UCSF

UC San Francisco Previously Published Works

Title

Periconceptional folate consumption is associated with neonatal DNA methylation modifications in neural crest regulatory and cancer development genes

Permalink

https://escholarship.org/uc/item/0ck1g8gs

Journal

Epigenetics, 10(12)

ISSN

1559-2294

Authors

Gonseth, Semira

Roy, Ritu

Houseman, E Andres

et al.

Publication Date

2015-12-02

DOI

10.1080/15592294.2015.1117889

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Epigenetics 10:12, 1166–1176; December 2015; © 2015 Taylor & Francis Group, LLC

Periconceptional folate consumption is associated with neonatal DNA methylation modifications in neural crest regulatory and cancer development genes

Semira Gonseth, MD[1,*], Ritu Roy, MA[2], E Andres Houseman, ScD[3], Adam J de Smith, PhD[1], Mi Zhou, MD[1], Seung-Tae Lee, MD, PhD[4], Sébastien Nusslé, PhD[5], Amanda W Singer, PhD[6], Margaret R Wrensch, PhD[7], Catherine Metayer, MD, PhD[6], and Joseph L Wiemels, PhD[1]

[1]Department of Epidemiology and Biostatistics; Laboratory for Molecular Epidemiology; University of California; San Francisco, CA, USA; [2]Computational Biology Core; HDF Comprehensive Cancer Center; University of California; San Francisco, CA, USA; [3]College of Public Health and Human Sciences; Oregon State University; Corvallis, OR, USA; [4]Department of Laboratory Medicine; Yonsei University College of Medicine; Seoul, Republic of Korea; [5]Department of Environmental Science; Policy & Management; University of California; Berkeley, CA, USA; [6]School of Public Health; University of California; Berkeley, CA, USA; [7]Department of Neurological Surgery; University of California; San Francisco, CA, USA

Keywords: cancer prevention, developmental origin of health and disease, DNA methylation, epigenetics, folate, neural tube defects

Abbreviations: β, regression coefficient; CCLS, California Childhood Leukemia Study; DFE, dietary folate equivalents; DMR, differentially methylated region; GEO, gene expression omnibus; IQR, inter-quartile range; LEF1, lymphoid enhancer-binding factor 1; MAZ, *MYC*-associated zinc finger protein; NTD, neural tube defect; TCF3, pre-B cell transcription factor 3

Folate deficiency during early embryonic development constitutes a risk factor for neural tube defects and potentially for childhood leukemia via unknown mechanisms. We tested whether folate consumption during the 12 months prior to conception induced DNA methylation modifications at birth in healthy neonates with a genome-wide and agnostic approach. We hypothesized that DNA methylation in genes involved in neural tube development and/or cancer susceptibility would be affected by folate exposure. We retrospectively assessed folate exposure at the time of conception by food-frequency questionnaires administered to the mothers of 343 healthy newborns. We measured genome-wide DNA methylation from neonatal blood spots. We implemented a method based on bootstrap resampling to decrease false-positive findings. Folate was inversely associated with DNA methylation throughout the genome. Among the top folate-associated genes that were replicated in an independent Gambian study were *TFAP2A*, a gene critical for neural crest development, *STX11*, a gene implicated in acute myeloid leukemia, and *CYS1*, a candidate gene for cystic kidney disease. Reduced periconceptional folate intake was associated with increased methylation and, in turn, decreased gene expression at these 3 loci. The top folate-sensitive genes defined by their associated CpG sites were enriched for numerous transcription factors by Gene Set Enrichment Analysis, including those implicated in cancer development (e.g., *MYC*-associated zinc finger protein). The influence of estimated periconceptional folate intake on neonatal DNA methylation levels provides potential mechanistic insights into the role of this vitamin in the development of neural tube defects and childhood cancers.

Introduction

Folate vitamin B9 is an essential nutrient that naturally occurs in many foods (including fresh fruits and vegetables, eggs, and almonds). Other common sources include fortified grains and cereals and vitamin supplementation. Folate intake is critical during early embryonic development; deficiency can cause neural tube defects (NTDs, a group of birth defects deriving from neural crest development aberrations)[1] and other birth defects (cardiovascular, urinary tract, and oral defects).[2,3] Prenatal folate supplementation was shown to have a protective effect for childhood acute lymphoblastic leukemia according to a recent analysis.[4]

The mechanisms by which folate influences disease risk are currently unknown, but emphasis has been placed on DNA methylation.[1,5] Previous candidate-gene studies have demonstrated that folate deficiency during fetal life impacts DNA methylation at birth in animals[6-21] and in humans.[22-32] As folate is a necessary cofactor for shuttling DNA methyl groups via the *1*-carbon metabolic cycle,[33] it could hypothetically act as a

*Correspondence to: Dr. Semira Gonseth; Email: semira.gonseth@gmail.com
Submitted: 08/26/2015; Revised: 10/07/2015; Accepted: 11/03/2015
http://dx.doi.org/10.1080/15592294.2015.1117889

limiting factor during embryonic development. However, this theoretical concept has been challenged by inconsistencies in the directions of association between folate exposure during fetal life and global DNA methylation at birth in folate-replete populations (assessed through LINE-1 DNA methylation).[24,31,34]

The most critical period for the establishment of DNA methylation patterns is shortly after fertilization between the morula and the blastocyst stages (i.e., at days 4 to 5 after fertilization), when the embryo's DNA methylation is almost completely reprogrammed.[35] Once DNA methylation is established, theoretically only slight modifications will occur.[36] Thus, if folate plays a critical role in DNA methylation, the main effects are likely to occur before the reprogramming of DNA methylation around the time of conception. Using retrospectively estimated periconceptional folate intake from food frequency questionnaires and a genome-wide methylation array applied to newborn blood spots, we aimed at agnostic discovery of associations between folate exposure and methylation at birth. Given the epidemiological evidence of associations between folate exposure at conception and NTDs and childhood leukemias, we hypothesized that DNA methylation in genes involved in neural tube development and/or cancer susceptibility would be associated with folate exposure, thus providing potential mechanistic insights.

Results

A total of 343 healthy subjects were included in the analysis (**Table 1**) and were run as 2 separate batches (set 1 and set 2). As shown, the mean folate was significantly higher in set 1 and the estimated proportions of white blood cells slightly but significantly varied between sets. The two sets were otherwise similar.

Defining folate-associated CpGs

The initial number of CpGs assessed was 485, 512 in both sets. We excluded 692 CpG sites in set 1 and 447 CpG sites in set 2 because they did not meet quality control criteria (see Methods). In accordance with previous recommendations,[37,38] we excluded all CpGs that were in or close to a SNP, on a sex chromosome and polymorphic CpG sites or CpGs resulting from bad probes (n=165,556). After these exclusions, 319,264 CpGs were included in the analysis. Five samples in set 1 and 4 samples in set 2 were excluded due to aberrant distributions of their raw $\beta-$values, suggesting a bisulfite treatment error. One sample in set 1 and 2 in set 2 had mismatches between reported and predicted gender and were therefore also excluded.

Folate exposure prior to conception appeared to have genome-wide effects on DNA methylation. There were more significantly associated CpG sites than statistically expected by chance, as illustrated by the left deviation of the histograms of *P*-values and the upper deviations of the qq-plots (**Fig. S1 and Fig. 1**). The qq-plots presented a general inflation (inflation factor based on median chi-squared: 1.295 for set 1 and 1.071 for set 2). As illustrated in **Table 2**, there were many more negative than positive regression coefficients for the association of DNA methylation with folate in both sets among concordant loci (McNemar test $P<0.001$). This suggests an inverse association of folate consumption and methylation. Mean β-values for all the analyzed CpG sites stratified by categories of folate consumption are given in **Table S1**.

Second, we implemented a method for the selection of CpG sites within an epigenome-wide association study that decreases the number of false-positive findings. The method is based on bootstrap resampling in 2 different sets. We selected the sign concordant and most significant CpG sites in both sets with an interquartile range of the distribution of *P*-values from bootstrap

Table 1. Principal characteristics of healthy control participants from the California Childhood Leukemia Study.

