using a standard salting out procedure.

149 Approximately 1 μg of DNA was bisulfite treated using the EZ 96-DNA methylation kit

150 (Zymo Research, Irvine, CA, USA) using the manufacturer's standard protocol. The Infinium

151 MethylationEPIC BeadChips (age 10-year samples) and Infinium HumanMethylation450

152 (age 18-year samples) from Illumina (Illumina, San Diego, CA, USA) were used to obtain

153 methylation levels following the manufacturer's standard protocol. Methylation 450k and

154 EPIC data (β values) were pre-processed for quality control, following the CPACOR pipeline

155 (35), and batch effect was corrected using ComBat (36).

156

Candidate *Leptin* CpG selection

158 In Obermann-Borst *et al.* study, seven CpG sites in the *LEP* gene were analysed. The UCSC

159 Batch Coordinate Conversion tool (LiftOver, https://genome.ucsc.edu/cgi-bin/hgLiftOver)

160 was used to convert genome coordinates from the hg18 assembly, used by Obermann-Borst

161 *et al.*, to the human genome hg19 assembly (37). Four of these seven CpG sites were present

162 in our analysis. For all CpGs annotated near *LEP* gene, for the 18-year samples a total of 16

163 CpG sites were analysed in *LEP*, whereas 23 CpG sites in *LEP* were analysed in the 10-year

164 samples, where the more comprehensive EPIC beadchips were used (See Supplementary

165 Table S1).

166

Confounding factors

The same covariates used by Obermann-Borst *et al*. were used in the analysis, namely child sex, birth weight, gestational age, serum leptin (ng/mL at either 10 or 18 years), BMI (at either 10 or 18 years), maternal socioeconomic status (defined using maternal socioeconomic cluster information: high, low, low-low, low-mid, and mid, using household income, number of rooms and maternal education), and maternal smoking at birth. However, unlike the original study, birth weight and gestational age were kept as separate variables and z-scores were not considered to be appropriate (38). Parity and maternal age were added as covariates for these analyses, having been previously found to be associated with breastfeeding duration (39). Additionally, breastfeeding duration was used as a continuous variable rather than categorical. Cell proportions (CD8T, CD4T, NK, B Cells, Monocytes, Granulocytes) were estimated using the minfi package (40) and cell composition coefficients were derived using the Houseman method (41).

Commented [RF1]: Reviewer #1 comment on maternal socio-economic status

BMI Trajectories

Four BMI trajectories (normal, early persistent obesity, delayed overweight, and early transient overweight) covering the period from 1 to 18 years of age, as proposed by Ziyab *et al*. (42), were utilised to assess the association between breastfeeding duration and BMI trajectories. These four trajectories were categorized using age-specific BMI thresholds for 'obesity' and 'overweight', as defined by the International Obesity Taskforce for ages 4, 10 and 18 years (43), and WHO standards for age 1 year (44). They were defined in 1,240 participants of the IOWBC and the trajectories for 10 and 18-year samples are described in Table 1.

190

Statistical analyses

192 Intra-class correlations (ICC) were estimated among 14 common *LEP* CpGs in 162 matched

193 samples (common between 10 and 18 years) to assess the DNAm stability over time. Four

194 models were run exploring the association between exclusive breastfeeding duration or total

195 breastfeeding duration and *LEP* DNAm at either 10 or 18 year using linear regression

196 adjusted for covariates and cell types.

197

198 Further, we performed an epigenome-wide association study (EWAS) to identify

199 differentially methylated positions (DMPs) that are associated with breastfeeding durations

200 (both exclusive and total), using linear regression adjusted for confounding factors and cell

201 types, in both 10 and 18 year samples as described for the *LEP* analyses. Multiple hypothesis

202 testing was accounted for by controlling the false discovery rate (FDR), using Benjamini and

203 Hochberg's algorithm (45). CpGs with FDR corrected P-value < 0.05, were considered

204 statistically significant. To identify differentially methylated regions (DMRs), composed of

205 multiple signals across individual CpG positions, comb-p (46) (Python v2.7) was used.

206 Comb-p identifies regions enriched with low unadjusted P-values (here, P-value < 0.05) from

207 the EWAS analysis, based on the probe location. For each region, the comb-p algorithm

208 adjusts the CpG P-values for auto-correction between probes by using the Stouffer-Liptak-

209 Kechris (slk) correction with multiple testing adjustment using a one-step Sidak correction

210 method. Regions with least two CpG probes within 200 base pairs, having a Sidak-corrected

211 P-value < 0.05, were considered statistically significant.

212

213 The association between breastfeeding duration (exclusive and total) and BMI trajectories

214 was tested in the subset of children, matched with 10-year old samples, using multi-nominal

215 logistic regression with 'normal' BMI trajectory as reference. Furthermore, the association

216 between BMI trajectory and *LEP* DNAm was investigated with only those CpGs found to be

217 significant with breastfeeding duration. P-value < 0.05 was used as the significance level. All

218 analyses were performed using R (version 3.3.2).

