Clinical Epigenetics

A randomized controlled trial of folic acid intervention in pregnancy highlights a putative methylation-regulated control element at ZFP57

--Manuscript Draft--

Manuscript Number:	CLEP-D-18-00378R1
Full Title:	A randomized controlled trial of folic acid intervention in pregnancy highlights a putative methylation-regulated control element at ZFP57
Article Type:	Research
Section/Category:	Environmental Epigenetics
Funding Information:	Economic and Social Research Council (ES/N000323/1) — Professor Colum Walsh Economic and Social Research Council (ES/N000498/1) — Prof Caroline L Relton
Abstract:	Abstract Background: Maternal blood folate concentrations during pregnancy have been previously linked with DNA methylation patterns, but this has been done predominantly through observational studies. We showed recently in an epigenetic analysis of the first randomized controlled trial (RCT) of Folic Acid Supplementation specifically in Second and Third Trimester (The EpiFASSTT Trial) that methylation at some imprinted genes was altered in cord blood samples in response to treatment. Here we report on epigenome-wide screening using the Illumina EPIC array (~850,000 sites) in these same samples (n=86). Results: The top-ranked differentially-methylated promoter region (DMR) showed a gain in methylation with folic acid (FA) and was located upstream of the imprint regulator ZFP57. Differences in methylation in cord blood between placebo and folic acid treatment groups at this DMR were verified using pyrosequencing. The DMR also gains methylation in maternal blood in response to FA supplementation. We also found evidence of differential methylation at this region in an independent RCT cohort, the AFAST trial. By altering methylation at this region in two model systems in vitro we further demonstrated that it was associated with ZFP57 transcription levels. Conclusions: These results strengthen the link between folic acid supplementation during later pregnancy and epigenetic changes and identify a novel mechanism for regulation of ZFP57. This trial was registered 15th May 2013 at www.isrctn.com as ISRCTN19917787. Keywords: Folic Acid; DNA methylation; Cord Blood; Offspring; Imprinting; ZFP57
Corresponding Author:	Colum Walsh University of Ulster UNITED KINGDOM
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	University of Ulster
Corresponding Author's Secondary Institution:	
First Author:	Rachelle E Irwin
First Author Secondary Information:	
Order of Authors:	Rachelle E Irwin Sara-Jayne Thursby Miroslava Ondicova Kristina Pentieva Helene McNulty Rebecca C Richmond

	Aoife Caffrey
	Diane J Lees-Murdock
	Marian McLaughlin
	Tony Cassidy
	Matthew Suderman
	Caroline L Relton
	Colum Walsh
Order of Authors Secondary Information:	
Response to Reviewers:	MS ID#: CLEP-D-18-00378 MS TITLE: ¬A randomized controlled trial of folic acid intervention in pregnancy highlights a putative methylation-regulated control element at ZFP57 Response to Reviewers We would like to thank the reviewers for their positive feedback and constructive criticisms provided, and we appreciate the time they have spent reviewing the manuscript. In regards to their concerns, we have made alterations in blue in the revised version, and responded to each point raised individually as further detailed below. In summary, particular changes and new data to note in the revised MS are:- -A revised Fig.1 showing exclusion/attrition from each arm of the study -A new table of the top 5 ranked sites (Fig.2D) including verification (Results, p7) -Further information on top-ranked regions including verification of one (Results, p8) -More experimental detail on the data handling methods used to find and correct for confounders etc (p21) ~~~~~~ Reviewer #1 We are very appreciative of the encouraging comments regarding the importance of the findings in the manuscript, the suitable study design, appropriate controls and the strength of our further comprehensive approach in terms of pyrosequencing, the use of in vitro models and the verification of our findings in another independent RCT. Major comments C1. -Throughout the manuscript, the authors did not strongly rely on p-values for the identification of differentially methylated regions, which can actually be considered a strength. Identification of differentially methylated regions is based on ranking (considering the RnBeads methylation measures and p-values). However, in order to be able (for the reader) to evaluate the effects observed, it would be helpful to provide some information on the p-values. Not providing this information (even when the statistical approaches are not driven by p-values) raises the question whether the finding were statistically significant. In Figure 3A, p-values were given for the top-5 differentially methylated regions at promoters. Were these p-values adjusted for multiple testing? R1. -The reviewer is correct that we are not reliant on p values for identifying the differentially methylated regions of interest, but instead use the RnBeads rank as our initial springboard for exploring regions with biologically meaningful differences. The p values were generated from linear models employed in the limma package and these were not adjusted, which we now state in the text and legend: these are provided for comparison purposes only and it can be seen that the top regions are not ordered from lowest to highest, since the second-ranked gene has a lower p-value than the top-ranked one. Just to expand on why ranking was used instead, this measure is a combination of the change in mean methylation (also known as the delta beta for EPIC arrays), the quotient of mean methylation and the combined p value as stated on p6 at first use, and in Methods under Statistical analysis (p22). To highlight this, we have included an extra sentence on p6 (L22ff) to further explain our choice of ranking over p value alone. The ranking approach in RnBeads was developed by Assenov and colleagues to provide a more integrated measure, since sites with very good p values often have very small differences in methylation which have little biological meaning; likewise, the quotient may be very high, e.g. between 5% and 1% methylation, but the absolute

change low. Where dealing with individual sites, the ranking system integrates the absolute change, the quotient and the p value to come up with a single measure, for example when looking at the top 1000 sites in Fig.2A and described on p6. We have now included a table of the top-ranked sites as Fig.2D and a brief discussion of these on p7 in response to both reviewers.
While ranking is our initial way of screening differentially methylated sites and regions, the second most important feature for us was the % change in methylation, as only sites showing differences which are 5% or greater can be reliably verified using pyrosequencing: others, no matter how good the p value or how high the rank, cannot be verified by a second lab method. We therefore find that using ranking first, then % change, we can reliably verify differential methylation in our systems; p value would then only be used as a third criteria if so wished. While examining the top-ranked sites also uncovered some of interest, analysis at a site level (whether by rank or p value) is less reliable, since these are more prone to artefacts due to the presence of SNPs and show poorer correlation when assayed using pyrosequencing as described in the new text (p7 L9ff)
Ranking can also be applied at the region level however as we had done in the case of Fig.3A and here was based on the combination of the average difference in means across all sites in the promoter regions of the sample groups, the mean of quotients in mean methylation and the combined p-value, which was calculated from all site p-values in the region using a generalization of Fisher's method as described in Methods p22. In response to the reviewers, we have now also discussed the other top-ranked regions in Fig.3A in more detail on p8 and explained why these were of less interest than ZFP57, due to having low absolute levels of change in methylation, having unknown functions, or similar (see p8, L6ff).

C2. -A major strength of the current study is that different approaches were used to confirm the findings for the upstream region of ZFP57. Among the approaches used, two different in vitro experiments were conducted to explore transcriptional consequences of methylation changes in the upstream region. These experiments may, however, not be well-equipped to draw firm conclusions about regulation of ZFP57 expression. In the first experiment, knockout (KO) cells were used with a mutation in DNMT1 and DNMT3b. As stated by the authors, these cells are hypomethylated at many loci. The authors concluded that observed expression changes of ZFP57 in these KO cells (as compared to wildtype) may be linked to hypomethylation of the upstream region. The title of this results section says "The upstream region is a methylation-dependent regulator of ZFP57 transcription". However, as not only this upstream region, but many loci were hypomethylated, it cannot be concluded whether (only the) upstream region regulates gene expression. Similarly, in the second experiment, 5-aza-dC was used to perturb methylation in a global manner. Also in this experiment, it is difficult to state whether the observed differences in gene expression were (only) resulting from hypomethylation of the identified upstream region. To conclude, both models lack specificity and may not be fully suitable to study specific regulation of ZFP57 expression by the upstream region identified.
R2. -While using isogenic cells with and without mutations in the methyltransferases has been the gold standard to demonstrate a role for methylation in gene control, for example in the case of imprinted genes (Li, Beard and Jaenisch Nature 366: 362 1993), X-inactivation (Beard, Li and Jaensich Gene Dev 9:2325 1995) and endogenous retroviruses (Walsh, Chaillet and Bestor Nature genet 20:116 1998), it is true that it is hard to formally exclude an indirect effect through change at some second locus. We minimised this risk by using biochemical inhibition of the methyltransferases in a second cell line as well, reducing the likelihood that indirect effects are responsible. However ideally we would need to alter methylation specifically at the DMR using an epiCRISPR approach (while screening for off-target effects) and also firmly establish the role of the DMR in control of ZFP57 transcription: these extensive studies have just begun and will form the basis as a follow-up manuscript as referred to on P16, L8 and also P17, L16ff. Acknowledging these limitations, we have reworded the title of this section to read “Demethylation of the upstream region was accompanied by increased ZFP57 transcription” (p9, L10ff) and the abstract to read “By altering methylation at this region in two model systems in vitro we further demonstrated that it was associated with ZFP57 transcription levels.” Likewise we have altered the MS title to read “…highlights a putative methylation-regulated….”

C3 Moreover, it would be of specific interest to study the effects of methylation gain (rather than loss) in the current models and to consider folic acid as a potential modulator of methylation at this region.
R3 - We agree with the reviewer that this would indeed be of real interest and have made some attempts to modulate levels of folate in a cell line. While we have generally seen some small methylation increase at this DMR upstream of ZFP57 upon folic acid treatment the results have been variable and not sufficiently well reproducible. This approach appears to be difficult and not representative of the physiological response to folic acid supplementation due to the complexity of the one carbon metabolism cycle in vivo, with its reliance on several essential cofactors and tissue interactions (see ref 2).

C4-The authors mentioned that the identified DMR upstream of ZFP57 was previously reported in an observational study by Amarasekera et al (ref 52), which, together with results from MoBa+ Generation R, support this region 'being a true folate-sensitive DMR'. However, the results by Amarasekera point towards hypomethylation in women with a high folate status, whereas in the current study a gain in methylation was found with higher folic acid intake. These results therefore seem contradictive. Do the authors have an explanation for this?
R4. -It is true that the Amarasekara study found mothers with higher FA intake had lower methylation at this DMR rather than gains in methylation: in the MoBa + Generation R study, sites at this region were confirmed as having differences, but no direction or magnitude of change was stated, it may have been gain or loss. These studies differed in sample size, design and analysis methods from ours, with the crucial difference being that our results are from a randomized controlled trial (RCT) which can look experimentally at whether increasing FA in mothers causes in methylation or not, whereas the previous studies were observational and so could not examine whether there was a causal link between increased maternal folate and increased offspring DNA methylation of specific genes, even with a large sample size. A more relevant comparison for us was the AFAST study, which was also an RCT where mothers were supplemented with folic acid: in this cohort too there was a gain in methylation at the ZFP57 DMR (Fig.5) in response to increasing the levels of folate in mothers, consistent with our results. Finally, we also checked the response to FA in maternal bloods in our own cohort, which is a semi-independent population. This again showed that supplementation gave a gain in methylation at the ZFP57 DMR between the start of the intervention (GW14) and the last time-point assessed (GW36) of +5.51% (Table 2); strikingly, this was even similar in magnitude to the change seen in the offspring (cord bloods). Given the good agreement in direction and magnitude of change at the ZFP57 DMR in the mothers, the offspring and in the AFAST offspring, all of which are RCT cohorts, we feel this strongly supports there being a gain in methylation rather than loss in response to FA supplementation.
While the design is the main and critical difference between the studies, we also note the very high levels of serum folate concentrations in the high intake group in the Amarasekara study (see next point below for more detail).

Minor comments
C5 General: Presumably, all women took folic acid before conception and during the first trimester (for the prevention of NTDs). For women in the placebo group, this means that they actively stopped taking folic acid at time of the current study. To what extent can observed differences be attributed to this acute stop (and potentially decline in folate status, see post-intervention values Table 1) rather than prolonged use (versus a placebo)? Could this explain some of the observed findings (e.g. decline in methylation for the placebo group in Table 2)? Could the authors speculate on this aspect in more detail?
R5. While it is interesting to view it this way, there are some solid reasons why it would be erroneous to do so. Although Reviewer 1 is quite correct that all the women included in this trial reported that they had taken folic acid supplements in the first trimester, this relied on retrospective self-reported usage of folic acid; in this way the first trimester of pregnancy is outside the period investigated in our RCT. This RCT of folic acid supplementation starts at the beginning of the second trimester (not before then); women at this point were randomized to receive either folic acid or placebo, with the active treatment (intervention) therefore being folic acid. We cannot consider this to be a trial of withdrawal of FA supplements in trimesters 2 and 3 because we did not provide and control the supplement usage during trimester 1 (before randomization), we simply recorded what participants reported to us.

The values in Table 1 do show a decline in folate levels without ongoing folic acid supplementation in the placebo group: this is a widely-reported effect observed during pregnancy in larger observational studies, which would have included women both taking and not taking FA in first trimester (see work cited in ref 42), so we are not seeing any unusual effect in the placebo group due to an abrupt stop in FA, but rather the results are in line with a range of previous studies showing that folate levels drop during pregnancy. What we can say with certainty, due to this being an RCT, is that 1)the levels of folate in the placebo and treated groups were not different at the start of the trial; 2)the levels drop in the placebo group (from 48.8 to 23.6 nmol/L in serum comparing GW14 and GW36), in line with all previous observational work and 3)the levels in the treated group are not elevated over normal but rather are protected from this drop by the dietary supplementation (45.8 nmol/L at GW14 and 46.5 nmol/L at GW36). We have added a few lines to the Results (p5 L24ff) to highlight this point. See also our previous papers on our RCT (refs 42 and 43) and literature cited therein for a fuller discussion of this point.

Just to note that the normal levels of folate in our treatment group at GW36 is another point of difference between our RCT and the observational study by Amarasekera et al, where the high folate group they chose represented an extreme of the population and had blood folate levels almost twice those seen in our treated group at the end of the intervention (74.59+/- 6.1 nmol Amarasekera vs 46.5 +/-19.5 nmol/L GW36 treated group (Table1)). We have added some sentences to the Discussion at the relevant point to highlight this and other differences with the named study (p14, L5ff).

C6-Results: did a Manhattan (or Volcano) plot reveal any (potential) interesting sites?
R6. - Manhattan or volcano plots normally are used to identify sites with high p values: as indicated above, we are using instead a ranking approach, but in response to this request, and a similar one from Reviewer 2, we have provided instead a table of the top-ranked differentially methylated sites (Fig.2D) as mentioned above. We describe these and also provide data on those which we have verified in the revised Results section (p7, L9ff).

C7-Abstract: page 2, line 7-8. "Maternal blood folate concentrations during pregnancy have been previously linked with DNA methylation changes". 'Changes' can be removed (or replaced by profiles / patterns).
R7 This has now been edited from ‘changes’ to ‘patterns’.

C8 Abstract: page 2, line 28-29. "Differences in methylation in cord blood between groups...". The definition of groups have not been given before in the abstract. Please explain what the groups are (FA versus placebo) in the abstract.
R8 This now reads as ‘Differences in methylation in cord blood between placebo vs folic acid treatment groups.’
C9 Abstract, page 2, line 31-32. "... were verified using pyrosequencing, and the region responds to FA supplementation in cord blood". The latter part of this sentence is not clear. Should this be maternal blood (although in that case, the sentence is still not very clear)?
R9 This has now been changed to “….were verified using pyrosequencing. The DMR also gains methylation in maternal blood in response to FA supplementation.”