Variables	Set 1 (n=167)	Set 2 (n=176)	*P*
Gestational age [mean (±S.D.)](weeks)	38.9 (±2 0.1)	39.0 (±2 0.6)	0.73†
White/Caucasian (%)	62.9	48.3	0.22∫
African American (%)	5.4	2.8	
Native American (%)	0.6	0.6	
Asian or Pacific Islander (%)	9.6	8.5	
Mixed or others (%)	21.6	39.8	
Male (%)	61.7	58	0.48†
Total dietary folate equivalents* [mean (±S .D.)] (μg/d)	732.8 (± 405.4)	503.8 (± 281.4)	<0.001†
Mean (sd) proportions of cells:			
T cells CD8+	0.097 (±0 0.03)	0.073 (±0 0.023)	<0.001†
T cells CD4+	0.208 (±0 0.059)	0.193 (±0 0.054)	<0.05†
B cells	0.104 (±0 0.02)	0.091 (±0 0.047)	<0.001†
Monocytes	0.134 (±0 0.029)	0.126 (±0 0.029)	<0.01†
Natural killer cells	0.038 (±0 0.029)	0.031 (±0 0.027)	<0.05†
Granulocytes *(not included in the model)*	0.513 (±0 0.096)	0.514 (±0 0.096)	0.97†

The CCLS is a case-control study that has recruited children with a diagnosis of acute childhood lymphoblastic leukemia and matched controls since 1996. A subset of participants with complete smoking information were selected from the whole study sample and randomly assigned to set 1 or set 2.
*Assessed by maternal food frequency questionnaire for diet one year prior to birth;
∫assessed by Chi-square test;
†assessed by Student t test.

<0.05 in both sets (**Table 3**). This selection process yielded 4 CpG sites: cg22664307 (chr6q24.2), which is located near the promoter of *STX11*; cg21039708 (chr14q22.3), within the promoter of *OTX2*; cg15219145 (chr6p24.3), within the promoter of *TFAP2A*; and cg13499966 (chr2p25.1), which is located near the promoter of *CYS1* (aka *cystin-1*). Because folate intake was a continuous measure with a large range among participants, regression coefficients appear small, though they still represent significant modifications. For instance, the difference in DNA methylation at cg15219145 for a hypothetical "average participant" in set 1 between the minimal value of folate (118.8 μg/d) and the maximal one (2056 μg/d)—all other parameters being equal—is 0.055 and 0.041 [β-values], respectively, i.e., a 25% difference. Multiple-testing correction was performed by a permutations-based *P*-values test,[39,40] and the top 4 CpG sites survived this approach with a corrected *P*-value <0.05 in both sets (**Table 3**). The top 4 CpG sites and their DNA methylation levels as well as the DNA methylation levels of the CpG sites that were allocated to the same genes were globally concordant between the sets (as illustrated in **Figs. S2–5**), suggesting regional replication beyond the single CpG site identified in our bootstrap analyses.

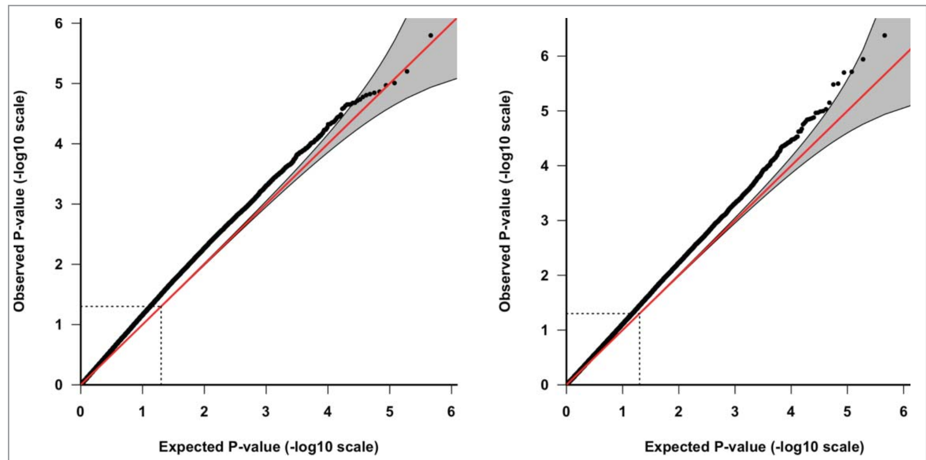

Figure 1. QQ-plots of observed vs. expected *P*-values in set 1 and set 2, respectively. In black: initial locus-by-locus model analysis; in gray: 95% confidence interval. Note the global slight left deviation of the initial analysis from the null effect line (represented in red) suggesting that we observe a larger number of low *P*-values than what one can expect to find by chance. The deviation is more pronounced for set 2 than for set 1.

Impact of DNA methylation on gene function

Gene expression of the top 4 genes was significantly and negatively associated with methylation at the 4 CpG sites [*P*<0.001 for the *STX11* transcript; *P*<0.001 for 2 *OTX2* transcripts (ILMN_1688075 and ILMN_1722959) and *P*<0.01 for a third transcript (ILMN_2323465); *P*<0.001 for 2 *TFAP2A* transcripts (ILMN_1733135 and ILMN_1765574) and *P*<0.05 for a third *TFAP2A* transcript (ILMN_2374115), the fourth one (ILMN_2282477) was not significant; and *P*<0.05 for the only *CYS1* transcript, as illustrated in **Fig. 2**].

Gene Set Enrichment Analysis

Gene set enrichment analysis including the genes listed in **Table S2** (number of CpG sites=429, number of corresponding associated genes=365, selected with a median *P-value* from bootstrap <0.1 and a significant multiple-testing corrected *P*-value <0.05), revealed significant enrichment of multiple transcription factor pathways, in addition to specific cancer and erythrocyte development pathways. Eighty genes had a transcription start motif that matches annotation for the pre-B cell transcription factor 3 (TCF3, E2A immunoglobulin enhancer binding factors E12/E47, q-value <10 −<τβρ>−23 </tbr>). Seventy-one genes had a transcription start motif for the transcription factor *MYC*-associated zinc finger protein (MAZ, purine-binding transcription factor, q-value 10−<τβρ>−19 </tbr>). Furthermore, 78 genes had a transcription start motif for the Sp1 transcription factor (SP1, q-value 10−<τβρ>−17 </tbr>), and 54 genes for the lymphoid enhancer-binding factor 1 (LEF1, q-value <10 −<τβρ>−11 </tbr>). Finally, 53 genes were implicated in erythroid progenitor regulation (q-value <10 −<τβρ>−11 </tbr>).

Replication of Candidate Folate-Associated CpG Sites

Initially, we focused specifically on 4 "metastable epialleles" that were recently associated with folate intake in humans at conception (see details in **Table S3**).[23] While not reaching

Table 2. Directions of the association between folate exposure around time of conception and DNA methylation in healthy control participants from the California Childhood Leukemia Study.

All CpG sites			CpG sites with *P*<0.05 (initial analysis)		
Set1/ Set2	**Positive β**	**Negative β**	**Set1/ Set2**	**Positive β**	**Negative β**
Positive β	54, 153	64, 103	Positive β	118	151
Negative β	89, 061	111, 946	Negative β	425	862

We found that the association was significantly more often negative (i.e., more folate was associated with less methylation), among all the CpG sites and also among statistically significant CpG sites (McNemar test of repartition *P*<0.001 in both tables). Participants were randomly allocated in 2 datasets [n(set 1)=167, n(set 2)=176]. β=regression coefficient.

Table 3. Association between folate exposure around time of conception and DNA methylation in healthy control participants from the California Childhood Leukemia Study.