219

Mediation Analysis

221 Model-based causal mediation analysis, using the R-package "mediation" (47), was

222 performed to explore whether the association between breastfeeding duration and BMI

223 trajectory is mediated by DNAm (Figure 1). Only the *LEP* CpGs significantly associated with

224 BMI trajectory were utilised for the mediation analysis as the mediator. Firstly, we estimated

225 the effect of breastfeeding duration (exposure) on *LEP* DNAm (mediator) and then, we

226 estimated combined effect of breastfeeding duration (exposure) and *LEP* DNAm (mediator)

227 on the BMI trajectory (outcome). Both of these effects were adjusted for the covariates

228 mentioned above. Bootstrapping was used to test significance level (P-value < 0.05).

229

230

Results

Stability of DNA methylation

233 Intra-class correlation shows that 64.3% of the common CpGs (9 out of 14) were fairly-

234 moderately stable (ICC between 0.05 to 0.75) over eight years (Supplementary Table S2). A

235 moderate-good degree of stability was found in three CpGs between the two measurements.

236 The highest ICC observed (in cg19594666) was 0.69 with a 95% confidence interval from 237 0.58 to 0.77 (F(161,161) = 3.22, P-value = 2.79×10^{-13}).

Breastfeeding duration and *LEP* DNA methylation

239 At 10 years, the total breastfeeding duration was significantly associated with DNAm at four 240 CpG sites in *LEP* (P-value < 0.05, Table 2). Two of the CpGs (cg03084214 and cg23753947) 241 were available in 18-year samples, which showed significant differences with 10-year 242 samples (Supplementary Figure 1). However, for the 18-year samples, there were no 243 significant associations seen between either total breastfeeding or exclusive breastfeeding 244 duration and DNAm at any of the 16 *LEP* CpG sites interrogated (Supplementary Tables S3 245 and S4).

246

247 **Table 2** Association between total breastfeeding duration and *LEP* DNAm at 10 years

Probe name	Estimate	SE	P-value
cg03084214	5.23E-04	2.33E-04	0.026
cg11005360	-2.39E-04	1.07E-04	0.027
cg23381058	-4.91E-04	1.96E-04	0.013
cg23753947	1.73E-04	8.20E-05	0.036

248 Columns: Probe name = *LEP* CpG sites from Illumina Infinium MethylationEPIC Beadchip 249 significantly associated with total breastfeeding duration; Estimate = coefficient denoting the 250 increase or decrease in methylation by one unit (in week) of increase of total breastfeeding 251 duration; SE = standard error of coefficient; P-value < 0.05 was considered significant. Only 252 the four CpG sites considered significant are included, see Supplementary Table S5, for all 253 23 CpG sites tested.

Similarly, the duration that a child was exclusively breastfed was associated with significant changes (P-value < 0.05) in DNAm at four CpG sites, in samples taken at 10 years (Table 3). Lower methylation of one site (cg23381058) was significantly associated with both exclusive and total breastfeeding duration. Another CpG (cg18603538) was available in 18 year samples, which showed no significant differences with 10 year samples (Supplementary Figure 1).

Table 3 Association between exclusive breastfeeding duration and *LEP* DNAm at 10 years

Probe name	Estimate	SE	P-value
cg05091920	4.23E-04	2.12E-04	0.048
cg15792829	5.87E-04	2.37E-04	0.015
cg18603538	8.07E-04	3.69E-04	0.030
cg23381058	-8.17E-04	4.00E-04	0.043

Columns: Probe name = *LEP* CpG sites from Illumina Infinium MethylationEPIC Beadchip significantly associated with exclusive breastfeeding duration; Estimate = coefficient denoting the increase or decrease in methylation by one unit (in week) of increase of exclusive breastfeeding duration; SE = standard error of coefficient; P-value < 0.05 was considered significant. Only the four CpG sites considered significant are included, see Supplementary Table S6, for all 23 CpG sites tested.

Genome-wide DMPs and DMRs identifications

The EWAS analyses for both exclusive and total breastfeeding did not identify any significant DMPs (FDR < 0.05) at either age 10 or 18 years (Supplementary Figures 2 - 5).

272 Analysis of differentially methylated regions using comb-p identified five significant DMRs

273 (Sidak-corrected p-values: 9.6 x 10^{-9} - 0.005) (Table 4). We identified two DMRs for 10-year

274 samples and one for 18-year samples in the case of total breastfeeding duration. For exclusive

275 breastfeeding duration, we identified two significant DMRs for 18-year samples and none for

276 the 10-year samples. There was no overlap between the DMRs identified for exclusive and

277 total breastfeeding.

278

279 **Table 4** Statistically significant differentially methylated regions (DMRs) (Sidak P-value <

280 0.05)

281 (a) Total breastfeeding duration

Age	Location	No. probes	Slk P-value	Sidak P-value	Ref Gene name and genomic feature		CpG feature
10	chr10:104535854-104535920	2	4.62E-07	0.005	WBP1L	TSS200; Body;5'UTR; 1stExon	NA
10	chr19:54567060-54567128	3	3.74E-07	0.004	VSTM1	5'UTR;1stExon;Body	NA
18	chr22:37678728-37678791	2	2.49E-07	0.002	CYTH4	Body	NA

282

283 **(b)** Exclusive breastfeeding duration

Age	Location	No. probes	Slk P-value	Sidak P-value	Ref Gene name and genomic feature		CpG feature
18	chr1:173837002-173837197	7	4.13E-12	9.62E-09	GAS5	Body;TSS150; TSS200	Island, Shore
18	chr2:24397787-24397845	3	1.43E-07	0.001	FAM228A	TSS200	Island

284 Both Slk = Uncorrected Stouffer-Liptak-Kechris p-values, and Sidak p-values corrected for

285 multiple testing are reported. Here, 5'UTR = 5 prime untranslated region; TSS200 = 0–200

286 bases upstream from the transcription start sites; TSS1500 = 200–1500 bases upstream from

287 the transcriptional start site; body = gene body.