C10 Abstract: page 2, line 38-39. "By altering methylation at this region in two model systems in vitro, we further demonstrated that it regulated ZFP57 transcription." Based on the current experiments, it cannot be concluded whether this region regulated ZFP57 transcription, only whether or not there was an association with ZFP57 transcription.
R10 This has now been changed to "By altering methylation at this region in two model systems in vitro, we further demonstrated that it was associated with ZFP57 transcription."

C11 Methods: Were any potential confounders considered in the differential methylation analyses?
R11 We controlled for both known and unknown confounders using a number of approaches: we have expanded this section of the methods to elaborate (p21, L2ff) as follows:

“Initial data exploration in RnBeads used principle components analysis (PCA) to explore potential correlations between the groups and known confounders such as BMI, smoking, gender etc. In addition, in order to account for any hidden confounding variables in the dataset, surrogate variable analysis was carried out using the sva package with the Buja and Eyboglu algorithm from (1992) [68]. Briefly, potential surrogate variables such as age, sample plate, Sentrix ID and Sentrix Position were tested for association with the target variable sample group using PCA and any surrogate variable with a high correlation to sample group was adjusted for and incorporated into the making of the limma based linear model.“

C12 Results: page 5, line 48-50. Minor suggestion: Red Cell Folate without capitals. 'also' can be removed from this sentence.
R12 This has been amended.

C13 Results: page 7, line 37-40. Not clear why the authors refer to figure 3C in this sentence. The numbers mentioned in this sentence (6.23%) refer to panel 3A.
R13 Corrected to Fig.3A.

C14 Results: page 9, line 39-40. 'samples' is presumably subjects (or participants / women).
R14 This has been amended to read ‘participants’

C15 Discussion: page 13, line 21-22. Minor typo: examined = examine.
R15 Corrected

C16 . -Results: why is a p-value given for the methylation differences assessed by the pyrosequencing assay (in Figure 3C), but not for the maternal methylation values (in Table 2)?
R16. -The p values for the maternal methylation values in Table 2 were not significant: this is now mentioned in the text on P11, L24.

C17-Figure 1: Presuming that some of the excluded women had actually started with the intervention, it would be informative to present the reasons for exclusion (and corresponding numbers) for both study arms separately.
R17. -Figure 1 now includes the numbers of women excluded from the study for each arm (as below) and this has been further detailed in the methods.

C18 Figure 2a: label for the x-axis shifted to the right.
R18 This has been amended

~~~~~~
Reviewer #2:

This is an interesting and well-written study by Irwin and colleagues reporting a DMR upstream of ZFP57 with increased methylation after folic acid supplementation. Differences in methylation were verified using pyrosequencing. For this study epigenome-wide screening was performed on samples from the EpiFASSTT trial. They also found evidence of differential methylation at this region in an independent cohort, the AFAST trial.

Just a few comments/suggestions:

C1 Background
Page 3, Line 22: The authors report some health benefits of FA supplementation but is missing important background information. The maternal folate status was associated with neural tube defect (NTD) risk as early as 40 years ago and up to 70% of NTDs can be prevented by an optimal maternal red blood cell folate concentration, information on neural tube defects and the benefits of FA supplementation is lacking. [Greene et al, PMID: 21613818; Blom et al, PMID: 16924261; Rochtus et al, PMID: 26349489]
R1. -We thank the reviewer for bringing this to our attention. We have now included these important and interesting references: “Despite the identification of a relationship between maternal folate status and NTDs as early as 40 years ago, information on the mechanism behind between the benefits of FA supplementation with respect to NTDs
~~~~~~

remains to be fully elucidated (reviewed in [8]), as does the relationship of FA, NTDs and DNA methylation [9]. There is however little dispute as regards the protective effect of folic acid supplementation before and in early pregnancy, which was proven in clinical trials to reduce NTDs by approximately 70% [10]. Furthermore"

C2 Page 3, Line 42: They write "At a molecular level, there is some evidence in human that epigenetic changes could be the mechanism underpinning some of the effects of folate, particularly in second and third trimester, as reviewed elsewhere." Several lines of evidence support the link between NTDs and epigenetics; especially with an impaired methylation cycle [Greene et al, PMID:21613818]. The neural tube closes in the first trimester. Please elaborate.
R2. – We have added some text and cited this excellent review at the relevant point as follows: "At a molecular level, there is some evidence in human that epigenetic changes could be the mechanism underpinning some of the effects of folate, both in the first trimester [8] in the prevention of NTDs, and also in second and third trimester, as reviewed elsewhere [2]."

C3 Please elaborate how differential methylation of a region upstream of the imprinted gene ZFP57 can influence transcription.
R3. –The region where the upstream DMR is located has some features characteristic of a control element, as from examining publicly-available datasets on the UCSC genome browser there are DNAse I hypersensitive sites present here and data suggesting transcription factors may bind here. The potential function for this genomic region is currently being explored in the lab in more detail. We have added a sentence to the Discussion to highlight this information (p17, L16ff). In the Amarasekera paper they found an association between hypomethylation of this DMR and increased transcriptional activity of the ZFP57 gene, but did not test the effects of removing methylation using cell lines and inhibitors as we have.

C4 Cell experiments were performed with HCT116 and HCT116 DKO cell lines. It would be interesting to perform cell experiments to look for the influence of FA supplementation? Were these experiments performed or can they be performed?
R4- See the response to C3 for Reviewer 1 above.

C5 Page 12, line 10: give reference of study that reported that maternal folate concentrations in the third trimester were associated with changes at a DMR at a similar location. What do you mean with similar, are these regions overlapping? Please explain further.
R5. -We understand and apologise that this sentence is confusing, and for clarity we have added the reference here and confirmed we are looking at the same DMR: 'maternal folate concentrations in the third trimester were associated with changes at a DMR at the same genomic location [56]'.

C6 Page 12, line 24: how can you conclude that this is a folate-sensitive DMR?
R6. There are a number of lines of evidence presented in the paper and from the literature to support this point: 1)this region is the top differentially-methylated region in cord blood in this randomized control trial of folic acid intake; 2)the region shows the same direction and similar magnitude of response in offspring in a second RCT, the AFAST cohort; 3)this DMR responds directly to FA in the treated mums; 4)the region has been identified in a previous observational study (Amarasekera) and confirmed in a meta-analysis (MoBa + GenR). This all supports it being a true folate-sensitive DMR.

C7Validation with pyrosequencing: please give absolute values and p-values of the gain and loss of methylation.
R7 We have now included the gain of methylation and magnitude (p14 L18ff): "We could also verify using a separate biological assay the magnitude and direction of change in methylation, a gain of 5.44% in the treatment group, at the DMR in cord blood by using pyrosequencing ($p = 0.172$)."

C8 Figure 2B: "Probe methylation density plot comparing the distributions of methylation values per sample group. In the treatment group there is a decrease in the number of fully methylated sites ($\beta > 0.75$)." How can you explain that there is a decrease in the number of fully methylated sites after FA supplementation? You would expect the opposite.

	R8 –Although a decrease may seem counterintuitive, we and others have observed this previously. For example, in the meta-analysis performed by Joubert et al. (2016) regarding the MoBa and Generation R studies, there were more CpG with significantly decreased methylation than increased (94% vs 6%). Likewise, we found decreased methylation in response to FA at the high copy number interspersed element LINE1, often used as a litmus test for genome-wide methylation (Caffrey et al, ref 43). The biochemical reasons behind this are currently unclear, however literature suggests that higher levels of FA may cause feedback inhibition by altering the SAM:SAH ratio and therefore the intracellular methylation potential (Christensen et al., 2015). We have added some sentences to the Discussion clarifying this (p13 L8ff). C8 Figure 3A: Top 5 differentially methylated regions: did the authors perform cell experiments with the other genes? Please give more background information on these genes and why they did (not) elaborate these genes. R8. In response to the reviewers, we have now also discussed the other top-ranked regions in Fig.3A in more detail on p8, L6ff and explained why these were of less interest than ZFP57: “For the top 5 regions, ZFP57 was of particular interest and is dealt with below. Two others (CES1, a liver carboxylesterase, and ANKRD20A11P, a pseudogene) show less than 5% change in methylation and so could not be verified: DUSP22 which has a larger change is also a pseudogene and so of less interest. The last DMR is located at a microRNA cluster MIR4520A/B and loses approximately 7.22% overall in the treatment group, averaged over a number of well-spaced CG. Due to pyrosequencing assay design constraints, we could only cover one site (cg08750459) from the array at this locus but that site showed reasonable concordance (loss of 12.24% (p = 0.008) in array and 9.45% (p = 0.006) by pyroassay). The function of these microRNA remains obscure however. “ We now also present data on the top 5 ranking sites in Fig.2D, with data on their verification and some discussion of their function on p7, L9ff “We examined the top-ranking sites as identified by RnBeads (Fig.2D): of these, the CG site in the ATP11A gene contained a single nucleotide polymorphism (SNP) missed by the quality control routines: the same was true of the CG at the MAGI2 gene. The presence of the SNPs at these CGs leads to the erroneous appearance of a change in methylation, so these were discounted. Two of the other top-ranked sites were at the PRKAR1B locus, which encodes a regulatory subunit of cyclic AMP-dependent protein kinase A, and one was at NXN, a member of the thioredoxin superfamily: however all three were listed as located in the respective gene body and so are less likely to contribute to transcriptional control. Nevertheless to verify these we used a second method utilizing commercial pyrosequencing methylation assays (pyroassays) designed to query the same CGs. These reported smaller average differences in methylation between treated and placebo groups than seen with the array of 6.6% for cg08104960 at NXN, and 4.2% (cg06242242) and 2.2% for (cg05729249) for the sites at PRKAR1B: only the site at NXN was significant (p=0.002, t-test). “ As the other targets are either in gene bodies, and thus unlikely to exert significant transcriptional control, or have unknown function, we therefore concentrated on ZFP57, which in addition to being the top hit and validated in another cohort, also is strongly functionally linked to imprinting, and we have found differences in methylation at imprinted loci in this cohort (this report and ref 43). This rationale is emphasized in the Discussion (p13). C9 All Figures: the layout of the Figures is nice, but please make them more uniform: use same letter type for all Figures. R9. -We would like to thank the reviewer for their kind comments regarding the figures of the manuscript, and we have now made these more uniform as suggested.
Additional Information:	
Question	**Response**
Is this study a clinical trial?<hr><i>A clinical trial is defined by the Word Health Organisation as 'any	Yes

research study that prospectively assigns human participants or groups of humans to one or more health-related interventions to evaluate the effects on health outcomes'.	
We require registration of all clinical trials that are reported in manuscripts submitted to the journal. More information about trial registration, including the trial registries that currently meet all of the ICMJE guidelines, can be found in the FAQ section of "About ICMJE" at http://www.icmje.org/about-icmje/faqs/clinical-trials-registration/.<p>Please provide the following information where prompted:<hr>Enter the Trial Registration Number: as follow-up to "Is this study a clinical trial?<hr>A clinical trial is defined by the Word Health Organisation as 'any research study that prospectively assigns human participants or groups of humans to one or more health-related interventions to evaluate the effects on health outcomes'."	ISRCTN19917787
Enter the name of the registry: as follow-up to "Is this study a clinical trial?<hr>A clinical trial is defined by the Word Health Organisation as 'any research study that prospectively assigns human participants or groups of humans to one or more health-related interventions to evaluate the effects on health outcomes'."	ISRCTN
Enter the URL of the trial registry record: as follow-up to "Is this study a clinical trial?<hr>A clinical trial is defined by the Word Health Organisation as 'any research study that prospectively assigns human participants or groups of humans to one or more health-related interventions to evaluate the effects on health outcomes'."	http://www.isrctn.com/ISRCTN19917787
Enter the date that you registered your trial (in mm/dd/yyyy format): as follow-up to "Is this study a clinical trial?<hr>A clinical trial is defined by the Word Health Organisation as 'any research study that prospectively assigns human participants or groups of humans to one or more health-related interventions to evaluate the effects on health outcomes'."	05-15-2013
Enter the date of enrolment of the first participant to the trial (in mm/dd/yyyy format): as follow-up to "Is this study a clinical trial?<hr>A	06-15-2013

clinical trial is defined by the Word Health Organisation as 'any research study that prospectively assigns human participants or groups of humans to one or more health-related interventions to evaluate the effects on health outcomes'."	
Was your trial registered before the first participant was enrolled? (i.e. prospectively registered) as follow-up to "Is this study a clinical trial?<hr>A clinical trial is defined by the Word Health Organisation as 'any research study that prospectively assigns human participants or groups of humans to one or more health-related interventions to evaluate the effects on health outcomes'."	Yes
Within your manuscript, have you also included details of your trial registration at the end of your abstract? • Name of the registry • Trial registration number • Date of registration • URL of trial registry record *Example: Trial registration: ISRCTN, ISRCTN12345678. Registered 28 September 2014, http://www.isrctn.com/ISRCTN12345678* as follow-up to "Was your trial registered before the first participant was enrolled? (i.e. prospectively registered)"	I confirm I have provided trial registration details at the end of the abstract

A randomized controlled trial of folic acid intervention in pregnancy highlights a putative methylation-regulated control element at *ZFP57*

Rachelle E. Irwin[1] (r.irwin@ulster.ac.uk), Sara-Jayne Thursby[1] (thursby-sj@ulster.ac.uk), Miroslava Ondičová[1] (ondicova-m@ulster.ac.uk), Kristina Pentieva[2] (k.pentieva@ulster.ac.uk), Helene McNulty[2] (h.mcnulty@ulster.ac.uk), Rebecca C. Richmond[4] (Rebecca.Richmond@bristol.ac.uk), Aoife Caffrey[2] (caffrey-a@ulster.ac.uk), Diane J. Lees-Murdock[1] (dj.lees@ulster.ac.uk), Marian McLaughlin[3] (m.mclaughlin@ulster.ac.uk), Tony Cassidy[3] (t.cassidy@ulster.ac.uk), Matthew Suderman[4] (matthew.suderman@bristol.ac.uk), Caroline L. Relton[4] (Caroline.Relton@bristol.ac.uk) and Colum P. Walsh[1,5] (cp.walsh@ulster.ac.uk)

[1]Genomic Medicine Group and [2]Nutrition Innovation Centre for Food and Health, School of Biomedical Sciences, Ulster University, Coleraine, UK, [3]Psychology Institute, Ulster University, Coleraine, UK and [4]MRC Integrative Epidemiology Unit, Bristol Medical School, University of Bristol, Bristol, United Kingdom.

[5]To whom correspondence should be addressed at: School of Biomedical Sciences, Ulster University, 1 Cromore Road, Coleraine BT52 1SA, UK. Phone: +44 28 7012 4484; Email: cp.walsh@ulster.ac.uk

Abstract

Background: Maternal blood folate concentrations during pregnancy have been previously linked with DNA methylation patterns, but this has been done predominantly through observational studies. We showed recently in an epigenetic analysis of the first randomized controlled trial (RCT) of Folic Acid Supplementation specifically in Second and Third Trimester (The EpiFASSTT Trial) that methylation at some imprinted genes was altered in cord blood samples in response to treatment. Here we report on epigenome-wide screening using the Illumina EPIC array (~850,000 sites) in these same samples (*n*=86).

Results: The top-ranked differentially-methylated promoter region (DMR) showed a gain in methylation with folic acid (FA) and was located upstream of the imprint regulator *ZFP57*. Differences in methylation in cord blood between placebo and folic acid treatment groups at this DMR were verified using pyrosequencing. The DMR also gains methylation in maternal blood in response to FA supplementation. We also found evidence of differential methylation at this region in an independent RCT cohort, the AFAST trial. By altering methylation at this region in two model systems *in vitro* we further demonstrated that it was associated with *ZFP57* transcription levels.