	Set 1					Set 2								
	Initial analysis		1, 000 bootstrap			Initial analysis		1, 000 bootstrap			1, 000 permutations			
CpG ID	Regression coef.	*P*	Median *P*	IQR		Regression coef.	*P*	Median *P*	IQR		corr. *P*-values	Gene	Location	CpG location
cg22664307	-0.00013	0.0005	0.0072	0.002	0.021	-0.00026	0.0010	0.0113	0.003	0.034	0.033	*STX11*	6q24.2	89bp upstream of promoter
cg21039708	-0.00021	0.0013	0.0118	0.004	0.038	-0.00034	0.0001	0.0019	0.000	0.008	0.023	*OTX2*	14q22.3	within promoter
cg15219145	-0.00016	0.0017	0.0148	0.004	0.043	-0.00023	0.0016	0.0138	0.004	0.041	0.025	*TFAP2A*	6p24.3	within promoter
cg13499966	-0.00012	0.0024	0.0167	0.005	0.049	-0.00021	0.0004	0.0062	0.002	0.020	0.016	*CYS1*	2p25.1	369bp upstream of promoter

The four top CpG sites from both sets 1 and 2 are displayed: Median *P*-values and inter-quartile ranges (IQR) were derived from 1, 000 bootstrap resampling. Only the CpG sites with an IQR of *P-values* $P < 0.05$ from bootstrapping in both sets were selected. Note that all 4 CpG sites are located in or near the genes promoters, and that they all are negatively associated with DNA methylation. Participants were randomly allocated in 2 data sets in order to replicate the findings [n(set 1)=167, n(set 2)=176]. Permutation tests *P*-values are derived from 1, 000 permutations when combining set 1 and set 2 together in order to account for multiple testing.

statistical significance individually in our study, 87.5% of the CpG sites in the 4 epialleles had the same negative direction of association with folate as demonstrated by Dominguez-Salas et al. (2014) (**Fig. S7**, $P<0.05$). Collectively, the median methylation Z-score of these genes was negatively associated with folate intake when dichotomized into high vs. low folate groups (± 515 μg/d, regression coefficient= -0.02, $P<0.05$, see **Fig. S8**). We subsequently investigated whether 7 differentially methylated regions (DMRs) previously found to be associated with folate levels in cord blood, i.e., at the end of pregnancy,[34] were also associated with periconceptional folate levels in our study. DNA methylation information was available for 62 CpG sites located in those DMRs (see details in **Table S4**). Although the associations between DNA methylation and folate exposure at conception did not reach statistical significance apart from 2 CpG sites, we observed the same direction of association as Amarasekera et al. for one DMR (*C21orf56*); we observed an inverse association for 5 DMRs (*ACADM*, *FZD7*, *LASP1*, *WNT9A*, *ZFP57*); and the direction of association was equivocal for the last DMR (*LY6E*, see **Fig. S9**).

Subgroup Analyses

In order to test the hypothesis that folate has a limiting-factor effect on DNA methylation in folate-deficient but not in folate-replete populations, we performed subgroup analyses by 3 categories of folate consumption. We observe that the category with the lowest consumption (<200 μg/d) has more positive than negative associations, and this seems to reverse in categories with more folate consumption.(see **Fig. 3**)

Validation of the Top Hits in an Independent Data Set

We used publicly available data of the "The Gambia / Aflatoxin / Seasonality study" to validate our top hits.[41,42] Briefly, this study population represents a *natural experiment*: Gambian mothers are exposed to high vs. low levels of folate and other methyl-donor groups depending on the season of the year, either dry or rainy.[43] Therefore, babies that were conceived during one or the other season were naturally exposed to high or low levels of folate at the time of conception. Babies' DNA methylation was measured at 3–6 months old in peripheral blood cells. All of our 4 top hits were associated in the same direction with folate exposure as in our initial analysis, i.e., with negative associations between folate exposure and DNA methylation. For three of the 4 top hits, this association was significant. (<4.6*P*x10−<τβρ>−5 </tbr> for cg15219145 in the promoter of *TFAP2A*, and $P<0.05$ for cg13499966 in *CYS1* and cg22664307 in *STX11*, **Table 4**).

Discussion

Using a position-agnostic genome-wide approach our results show, for the first time, that maternal self-reported folate exposure prior to the time of conception is associated with DNA methylation at birth in the offspring's blood in sites throughout the genome. Two of the top associated sites were located at or near the promoter regions of the neuro-facial developmental genes *TFAP2A* and *OTX2*, which are implicated in neural crest development and are responsible for neuro-facial birth defects when mutated in the germline.[44-46] This suggests a potential mechanism whereby folate may affect susceptibility to neural tube defects via its impact on the function of these genes. *OTX2* has also been involved in medulloblastoma development.[47] Moreover, among the top folate-sensitive genes were *STX11*, a gene implicated in acute myeloid leukemia[48] and *CYS1* which is a candidate for cystic kidney disease.[49] We successfully validated 3 of our 4 top hits in an independent study, and the association of the site located in the promoter of *TFAP2A* was very strong. In addition, all 4 hits had a concordant direction of association with folate exposure. The top folate-sensitive genes were largely enriched for numerous transcription factor pathways, some of which are implicated in cancer development (such as *MYC*-associated zinc finger protein[50,51] and in blood cell development and leukemia (e.g., TCF3 and LEF1).

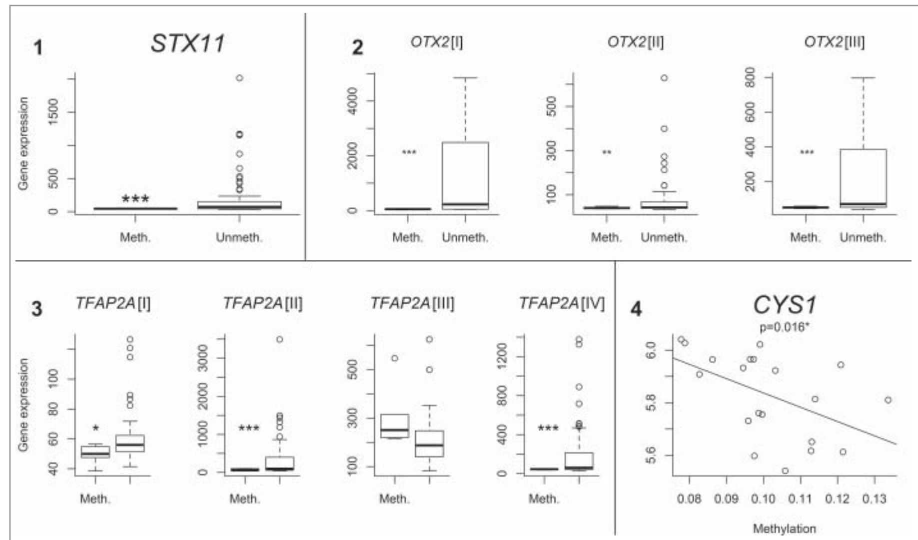

Figure 2. Panel 1: *STX11*, gene expression in relation to DNA methylation in its promoter (cg22664307) in 86 pluripotent stem cells and their derivatives (data source: GEO, GSE30654). Note that DNA methylation significantly repressed gene expression. DNA methylation has been dichotomized into "Methylated" and "Unmethylated" categories, with the cutoff being β-value ± 0.032. **Panel 2**: *OTX2* gene expression in relation to DNA methylation in its promoter (cg21039708) in 86 pluripotent stem cells and their derivatives (data source: GEO, GSE30654). Note that for all the 3 transcripts, DNA methylation significantly repressed gene expression. DNA methylation has been dichotomized into "Methylated" and "Unmethylated" categories, with the cutoff being β-value ± 0.1. **Panel 3**: *TFAP2A* gene expression in relation to DNA methylation in its promoter (cg15219145) in 86 pluripotent stem cells and their derivatives (data source: GEO, GSE30654). Note that for 3 out of the 4 transcripts, DNA methylation significantly repressed gene expression. DNA methylation has been dichotomized into "Methylated" and "Unmethylated" categories, with the cutoff being β-value ± 0.07. **Panel 4**: Correlation between DNA methylation at 369 bp ahead of the promoter (cg13499966) of *CYS1* and its expression in peripheral mononuclear blood cells of 20 healthy adult men (data source: GEO, GSE49065). Note that the association is significant and has a positive slope. The p-*value* is derived from a Spearman's test of correlation. Footnote: * for $P<0.05$, ** for $P<0.01$, and *** for $P<0.001$; *P*-values were derived from Student t tests (unless otherwise specified).