288

289

Breastfeeding duration and BMI Trajectories

Total duration of breastfeeding was significantly associated with the early transient

overweight BMI trajectory (Estimate = -0.02, P-value = 0.002) (Table 5). However, there was

no association with the early persistent obesity or delayed overweight trajectories. A similar

pattern was seen when testing exclusive breastfeeding duration and obesity (Table 5), though

the association between exclusive breastfeeding duration and early transient overweight

narrowly missed our threshold for significance (Estimate = -0.03, P-value = 0.05).

297

Table 5 Association between duration of total and exclusive breastfeeding and BMI

trajectory

BMI trajectory	Total breastfeeding			Exclusive breastfeeding		
	Estimate	SE	P-value	Estimate	SE	P-value
2	0.007	0.007	0.29	0.002	0.014	0.88
3	-0.02	0.008	0.002*	-0.03	0.015	0.05
4	0.012	0.014	0.40	-0.02	0.030	0.43

BMI trajectory defined by Ziyab *et al*. (42). Columns: trajectory 2 = early persistent obesity,

trajectory 3 = early transient overweight, trajectory 4 = delayed overweight; Estimate =

coefficient denoting one-unit (in week) increase in breastfeeding duration is associated with

the increase in the log odds of being in BMI trajectory; Standard Error = standard error of

304 coefficient. P-value < 0.05 was considered significant and is indicated by *. Each trajectory

305 was tested against the control group (BMI Trajectory 1).

306

DNA methylation and BMI Trajectories

308 Of the four CpG sites where total breastfeeding duration was associated with DNAm, one site

309 (cg23381058) also had a significant association between DNAm and the early transient

310 overweight trajectory (P-value = 0.034) (Table 6). Note this CpG site was associated with

311 both total and exclusive breastfeeding duration. Of the four sites where exclusive

312 breastfeeding was associated with DNAm, methylation at another site (cg05091920) was

313 significantly associated with the early persistent obesity trajectory (P-value = 0.028). These

314 results have been placed in Supplementary Table S7 along with other non-sginificant CpGs

315 for all of the trajectories.

316

317 **Table 6** Association between early transient overweight (BMI trajectory 3) and *LEP* DNAm

318 at 10 years previously associated with breastfeeding duration (associated P-values are shown)

	Total breastfeeding				Exclusive breastfeeding		
Probe name	Estimate	SE	P-value	Probe name	Estimate	SE	P-value
cg03084214	6.585	7.330	0.349	cg05091920	-15.843	10.651	0.125
cg11005360	2.986	15.504	0.366	cg15792829	-17.064	9.602	0.155
cg23381058	3.183	4.277	0.034*	cg18603538	-1.101	6.033	0.292
cg23753947	-23.567	17.710	0.946	cg23381058	4.805	3.187	0.034*

319 Columns: Probe name = *LEP* CpG sites from Illumina Infinium MethylationEPIC Beadchip,

320 which are significantly associated with either total or exclusive breastfeeding; Estimate =

321 coefficient denoting one-unit increase in the methylation level for specific CpG is associated

322 with the decrease in the log odds of being in the BMI trajectory 3; Standard Error = standard

323 error of coefficient. DNA methylation analysed against BMI trajectories with BMI trajectory

324 1 (normal) as a control. P-value < 0.05 was considered significant and indicated by *.

325

Mediation analysis

327 There were no statistically significant effects of cg23381058 methylation as a mediator

328 between total breastfeeding duration and the early transient overweight trajectory (Table 7).

329 Similarly, there was no association when looking at exclusive breastfeeding duration (Table

330 7). In the case of total breastfeeding duration, no significant direct effect (P-value = 0.06) or

331 total effect (P-value = 0.07) on early transient overweight trajectory were observed.

332

333

Discussion

335 We examined the association between breastfeeding duration and DNAm on the *LEP* gene,

336 finding a significant association at 10 years but not at 18 years. Exclusive breastfeeding

337 duration was significantly associated with DNAm at four CpGs and total breastfeeding was

338 also associated with DNAm at four sites, and among them one site, cg23381058, was in

339 common. Previous studies investigating the association between breastfeeding and DNAm, in

340 animal and human models, have used total breastfeeding duration as their exposure (48).

341 However, there are concerns about the variability of breastfeeding within these parameters.

342 For example, infants could be largely formula-fed, yet still classified as breastfeeding. Using

343 an exclusive breastfeeding category, in comparison with total breastfeeding, gives more

confidence over the exposure variable. The importance of this distinction is highlighted by ongoing research into impact of breastfeeding on infant outcomes, shaping health policy (1).