Conclusions: These results strengthen the link between folic acid supplementation during later pregnancy and epigenetic changes and identify a novel mechanism for regulation of *ZFP57*. This trial was registered 15th May 2013 at www.isrctn.com as ISRCTN19917787.

Keywords: Folic Acid; DNA methylation; Cord Blood; Offspring; Imprinting; ZFP57

Background

Folate is an essential B-vitamin required for viable embryonic and fetal development and as an important dietary constituent throughout life, fundamental in cellular biosynthesis and DNA methylation pathways [1,2]. Folic acid (FA) is the oxidized, and more stable, synthetic form of folate which is exclusively found in supplements and fortified foods [3]. Well-established evidence from randomized controlled trials [4,5] has led to recommendations, in place globally, that women should consume 400µg/d FA from prior to conception until the end of the first trimester in order to protect against neural tube defects (NTDs) [6,7]. Despite the identification of a relationship between maternal folate status and NTDs as early as 40 years ago, information on the mechanism behind the benefit of FA supplementation with respect to NTDs remains to be fully elucidated (reviewed in [8]), as does the relationship of FA, NTDs and DNA methylation [9]. There is however little dispute as regards the protective effect of folic acid supplementation before and in early pregnancy, which was proven in clinical trials to reduce NTDs by approximately 70% [10]. Furthermore there remains a lack of evidence as to whether it is beneficial to mother and/or child to continue this supplementation throughout the entire pregnancy [11,12]. FA supplementation during pregnancy has been associated with health benefits such as reduced risk of low birth weight [13], language delay [14], autism [15], reduced risk of psychosis [16] and other pediatric problems [17]. In addition, observational studies have indicated that FA supplement use by mothers during pregnancy is associated with better cognitive health and brain development in the child [14,18,19], possibly related to the fact that there is a brain growth spurt at the end of the second trimester [20,21]. However there may also be potential adverse effects from excess folate in later pregnancy, an aspect which would also benefit from further exploration [12].

At a molecular level, there is some evidence in human that epigenetic changes could be the mechanism underpinning some of the effects of folate, both in the first trimester [8] in the

prevention of NTDs, and also in second and third trimester, as reviewed elsewhere [2]. Folate is essential for the production of S-adenosylmethionine (SAM), which provides the methyl group to the DNA methyltransferases (DNMTs), which carry out DNA methylation. DNA methylation is an essential means of maintaining transcriptional silencing at many different classes of genes when it occurs at promoter and enhancer elements, including endogenous retroviruses, genes on the inactive X, and imprinted genes [22] but can also facilitate transcription when occurring in the gene body [23–25]. DNA methylation is vital for embryonic survival and development, as mice carrying mutations in the DNA methyltransferases die *in utero* or shortly after birth [26,27]. Some DNA methylation marks are inherited from the parents in the form of differential methylation on the paternal or maternal copy. This includes both the canonical imprinted loci, as well as some germline and neuronal genes [25,28,29], at all of which methylation plays a direct role in controlling transcription. Both animal and human studies have indicated that the fetal epigenome is vulnerable to environmental exposures, such as methyl group availability from the maternal diet [30–36].

Imprinted genes are a paradigm for transmission of epigenetic information across generations. Methylation differences between the paternal and maternal copies of imprinted genes are established in the germ cells and are known to be important for transcriptional regulation. Accordingly, inappropriate loss or gain of methylation at imprint control regions (ICR) is a diagnostic feature for several human disorders. These regions are protected from the wave of demethylation which occurs prior to implantation by several factors, such as PGC7/STELLA [37] and ZFP57, a Krueppel-associated box (KRAB) domain zinc finger protein [38,39]. Several studies to date have centered on analyzing the effects of nutrition in particular on imprinted genes [31–33,40] and have shown that not only can altered diet result in an altered epigenotype, but it can also affect phenotype and predisposition to childhood and adulthood disease [41].

We have previously reported data from a randomized controlled trial of Folic Acid Supplementation in the Second and Third Trimester (The FASSTT Trial; ISRCTN19917787) where we found supplementation led to significant protection against folate depletion in mothers and offspring [42], and more recently that this led to differences in DNA methylation at some imprinted loci by using a candidate gene approach [43]. Here we used the Infinium Methylation EPIC Beadchip Array to profile genome-wide DNA methylation levels in cord blood in an unbiased screen for regions susceptible to DNA methylation changes in response to altered FA levels. We report here that the top candidate region affected is a differentially methylated region (DMR) upstream of the gene encoding ZFP57. We verified our finding using pyrosequencing in cord blood, and also show that the region responds to FA supplementation in maternal blood. Additionally, we confirm that altering methylation results in changes in *ZFP57* transcription.

Results

Maternal FA supplementation significantly improves folate status in mother and baby

For the current analysis, the same 86 cord blood samples from the FASSTT trial (outlined in Fig. 1) which had been analyzed previously for candidate gene methylation [43] were used: a summary of the most pertinent characteristics are given in Table 1 for convenience. At baseline (gestational week 14 (GW14)), there were no detectable differences between the treatment and placebo groups in maternal characteristics, dietary folate intakes, serum or red blood cell (RBC) folate concentrations, or in *MTHFR* status, as expected following randomization. There were also no significant differences in neonatal characteristics such as weight, length, head circumference etc (Table 1). However as a result of treatment with FA during trimesters 2 and 3, maternal serum and RBC folate became significantly different between placebo and treated group, as previously reported from this trial. The normal decline in maternal folate biomarkers previously reported from observational studies during pregnancy are mirrored in the placebo

group where serum folate decreased from 48.8 to 23.6 nmol/L between GW14 and GW36 (Table 1). FA supplementation served to protect the mothers in the treatment group, where folate concentrations remained stable over the course of pregnancy (i.e. serum folate 45.8 nmol/L at GW14 and 46.5 nmol/L at GW36). Cord serum and RBC folate concentrations were also significantly higher in infants of the mothers supplemented with FA compared with those from the placebo mothers (Table 1). RBC folate concentrations in mothers and offspring were strongly correlated (r = 0.619; *P* = <0.001, Suppl. Fig. 1).

Widespread alterations to DNA methylation levels in cord blood in response to late gestation maternal FA supplementation

DNA was purified from cord blood and quantified prior to bisulfite conversion and hybridization to the Infinium Methylation EPIC Beadchip Array, which covers more than 850,000 CpG sites distributed across the genome. Methylation values are expressed as a decimal value β between 0.0 (no methylation) and 1.0 (fully methylated). Data were analyzed and visualized using the RnBeads package in RStudio (see methods section). As a control, a quantile-quantile (QQ) plot of observed versus expected Chi-squared values was generated and showed no evidence of population substructure effects (Suppl. Fig.2). Figure 2A is a scatterplot showing mean β value for each CpG site analyzed in treated versus placebo samples. Overall, methylation at individual CpG remains closely correlated (ρ=0.998) between the two groups as expected, with most sites falling along the diagonal. Sites which differed in methylation between placebo and treatment groups were automatically ranked by RnBeads, which uses a combination of the change in mean methylation, the quotient of mean methylation and the combined *p* value, and the 1000 top-ranking sites are highlighted in red in Figure 2A. This metric was developed to take into account not only *p* value but the magnitude of the change in methylation and in our experience is a more reliable indicator of biologically meaningful differences than *p* value alone. Sites falling along either side of the diagonal, representing gains

and losses in methylation after treatment, can both be seen, with a tendency to greater numbers of sites losing. Consistent with this, a methylation density distribution plot shows that after treatment there was a clear decrease in the numbers of sites in the top quartile for methylation (β=0.75-1.00; Fig. 2B). Taking the top 1000 ranking sites overall, approximately 2/3 (*n*=658) lost and 1/3 (*n*=342) gained methylation (Fig. 2C). However, the magnitude of these changes was generally modest, with only 302 (193+109) losing or gaining more than 5% methylation, the minimum change which we could potentially verify using pyrosequencing, and only 76 sites losing or gaining more than 10% (Fig. 2C).

We examined the top-ranking sites as identified by RnBeads (Fig.2D): of these, the CG site in the *ATP11A* gene contained a single nucleotide polymorphism (SNP) missed by the quality control routines: the same was true of the CG at the *MAGI2* gene. The presence of the SNPs at these CGs leads to the erroneous appearance of a change in methylation, so these were discounted. Two of the other top-ranked sites were at the *PRKAR1B* locus, which encodes a regulatory subunit of cyclic AMP-dependent protein kinase A, and one was at *NXN*, a member of the thioredoxin superfamily: however all three were listed as located in the respective gene body and so are less likely to contribute to transcriptional control. Nevertheless to verify these we used a second method utilizing commercial pyrosequencing methylation assays (pyroassays) designed to query the same CGs. These reported smaller average differences in methylation between treated and placebo groups than seen with the array of 6.6% for cg08104960 at *NXN*, and 4.2% (cg06242242) and 2.2% for (cg05729249) for the sites at *PRKAR1B*: only the site at *NXN* was significant (p=0.002, t-test).

Identification and verification of a differentially methylated region upstream of *ZFP57*

Given that single sites are more susceptible to confounders such as the presence of SNPs and show only moderate accuracy on verification, and to maximize our chances of finding biologically significant changes, we also looked for genomic intervals showing coherent

alterations in methylation across multiple neighboring sites [44], rather than isolated CpGs. Figure 3A lists the top 5 differentially methylated regions (DMR) found at promoters, ordered by RnBeads ranking which is here computed by combining measures at adjacent sites using a linear hierarchical model as described in Methods: uncorrected *p* value and % change in methylation are also shown for comparison. For the top 5 regions, *ZFP57* was of particular interest and is dealt with below. Two others (*CES1*, a liver carboxylesterase, and *ANKRD20A11P*, a pseudogene) showed less than 5% change in methylation and so could not be verified: *DUSP22* which has a larger change is also a pseudogene. The last DMR is located at a microRNA cluster *MIR4520A/B* and loses approximately 7.22% overall in the treatment group, averaged over a number of well-spaced CG. Due to pyrosequencing assay design constraints, we could only cover one site (cg08750459) from the array at this locus but that site showed reasonable concordance (loss of 12.24% ($p = 0.008$) in array and 9.45% ($p = 0.006$) by pyroassay). The function of these microRNA remains obscure however.

Of more interest in the context of this cohort was the highest ranking promoter DMR identified using RnBeads [45], which was located on chromosome 6, the closest gene being the known regulator of genomic imprinting *ZFP57*. The identified DMR consisted of 15 CpG sites and mapped approximately 3kb upstream of the first exon of the gene, a region containing additional adjacent sites also gaining methylation. Figure 3B shows a genomic map of the first exon of *ZFP57* and the upstream region, overlaid with a track showing the locations of EPIC probes and whether they gained or lost methylation. Also shown is a graph of averaged methylation values at the numbered CpG probes from the array in placebo and treatment groups, showing a clear difference in methylation extending beyond the DMR. To confirm these results using a second method we designed a pyrosequencing methylation assay (pyroassay) to cover some of these CpG sites, as shown in Fig.3B. Due to the CpG density of this region, thus difficulty in pyrosequencing primer design, our pyroassay is not directly

overlapping all CpGs identified by RnBeads as the DMR but is inside the area showing methylation differences. We then carried out PCR and pyrosequencing for all the samples. Methylation levels at each array site (*n*=15) in the DMR, averaged across each group are shown in Fig.3C. This confirmed clear differences between the two groups with respect to methylation across this region, with cord bloods from the children born to FA-supplemented mothers showing higher methylation at all sites, and 6.23% more methylation overall for the significant DMR probes (Fig.3A). The overall gain in methylation at the CpGs covered by the pyroassay (*n*=6) was very similar in magnitude and direction to that seen over the neighboring CpG by the array (+5.44% vs +6.23%, respectively- Fig.3A, C).

Demethylation of the upstream region was accompanied by increased *ZFP57* transcription

Having established that methylation differences at the upstream DMR are evident between FA-supplemented and placebo-treated controls, we wished to test mechanistically if such differences could impact on transcription from the downstream gene. To do this, we first used a well-established model, the paired colorectal cancer lines HCT116 and its derivative HCT116 DKO (double knockout), which carries mutations in two of the methyltransferase genes *DNMT1* and *DNMT3B* and is known to be hypomethylated at many loci [46]. Methylation array data available in-house showed differential methylation between the parental or wild type HCT116 (WT) and paired DKO cells at the same region upstream of *ZFP57* found in the FASSTT cohort, indicated by red coloured bars whose height is proportional to the loss of methylation (Fig.4A); this indicates that DNMT1 and DNMT3B are required for methylation at this locus. We confirmed these results using our pyroassay, which showed >80% methylation in WT HCT116 cells and a drop to <20% in DKO cells ($p<0.001$) (Fig.4B).

To determine if methylation at this upstream region can regulate transcription at the *ZFP57* gene 3kb downstream, we designed primers to cover part of the transcript as shown in

Fig.5A (FW/RV) and carried out reverse transcription on mRNA from the cells followed by polymerase chain reaction (RT-PCR). While minimal transcript could be detected in the HCT116 WT cells, which are heavily methylated, signal was readily apparent in the demethylated DKO cells (Fig. 4C). We confirmed this expression pattern quantitatively using RT-qPCR (Fig.4D). While these results show that the gene can be de-repressed in response to loss of methylation, it is normally not expressed in colon cells, from which HCT116 were derived, so we used the neuroblastoma cell line SH-SY5Y to test the effect of methylation changes on transcription in a neural cell type. *ZFP57* is normally transcribed in neural tissue as well as early embryo [47], but shows some methylation in the SH-SY5Y cells, which may be due to differences among neural cell types, or reflect accumulation of methylation during culture: however these cells are likelier than HCT116 to contain neural-specific transcription factors. Here we used a second method to perturb methylation, namely treatment with the DNA methyltransferase inhibitor 5'aza-2'deoxycytidine (5-aza-dC). Exposure of the cells to this small molecule inhibitor caused loss of methylation at the upstream region (Fig.4E). RT-PCR confirmed that *ZFP57* was de-repressed upon treatment with 5-aza-dC (Fig.4F). Quantification of mRNA levels with RT-qPCR again indicated a substantial increase in transcription from the gene in response to loss of methylation (Fig. 4G).

Greater variability at imprinted DMR in folate-treated samples

These results suggest that the increased methylation seen at the *ZFP57* upstream region will lead to a decreased transcription. Since ZFP57 plays a role in maintaining methylation specifically at imprinted genes, we examined methylation levels at these regions using data from the EPIC array. We used germline differentially methylated regions as defined by [48] and assessed average methylation across all probes which fell within these intervals. We excluded DMR which were flagged as acquiring methylation differences somatically, and also germline DMR where methylation as assessed by the array fell outside the 35-65% methylation

range defined as normal in that study. This left 15 imprinted germline DMR for which the median methylation level fell within the normal range in the placebo group (Suppl. Fig.3A). Comparing the samples from the folate supplemented group, only the maternally imprinted Neuronatin gene (*NNAT*) showed a small but significant loss of methylation in the treatment group (p=0.022, Mann-Whitney U test (MWU)) but there was no significant difference between placebo and treatment for any other DMR. However it was notable that 11/15 DMR showed a significantly greater variability in methylation in treated participants (p<0.001, Chi-squared test), which can be seen from the greater interquartile range (IQR -see Suppl. Fig.3A). Along with this greater variability in the treatment group, the median methylation levels trended lower than the placebo group for almost all imprinted genes. (Suppl. Fig. 3A). We repeated this analysis using imprinted DMR as defined by Court et al [49], which defines slightly larger DMR based on an analysis of Illumina 450K data. After applying similar criteria as above, this left 14 DMR suitable for comparison. Using these genomic intervals, again only *NNAT* showed a significantly different level of methylation in treated samples (p=0.022, MWU; Supp.Fig.3B), although *PLAG1* was also close to significant (p=0.072, MWU). Again, the IQR for the imprints showed greater variability in the treated than placebo groups (p<0.001, Chi-squared test) and medians tended to be lower in the FA-treated group (Suppl.Fig.3B).