We observed that folate exposure was primarily *inversely* associated with DNA methylation. Our population was predominantly constituted of individuals that were well nourished in folate (about 2/3 of the participants met the recommendations of 600 μg/d during pregnancy).[52] The observed inverse relation is in concordance with a previous study by Boeke et al. that, using similar methods, found a significant negative association between global LINE-1 DNA methylation at birth and folate exposure at the time of conception in a folate-replete US population.[53] Two other UK studies and one Australian study, which only considered folate from supplements or total folate at the *end* of the pregnancy in their assessments of global DNA methylation, have produced inconsistent results (either a negative association or no association at all).[24,31,34] Our finding in concert with Boeke et al. suggests that folate may not produce a limiting-factor effect on DNA methylation in folate-replete populations. Our subgroup analyses suggest that folate-deprived populations have more positive associations between folate and DNA methylation, and that this reverses in folate-replete populations. This is consistent with the hypothesis that folate at conception may be acting as a limiting factor on DNA methylation but only below a certain threshold (around 200 μg/d). Although the reason why folate is negatively associated with DNA methylation in folate-replete populations is currently unknown, we can hypothesize that folate levels produce negative feedback controls on DNA methylation above a certain threshold, by yet unknown mechanisms. However, we must use some caution regarding our subgroup analyses due to the low number of participants in the lowest category of folate consumption (n=21). Moreover, in addition to the negative associations that we found between DNA methylation and folate exposure, a global inflation on the qq-plots was observed. This inflation can suggest either an unaccounted population stratification factor (such as ancestry, for instance) or a global effect of folate exposure on epigenome wide DNA methylation. We believe that the latter is more probable given that a recent before/after study on humans showed that high folate supplementation was associated with global hypomethylation in spermatozoids.[54] In this study, the association between DNA methylation and folate exposure was also negative, concordantly to our findings.

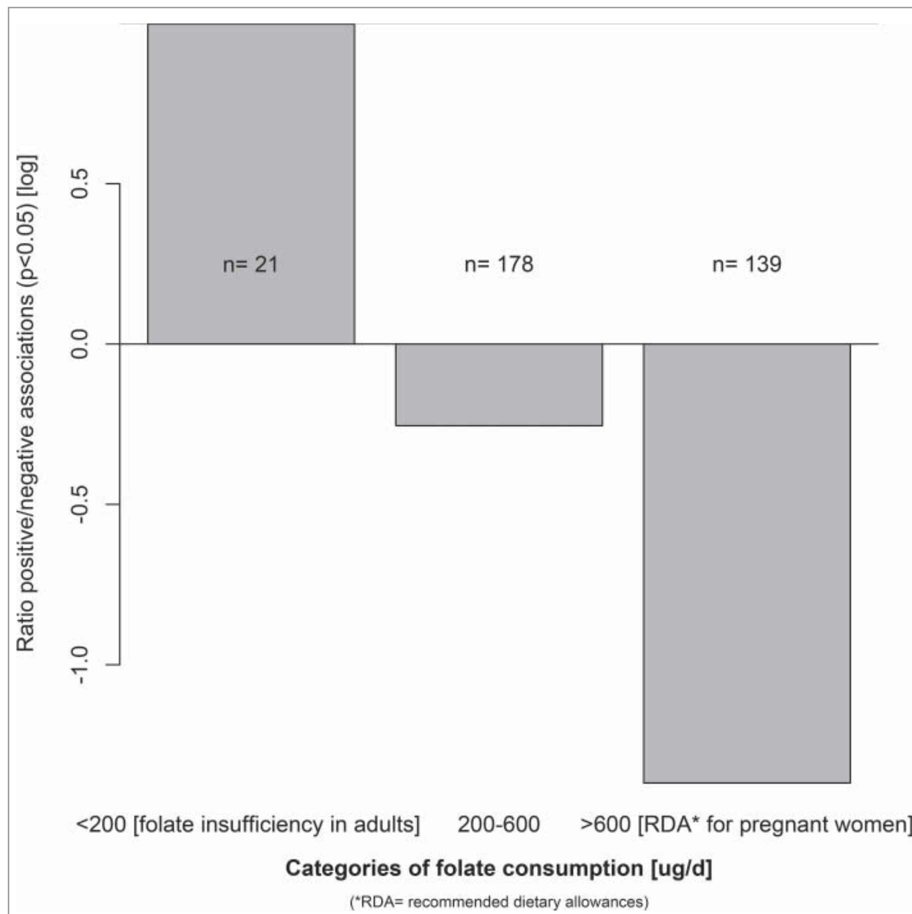

Figure 3. Ratios of positive/negative associations between DNA methylation and folate (as a linear term) by categories of maternal folate consumption. We observe that the category with the lowest consumption (< 200 μg/d) has more positive than negative associations, and this seems to reverse in categories with more folate consumption.

In one previous study, Amarasekera et al.[34] examined the association between very low or very high folate exposure at the 3rd trimester of pregnancy and genome-wide methylation at birth. Seven DMRs were found that were either hypo- or hypermethylated in the high- vs. low-folate group. Among the 7 genes associated with the DMRs, almost no significant association with folate exposure at conception was found in our analysis. Moreover, the direction of methylation was opposite of the results from Amarasekera et al. for five DMRs (*ACADM*, *LASP1*, *FZD7*, *WNT9A*, and *ZFP57*). In contrast to the current study, Amarasekera et al. estimated folate exposure at the end of pregnancy, once the establishment of the epigenetic state—which principally occurs during early embryogenesis and persists through life[35,36] has already been established, and thus constitutes a crucial difference with our study, potentially explaining the differences found between the studies. Indeed, maternal blood folate levels at the 3rd trimester of pregnancy are likely to be higher than in the preconception period, as it was recently reported in the US that 89% of pregnant women take folate supplements in their 3rd trimester vs. only 55–60% in the 1st trimester.[55] In a candidate-genes study, Dominguez-Salas et al. observed significantly less DNA methylation across 6 metastable epialleles at birth in children exposed to high vs. low folate during the periconception period.[23] We tried to reproduce these findings in our data by measuring the association between DNA methylation and folate at 8 CpGs located in 4 of the metastable epialleles investigated by Dominguez-Salas et al. About 90% of the associations were negative, and mean DNA methylation across the 8 CpGs was significantly and negatively associated with folate exposure at conception, thus supporting the findings as Dominguez-Salas et al. The concordance between our data and that from a previous analysis using folate at the time of conception, but not consistently with studies that assessed folate exposure later in pregnancy or at birth, highlights the importance of the timing of the folate exposure variable when assessing its impact on DNA methylation.

Table 4. Validation of our top hits in whole peripheral blood of 118 infants aged 3–6 months participants to the Gambia / Aflatoxin / Seasonality study. DNA methylation at the 4 top hits was regressed against the season of the babies' conception, either the dry season (i.e., high folate exposure; coded 1), or the rainy season (i.e., low folate exposure; coded 0). All models were adjusted for gender.