In addition, epigenome-wide association analysis was undertaken to identify other DMPs and DMRs associated with the breastfeeding duration (exclusive and total) and for both 10 and 18 year samples. To our knowledge this is the first time such an analysis, investigating the association between breastfeeding and DNAm, has been conducted using an EWAS approach. In the case of DMPs, no genome-wide significant (FDR < 0.05) hits were found. This is likely due to lack of power and follow up studies in larger cohorts and/or with DNAm measurements earlier in childhood are warranted. We did however identify five significant DMRs (Sidak-corrected P-values < 0.05) associated with breastfeeding duration. One of these genes, *GAS5*, was reported in a recent paper exploring whether Long non-coding RNAs (lncRNAs), important in development and gene expression, were encapsulated in extracellular vesicles within breastmilk (49). Amongst others, *GAS5* was detected and indeed highly correlated. Involved in metabolic functions and immune system programming, and as a significant epigenetic regulator, *GAS5* is a potential candidate breast-milk compound, that could be important in the effect of breastfeeding. We are not aware of any previous research linking breastfeeding with the remaining DMRs; *WBP1L*, *CYTH4, VSTM1* and *FAM228A*. Interestingly though, a recent study by Rzehak *et al.* (50), found an association between DNAm at cg14518658 near *CYTH4* gene and body composition in pre-school children, adjacent to one of the significant DMRs (chr22:37678728-37678791) identified in this study.

The study by Obermann–Borst *et al.* demonstrated that breastfeeding is associated with DNAm in *LEP* gene at 17 months of age. Obermann-Borst *et al*. hypothesised that induced epigenetic regulation of leptin expression may be mechanistic in the protective effect of breastfeeding against obesity (Figure 1). Our results support Obermann-Borst *et al*. findings, at a later stage in childhood (10 years) but not into young adulthood (18 years). However,

369 though significant, these effect sizes were small and we saw varying direction of effect

370 between probes. Stability of DNAm at cg23381058 over time could not be assessed as it was

371 not present on both DNAm Bead Chip platforms. However, the majority of CpGs exhibited

372 moderate stability over time. This change in DNAm may be a response to other

373 environmental stimuli by 18 years, and indeed it is recognised that DNAm levels in an

374 individual alter throughout childhood (51).

375 The next stage was to examine whether breastfeeding duration is associated with obesity in

376 the IOWBC. The BMI trajectories, devised by Ziyab *et al*. (42), incorporate the different age-

377 specific (BMI) thresholds of obesity proposed by the International Obesity Task Force and

378 WHO, allowing analysis of the maintenance of this association throughout childhood. A

379 significant association between total breastfeeding duration and the 'early transient

380 overweight' BMI trajectory was seen in the subset of children, matched to 10–year-old

381 samples. It should be noted that the sample size in this trajectory was considerably smaller

382 (n=33 at age 10) than in the normal reference category (n=211). No association was seen

383 between total breastfeeding duration and early persistent obesity or delayed overweight

384 trajectory. Again the sample sizes in these trajectories were far smaller (n=12 at age 10 in

385 early persistent obesity group) than the reference trajectory, therefore, a significant effect of

386 breastfeeding may not have been seen due to low power. Nonetheless, these results suggest

387 breastfeeding duration may be involved in a short-term inverse association with childhood

388 obesity. Another independent study examining the BMI trajectories of 276 infants, found that

389 formula-fed infants gain weight more rapidly and disproportionatly than breastfed infants

390 (52). This was attributed to increased lean mass, suggesting that delayed obesity associated

391 with breastfeeding is not related to increased adiposity in infancy. This evidence contradicts

392 our results, which indicates that breastfeeding is independently associated with both DNAm

393 and obesity in childhood, but not long-term.

394 Multiple studies have demonstrated that DNAm at a number of loci is significantly associated

395 with childhood obesity, for example in *HIF3A* gene (53), as well as leptin and *adiponectin*,

396 where DNAm is also associated with Low-density lipoprotein cholesterol levels (54,55).

397 However, in this case *LEP* DNAm levels have been reported to be negatively associated with

398 BMI, contrasting with the hypothesis by Obermann-Borst *et al*., who suggested that hypo-

399 methylation on the *LEP* gene may protect against obesity, i.e. reduce BMI. Indeed, we found

400 a mixture of increased and decreased methylation indicating the relationship may not be

401 simple. We also sought to test the association between DNAm and BMI trajectory in the

402 IOWBC. DNAm at the CpG site cg23381058, which had shown a significant negative

403 association with total and exclusive breastfeeding duration, was found to be negatively

404 associated with early transient overweight, though mediation analysis of this CpG site was

405 not significant (p<0.05). This CpG site is located within an intron of *LEP*, in a transcription

406 factor binding site. Conversly, an effect was seen at cg05091920, though here a decrease in

407 the methylation level was associated with the decrease in the log odds of being in early

408 persistant obesity trajectory.