Increased *ZPF57* methylation in response to FA in maternal blood samples

In order to investigate the effects of FA in maternal tissue, and to elucidate if this differentially methylated region upstream of *ZFP57* was directly responsive, we carried out pyrosequencing on matched maternal buffy coat samples at GW14 (n=24) and GW36 (n=24) (i.e. comparing the same mother's blood sample taken before and after intervention). Pyrosequencing analysis confirmed that FA-supplemented mothers show a 5.51% increase in DNA methylation levels at this DMR after late gestation supplementation (p value not signif.), in contrast to non-supplemented mothers, whose methylation levels decreased 1.51% at GW36 (Table 2).

Effect of FA at the *ZFP57* DMR in a second cohort

In order to test the generality of the effect of folic acid intervention on this genomic region, we examined data from a second randomized-controlled trial. The Aberdeen Folic Acid Supplementation Trial (AFAST) was an RCT using two doses of folic acid (0.2 and 5mg/day vs placebo) during pregnancy, with intervention starting at antenatal booking at <30weeks gestational age [50]. The study was conducted in the late 1960s and recently Richmond and colleagues [35] followed-up on the offspring born to the mothers who had participated in the trial, mean present age 47 years. Saliva samples were collected from those who could be identified and consented, with subsequent 450k array analysis conducted using modelling approaches to correct for hidden variables such as cell counts [35]. Examination of the CpG in the *ZFP57* DMR which we had identified in the EpiFASSTT cohort showed the same trends in the AFAST high-folate cohort versus placebo, with change in a positive direction across the whole region (Fig.5A), although effect size was lower at each site in the AFAST study (Fig.5B).

Discussion

We have previously reported DNA methylation differences at imprinted loci using cord blood from the EpiFASSTT trial of folic acid (FA) supplementation in later pregnancy by using a candidate gene approach. Here we used the same samples to carry out an unbiased genome-wide screen for methylation differences using the EPIC array. The top hit was a differentially methylated region upstream of the imprint controller *ZFP57* and we separately verified methylation differences by pyroassay. This region responded to FA supplementation in maternal blood as well as in cord blood, and showed differences between FA-treated and untreated in an independent cohort [50]. Altered methylation at *ZFP57* was associated with increased variation in methylation at imprinted loci in cord blood. We also showed using two

separate cell line models that altering methylation at the *ZFP57* upstream region can affect transcription, indicating a potential feedback mechanism may be operating here. We were also able to identify and verify methylation changes at a number of other individual CpG sites including some in the gene bodies of the *NXN* and *PRKAR1B* genes and at the start of the *MIR4520A/B* gene, but these were less likely to have functional consequences. It is notable also that we found more decreases in methylation genome-wide than increases, which may seem counter-intuitive; however we and others have reported similar response to FA previously [43, 56]. It is has been suggested that FA may cause feedback inhibition by altering the SAM:SAH ratio and therefore the intracellular methylation potential [51].

Uncovering a DMR at a region controlling *ZFP57* transcription as the top hit in an unbiased screen was particularly striking in the EpiFASSTT randomized controlled trial, where we have already shown, using a candidate gene approach, that methylation levels were perturbed at some imprinted loci. The primary importance of ZFP57, as described in the literature from mechanistic work, is in maintaining imprinting, and it is currently the only protein known to be dedicated solely or largely to this epigenetic process [52]. ZFP57 was discovered as a maternal-zygotic effect gene which was required in mice for establishing methylation at some imprints in the oocyte, and for maintaining all imprints, both maternal and paternal, in the preimplantation embryo [38]. It does this by binding to a conserved hexamer consensus sequence (5'-TGCme5CGC-3) found at all imprinting control regions (ICRs) [53,54], recognizing the methylated CpG in this motif, as shown in a crystallographic study [55]. Deletion of mouse *Zfp57* causes a loss of methylation from the modified parental allele by mid-gestation, with subsequent dysregulation of transcription at imprinted loci and embryonic lethality [54]. Importantly, mutations in the human homologue *ZFP57* are also associated with hypomethylation of multiple imprinted loci, indicating a conserved role in human for this gene in maintaining imprints [39].

Although this is the first report, to our knowledge, from a randomized controlled trial of FA intervention which implicates methylation changes at *ZFP57*, it was previously reported from a small observational study (*n*=23) that maternal folate concentrations in the third trimester were associated with changes at a DMR at the same genomic location [56] when cord blood DNA methylation levels at birth were profiled. While that study reported a loss rather than gain of methylation, it was not an RCT but an observational study, and so could not test the effects of folate supplementation directly in a controlled fashion: there were many other differences in study design, numbers of participants and analysis methods. It should also be noted that the high folate group in that study had levels of serum folate almost twice those seen in our treated samples (74.59+/- 6.1 nmol/L Amarasekera *et al.* vs 46.5 +/-19.5 nmol/L GW36 treated group this study), highlighting that we are protecting normal folate levels rather than elevating them. Although the largest-to-date observational study, comprising a meta-analysis of the MoBa (*n*=1275) and Generation R (*n*=713) cohorts, did not identify this region as a top hit, they could confirm that 5 CpG sites within this 923bp region were significantly altered, though not the direction of change [57]. These two papers reporting changes from different observational studies nevertheless lend considerable support to this being a true folate-sensitive DMR. We could also verify using a separate biological assay the magnitude and direction of change in methylation, a gain of 5.44% in the treatment group, at the DMR in cord blood by using pyrosequencing (p = 0.172). Furthermore, by comparing the mothers pre- and post-intervention we could show that this region also gained methylation in the treated mothers, but lost methylation in the placebo group, providing a further degree of validation.

To extend our findings, we also used data from one of the few other RCT testing the role of folic acid during pregnancy, the AFAST study [50]. We found a small effect (Cohen's D<0.2) at all the CpG across the *ZFP57* DMR, whereas there was a medium effect (Cohen's D <0.5) seen at the same region in the EpiFASSTT study. The effect in AFAST was only seen

with the high dose of FA (5mg/day) vs placebo, rather than the lower dose (200μg/day) which was closer to that used in EpiFASSTT (400μg/day), and the effect size was smaller than that seen in EpiFASSTT. There may be a number of reasons why effect size was smaller in AFAST: 1)the time between exposure and measurement is much greater, with median age 47 years in AFAST, vs newborns in EpiFASSTT; 2)the AFAST participants used were recruited significantly later than other groups (20.2 weeks for high dose vs 16.3 for low dose), meaning that there was less time spent exposed to the additional FA while in the womb 3)the AFAST DNA samples were derived from saliva, while the EpiFASSTT DNA samples are from cord blood and 4)final numbers for the AFAST comparisons were very low (5mg/day n=23; placebo n=43). Notwithstanding these limitations, the AFAST study showed a similar effect in terms of direction and magnitude at the same region upstream of *ZFP57*, providing further evidence that this is a bona fide FA sensor.

Given the role of ZFP57 in imprint maintenance, we also took advantage of the array to examine imprinted genes in our samples. Of these, only the maternal imprint *NNAT* (neuronatin) showed a small but significant loss of methylation in the treatment group, consistent with other evidence [58]. *NNAT* is highly expressed in brain and placental tissue and functions during brain development to regulate ion channels and maintain hindbrain and pituitary segment identity [59]. *ZFP57* is essential for the maintenance of this imprint [38]. Induction of increasing mRNA levels of *NNAT* commences at midgestation in association with neurogenesis, and peaks upon neuroepithelial proliferation and neuroblast formation [60], which would coincide with when folate concentrations increased in the treated group. Although we previously reported significant differences overall at *IGF2*, and at some CpG for *GRB10* in our candidate gene approach using these samples [43], that was based on pyroassays which covered smaller regions of the imprinted DMR, whereas the probes from the array are more dispersed and cover a larger area. It was also notable that, while there was little change

at other imprinted DMR as assessed by the array, there did appear to be an increase in the variability of methylation at these regions, an effect which was small but statistically significant and consistent with findings from a mouse model where FA supplementation increased variance in methylation levels across generations [61]. Given that *ZFP57* has a role in maintaining imprints, increased methylation at the upstream controller as seen in our FA-treated samples should lead to decreased transcription of *ZFP57*, which could potentially lead to reduced ability to maintain imprints and increased variability in methylation at the ICR. These possibilities can be further explored using our *in vitro* cell models.

It remains to be established from mechanistic studies in mouse whether ZFP57 plays any role in maintaining methylation *in vivo* in the post-implantation embryo. It is also possible that methylation of the DMR in human blood may not reflect the methylation levels seen at earlier stages, or in tissues which normally express the gene, which includes oocytes and some neural cells. It may be that methylation levels at the *ZFP57* DMR reported here reflect changes which have occurred in cord and maternal bloods independently of what is occurring in germline, and this would need to be assessed. It is also quite likely, given that imprints are thought to be established much earlier during development that it would not be until the next generation that effects at imprinted germline DMRs could be seen. In this context, several studies have pointed to transgenerational rather than intergenerational effects at imprinted loci [62,63]. It should also be noted that methylation levels varied substantially across the *ZFP57* DMR and between individuals (max = 94.97, min = 20.95), unlike the imprinted DMR which vary much less and may be buffered against methylation changes by multiple mechanisms.

In addition to its well-established role in imprinting, ZFP57 has also been proposed to act as a transcriptional repressor in Schwann cells, which comprise the principal glia of the peripheral nervous system [47]. Recent work from our group has indicated children born from mothers supplemented with FA in late gestation have psychosocial developmental benefits,

scoring significantly higher for emotional intelligence and resilience in comparison with children not exposed to FA supplementation in later pregnancy [64]. Further work needs to be carried out to check if there are any other, novel targets of ZFP57 which may be affected in later childhood and adulthood.

We sought to clarify whether an increase in methylation at the *ZFP57* DMR as seen in this RCT would have a substantial effect on production of the protein. In order to explore whether changes in methylation can alter transcription we utilized cell lines where the only variable was the presence or absence of DNA methylation. Our results from these two systems (HCT116 cells with methyltransferase deficiency and SH-SY5Y cells treated with an inhibitor) showed that altering methylation alone can cause changes in transcription at the *ZFP57* locus, and that this is linked to changes in methylation at the DMR. Our results therefore support the hypothesis that the DMR represents an upstream control element for the gene, which we have shown from the RCT is sensitive to methyl donor status in the diet. Little is currently known about the factors controlling *ZFP57* transcription. Interestingly, the region containing the DMR does not appear to be conserved in mice, and so may represent a human-specific element. However it has features characteristic of a control element, as from examining publicly-available datasets on the UCSC genome browser there are DNAse I hypersensitive sites present here and data suggesting transcription factors may bind. We are currently exploring these aspects of the work further.

Conclusions

Despite the limitations discussed above, we have nevertheless shown conclusively that a region upstream of the imprint controller *ZFP57* shows changes in methylation in mothers in response to intervention during later pregnancy with FA, a methyl donor, and that this effect is also evident in the cord blood in their offspring. Our findings are borne out by other observational studies as well as an independent RCT [50]. We have also clearly demonstrated that altering

methylation is sufficient in itself to cause changes in transcription of the gene. These results have implications for the control of imprinting by environmental inputs and uncover a novel transcriptional control element which may be involved in this process.

Methods

Study design and sample collection

Samples were acquired from the FASSTT (Folic acid supplementation in the second and third trimester) study cohort, a previously-conducted double-blinded, randomized controlled trial in Northern Ireland described in full previously [42,43]. To summarize in brief, women with singleton pregnancies were recruited at approximately 14 weeks of gestation from antenatal clinics at the Causeway Hospital, Coleraine (n = 226; Fig.1). Women were excluded from participation if they were taking medication known to interfere with B-vitamin metabolism or if they had any vascular, renal, hepatic or gastrointestinal disease, epilepsy or had a previous NTD-affected pregnancy. Prior to randomization, n=36 women withdrew from the study. The remaining eligible participants at the end of their first trimester were randomized into two groups; one group received 400µg/d folic acid (n = 96) and the other a placebo in pill form (n = 94) until the end of their pregnancy. Randomization was done on a double-blind basis. Maternal non-fasting blood samples were taken at gestational week 14 (GW14), prior to intervention commencement, and at GW36, towards the end of the intervention. The study was completed by 119 women, as 71 participants were excluded during the study (see Fig.1). A total of n=37 women were excluded from the folic acid group for the following reasons: participant withdrawal n=11, pregnancy complications n=13, prescribed folic acid n=6, foetal death n=6, non-compliance n=6. A total of n=34 women were excluded from the folic acid group for the following reasons: participant withdrawal n=14, pregnancy complications n=8, prescribed folic acid n=5, foetal death n=2, non-compliance n=3, hospital transfer n=2.

Umbilical cord blood samples were collected after the expulsion of the placenta at delivery, along with birth weight, length, head circumference, mode of delivery and Apgar score.

Blood sample processing and B-vitamin biomarker determination

Blood samples were collected in EDTA-lined tubes, kept refrigerated and processed within 4h (excepting cord blood, processed within 24h). Blood samples were analyzed for serum and red blood cell folate and vitamin B12 via microbiological assay as previously described [65,66] . The buffy coat was used for methylenetetrahydrofolate reductase (*MTHFR*) 677C>T genotyping as described (Frosst *et al.* 1995). Quality control was affirmed by repeated analysis of stored batches of pooled samples. Intra- and inter-assay CVs were ≤8.2% for serum and RBC folate and ≤10.4% for serum vitamin B12.

Maternal dietary analysis

Dietary data was collected using a 4d food diary in combination with a food-frequency questionnaire during the second trimester of pregnancy, with particular emphasis on a of B-vitamin-fortified food intake. Dietary analysis was carried out using WISP version 3.0 (Tinuviel Software, UK) modified to segregate naturally-occurring folate in foods versus folic acid fortification of foods; these were combined to enable calculation of dietary folate equivalents.

Cell culture

HCT116 and double knockout (DKO) cells (Rhee *et al.* 2002) were cultured in 1g/L glucose DMEM supplemented with 10% FBS and 1x NEAA (Thermo Scientific, Loughborough, UK). SH-SY5Y cells were cultured in DMEM/F12 medium supplemented with 10% FBS (Thermo Scientific) For treatment with 5'aza-2-deoxycytidine (5-aza-dC) (Sigma-Aldrich, Dorset, UK), SH-SY5Y cells were seeded onto a 90mm plate in complete medium, and the following day medium was replaced and supplemented with 5-aza-dC at a final concentration of 1μM, which

was renewed at 24-hour intervals up to 72 hours. Cells were then harvested for DNA and RNA extraction.

Transcriptional analysis

RNA was extracted using the RNeasy Mini kit (Qiagen, Crawley, UK) according to manufacturer's instructions. Complementary DNA (cDNA) was synthesized and RT-qPCR/RT-PCR were carried out as previously [29]. Primer sequences are listed in Supplementary Table 1. Human reference total RNA was used as a positive control for expression (Clontech, UK).