CpG ID	Regression coef.	*P*	Gene	Location
cg22664307	-0.0879	0.0469	*STX11*	6q24.2
cg21039708	-0.0328	0.6337	*OTX2*	14q22.3
cg15219145	-0.2009	$4.6x10^{-5}$	*TFAP2A*	6p24.3
cg13499966	-0.0896	0.0488	*CYS1*	2p25.1

Our study has several limitations. Maternal folate intake in the year prior to pregnancy was assessed by food frequency questionnaires at a median of 4 y after the birth of the child. This long time-lapse presents a risk of non-differential misclassification of folate exposure due to a lack of precision in recall, although previous work suggests that participants are accurately and/or reliably able to recall their past diets. For instance, nutrient intake derived from food-frequency questionnaires of participants about their usual diet 10–15 y ago were within 10% of their actual nutrient intake when derived from their past diet records.[56] In addition, the length of the recall period ($\leq$ 4 or $\geq$ 7 years) did not impact the repeatability of the FFQ answers related to frequency of fruit consumption, quantity of cured meat consumed, and vitamin supplement use in an study where 5 repeated questions were administered to a subset of 85 CCLS participants in the year

following their initial answer to the FFQ (unpublished results). However, such bias could have contributed to an underestimation of the strength of the associations. Cell-mixture estimation was performed using a reference-set derived from an adult population, thus potentially underestimating certain neonatal-specific cell types. However, we attempted to overcome this potential issue by adjusting for possible residual cell-mixture effect using methodology of Teschendorff et al. and Leek et al.[57,58]

The finding that 2 out of the top 4 genes in our analyses are ones implicated in neural tube defects and in other birth defects (*TFAP2A* and *OTX2*) supports the hypothesis that alterations in DNA methylation could be a mechanistic link between folate deficiency and neural tube defects.[1,5] CpG cg15219145 (chr6p24.3) is located in one of the *TFAP2A* promoters. *TFAP2A* encodes a transcription factor implicated in epidermal and neural crest development[59] and is essential for initiation of the normal developmental program in arches 2–7 and in the trunk neural crest.[44] Furthermore, mutations in *TFAP2A* cause the branchio-oculo-facial syndrome, which is a rare autosomal dominant birth defect.[45] According to our results, we can presume that a lower folate exposure at the time of conception increases methylation at the promoter of *TFAP2A*, repressing its expression. Therefore, in profound folate deficient individuals this mechanism may perhaps lead to defects in neural tube development without any changes in DNA sequence. A similar mechanism could apply to CpG cg21039708 (chr14q22.3), which is located in the promoter of *OTX2* and for which DNA methylation was strongly and negatively associated with gene expression. Mutations in this gene are associated with a micro/anophtalmia syndrome.[46] The folate-sensitivity of DNA methylation levels at the promoters of these 2 genes provides promising evidence toward the understanding of the pathogenesis of neural tube defects, and further studies should be carried out to fully elucidate the role of methylation in development of this condition.

Our findings are compatible with the hypothesis that an early onset of cancers—such as acute childhood leukemia may be triggered by low folate exposure through DNA methylation mechanisms. Indeed, *OTX2* is described as an oncogene that is involved in the development of medulloblastoma.[47] Interestingly, intracranial childhood tumors (including medulloblastomas) may be prevented by folate intake during pregnancy.[60] Among our top genes was *STX11*, which is involved in cell-cycle regulation and acute myeloid leukemia pathogenesis.[48] Our CpG hit—cg22664307 (6q24.2)—is located 89 bp upstream of the promoter of *STX11*, and methylation at this locus was associated with decreased expression. Moreover, pathway analysis of our most significant genes revealed overlap with several transcription factors, including *MAZ*, *TCF3*, *SP1*, *LEF1*, which are all implicated at various degrees in cell-cycle regulation and cancer development.[50,51,61-63] Further studies should investigate the putative mechanisms of folate on these transcription factors' activity, as they might be implicated in the pathogenesis of cancer development, especially those that develop *in utero*.

In sum, this study adds to a body of literature that increasingly suggests that folate nutrition in the periconceptional period is inversely associated with DNA methylation in folate-replete populations. We identified and validated in an independent study key genes exhibiting this profile that could potentially prove to be a mechanistic link between folate and birth defects, as well as cancer. Further research should consider the levels of other micronutrients, key intermediary metabolites in the folate pathway such as homocysteine and methionine, with regards to methylation levels and the genetics of folate metabolism.

Material and Methods

Study design and study population

Healthy control participants of the California Childhood Leukemia Study (CCLS)—a Northern California population-based case-control study that has recruited children with a diagnosis of leukemia and matched healthy controls since 1996—were included in the study. Case children were recruited in major clinical centers of California at time of the initial diagnosis. At the same time, a healthy control was randomly selected from the general population, using the statewide birth certificate files. Controls were matched by frequency, month and year of birth, gender, race, Hispanic status, and the mother's county of residence at birth. A subset of participants with complete smoking information were selected from the whole study sample and randomly assigned to set 1 or set 2. The Institutional Review Boards of all collaborating institutions approved the study protocol. Written informed consents were obtained for all participating subjects.

Dried blood spots

Archived neonatal dried blood spots (DBS) at birth were retrospectively obtained from the California Department of Public Health. Five 14-mm diameter blood spot specimens were collected from infants on S&S filter paper (also known as Guthrie cards) by heel-stick after 12 hand usually no later than 6 d of age.

Early-development folate exposure

Extensive information on maternal diet one year prior to conception was retrospectively collected through a modified version of the Block food frequency questionnaire (FFQ).[64] The questionnaire was administered to the mother at time of her child's diagnosis (or at the equivalent age for controls, median of 4 y after the child's birth). An analysis of the answers to the food frequency questionnaire allows calculation of maternal Dietary Folate Equivalents (DFEs). Briefly, the FFQ took approximately 20 minto administer. Frequency (by category from "never or less than once per month," to "2 or more per day") and quantity (i.e., an estimation of the portions sizes) of the consumption of 76 food items were assessed, as well as vitamin supplement use. We calculated total DFEs by multiplying the frequency and the reported quantity of each food by its DFE content, and summing over all foods. DFEs take into account any differences in folate bioavailability according to the source and to the chemical form (e.g., folic acid or folate), including natural, supplemental, and food fortification sources.

Epigenetic material

Approximately 300–500 ng of high molecular weight DNA was extracted from a ¼ section of a 1.5 cm^2 archived neonatal DBS (stored at -20°C from the time of birth) using Qiagen blood card extraction protocol and bisulfite treated using the EZ DNA Methylation-Direct™ Kit (Zymo). Genome-wide DNA methylation was then measured in these bisulfite converted DNA samples using Illumina© Infinium HumanMethylation450 BeadChip arrays. CpG sites with detection *P*-values > 0.01 were defined as bad CpG sites and discarded. CpG sites with >15 % of absence of information (i.e., >15% of total samples) were totally excluded from the analysis. Samples with >15 % of bad CpG sites (of the 450K loci) were also excluded from the analysis. The DNA methylation data preprocessing consisted of functional normalization according to Fortin et al.[65] to control for batch and position effects, and removal of SNP- and sex-related CpGs, as well as polymorphic CpG sites and cross-hybridizing probes.[38] Following this, there were 319, 264 CpG sites included in the analysis. Joo et al. have demonstrated that DNA methylation measured by the HM450k array on archived dried blood spots is fully correlated with DNA methylation measured by the same platform on same individuals' matched frozen buffy coats (correlation coefficient = 0.99), therefore proving that this material is suitable for DNA methylation analyses.[66]

Statistical analyses

We examined association between maternal folate intake and DNA methylation intensity of 319, 264 CpG sites in a locus-by-locus analysis. *Logit*-transformed β-values (M-values) were used as the outcome variable, and total DFE (folate) as the continuous exposure variable. The following covariates were included in the locus-by-locus models: gender, gestational age, proportions of monocytes, B-cells, T-cells (CD4T and CD8T), and natural killer cells (estimated on reference-based methods according to Houseman et al.[67]). We also included the first 2 principal components generated by principal component analysis computed on the matrix of residuals of the models, in order to adjust for possible residual cell-mixture effect that would not have been estimated with the reference-based technique, according to Teschendorff et al. and Leek et al.[57,58] Such residual cell-mixture effects in our experiment may be due to blood cell types specific to neonates (e.g., a high proportion of reticulocytes). The same procedure was implemented in 2 independent sample sets (n_{set1}=167, n_{set2}=176). Genomic inflation factors were calculated based on median chi-square distribution.[68] In order to decrease the number of chance findings (i.e., false-positive findings), the data were resampled 1, 000 times with replacement (i.e., bootstrapping) in each set, with the locus-by-locus model run on each bootstrap sample. This procedure resulted in a distribution of *P*-values for each CpG site. We then selected all CpG sites with an inter-quartile range of *P*-values from bootstrapping <0.05 in both sets for further analysis (i.e., the most significant CpG sites that replicated across set 1 and set 2). We also chose a less stringent median bootstrapping *P-value* cut-off of $P<0.1$ in both sets to generate a larger list of CpG sites for pathway analysis.