409 We recognise that there are some limitations to this study. Firstly, we had no measurements

410 of DNAm in infancy and thus our mediation analysis is based on the assumption that DNAm

411 at age 10 years associated with breastfeeding duration is stable from infancy. Most

412 importantly, it is difficult to establish a causal relationship between breastfeeding, DNAm

413 and obesity. It has been previously reported that changes in DNAm are a consequence of

414 obesity (56–58), however these studies were largely conducted in adult populations, and

415 inter-study comparison is difficult. Furthermore, breastfeeding may alter DNAm in a tissue

416 specific manner, and what effect this DNAm has on *LEP* expression as well as tissue

417 specificity must be explored. Whilst this study and that of Obermann-Borst *et al*. used

418 peripheral blood samples, and leptin hormone circulates through the blood, the *LEP* gene is

known to be expressed in adipocytes, questioning the biological relevance. However, it has been demonstrated that epigenetic markers in blood accurately reflect those in adipose tissue (54,55). Peripheral blood remains an effective tissue from which to measure DNAm in childhood, though the investigation of alternative tissues is of interest.

The findings of this study require replication in other cohorts with childhood data and larger sample sizes to improve our understanding of how the relationship between breastfeeding, DNAm and obesity varies throughout childhood. Cohorts with methylation data from cord blood and matched participants later in childhood could also help to establish causality. Epigenome-wide association studies have helped identify thousands of differentially methylated genes, associated with disease processes and/or environmental processes. This will be an important step to identify the importance of DNAm as a mechanism propagating the long-term protective effects of breastfeeding.

Conclusion

In line with previous studies, significant association of breastfeeding duration with altered methylation at seven CpG sites on *LEP* gene was observed and association of breastfeeding duration with childhood obesity was confirmed. Future studies in larger population samples are warranted to further explore the potential of DNAm at *LEP* and other sites in the genome in mediating the effect of breastfeeding on childhood obesity.

List of abbreviations

CpG: Cytosine-phosphate-Guanine

DNAm: DNA methylation

IOWBC: Isle of Wight birth cohort

LCPUFAs: Long-chain polyunsaturated fatty acids

LEP: *Leptin*

WHO: World Health Organisation

Declarations:

Ethics approval and consent to participate

Ethical approvals for both IOWBC was obtained from the relevant authorities as follows:

IOWBC – 10-year follow-up: Isle of Wight Local Research Ethics Committee, Isle of Wight

Health Authority, Protocol No. 18/98. DATE: 20/07/1998. 18-year follow-up: Isle of Wight,

Portsmouth & Southeast Hampshire Research Ethics Committee. REC reference number:

06/Q1701/34, Date: 21/06/2006.

Consent for publication

Not applicable.

Availability of data and material

The datasets used and/or analysed during the current study are available from the

corresponding author on reasonable request. For access to the full Isle of Wight Cohort data

please see: http://www.allergyresearch.org.uk/studies/birth-cohort/ - cohort-data-use

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

JWH and FIR conceived and designed the study; WBS analysed the data and wrote the initial

draft. VB, GAL, AHZ, NS and NM analysed phenotypes in the cohort. RJK and SHA

collected the clinical data; JWH, SE, HZ and WK deigned and oversaw the epigenetic studies

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481

in the cohort. SHA is the Chief Investigator for the IOWBC. All authors read and approved

the final manuscript.

Acknowledgements

References

1.   WHO. Essential Nutrition Actions: improving maternal, newborn, infant and young

child health and nutrition. World Health Organization; 2013.

2.   McAndrew F, Thompson J, Fellows L, Large A, Speed M, Renfrew MJ. Infant

Feeding Survey 2010. Health and Social Care Information Centre; 2012.

3.   Victora CG, Bahl R, Barros AJD, França GVA, Horton S, Krasevec J, et al.

Breastfeeding in the 21st century: epidemiology, mechanisms, and lifelong effect.

Lancet. 2016 Jan;387(10017):475–490.

4.   Sankar MJ, Sinha B, Chowdhury R, Bhandari N, Taneja S, Martines J, et al. Optimal

breastfeeding practices and infant and child mortality: a systematic review and meta-

analysis. Acta Paediatr. 2015 Nov;104:3–13.

5.   Horta B, Victora CG. Short-term effects of breastfeeding: a systematic review of the

benefits of breastfeeding on diarrhoea and pneumonia mortality. World Health

504   Organization; 2013.

505   6.   Horta BL, Mola CL de, Victora CG. Long-term consequences of breastfeeding on

506   cholesterol, obesity, systolic blood pressure and type 2 diabetes: a systematic review

507   and meta-analysis. Acta Paediatr. 2015 Nov;104:30–37.

508   7.   Horta BL, Mola CL de, Victora CG. Breastfeeding and intelligence: a systematic

509   review and meta-analysis. Acta Paediatr. 2015 Nov;104:14–19.

510   8.   Amitay EL, Keinan-Boker L. Breastfeeding and Childhood Leukemia Incidence.