DNA extraction, bisulfite conversion and Infinium MethylationEPIC Beadchip Array

Genomic DNA was extracted from cultured cells as previously described [25] and from cord blood using the QiAMP DNA Blood Mini kit (Qiagen), according to manufacturer's instructions. Purity and integrity of DNA were assessed by agarose gel electrophoresis and using the Nanodrop 2000 spectrophotometer (Labtech International, Ringmer, UK). DNA quantification was determined using Quant-IT PicoGreen dsDNA Assay Kit (Invitrogen, Paisley, UK). The DNA at a concentration of 50ng/µl was sent to Cambridge Genomic Services (Cambridge, UK), who bisulfite converted the DNA in-house using the EZ DNA Methylation Kit (Zymo Research, California, USA) prior to hybridization to the Infinium Human Methylation EPIC BeadChip Array and scanning with the Illumina iScan according to manufacturer's instructions (Illumina, Chesterford, UK).

Bioinformatic analysis

GenomeStudio (Illumina v3.2) was used for initial data processing. Subsequently *idat* files were imported into the *RnBeads* package (version 1.6.1) [45] in the freely available statistical software platform R (version 3.1.3) using R Studio interface (Version 0.99.903). Samples were quality control checked including removal of probes with missing values, containing SNPs, or of poor quality using the *greedycut* algorithm, then sex chromosomes were removed from the

analysis. Background correction was carried out using *methylumi.noob* and the methylation values of the remainder probes were normalized using *bmiq* [67]. Initial data exploration in RnBeads used principle components analysis (PCA) to explore potential correlations between the groups and known confounders such as BMI, smoking, gender etc. In addition, in order to account for any hidden confounding variables in the dataset, surrogate variable analysis was carried out using the *sva* package with the Buja and Eyboglu algorithm from (1992) [68]. Briefly, potential surrogate variables such as age, sample plate, Sentrix ID and Sentrix Position were tested for association with the target variable sample group using PCA and any surrogate variable with a high correlation to sample group was adjusted for and incorporated into the making of the *limma* based linear model. The methylation intensities for each probe, each representing a CpG site, were represented as β-values (ranging from 0, unmethylated, to 1, fully methylated) and these were plotted against genomic loci (based on *hg19* -Human Genome Build 19) using *GALAXY* software (https://usegalaxy.org/) [69] in order to visualize changes in DNA methylation on the University of California at Santa Cruz genome browser (https://genome.ucsc.edu/) as described previously [70].

Bisulfite pyrosequencing

Primers spanning the probes of interest from the array were designed using the PyroMark Assay Design Software 2.0 and bisulfite-treated DNA PCR-amplified using the PyroMark PCR kit prior to analysis on a PyroMark Q24 according to manufacturer's instruction (Qiagen). The primer sequences are summarized in Supplementary Table 2. Amplification was carried out as follows: 95°C for 15min, followed by 45 cycles of 95°C for 30sec, 56°C for 30 sec, and 72°C for 30sec, with a final elongation step at 72°C for 10min. Products were verified via gel electrophoresis prior to pyrosequencing analysis.

Statistical analysis

Statistical analysis was performed using the Statistical Package for the Social Sciences software (SPSS) (Version 22.0; SPSS UK Ltd., Chertsey, UK). The results are expressed as mean ± SD, except where otherwise stated. For normalization purposes, variables were log transformed before analysis, as appropriate. Differences between treatment groups for participant characteristics were assessed using an independent *t* test for continuous variables or chi-square for categorical variables. Pyrosequencing data and RT-qPCR data were analyzed using Student's *t* test to identify statistical differences between intervention groups. A *p* value <0.05 was considered significant. Differential methylation analysis was conducted in *RnBeads* (see above) on a site and region level. The normalized β-values were converted into M-values ($M = \log2(\beta/(1-\beta))$) and differential methylation between samples (placebo vs. treatment) was estimated with hierarchical linear models using *limma*. Ranking was automatically carried out in *RnBeads* and was based on the combination of the average difference in means across all sites in the promoter regions of the sample groups, the mean of quotients in mean methylation and the combined p-value, which was calculated from all site p-values in the region using a generalization of Fisher's method [71]. The smaller the combined rank for a region, the more evidence for differential methylation it exhibits.

Abbreviations

AFAST: Aberdeen Folic Acid Supplementation Trial; 5-aza-dC: 5'aza-2'deoxycytidine; BMI: Body mass index; DKO: Double knockout; DMR: Differentially methylated region; DNMT: DNA methyltransferases; ESRC: Economic and Social Research Council; BBSRC: Biotechnology and Biological Sciences Research Council; FA: Folic acid; FASSTT: Folic acid supplementation in second and third trimester; GW: Gestational week; ICR: Imprint control region; IQR: Interquartile range; KRAB: Krueppel-associated box; MRC: Medical Research Council; MTHFR: Methylene tetrahydrofolate reductase; MWU: Mann-Whitney U test; NTD:

Neural tube defects; NNAT: Neuronatin; Pyroassay: pyrosequencing methylation assay; ORECNI: Office for Research and Ethics Committees Northern Ireland; QQ: Quantile-quantile; RT-PCR: Reverse transcription- polymerase chain reaction; RCT: Randomized controlled trial; SAM: S-adenosylmethionine; SPSS: Statistical Package for the Social Sciences; WT: Wild type; ZFP57: Zinc finger protein 57

Declarations

Ethics approval and consent to participate

The Office for Research and Ethics Committees Northern Ireland (ORECNI) granted ethical approval (reference 05/ Q2008/21) and each participant gave written informed consent upon recruitment.

Consent for Publication

Not applicable

Availability of data and materials

The datasets used and analyzed during the current study are available where appropriate from the corresponding author on reasonable request and subject to governance regulations at Ulster (EpiFASSTT): for data from the AFAST study contact C. Relton.

Competing Interests

The authors declare that they have no competing interests.

Funding

Work was supported by grants jointly funded by the Economic and Social Research Council (ESRC) and Biotechnology and Biological Sciences Research Council (BBSRC), grant refs: ES/N000323/1 (CPW) and ES/N000498/1 (CLR). RCR, MS and CLR work in a unit supported by the Medical Research Council (MC_UU_12013/1, MC_UU_12013/2 and MC_UU_12013/8).

Authors' contributions

CPW, KP and HM designed and planned the work; REI and MO carried out the lab work; REI and SJT performed the bioinformatics analysis for EpiFASSTT; RR carried out the analysis for AFAST; AC and DLM helped with EpiFASSTT samples and statistics; MM and TC advised on biopsychosocial correlations; MS advised on bioinformatics approaches for both cohorts; CLR advised on overall approaches and coordinated the AFASST comparison; REI and CPW wrote the paper; all authors commented on the final MS.

Acknowledgements

The authors are grateful to the other members of the Walsh and Relton labs who provided valuable feedback on the work.

References

1. Crider KS, Yang TP, Berry RJ, Bailey LB. Folate and DNA Methylation : A Review of Molecular Mechanisms and the Evidence for Folate's Role. Am Soc Nutr. 2012;3:21–38.

2. Irwin RE, Pentieva K, Cassidy T, Lees-Murdock DJ, McLaughlin M, Prasad G, et al. The interplay between DNA methylation, folate and neurocognitive development. Epigenomics 2016;8(6);863–79.

3. Bailey LB, Stover PJ, McNulty H, Fenech MF, Gregory JF, Mills JL, et al. Biomarkers of Nutrition for Development--Folate Review. J Nutr. 2015;145(7):1636S-1680S.

4. MRC Vitamin Study Research Group. Prevention of neural tube defects: results of the Medical Research Council Vitamin Study. MRC Vitamin Study Research Group. Lancet. 1991;338:131–7.

5. Czeizel AE, Dudás I. Prevention of the first occurrence of neural-tube defects by periconceptional vitamin supplementation. Obstet Gynecol Surv. 1993;48:395–7.

6. Centers for Disease Control. Recommendations for the use of folic acid to reduce the number of cases of spina bifida and other neural tube defects. Morb Mortal Wkly report. 1992;41:1–8.

7. Department of Health. Folic acid and the prevention of neural tube defects. Heywood, United Kingdom; 1992.

8. Greene NDE, Stanier P, Moore GE. The emerging role of epigenetic mechanisms in the etiology of neural tube defects. Epigenetics. 2011;6(7):875-83.

9. Rochtus A, Jansen K, Geet C, Freson K. Nutri-epigenomic Studies Related to Neural Tube Defects: Does Folate Affect Neural Tube Closure Via Changes in DNA Methylation? Mini-Reviews Med Chem. 2015;15(13):1095-102.

10. Blom HJ, Shaw GM, Den Heijer M, Finnell RH. Neural tube defects and folate: Case far from closed. Nat. Rev. Neurosci. 2006;7(9):724-31.

11. Mills JL, Molloy AM, Reynolds EH. Do the benefits of folic acid fortification outweigh

the risk of masking vitamin B12 deficiency? BMJ. 2018;360:k724.

12. Schrott R, Murphy SK. Folic acid throughout pregnancy: Too much? Am J Clin Nutr. 2018;107:497–8.

13. Hodgetts V, Morris R, Francis A, Gardosi J, Ismail K. Effectiveness of folic acid supplementation in pregnancy on reducing the risk of small-for-gestational age neonates: a population study, systematic review and meta-analysis. BJOG. 2015;122:478–90.

14. Roth C, Magnus P, Schjølberg S, Stoltenberg C, Surén P, McKeague IW, et al. Folic acid supplements in pregnancy and severe language delay in children. JAMA. 2011;306:1566–73.

15. Wang M, Li K, Zhao D, Li L. The association between maternal use of folic acid supplements during pregnancy and risk of autism spectrum disorders in children: A meta-analysis. Mol Autism. 2017;8:51.

16. Eryilmaz H, Dowling KF, Huntington FC, Rodriguez-Thompson A, Soare TW, Beard LM, et al. Association of Prenatal Exposure to Population-Wide Folic Acid Fortification With Altered Cerebral Cortex Maturation in Youths. JAMA Psychiatry. 2018; 75(9):918-928.

17. Barua S, Kuizon S, Junaid MA. Folic acid supplementation in pregnancy and implications in health and disease. J. Biomed. Sci. 2014;21:77.

18. Julvez J, Fortuny J, Mendez M, Torrent M, Ribas-Fitó N, Sunyer J. Maternal use of folic acid supplements during pregnancy and four-year-old neurodevelopment in a population-based birth cohort. Paediatr Perinat Epidemiol. 2009;23:199–206.

19. Villamor E, Rifas-Shiman SL, Gillman MW, Oken E. Maternal intake of methyl-donor nutrients and child cognition at 3 years of age. Paediatr Perinat Epidemiol. 2012;26:328–35.

20. de Graaf-Peters VB, Hadders-Algra M. Ontogeny of the human central nervous system: What is happening when? Early Hum Dev. 2006;82:257–66.

21. Nyaradi A, Li J, Hickling S, Foster J, Oddy WH. The role of nutrition in children's neurocognitive development, from pregnancy through childhood. Front Hum Neurosci.

2013;7:97.

22. Smith ZD, Meissner A. DNA methylation: roles in mammalian development. Nat Rev. 2013;14:204–20.

23. Wu H, Coskun V, Tao J, Xie W, Ge W, Yoshikawa K, et al. Dnmt3a-dependent nonpromoter DNA methylation facilitates transcription of neurogenic genes. Science (80-). 2010;329:444–7.

24. Neri F, Krepelova A, Incarnato D, Maldotti M, Parlato C, Galvagni F, et al. XDnmt3L antagonizes DNA methylation at bivalent promoters and favors DNA methylation at gene bodies in ESCs. Cell. 2013;155(1):121-34.

25. Irwin RE, Thakur A, O' Neill KM, Walsh CP. 5-Hydroxymethylation marks a class of neuronal gene regulated by intragenic methylcytosine levels. Genomics. 2014;104:383–92.

26. Li E, Bestor TH, Jaenisch R. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. Cell. 1992;69:915–26.

27. Okano M, Bell DW, Haber DA, Li E. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. Cell. 1999;99:247–57.

28. Borgel J, Guibert S, Li Y, Chiba H, Schübeler D, Sasaki H, et al. Targets and dynamics of promoter DNA methylation during early mouse development. Nat Genet. 2010;42:1093–100.

29. Rutledge CE, Thakur A, O'Neill KM, Irwin RE, Sato S, Hata K, et al. Ontogeny, conservation and functional significance of maternally inherited DNA methylation at two classes of non-imprinted genes. Development. 2014;141:1313–23.

30. Dolinoy DC. The agouti mouse model: An epigenetic biosensor for nutritional and environmental alterations on the fetal epigenome. Nutr Rev. 2008;66:S7-11.

31. Steegers-Theunissen RP, Obermann-Borst SA, Kremer D, Lindemans J, Siebel C, Steegers EA, et al. Periconceptional maternal folic acid use of 400 μg per day is related to increased methylation of the IGF2 gene in the very young child. PLoS One. 2009;4(11):e7845.

32. Heijmans BT, Tobi EW, Stein AD, Putter H, Blauw GJ, Susser ES, et al. Persistent epigenetic differences associated with prenatal exposure to famine in humans. Proc Natl Acad Sci. 2008;105:17046–9.

33. Haggarty P, Hoad G, Campbell DM, Horgan GW, Piyathilake C, McNeill G. Folate in pregnancy and imprinted gene and repeat element methylation in the offspring. Am J Clin Nutr. 2013;97:94–9.

34. Dominguez-Salas P, Moore SE, Baker MS, Bergen AW, Cox SE, Dyer RA, et al. Maternal nutrition at conception modulates DNA methylation of human metastable epialleles. Nat Commun. 2014;5:3746.

35. Richmond RC, Sharp GC, Herbert G, Atkinson C, Taylor C, Bhattacharya S, et al. The long-term impact of folic acid in pregnancy on offspring DNA methylation: follow-up of the Aberdeen Folic Acid Supplementation Trial (AFAST). Int J Epidemiol. 2018;47(3): 928–937.

36. Pauwels S, Ghosh M, Duca RC, Bekaert B, Freson K, Huybrechts I, et al. Maternal intake of methyl-group donors affects DNA methylation of metabolic genes in infants. Clin Epigenetics. 2017;9:16.

37. Nakamura T, Arai Y, Umehara H, Masuhara M, Kimura T, Taniguchi H, et al. PGC7/Stella protects against DNA demethylation in early embryogenesis. Nat Cell Biol. 2007;9:64–71.

38. Li X, Ito M, Zhou F, Youngson N, Zuo X, Leder P, et al. A Maternal-Zygotic Effect Gene, Zfp57, Maintains Both Maternal and Paternal Imprints. Dev Cell. 2008;15:547–57.

39. Mackay DJG, Callaway JLA, Marks SM, White HE, Acerini CL, Boonen SE, et al. Hypomethylation of multiple imprinted loci in individuals with transient neonatal diabetes is associated with mutations in ZFP57. Nat Genet. 2008;40:949–51.

40. Cooper WN, Khulan B, Owens S, Elks CE, Seidel V, Prentice AM, et al. DNA methylation profiling at imprinted loci after periconceptional micronutrient supplementation in humans: Results of a pilot randomized controlled trials. World Rev. Nutr. Diet. 2014;26(5):1782-90.

41. Barker DJP. The developmental origins of chronic adult disease. Acta Paediatr Int J Paediatr Suppl. 2004;93(446):26-33.

42. McNulty B, McNulty H, Marshall B, Ward M, Molloy AM, Scott JM, et al. Impact of continuing folic acid after the first trimester of pregnancy: Findings of a randomized trial of folic acid supplementation in the second and third trimesters. Am J Clin Nutr. 2013;98:92–8.