For the top concordant CpG sites resulting from the bootstrap, corrections for multiple testing were performed by permutation tests.[39,40] In order to draw an empirical distribution of the regression coefficients under the null hypothesis that there is no association between DNA methylation and folate exposure, we randomly resampled our exposure variable (i.e., folate exposure) 1, 000 times while keeping the outcome variable and covariates fixed. *P*-values from permutations tests were ranked and compared to the ranked *P*-values from the initial analysis, and the permutation-corrected *P*-values were calculated as the number of higher *P*-values from permutations / total number of *P*-values for each rank. Both sets were combined for this analysis. Analyses were carried out using R[69]: the 'limma' package[70] was used to run the locus-by-locus analysis; the 'Gviz' package[71] was used to create gene location figures; and the 'Haplin' package[72] was used to draw qq-plots.

We also performed subgroups analyses. We categorized our participants in 3 subgroups with the following total DFE consumptions: (1) < 200 μg/d (n=21), (2) between 200 and 600 μg/d (n=178), and (3) > 600 μg/d (n=139). Categories were based on clinical references from the US. Institute of Medicine.[73] Less than 200 μg/d of DFE was associated with persistent erythrocytopenia in otherwise healthy adults, and the recommended dietary allowances for pregnant women is 600 μg/d of DFE. We modeled the associations between folate and DNA methylation (according to the above mentioned locus-by-locus model) separately in each category of folate consumption. We then calculated the ratio of the number of positive associations on the number of negative associations and we plotted the logarithm of this ratio for each category.

Gene expression

Our top CpG sites were located in or near gene promoters (<370 bp). To estimate whether DNA methylation at these loci may be associated with gene expression, we used 2 publicly available databases from Gene Expression Omnibus (GEO) containing both DNA methylation data and gene expression data from the same samples. The first and main dataset [accession: GSE30654] consisted of 86 pluripotent stem cells and their derivatives; the corresponding assays were HM450K methylation arrays and the Illumina HumanHT-12 V3.0 expression beadchip (*NB. Did not include* CYS1). Due to the large sample size, we assessed all the available CpG sites/genes in this data set. A second dataset [accession: GSE49065] was used to estimate the expression of *CYS1* only in function of its methylation, in peripheral mononuclear blood cells of 20 healthy adult men (HM450K and Affymetrix Human Gene 1.1 ST Array). Gene expression as a function of DNA methylation was plotted for each gene (see **Fig. S6**), and if a bimodal pattern was observed, the DNA methylation variable was dichotomized to test for an association with gene expression by performing a Student t test; otherwise, a Spearman test of correlation was performed with continuous variables.

Gene pathway analysis

Pathway analysis was carried out on the longer list of the significant ($P<0.1$) and sign concordant CpG sites from the 2 data

sets using the Gene Set Enrichment Analysis web tool.[74,75] For this analysis, we retrieved the gene identifiers for all the CpG sites of our list that were associated with a gene. We report *P*-values corrected for multiple testing (False Discovery Rate q-values) <10 −<τβρ>−10 </tbr>.

Replication of candidate folate-sensitive sites

We estimated DNA methylation variations in function of folate intake and covariates for candidate-probes from our array that were located in previously investigated folate-sensitive regions (as described by[23,34]); Genome addresses are listed in **Tables S3 and S4**. We combined set 1 and set 2 in order to increase sample size and statistical power. DNA methylation was assessed in function of folate intake; covariates of the model were: gender, race, gestational age, cell-mixture, and set number. We dichotomized the folate variable into high vs. low groups by taking the median of folate intake in our data. Similar to the methods described in Dominguez-Salas et al., we pooled the β-values of all the CpG sites of interest, by taking the median β-value for each participant.

Validation of the top hits

We used publicly available data of the "The Gambia / Aflatoxin / Seasonality study" from the GEO database to replicate our top hits (accession: GSE59592). The study analyzed DNA methylation of peripheral whole blood of 118 infants aged 3–6 months with the Illumina© Infinium HumanMethylation450 BeadChip arrays. The season of conception was recorded as being either dry or rainy. Previous analyses of this study population have shown that the mothers had lower levels of folate during the dry season than during the rainy season.[42,43] DNA methylation β-values were *logit*-transformed in M-values, and multiple linear regression tests were performed to assess DNA methylation at our 4 top hits as the outcome, in function of the season of conception as the exposure (the dry season corresponding to high folate exposure was coded 1, and the rainy season corresponding to low folate exposure was coded 0), while adjusting for gender.

Funding

This work was supported by the National Institute of Environmental Health Sciences and the Environmental Protection Agency [grant number: P01ES018172 to C.M. and J.L.W.]; The National Institute of Environmental Health Sciences [grant number: R01ES09137 to C.M. and J.L.W.]; The National Cancer Institute [grant number: R01CA155461 to J.L.W.]; The Swiss Cancer League [grant number: BIL KLS 3124–02–2013 to S. G.]; The Swiss Science National Foundation [grants numbers: P2LAP3_158674 to S.G., P2LAP3_148434 to S.N.]; The SICPA Foundation to S.G.; And the Sutter-Stöttner Foundation to S.G. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health or the Environmental Protection Agency.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Acknowledgments

We acknowledge Prof. Steve Selvin for his statistical advice regarding resampling methods. The CCLS is supported by our clinical collaborators and participating hospitals, which includes: University of California Davis Medical Center (Dr. Jonathan Ducore), University of California San Francisco (Dr. Mignon Loh and Dr. Katherine Matthay), Children's Hospital of Central California (Dr. Vonda Crouse), Lucile Packard Children's Hospital (Dr. Gary Dahl), Children's Hospital Oakland (Dr. James Feusner), Kaiser Permanente Sacramento (Dr. Vincent Kiley), Kaiser Permanente Santa Clara (Dr. Carolyn Russo and Dr. Alan Wong), Kaiser Permanente San Francisco (Dr. Kenneth Leung), and Kaiser Permanente Oakland (Dr. Stacy Month), and the families of the study participants. We thank the staff and faculty at San Francisco General Hospital Women's Options Center for assistance in the collection of human fetal tissues.

Supplemental Material

Supplemental data for this article can be accessed on the publisher's website.