511   JAMA Pediatr. 2015 Jun;169(6):e151025.

512   9.   Lodge CJ, Tan DJ, Lau MXZ, Dai X, Tham R, Lowe AJ, et al. Breastfeeding and

513   asthma and allergies: a systematic review and meta-analysis. Acta Paediatr. 2015

514   Nov;104:38–53.

515   10.   Yang YW, Tsai CL, Lu CY. Exclusive breastfeeding and incident atopic dermatitis in

516   childhood: a systematic review and meta-analysis of prospective cohort studies. Br J

517   Dermatol. 2009 Aug;161(2):373–383.

518   11.   Bion V, Lockett GA, Soto-Ramírez N, Zhang H, Venter C, Karmaus W, et al.

519   Evaluating the efficacy of breastfeeding guidelines on long-term outcomes for allergic

520   disease. Allergy. 2016 May;71(5):661–70.

521   12.   Matheson MC, Allen KJ, Tang MLK. Understanding the evidence for and against the

522   role of breastfeeding in allergy prevention. Clin & Exp Allergy. 2012

523   Jan;42(6):827–851.

524   13.   Beyerlein A, Kries R von. Breastfeeding and body composition in children: will there

525   ever be conclusive empirical evidence for a protective effect against overweight? Am J

526   Clin Nutr. 2011 Apr;94(suppl_6):1772S–1775S.

527 14. Uwaezuoke SN, Eneh CI, Ndu IK. Relationship Between Exclusive Breastfeeding and

528   Lower Risk of Childhood Obesity: A Narrative Review of Published Evidence. Clin

529   Med Insights Pediatr. 2017 Jan;11:117955651769019.

530 15. Yan J, Liu L, Zhu Y, Huang G, Wang PP. The association between breastfeeding and

531   childhood obesity: a meta-analysis. BMC Public Health. 2014 Dec;14(1).

532 16. Robinson SM. Infant nutrition and lifelong health: current perspectives and future

533   challenges. J Dev Orig Health Dis. 2015 Jun;6(05):384–389.

534 17. Ballard O, Morrow AL. Human milk composition: nutrients and bioactive factors.

535   Pediatr Clin North Am. 2013 Feb;60(1):49–74.

536 18. Micha R, Mozaffarian D. Saturated Fat and Cardiometabolic Risk Factors, Coronary

537   Heart Disease, Stroke, and Diabetes: a Fresh Look at the Evidence. Lipids. 2010

538   Mar;45(10):893–905.

539 19. Stroobant W, Braun KV, Jong JCK, Moll HA, Jaddoe VW, Brouwer IA, et al. Intake

540   of Different Types of Fatty Acids in Infancy Is Not Associated with Growth,

541   Adiposity, or Cardiometabolic Health Up to 6 Years of Age. J Nutr. 2017

542   Jan;jn241018.

543 20. Voortman T, Hooven EH van den, Braun KVE, Broek M van den, Bramer WM,

544   Chowdhurry R, et al. Effects of polyunsaturated fatty acid intake and status during

545   pregnancy, lactation, and early childhood on cardiometabolic health: A systematic

546   review. Prog Lipid Res. 2015 Jul;59:67–87.

547 21. Marincola FC, Dessì A, Corbu S, Reali A, Fanos V. Clinical impact of human breast

548   milk metabolomics. Clin Chim Acta. 2015 Dec;451:103–106.

549   22.   Verduci E, Banderali G, Barberi S, Radaelli G, Lops A, Betti F, et al. Epigenetic

550   Effects of Human Breast Milk. Nutrients. 2014 Apr;6(4):1711–1724.

551   23.   Pannaraj PS, Li F, Cerini C, Bender JM, Yang S, Rollie A, et al. Association Between

552   Breast Milk Bacterial Communities and Establishment and Development of the Infant

553   Gut Microbiome. JAMA Pediatr. 2017 Jul;171(7):647.

554   24.   Canani RB, Costanzo M Di, Leone L, Bedogni G, Brambilla P, Cianfarani S, et al.

555   Epigenetic mechanisms elicited by nutrition in early life. Nutr Res Rev. 2011

556   Oct;24(02):198–205.

557   25.   Peter CJ, Fischer LK, Kundakovic M, Garg P, Jakovcevski M, Dincer A, et al. DNA

558   Methylation Signatures of Early Childhood Malnutrition Associated With Impairments

559   in Attention and Cognition. Biol Psychiatry. 2016 Nov;80(10):765–774.

560   26.   Choi S-W, Friso S. Epigenetics: A New Bridge between Nutrition and Health. Adv

561   Nutr. 2010 Oct;1(1):8–16.

562   27.   Hoile SP, Clarke-Harris R, Huang R-C, Calder PC, Mori TA, Beilin LJ, et al.

563   Supplementation with N-3 Long-Chain Polyunsaturated Fatty Acids or Olive Oil in

564   Men and Women with Renal Disease Induces Differential Changes in the DNA

565   Methylation of FADS2 and ELOVL5 in Peripheral Blood Mononuclear Cells. Schunck

566   W-H, editor. PLoS One. 2014 Oct;9(10):e109896.

567   28.   Obermann-Borst SA, Eilers PHC, Tobi EW, Jong FH de, Slagboom PE, Heijmans BT,

568   et al. Duration of breastfeeding and gender are associated with methylation of the

569   LEPTIN gene in very young children. Pediatr Res. 2013 Jun;74(3):344–349.

29. Rossnerova A, Tulupova E, Tabashidze N, Schmuczerova J, Dostal M, Rossner P, et al. Factors affecting the 27K DNA methylation pattern in asthmatic and healthy children from locations with various environments. Mutat Res. 2013 Feb;741–742:18–26.