43. Caffrey A, Irwin RE, McNulty H, Strain JJ, Lees-Murdock DJ, McNulty BA, et al. Gene-specific DNA methylation in newborns in response to folic acid supplementation during the second and third trimesters of pregnancy: Epigenetic analysis from a randomized controlled trial. Am J Clin Nutr. 2018;107:566–75.

44. Teschendorff AE, Relton CL. Statistical and integrative system-level analysis of DNA methylation data. Nat. Rev. Genet. 2018;19(3):129-147.

45. Assenov Y, Müller F, Lutsik P, Walter J, Lengauer T, Bock C. Comprehensive analysis of DNA methylation data with RnBeads. Nat Methods. 2014;11:1138–40.

46. Rhee I, Bachman KE, Park BH, Jair KW, Yen RWC, Schuebel KE, et al. DNMT1 and DNMT3b cooperate to silence genes in human cancer cells. Nature. 2002;416:552–6.

47. Alonso MBD, Zoidl G, Taveggia C, Bosse F, Zoidl C, Rahman M, et al. Identification and characterization of ZFP-57, a novel zinc finger transcription factor in the mammalian peripheral nervous system. J Biol Chem. 2004;279:25653–64.

48. Woodfine K, Huddleston JE, Murrell A. Quantitative analysis of DNA methylation at all human imprinted regions reveals preservation of epigenetic stability in adult somatic tissue. Epigenetics and Chromatin. 2011;4(1):1.

49. Court F, Tayama C, Romanelli V, Martin-Trujillo A, Iglesias-Platas I, Okamura K, et al. Genome-wide parent-of-origin DNA methylation analysis reveals the intricacies of human imprinting and suggests a germline methylation-independent mechanism of establishment. Genome Res. 2014;24:554–69.

50. Charles DHM, Ness AR, Campbell D, Smith GD, Whitley E, Hall MH. Folic acid supplements in pregnancy and birth outcome: Re-analysis of a large randomised controlled trial and update of Cochrane review. Paediatr. Perinat. Epidemiol. 2005;19:2.

51. Christensen KE, Mikael LG, Leung KY, Lévesque N, Deng L, Wu Q, et al. High folic acid consumption leads to pseudo-MTHFR deficiency, altered lipid metabolism, and liver injury in mice. Am J Clin Nutr. 2015;101:646–58.

52. Mackin SJ, Thakur A, Walsh CP. Imprint stability and plasticity during development. Reproduction. 2018;156:43–55.

53. Strogantsev R, Krueger F, Yamazawa K, Shi H, Gould P, Goldman-Roberts M, et al. Allele-specific binding of ZFP57 in the epigenetic regulation of imprinted and non-imprinted monoallelic expression. Genome Biol. 2015;16:112.

54. Quenneville S, Verde G, Corsinotti A, Kapopoulou A, Jakobsson J, Offner S, et al. In embryonic stem cells, ZFP57/KAP1 recognize a methylated hexanucleotide to affect chromatin and DNA methylation of imprinting control regions. Mol Cell. 2011;44:361–72.

55. Liu Y, Toh H, Sasaki H, Zhang X, Cheng X. An atomic model of Zfp57 recognition of CpG methylation within a specific DNA sequence. Genes Dev. 2012; 26(21):2374-9.

56. Amarasekera M, Martino D, Ashley S, Harb H, Kesper D, Strickland D, et al. Genome-wide DNA methylation profiling identifies a folate-sensitive region of differential methylation upstream of ZFP57-imprinting regulator in humans. FASEB J. 2014;28:4068–76.

57. 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. 2016;7: 10577.

58. Hoyo C, Daltveit AK, Iversen E, Benjamin-Neelon SE, Fuemmeler B, Schildkraut J, et al. Erythrocyte folate concentrations, CpG methylation at genomically imprinted domains, and birth weight in a multiethnic newborn cohort. Epigenetics. 2014;9:1120–30.

59. Dou D, Joseph R. Cloning of human neuronatin gene and its localization to chromosome-20q11.2-12: the deduced protein is a novel "proteolipid." Brain Res. 1996;723:8–22.

60. Kikyo N, Williamson CM, John RM, Barton SC, Beechey C V., Ball ST, et al. Genetic and functional analysis of neuronatin in mice with maternal or paternal duplication of distal Chr 2. Dev Biol. 1997;190:66–77.

61. Li CCY, Cropley JE, Cowley MJ, Preiss T, Martin DIK, Suter CM. A sustained dietary change increases epigenetic variation in isogenic mice. PLoS Genet. 2011;7(4): e1001380.

62. Serpeloni F, Radtke K, de Assis SG, Henning F, Nätt D, Elbert T. Grandmaternal stress during pregnancy and DNA methylation of the third generation: an epigenome-wide association study. Transl Psychiatry. 2017;7:e1202.

63. Bygren LO, Tinghög P, Carstensen J, Edvinsson S, Kaati G, Pembrey ME, et al. Change in paternal grandmotherśearly food supply influenced cardiovascular mortality of the female grandchildren. BMC Genet. 2014;15.

64. Henry LA, Cassidy T, McLaughlin M, Pentieva K, McNulty H, Walsh CP, et al. Folic Acid Supplementation throughout pregnancy: psychological developmental benefits for children. Acta Paediatr Int J Paediatr. 2018.

65. Molloy AM, Scott JM. Microbiological assay for serum, plasma, and red cell folate using cryopreserved, microtiter plate method. Methods Enzymol. 1997;281:43–53.

66. Kelleher BP, Broin SD. Microbiological assay for vitamin B12 performed in 96-well microtitre plates. J Clin Pathol. 1991;44:592–5.

67. Teschendorff AE, Marabita F, Lechner M, Bartlett T, Tegner J, Gomez-Cabrero D, et al. A beta-mixture quantile normalization method for correcting probe design bias in Illumina Infinium 450 k DNA methylation data. Bioinformatics. 2013;29:189–96.

68. Leek JT, Johnson WE, Parker HS, Jaffe AE, Storey JD. The SVA package for removing batch effects and other unwanted variation in high-throughput experiments. Bioinformatics.

2012;28:882–3.

69. Giardine B, Riemer C, Hardison RC, Burhans R, Elnitski L, Shah P, et al. Galaxy: A platform for interactive large-scale genome analysis. Genome Res. 2005;15:1451–5.

70. Mackin SJ, O'Neill KM, Walsh CP. Comparison of DNMT1 inhibitors by methylome profiling identifies unique signature of 5-aza-2′deoxycytidine. Epigenomics. 2018;10(8):1085-1101.

71. Makambi KH. Weighted inverse chi-square method for correlated significance tests. J Appl Stat. 2003;30:225–34.

Table 1 General characteristics of participants from the EpiFASSTT trial

Characteristic	Placebo		Folic Acid		*P* value[1]
N	45		41		
Maternal characteristics (GW14)	**Mean**	**SD**	**Mean**	**SD**	
Age (y)	28.9	3.5	29.4	3.9	0.513
BMI (kg/m^2)	25.2	3.9	24.9	4.6	0.768
Smoker *n* (%)	8 (18)		6 (15)		0.693
Alcohol *n* (%)	3 (7)		1 (2)		0.618
Parity (n)	1 (1.1)		1 (1.0)		0.915
MTHFR 677TT genotype *n* (%)	5 (11)		2 (5)		0.291
Dietary Intakes					
Energy (MJ/d)	8.170	1.717	7.732	1.595	0.280
Dietary Folate Equivalents (µg/d)	364	172	387	152	0.582
Vitamin B12 (µg/d)	4.1	1.9	3.9	1.8	0.791
Neonatal characteristics					
Gestational age (wk)	40.1	1.3	40.0	1.1	0.540
Sex, Male *n* (%)	22 (49)		22 (54)		0.659
Birth weight (g)	3610	475	3557	465	0.601
Birth length (cm)	51.5	2.6	51.1	2.2	0.499
Head circumference (cm)	34.9	1.2	34.8	1.4	0.907
Apgar score at 5 min	8.4	0.4	9.0	0.3	0.220
Caesarian *n* (%)	11 (24)		10 (24)		0.995
B-vitamin Biomarkers					
Maternal pre-intervention (GW14)					
Serum folate (nmol/L)	48.8	19.8	45.8	19.5	0.469
RBC folate (nmol/L)	1185	765	1181	649	0.978
Serum B12 (pmol/L)	224	79	217	79	0.601
Maternal post-intervention (GW36)					
Serum folate (nmol/L)	23.6	17.9	46.5	24.8	**<0.001***
RBC folate (nmol/L)	991	404	1556	658	**<0.001***
Serum B12 (pmol/L)	168	51	157	60	0.229
Cord Blood					
Serum folate (nmol/L)	68.3	24.8	91.7	36.7	**0.004***
RBC folate (nmol/L)	1518	597	1877	701	**0.024***
Serum B12 (pmol/L)	276	155	251	107	0.776

Statistical comparisons by independent *t* test (continuous variables) or χ^2 test (categorical variables). *$P < 0.05$. GW, gestational week; BMI, body mass index; RBC, red blood cell.

Table 2 *ZFP57* methylation in maternal blood pre- and post-intervention

DNA methylation levels of *ZFP57* DMR in maternal blood samples at GW14 and GW36. GW, gestational week; SD, standard deviation

Sample Group	Gestational Week (GW)	Mean Methylation (%)	Standard Deviation (SD)	Change in methylation (%)
Treatment (*n*=24)	GW14	57.47	15.37	*+5.51*
	GW36	62.98	14.94	
Placebo (*n*=24)	GW14	64.36	6.58	*-1.51*
	GW36	62.85	7.13	

Figure Legends

Figure 1 FASSTT study outline for samples used in this study

Eligible pregnant women (*n*=226) were randomized into two groups: Placebo (*n*=94) and Folic acid (*n*=96). Women withdrew (*n*=25) or were excluded from the intervention for the reasons indicated. A total of 119 women completed the trial. Blood samples were taken at gestation weeks (GW) 14 (pre-intervention) and 36 (post-intervention). Cord blood samples (*n*=86) were taken at birth.

Figure 2 Widespread alterations to DNA methylation levels in cord blood in response to late gestation maternal folic acid supplementation

(A) Scatterplot comparing mean methylation levels (β values: 1=100%; 0=0% methylation) at individual probes in placebo and treated groups. The 1000 top ranking sites between groups are highlighted in red: ρ = correlation value. (B) Probe methylation density plot comparing the distributions of methylation values per sample group. In the treatment group there is a decrease in the number of fully methylated sites (β >0.75). (C) Split in top 1000 ranking sites losing or gaining methylation overall. Also shown are numbers of sites showing changes greater than 5% or 10%. (D) Top 5 differentially methylated sites overall, sorted by combined rank, the value being computed as the maximum (i.e. worst) value amongst the mean quotient log, mean difference in methylation and *p* value (*P*). No., number; Chr, chromosome; Position, coordinates in *hg19* human genome release; CG probe, identity number of the probe on the EPIC array; % change, difference in mean β value expressed as %; Gene, nearest gene; P, probability (uncorrected); Rank, RnBeads computed ranking value (lowest being best).

Figure 3 Top ranking promoter regions included imprint regulator gene *ZFP57*

((A) Top 5 differentially methylated regions (DMR) at promoters, sorted by combined RnBeads rank (smallest to largest) as for Fig.2D above, except combining values across all the CG sites in the DMR as detailed in Methods. Abbreviations as above except # probes, number

of probes on EPIC array included in DMR. (B) Top: Genome browser tracks showing the region around the DMR upstream of *ZFP57*, genomic coordinates in *hg19* human genome release and scale as shown. EPIC array probes showing differential methylation (blue, gain; red, loss) are indicated, with size indicating magnitude of change. The start of the *ZFP57* gene and the position of the pyrosequencing assay (Pyro) are also shown. Δβ, mean difference in β value between placebo and FA-treated groups; maximum gain and loss also shown (+0.09 β = 9% methylation). Bottom: Loess plot of β values across the region, with CpG identification numbers from array below; those forming the DMR defined by RnBeads are indicated, as well as sites analyzed by pyroassay. Each dot represents β value in an individual sample, with lines representing smoothed averages; color code is indicated at left. (C) Results of pyroassay covering the 6 sites indicated in B. Sample groups: cord blood DNA from placebo (n=45) and FA-treated (n=41). Mean, average of the individual means in that group; Max., largest of the mean methylation values in that group; Min, lowest mean in group; SD, standard deviation for the means; Change, difference in % methylation seen between groups; P, probability (Student's t-Test).

Figure 4 ZFP57 upstream region is a methylation-dependent regulator of transcription at this locus

(A) Schematic as in Fig.3 above but showing difference in methylation (Δβ) between HCT116 WT cells vs HCT116 DKO cells. The intron/exon structure and positions of the forward (FW) and reverse (RV) primers for RT-(q)PCR on the ZFP57 gene are also shown. (B) Methylation levels at individual CpG covered by the pyrosequencing assay in WT (HCT116) and knockout (DKO) cells. Values are shown as mean +/- SD for each site: **P<0.05; **P<0.01; ***P<0.001*. (C) RT-PCR showing up-regulation using the primers indicated in A, key as above. CTRL, positive control (human reference total RNA); NTC, negative control (no template control); 100bp, size standards ladder; *ACTB*, β-actin loading control. (D)

Confirmation of upregulation by RT-qPCR using the same primers, values normalized to *HPRT*; FC, fold change. (E) Methylation levels using pyroassay as in B but in 5-aza-dC treated SH-SY5Y cells (5-aza-dC), as compared to untreated (UT). (F) RT-PCR for 5-aza-dC treated cells from E. (G) RT-qPCR confirmation of *ZFP57* upregulation in 5-aza-dC-treated SH-SY5Y cells.

Figure 5 Comparison of AFAST and EpiFASSTT data for the DMR

(A)Effect size (Cohen's D) at each CpG in the *ZFP57* DMR was calculated by comparing high dose and placebo from the AFAST study and plotted against the locus (top track). A similar analysis was done for the EpiFASSTT data (bottom track). Maxima are indicated at right, scale bar and location at top: note- no other CpG outside the DMR are shown in this analysis. (B)The two sets of values from A are plotted on the same scale to give an indication of comparability.

Supplementary Table 1 Pyrosequencing and transcriptional primer sets used in this study

Pyroassay primers are given as bisulfite converted sequence. The same primers were used for both RT-PCR and RT-qPCR.

Supplementary Figure 1 Correlation between folate levels in cord blood and mother

Scatterplot shows log-converted red blood cell folate (RCF) levels in nanomoles per liter (nmol/l) at gestational week 36 (GW36) for mothers (post-intervention) and matched cord blood. The line of best fit shows significant correlation between mothers and offspring ($r = 0.619$; $P = <0.001$).

Supplementary Figure 2 QQ plot shows no evidence of population substructure effects

The observed Chi-squared (χ^2) values (open circles), plotted as –log10 of the *p* value for both sample groups, fit tightly to the expected χ^2 values (red line), indicating little evidence of association due to population substructure effects and that the top hits which deviate from the line (right-hand side) are likely to represent true differences due to loci with large effects.

Supplementary Figure 3 Median methylation levels at imprint control regions

Methylation levels at imprint control regions (ICR) were assessed by matching EPIC array probes to the imprint germline DMR intervals defined by [48] (A) or [49] (B) then taking the average (median) across each. The identities of each ICR and number of probes are indicated below. Boxes show the median and interquartile range for the individual averages from each group (Placebo n=45, Treated n=41), whiskers represent the range of values, dots indicate outliers.