References

1. Blom HJ, Shaw GM, den Heijer M, Finnell RH. Neural tube defects and folate: case far from closed. Nat Rev Neurosci 2006; 7:724-31; PMID:16924261; http://dx.doi.org/10.1038/nrn1986
2. Mossey PA, Little J, Munger RG, Dixon MJ, Shaw WC. Cleft lip and palate. Lancet 2009; 374:1773-85; PMID:19747722; http://dx.doi.org/10.1016/S0140-6736(09)60695-4
3. Hernández-Díaz S, Werler MM, Walker AM, Mitchell AA. Folic Acid Antagonists during Pregnancy and the Risk of Birth Defects. N Engl J Med 2000; 343:1608-14; http://dx.doi.org/10.1056/NEJM200011303432204
4. Metayer C, Milne E, Dockerty JD, Clavel J, Pombo-de-Oliveira MS, Wesseling C, Spector LG, Schüz J, Petridou E, Ezzat S, et al. Maternal Supplementation with Folic Acid and Other Vitamins and Risk of Leukemia in Offspring: A Childhood Leukemia International Consortium Study. Epidemiol Camb Mass 2014; 25:811-22; http://dx.doi.org/10.1097/EDE.0000000000000141
5. Beaudin AE, Stover PJ. Folate-mediated one-carbon metabolism and neural tube defects: Balancing genome synthesis and gene expression. Birth Defects Res. Part C Embryo Today Rev 2007; 81:183-203; http://dx.doi.org/10.1002/bdrc.20100
6. Alonso-Aperte E, Ubeda N, Achon M, Perez-Miguelsanz J, Varela-Moreiras G. Impaired methionine synthesis and hypomethylation in rats exposed to valproate during gestation. Neurol March 10 1999 1999; 52:750-6
7. Finnell RH, Spiegelstein O, Wlodarczyk B, Triplett A, Pogribny IP, Melnyk S, James JS. DNA Methylation in Folbp1 Knockout Mice Supplemented with Folic Acid during Gestation. J Nutr 2002; 132:2457S-61S; PMID:12163711
8. Waterland RA, Jirtle RL. Transposable Elements: Targets for Early Nutritional Effects on Epigenetic Gene Regulation. Mol Cell Biol 2003; 23:5293-300; http://dx.doi.org/10.1128/MCB.23.15.5293-5300.2003
9. Lillycrop KA, Phillips ES, Jackson AA, Hanson MA, Burdge GC. Dietary Protein Restriction of Pregnant Rats Induces and Folic Acid Supplementation Prevents Epigenetic Modification of Hepatic Gene Expression in the Offspring. J Nutr 2005; 135:1382-6; PMID:15930441
10. Waterland RA, Dolinoy DC, Lin JR, Smith CA, Shi X, Tahiliani KG. Maternalmethyl supplements increase offspring DNA methylation at Axin fused. Genesis 2006; 44:401-6; PMID:16868943; http://dx.doi.org/10.1002/dvg.20230
11. Maloney CA, Hay SM, Rees WD. Folate deficiency during pregnancy impacts on methyl metabolism without affecting global DNA methylation in the rat fetus. Br J Nutr 2007; 97:1090-8; PMID:17433124; http://dx.doi.org/10.1017/S0007114507670834