30. Al Maskari MY, Alnaqdy AA. Correlation between Serum Leptin Levels, Body Mass Index and Obesity in Omanis. Sultan Qaboos Univ Med J. 2006 Dec;6(2):27–31.

31. Yousefi M, Karmaus W, Zhang H, Ewart S, Arshad H, Holloway JW. The methylation of the LEPR/LEPROT genotype at the promoter and body regions influence concentrations of leptin in girls and BMI at age 18 years if their mother smoked during pregnancy. Int J Mol Epidemiol Genet. 2013;4(2):86–100.

32. Arshad SH, Holloway JW, Karmaus W, Zhang H, Ewart S, Mansfield L, et al. Cohort Profile: The Isle Of Wight Whole Population Birth Cohort (IOWBC). Int J Epidemiol. 2018 Mar;

33. Arshad SH, Karmaus W, Zhang H, Holloway JW. Multigenerational cohorts in patients with asthma and allergy. J Allergy Clin Immunol. 2017 Feb;139(2):415–421.

34. Arshad SH, Hide DW. Effect of environmental factors on the development of allergic disorders in infancy. J Allergy Clin Immunol. 1992 Aug;90(2):235–41.

35. Lehne B, Drong AW, Loh M, Zhang W, Scott WR, Tan S-T, et al. A coherent approach for analysis of the Illumina HumanMethylation450 BeadChip improves data quality and performance in epigenome-wide association studies. Genome Biol. 2015 Feb;16(1).

36. Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. Biostatistics. 2006 Apr;8(1):118–127.

37. Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, et al. The human genome browser at UCSC. Genome Res. 2002 Jun;12(6):996–1006.

38. Delbaere I, Vansteelandt S, Bacquer D De, Verstraelen H, Gerris J, Sutter P De, et al. Should we adjust for gestational age when analysing birth weights? The use of z-scores revisited. Hum Reprod. 2007 Jun;22(8):2080–2083.

39. Kitano N, Nomura K, Kido M, Murakami K, Ohkubo T, Ueno M, et al. Combined effects of maternal age and parity on successful initiation of exclusive breastfeeding. Prev Med Reports. 2016 Jun;3:121–126.

40. 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 May;30(10):1363–9.

41. Houseman E, Accomando WP, Koestler DC, Christensen BC, Marsit CJ, Nelson HH, et al. DNA methylation arrays as surrogate measures of cell mixture distribution. BMC Bioinformatics. 2012;13(1):86.

42. Ziyab AH, Karmaus W, Kurukulaaratchy RJ, Zhang H, Arshad SH. Developmental trajectories of Body Mass Index from infancy to 18 years of age: prenatal determinants and health consequences. J Epidemiol Community Health. 2014 Jun;68(10):934–941.

43. Cole TJ. Establishing a standard definition for child overweight and obesity worldwide: international survey. BMJ. 2000 May;320(7244):1240–1240.

44. WHO Child Growth Standards based on length/height, weight and age. Acta Paediatr (Oslo, Norw 1992) Suppl. 2006 Apr;450:76–85.

45. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and

Powerful Approach to Multiple Testing. J R Stat Soc Ser B [Internet].

1995;57(1):289–300. Available from: http://www.jstor.org/stable/2346101

46.   Pedersen BS, Schwartz DA, Yang I V., Kechris KJ. Comb-p: software for combining,

analyzing, grouping and correcting spatially correlated P-values. Bioinformatics. 2012

Nov;28(22):2986–8.

47.   Tingley D, Yamamoto T, Hirose K, Keele L, Imai K. mediation:RPackage for Causal

Mediation Analysis. J Stat Softw. 2014;59(5).

48.   Hartwig FP, Loret de Mola C, Davies NM, Victora CG, Relton CL. Breastfeeding

effects on DNA methylation in the offspring: A systematic literature review. PLoS

One [Internet]. 2017 Mar 3;12(3):e0173070-. Available from:

https://doi.org/10.1371/journal.pone.0173070

49.   Karlsson O, Rodosthenous RS, Jara C, Brennan KJ, Wright RO, Baccarelli AA, et al.

Detection of long non-coding RNAs in human breastmilk extracellular vesicles:

Implications for early child development. Epigenetics [Internet]. 2016 Oct

2;11(10):721–9. Available from: https://doi.org/10.1080/15592294.2016.1216285

50.   Rzehak P, Covic M, Saffery R, Reischl E, Wahl S, Grote V, et al. DNA-Methylation

and Body Composition in Preschool Children: Epigenome-Wide-Analysis in the

European Childhood Obesity Project (CHOP)-Study. Sci Rep. 2017 Oct;7(1):14349.

51.   Gervin K, Andreassen BK, Hjorthaug HS, Carlsen KCL, Carlsen K-H akon, Undlien

DE, et al. Intra-individual changes in DNA methylation not mediated by cell-type

composition are correlated with aging during childhood. Clin Epigenetics. 2016

Oct;8(1).