Figure 1

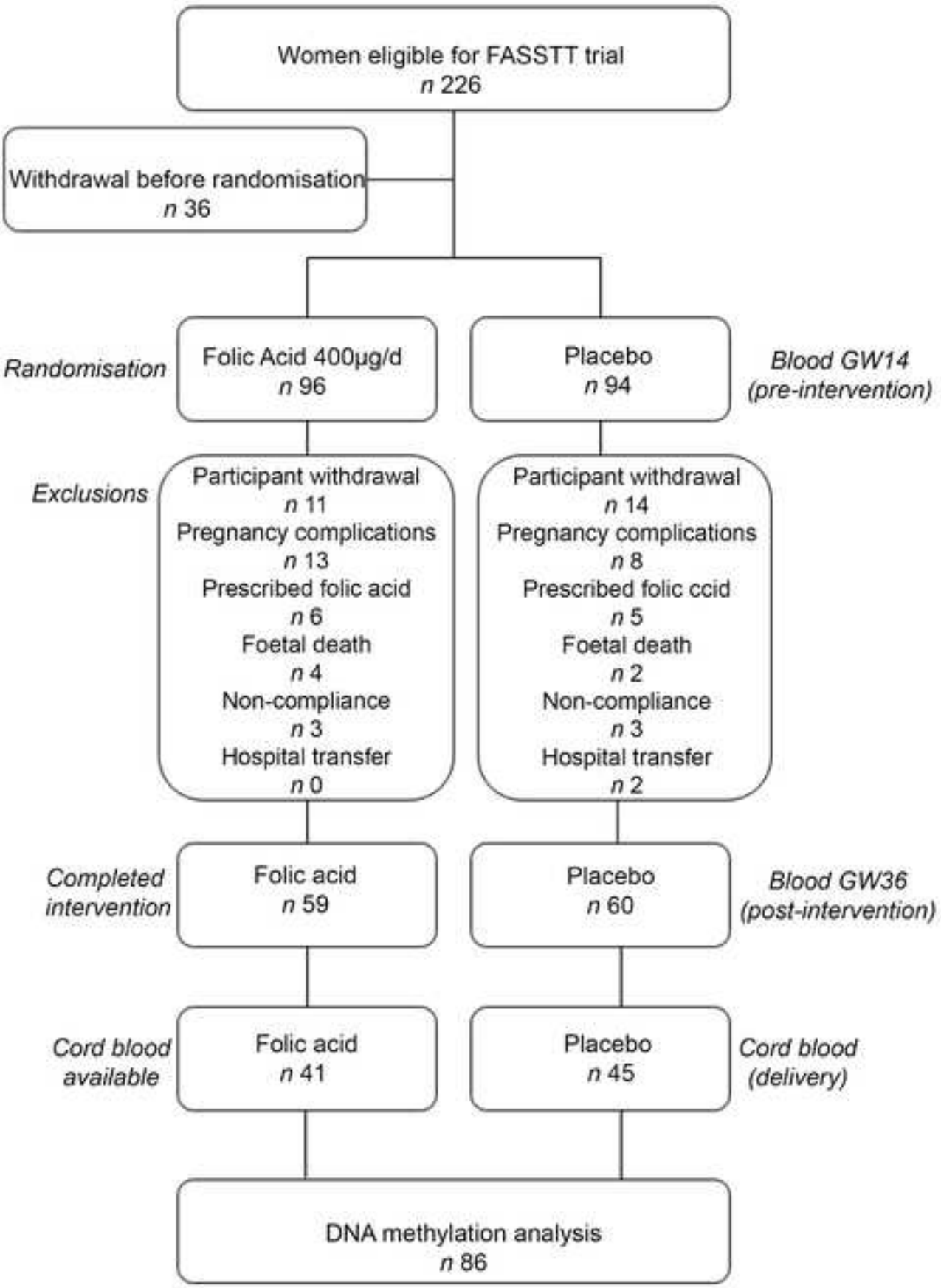
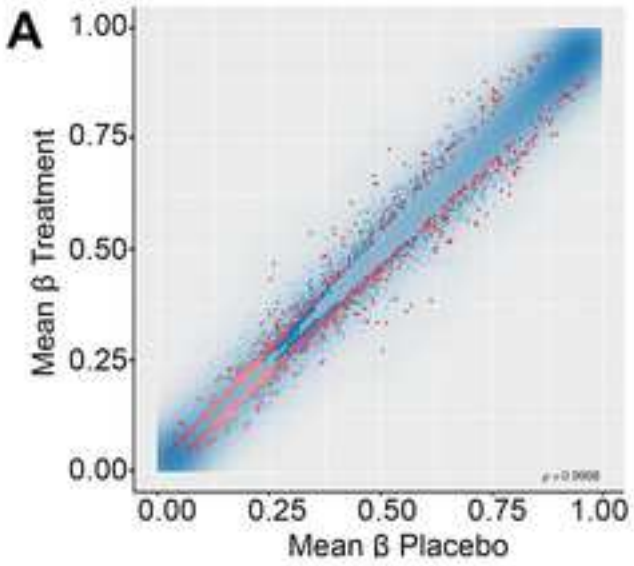
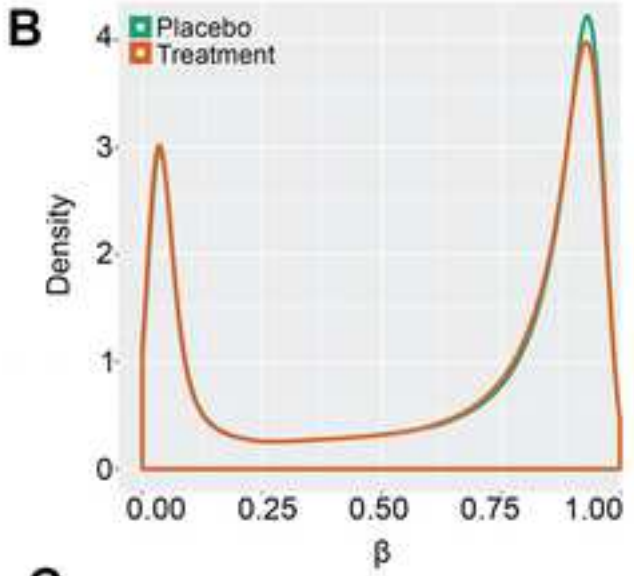

Women eligible for FASSTT trial
n 226
Withdrawal before randomisation
n 36
Randomisation
Folic Acid 400µg/d
n 96
Placebo
n 94
Blood GW14
(pre-intervention)
Exclusions
Participant withdrawal
n 11
Pregnancy complications
n 13
Prescribed folic acid
n 6
Foetal death
n 4
Non-compliance
n 3
Hospital transfer
n 0
Participant withdrawal
n 14
Pregnancy complications
n 8
Prescribed folic ccid
n 5
Foetal death
n 2
Non-compliance
n 3
Hospital transfer
n 2
Completed
intervention
Folic acid
n 59
Placebo
n 60
Blood GW36
(post-intervention)
Cord blood
available
Folic acid
n 41
Placebo
n 45
Cord blood
(delivery)
DNA methylation analysis
n 86

Figure 2





C

Top 1000 ranking sites	Losing methylation	Gaining methylation
Overall	658	342
>10%	44	32
>5%	193	109

D

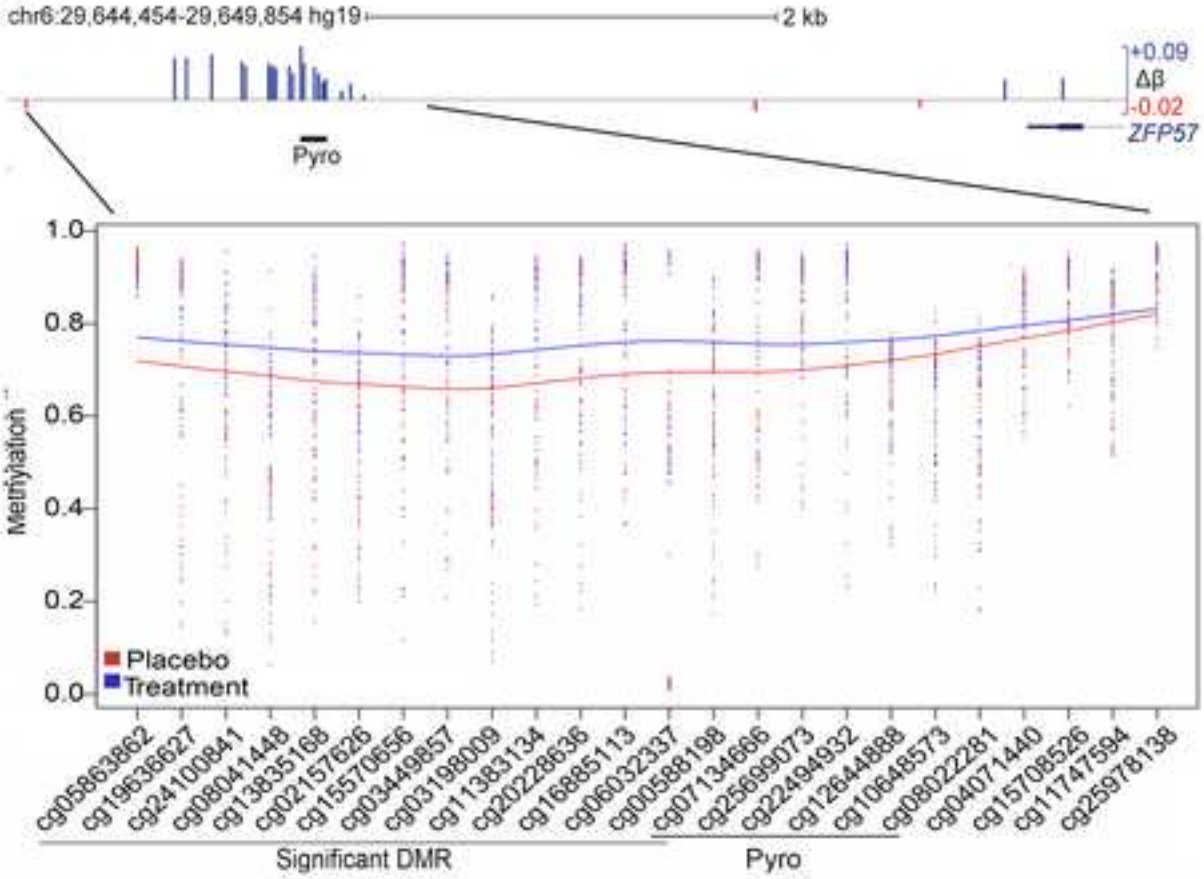
No.	Chr.	Position	CG probe	% change	Gene	P	Rank
1	13	113539522	cg21463262	-23.12	*ATP11A*	<0.001	150
2	7	766104	cg06242242	-8.98	*PRKAR1B*	<0.001	275
3	17	807555	cg08104960	-8.94	*NXN*	<0.001	280
4	7	766119	cg05729249	-7.51	*PRKAR1B*	<0.001	496
5	7	77740624	cg00401665	-9.85	*MAGI2*	<0.001	499

Figure 3

A

No.	Chr.	Position	# probes	% change in methylation	Gene	*P*	RnBeads computed rank
1	chr6	29648388-29650387	15	+6.23	*ZFP57*	0.012	79
2	chr21	15352259-15354258	10	+3.32	*ANKRD20A11P*	0.007	118
3	chr16	55866750-55868749	8	-3.86	*CES1*	0.011	154
4	chr6	290130-292129	6	+6.24	*DUSP22*	0.027	192
5	chr17	6558329-6560328	7	-7.22	*MIR4520A;MIR4520B*	0.029	205

B



C

Gene	Sample Group	Mean	Max.	Min.	SD	Change in methylation (%)	*P*
ZFP57	Placebo	61.15	84.96	24.37	16.67	*+5.44*	0.172
	Treatment	66.60	94.97	23.19	14.67		

Figure 4

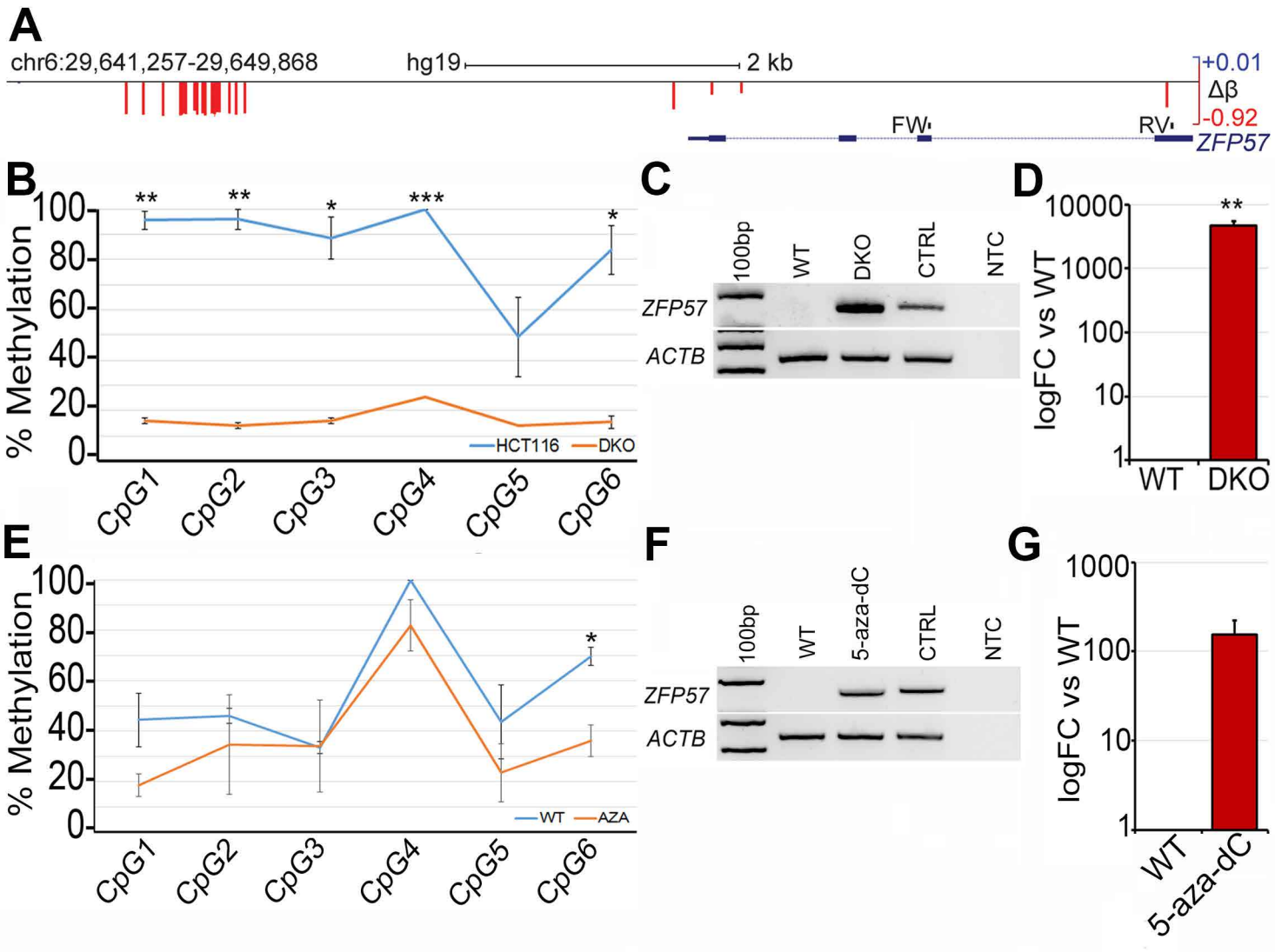
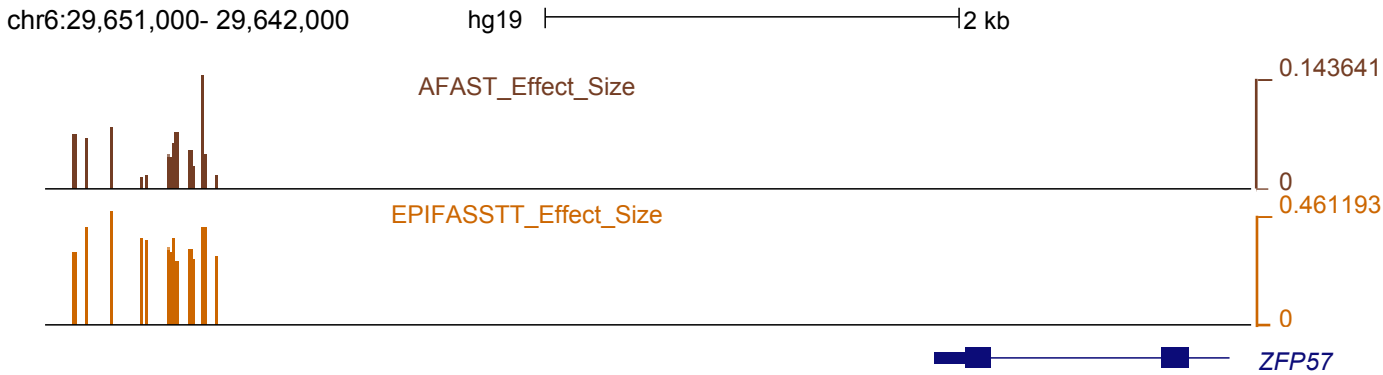
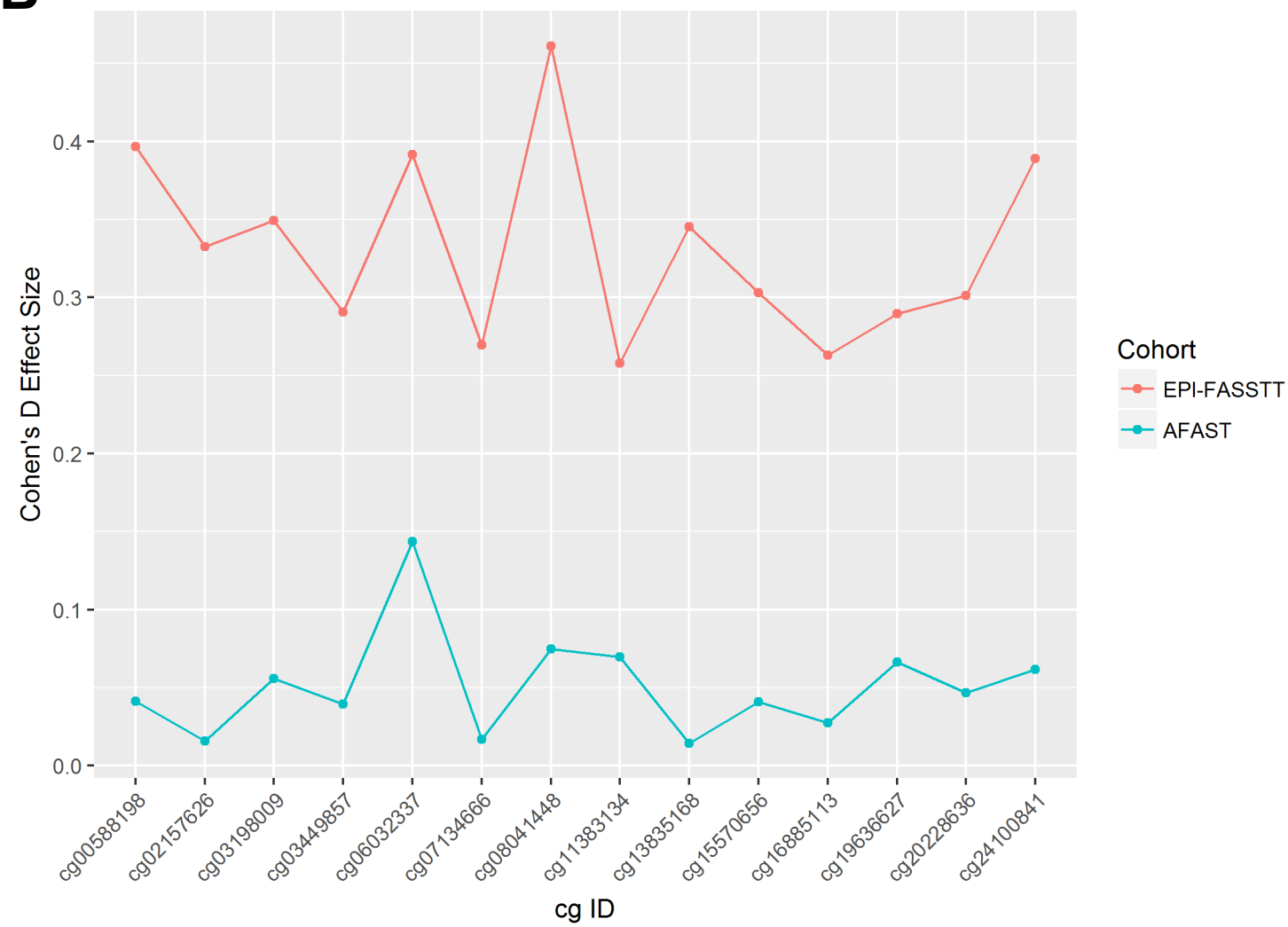


Figure 5

A



B



SuppFig2

SuppFig3

SuppTable1

SuppFig1

Fig4C ZFP57 raw image

Click here to access/download;Raw Image;Fig4C ZFP57.tif

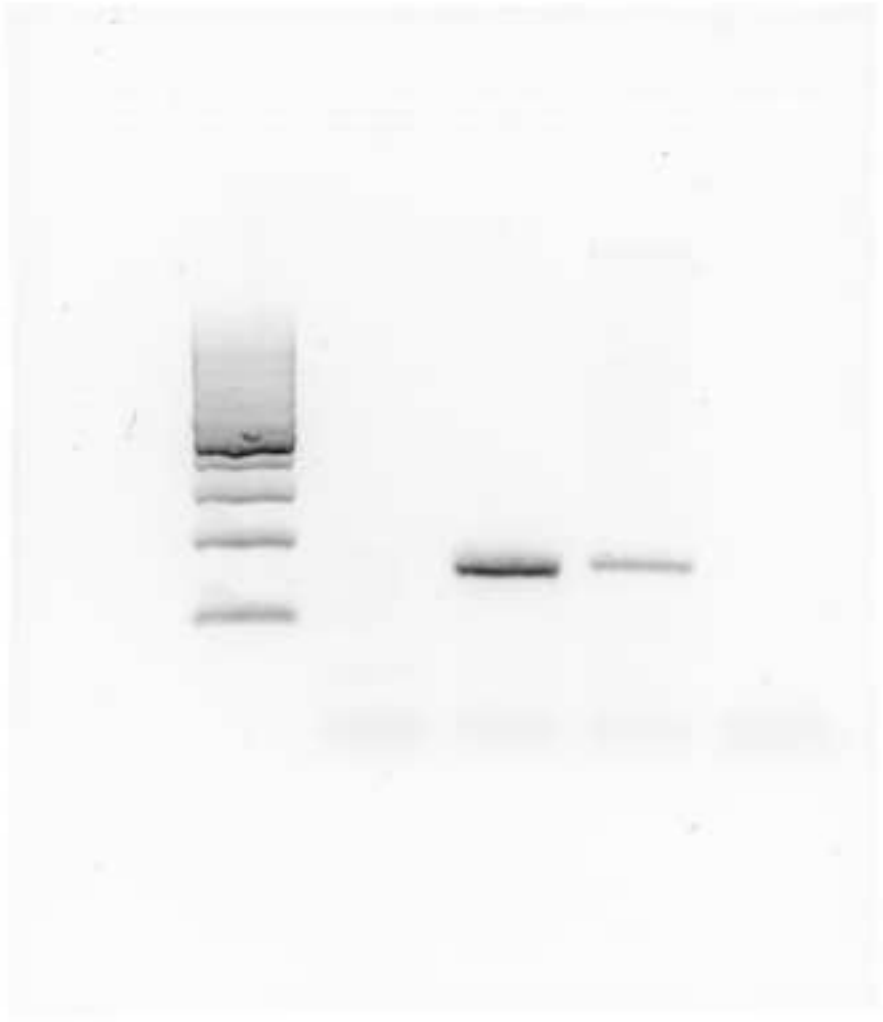


Fig4C ACTB raw image

Click here to access/download;Raw Image;Fig4C ACTB.tif

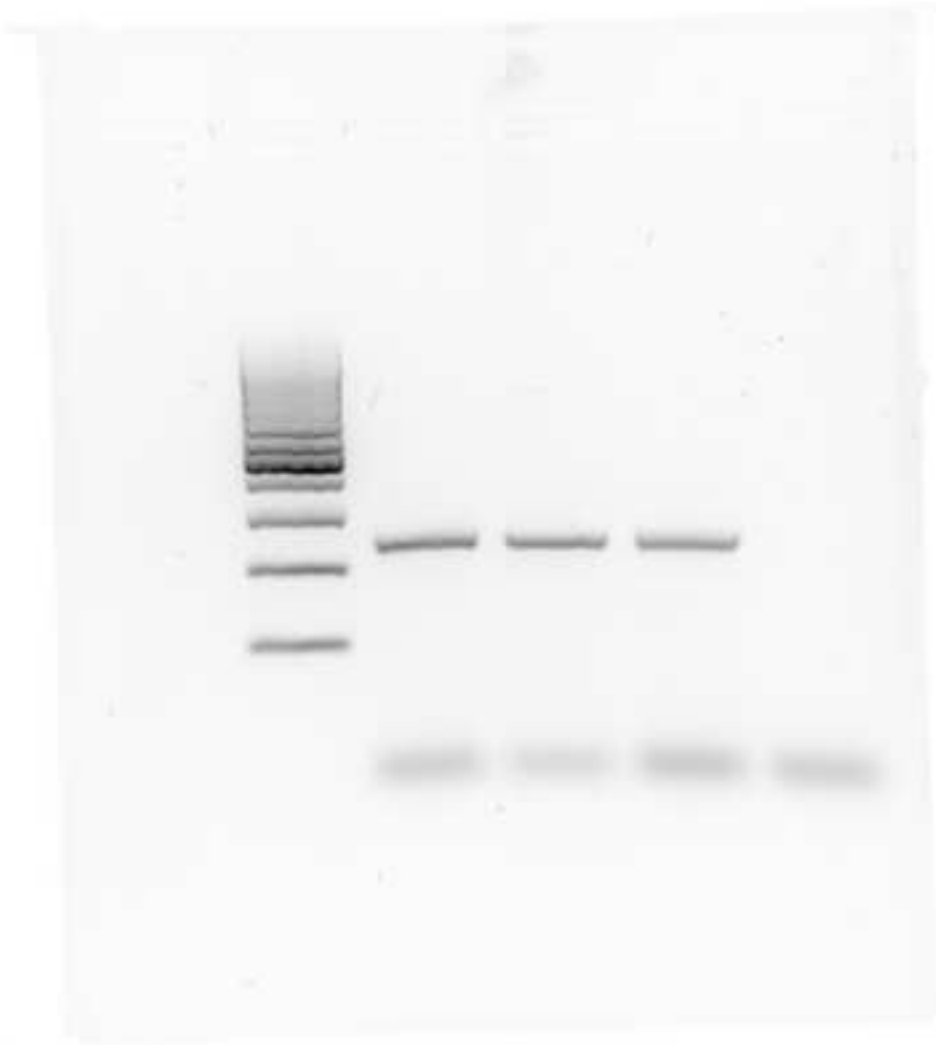


Fig4F ZFP57 raw image

Click here to access/download;Raw Image;Fig4F ZFP57.tif

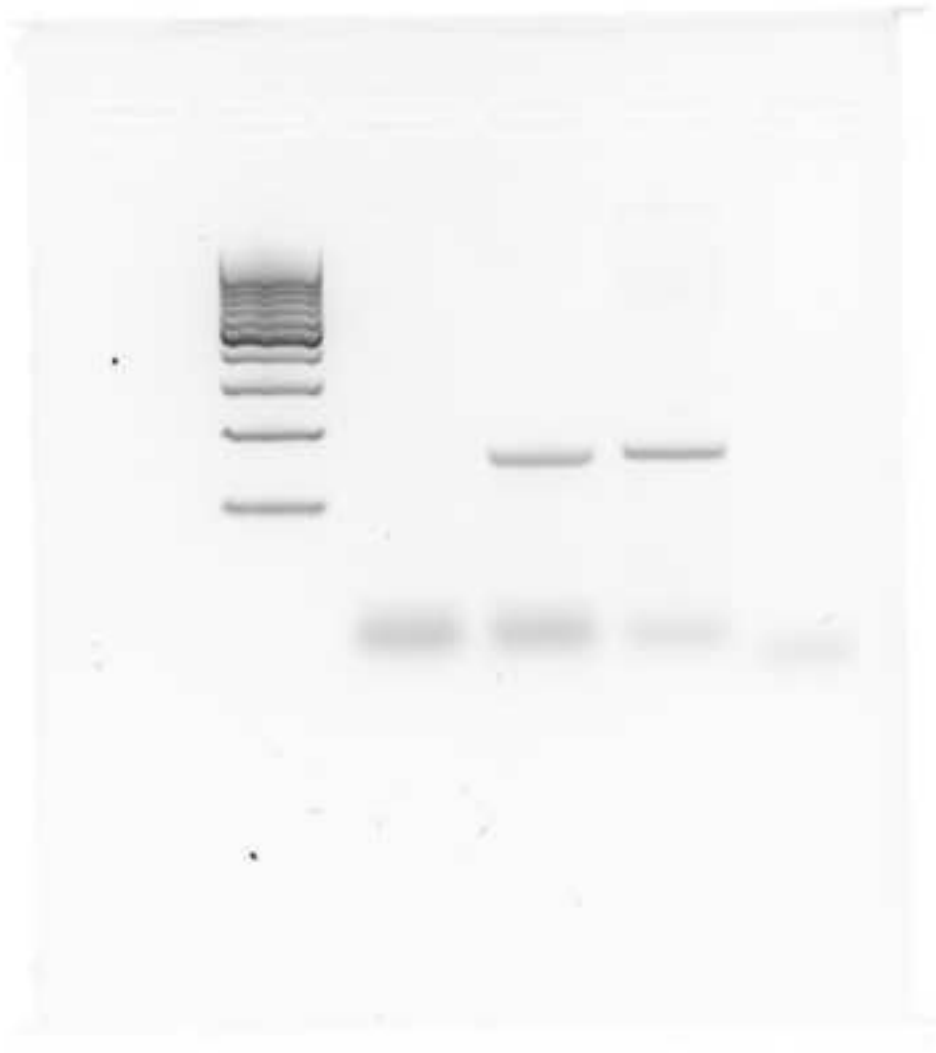


Fig4F ACTB raw image

Click here to access/download;Raw Image;Fig4F ACTB.tif