52.   Bell KA, Wagner CL, Feldman HA, Shypailo RJ, Belfort MB. Associations of infant

feeding with trajectories of body composition and growth. Am J Clin Nutr. 2017 Jun;106(2):491–498.

53. Wang S, Song J, Yang Y, Zhang Y, Wang H, Ma J. HIF3A DNA Methylation Is Associated with Childhood Obesity and ALT. Costa M, editor. PLoS One. 2015 Dec;10(12):e0145944.

54. Houde A-A, Légaré C, Biron S, Lescelleur O, Biertho L, Marceau S, et al. Leptin and adiponectin DNA methylation levels in adipose tissues and blood cells are associated with BMI, waist girth and LDL-cholesterol levels in severely obese men and women. BMC Med Genet. 2015 May;16(1).

55. Li R, Fein SB, Grummer-Strawn LM. Do Infants Fed From Bottles Lack Self-regulation of Milk Intake Compared With Directly Breastfed Infants? Pediatrics. 2010 May;125(6):e1386–e1393.

56. Dick KJ, Nelson CP, Tsaprouni L, Sandling JK, Aïssi D, Wahl S, et al. DNA methylation and body-mass index: a genome-wide analysis. Lancet. 2014 Jun;383(9933):1990–1998.

57. Ronn T, Volkov P, Gillberg L, Kokosar M, Perfilyev A, Jacobsen AL, et al. Impact of age, BMI and HbA1c levels on the genome-wide DNA methylation and mRNA expression patterns in human adipose tissue and identification of epigenetic biomarkers in blood. Hum Mol Genet. 2015 Apr;

58. Wahl S, Drong A, Lehne B, Loh M, Scott WR, Kunze S, et al. Epigenome-wide association study of body mass index, and the adverse outcomes of adiposity. Nature. 2017;541(7635):81–6.

1
2
3
4
5
6
7
661
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

662   **Table 1** Characteristics of Isle of Wight birth cohort participants at 10 and 18 years.

Characteristic	IOWBC Participants at 10 years (n=297)	IOWBC Participants at 18 years (n=305)
Child sex (female)	190 (64.0%)	209 (68.5%)
Maternal socioeconomic status		
1 (lowest)	41 (13.8%)	49 (16.1%)
2	60 (20.2%)	52 (17.0%)
3	89 (29.9%)	88 (28.9%)
4	82 (27.6%)	85 (27.9%)
5 (highest)	25 (8.4%)	31 (10.2%)
No maternal smoking during pregnancy	238 (80.1%)	250 (82.0%)
Gestational age <38 weeks	12 (4.0%)	15 (4.9%)
Overweight (BMI >19.8 at 10 years, >25 at 18)	62 (20.9%)	83 (27.2%)
Never breastfed	38 (12.8%)	48 (15.7%)
BMI Trajectory		
Normal (reference)	211 (71%)	215 (70.5%)
Delayed overweight (Trajectory 2)	41 (13.8%)	48 (15.8%)
Early transient overweight (Trajectory 3)	33 (11.1%)	30 (9.8%)
Early persistent obesity (Trajectory 4)	12 (4.1%)	12 (3.9%)

663   Descriptive statistics on exposure and covariates of participants included in study. This is

664   given as the number of participants in each category and as a percentage for each

665   characteristic. Note that for the overweight category, different BMI cut-offs are used at 10 or

666   18 years, as defined by the International Obesity Taskforce.

667 **Table 7** Mediation analysis between cg23381058, total breastfeeding duration and BMI

668 Trajectory 3 (Early transient overweight)

	Total breastfeeding duration				Exclusive breastfeeding duration			
		95% CI	95% CI			95% CI	95% CI	
	Beta	Lower	Upper	P-value	Beta	Lower	Upper	P-value
ACME	5.81E-05	-2.46E-04	4.31E-04	0.73	2.48E-04	-3.1E-04	1.19E-03	0.42
ADE	-2.43E-03	-4.94E-03	1.16E-04	0.06	-2.25E-03	-7.27E-03	2.76E-03	0.36
Total Effect	-2.37E-03	-4.91E-03	1.42E-04	0.07	-2.00E-03	-7.07E-03	3.06E-03	0.42
Prop. Mediated	-2.45E-02	-4.19E-01	2.72E-01	0.76	-0.12	-1.66	1.54	0.72

669 Columns: Effects - Average Causal Mediation Effects (ACME), Average Direct Effects

670 (ADE), Total effect = sum of a mediation (indirect) effect and a direct effect, Prop. Mediated

671 = Proportion of the total effect explained by the mediator.

672

673 **Figure 1** A mediation model for the association between breastfeeding duration (total and

674 exclusive) and childhood obesity.

675 β_M represent the effect estimate for breastfeeding duration on DNA methylation

676 (CpG ~ breastfeeding duration + covariates); β_Y represents the effect estimate for CpG on

677 childhood obesity (BMI trajectory ~ CpG + covariates; βc represent the direct effect (no

678 mediation) estimate for breastfeeding duration on childhood obesity (BMI trajectory ~

679 breastfeeding duration + covariates; βc′ represents the total effect estimate on childhood

680 obesity (BMI trajectory = breastfeeding duration + covariates + CpG).