12. Dolinoy DC, Huang D, Jirtle RL. Maternal nutrient supplementation counteracts bisphenol A-induced DNA hypomethylation in early development. Proc Natl Acad Sci USA 2007; 104:13056-61; PMID:17670942; http://dx.doi.org/10.1073/pnas.0703739104
13. Sinclair KD, Allegrucci C, Singh R, Gardner DS, Sebastian S, Bispham J, Thurston A, Huntley JF, Rees WD, Maloney CA, et al. DNA methylation, insulin resistance, and blood pressure in offspring determined by maternal periconceptional B vitamin and methionine status. Proc Natl Acad Sci USA 2007; 104:19351-6; PMID:18042717; http://dx.doi.org/10.1073/pnas.0707258104
14. Lillycrop KA, Phillips ES, Torrens C, Hanson MA, Jackson AA, Burdge GC. Feeding pregnant rats a protein-restricted diet persistently alters the methylation of specific cytosines in the hepatic PPAR? promoter of the offspring. Br J Nutr 2008; 100:278-82; PMID:18186951; http://dx.doi.org/10.1017/S0007114507894438
15. Kim JM, Hong K, Lee JH, Lee S, Chang N. Effect of folate deficiency on placental DNA methylation in hyperhomocysteinemic rats. J Nutr Biochem 2009; 20:172-6; PMID:18547797; http://dx.doi.org/10.1016/j.jnutbio.2008.01.010
16. Lawrance AK, Deng L, Rozen R. Methylenetetrahydrofolate reductase deficiency and low dietary folate reduce tumorigenesis in Apcmin/+ mice. Gut 2009; 58:805-11; PMID:19174418; http://dx.doi.org/10.1136/gut.2007.143107
17. Gong L, Pan YX, Chen H. Gestational low protein diet in the rat mediates Igf2 gene expression in male offspring via altered hepatic DNA methylation. Epigenetics Off J DNA Methylation Soc 2010; 5:619-26; http://dx.doi.org/10.4161/epi.5.7.12882
18. Downing C, Johnson TE, Larson C, Leakey TI, Siegfried RN, Rafferty TM, Cooney CA. Subtle Decreases in DNA Methylation and Gene Expression at the Mouse Igf2 Locus Following Prenatal Alcohol Exposure: Effects of a Methyl-Supplemented Diet. Alcohol Fayettev N 2011; 45:65-71; http://dx.doi.org/10.1016/j.alcohol.2010.07.006
19. Schaible TD, Harris RA, Dowd SE, Smith CW, Kellermayer R. Maternal methyl-donor supplementation induces prolonged murine offspring colitis susceptibility in association with mucosal epigenetic and microbiomic changes. Hum Mol Genet 2011; 20:1687-96; PMID:21296867; http://dx.doi.org/10.1093/hmg/ddr044
20. Kulkarni A, Dangat K, Kale A, Sable P, Chavan-Gautam P, Joshi S. Effects of Altered Maternal Folic Acid, Vitamin B12 and Docosahexaenoic Acid on Placental Global DNA Methylation Patterns in Wistar Rats. PLoS One 2011; 6:e17706; PMID:21423696; http://dx.doi.org/10.1371/journal.pone.0017706
21. Chmurzynska A, Stachowiak M, Pruszynska-Oszmalek E. Maternal protein and folic acid intake during gestation does not program leptin transcription or serum concentration in rat progeny. Genes Nutr 2012; 7:217-22; PMID:21735287; http://dx.doi.org/10.1007/s12263-011-0239-5
22. Cooper WN, Khulan B, Owens S, Elks CE, Seidel V, Prentice AM, Belteki G, Ong KK, Affara NA, Constância M, et al. DNA methylation profiling at imprinted loci after periconceptional micronutrient supplementation in humans: results of a pilot randomized controlled trial. FASEB J 2012; 26:1782-90; PMID:22267336; http://dx.doi.org/10.1096/fj.11-192708
23. Dominguez-Salas P, Moore SE, Baker MS, Bergen AW, Cox SE, Dyer RA, Fulford AJ, Guan Y, Laritsky E, Silver MJ, et al. Maternal nutrition at conception modulates DNA methylation of human metastable epialleles. Nat Commun 2014; 5:3746; PMID:24781383; http://dx.doi.org/10.1038/ncomms4746
24. Fryer AA, Nafee TM, Ismail K.M.K., Carroll, WD, Emes RD, Farrell WE. LINE-1 DNA methylation is inversely correlated with cord plasma homocysteine in man: a preliminary study. Epigenetics 2009; 4:394-8; PMID:19755846; http://dx.doi.org/10.4161/epi.4.6.9766
25. Fryer AA, Emes RD, Ismail KM, Haworth KE, Mein C, Carroll WD, Farrell WE. Quantitative, high-resolution epigenetic profiling of CpG loci identifies associations with cord blood plasma homocysteine and birth weight in humans. Epigenetics 2011; 6:86-94; PMID:20864804; http://dx.doi.org/10.4161/epi.6.1.13392
26. Steegers-Theunissen RP, Obermann-Borst SA, Kremer D, Lindemans J, Siebel C, Steegers EA, Slagboom PE, Heijmans BT. 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:e7845; PMID:19924280; http://dx.doi.org/10.1371/journal.pone.0007845
27. Ba Y, Yu H, Liu F, Geng X, Zhu C., Zhu, Q., Zheng, T., Ma, S., Wang, G., Li, Z., et al. Relationship of folate, vitamin B12 and methylation of insulin-like growth factor-II in maternal and cord blood. Eur J Clin Nutr 2011; 65:480-5; PMID:21245875; http://dx.doi.org/10.1038/ejcn.2010.294
28. Chang H, Zhang T, Zhang Z, Bao R, Fu C, Wang Z, Bao Y, Li Y, Wu L, Zheng X. Tissue-specific distribution of aberrant DNA methylation associated with maternal low-folate status in human neural tube defects. J Nutr Biochem 2011; 22:1172-7; PMID:21333513; http://dx.doi.org/10.1016/j.jnutbio.2010.10.003
29. Hoyo C, Murtha AP, Schildkraut JM, Jirtle R, Demark-Wahnefried W, Forman MR, Iversen ES, Kurtzberg J, Overcash F, Huang Z, et al. Methylation variation at IGF2 differentially methylated regions and maternal folic acid use before and during pregnancy. Epigenetics 2011; 6:928-36; PMID:21636975; http://dx.doi.org/10.4161/epi.6.7.16263
30. McKay JA, Groom A, Potter C, Coneyworth LJ, Ford D, Mathers JC, Relton CL. Genetic and Non-Genetic Influences during Pregnancy on Infant Global and Site Specific DNA Methylation: Role for Folate Gene Variants and Vitamin B12. Plos One 2012; 7:e33290; PMID:22479380; http://dx.doi.org/10.1371/journal.pone.0033290
31. 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; PMID:23151531; http://dx.doi.org/10.3945/ajcn.112.042572
32. Loke YJ, Galati JC, Morley R, Joo EJH, Novakovic B, Li X, Weinrich B, Carson N, Ollikainen M, Ng HK, et al Association of maternal and nutrient supply line factors with DNA methylation at the imprinted IGF2/H19 locus in multiple tissues of newborn twins. Epigenetics 2013; 8:1069-79; PMID:23917818; http://dx.doi.org/10.4161/epi.25908
33. Crider KS, Yang TP, Berry RJ, Bailey LB, Folate and DNA methylation: a review of molecular mechanisms and the evidence for folate's role. Adv Nutr Int Rev J 2012; 3:21-38; http://dx.doi.org/10.3945/an.111.000992
34. Amarasekera M, Martino D, Ashley S, Harb H, Kesper D, Strickland D, Saffery R, Prescott SL. Genome-wide DNA methylation profiling identifies a folate-sensitive region of differential methylation upstream of ZFP57-imprinting regulator in humans. FASEB J 2014; 28 (9):4068-76; PMID:24891518; http://dx.doi.org/10.1096/fj.13-249029
35. Kelsey G. Epigenome Changes During Development. In *Epigenetic Epidemiology*. Springer US, pp. 2012; 77-103
36. Rakyan VK, Blewitt ME, Druker R, Preis JI, Whitelaw E. Metastable epialleles in mammals. Trends Genet 2002; 18:348-51; PMID:12127774; http://dx.doi.org/10.1016/S0168-9525(02)02709-9
37. Price ME, Cotton AM, Lam LL, Farré P, Emberly E, Brown CJ, Robinson WP, Kobor MS. Additional annotation enhances potential for biologically-relevant analysis of the Illumina Infinium HumanMethylation450 BeadChip array. Epigenetics Chromatin 2013; 6:4-4; PMID:23452981; http://dx.doi.org/10.1186/1756-8935-6-4
38. Chen Y, Lemire M, Choufani S, Butcher DT, Grafodatskaya D, Zanke BW, Gallinger S, Hudson TJ, Weksberg R. Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray. Epigenetics 2013; 8:203-9; PMID:23314698; http://dx.doi.org/10.4161/epi.23470
39. Karin B. Michels Epigenetic Epidemiology
40. Westfall PH. Resampling-based multiple testing: Examples and methods for p-value adjustment John Wiley & Sons 1993
41. Hernandez-Vargas H, Castelino J, Silver MJ, Dominguez-Salas P, Cros MP, Durand G, Calvez-Kelm FL, Prentice AM, Wild CP, Moore SE, et al. Exposure to aflatoxin B1 in utero is associated with DNA methylation in white blood cells of infants in The Gambia. Int J Epidemiol 2015; 44(4):1238-48; PMID:25855716; http://dx.doi.org/10.1093/ije/dyv027
42. Silver MJ, Kessler NJ, Hennig BJ, Dominguez-Salas P, Laritsky E, Baker MS, Coarfa C, Hernandez-Vargas H, Castelino JM, Routledge MN, et al. Independent genomewide screens identify the tumor suppressor VTRNA2-1 as a human epiallele responsive to periconceptional environment. Genome Biol 2015; 16:118; PMID:26062908; http://dx.doi.org/10.1186/s13059-015-0660-y
43. Dominguez-Salas P, Moore SE, Cole D, da Costa KA, Cox SE, Dyer RA, Fulford AJ, Innis SM, Waterland RA, Zeisel SH, et al. DNA methylation potential: dietary intake and blood concentrations of one-carbon metabolites and cofactors in rural African women. Am J Clin Nutr 2013; 97:1217-27; PMID:23576045; http://dx.doi.org/10.3945/ajcn.112.048462
44. Barrallo-Gimeno A, Holzschuh J, Driever W, Knapik EW. Neural crest survival and differentiation in zebrafish depends on mont blanc/tfap2a gene function. Development 2004; 131:1463-77; PMID:14985255; http://dx.doi.org/10.1242/dev.01033
45. Meshcheryakova TI, Zinchenko RA, Vasilyeva TA, Marakhonov AV, Zhylina SS, Petrova NV, Kozhanova TV, Belenikin MS, Petrin AN, Mutovin GR. A Clinical and Molecular Analysis of Branchio-Oculo-Facial Syndrome Patients in Russia Revealed New Mutations in TFAP2A. Ann Hum Genet 2015; 79(2):148-52; PMID:25590586; http://dx.doi.org/10.1111/ahg.12098
46. Gerth-Kahlert C, Williamson K, Ansari M, Rainger JK, Hingst V, Zimmermann T, Tech S, Guthoff RF, van Heyningen V, FitzPatrick DR. Clinical and mutation analysis of 51 probands with anophthalmia and/or severe microphthalmia from a single center. Mol Genet Genomic Med 2013; 1:15-31; PMID:24498598; http://dx.doi.org/10.1002/mgg3.2
47. Di C, Liao S, Adamson DC, Parrett TJ, Broderick DK, Shi Q, Lengauer C, Cummins JM, Velculescu VE, Fults DW, et al. Identification of OTX2 as a Medulloblastoma Oncogene Whose Product can be Targeted by All-Trans Retinoic Acid. Cancer Res 2005; 65:919-24; PMID:15705891
48. Moritake H, Kamimura S, Nunoi H, Nakayama H, Suminoe A, Inada H, Inagaki J, Yanai F, Okamoto Y, Shinkoda Y, et al. Clinical characteristics and genetic analysis of childhood acute lymphoblastic leukemia with hemophagocytic lymphohistiocytosis: a Japanese retrospective study by the Kyushu–Yamaguchi Children's Cancer Study Group. Int J Hematol 2014; 100:70-8; PMID:24827398; http://dx.doi.org/10.1007/s12185-014-1591-1
49. Hou X, Mrug M, Yoder BK, Lefkowitz EJ, Kremmidiotis G, D'Eustachio P, Beier DR, Guay-Woodford LM. Cystin, a novel cilia-associated protein, is disrupted in the cpk mouse model of polycystic kidney disease. J Clin Invest 2002; 109:533-40;

PMID:11854326; http://dx.doi.org/10.1172/JCI0214099
50. Wang X, Southard RC, Allred CD, Talbert DR, Wilson ME, Kilgore MW. MAZ drives tumor-specific expression of PPAR gamma 1 in breast cancer cells. Breast Cancer Res Treat 2008; 111:103-11; PMID:17902047; http://dx.doi.org/10.1007/s10549-007-9765-7
51. Jiao L, Li Y, Shen D, Xu C, Wang L, Huang G, Chen L, Yang Y, Yang C, Yu Y, et al. The prostate cancer-up-regulated Myc-associated zinc-finger protein (MAZ) modulates proliferation and metastasis through reciprocal regulation of androgen receptor. Med Oncol Northwood Lond Engl 2013; 30:570; http://dx.doi.org/10.1007/s12032-013-0570-3
52. Institute of Medicine (US) Standing Committee on the Scientific Evaluation of Dietary Reference Intakes and its Panel on Folate, O.B.V. 1998 Folate
53. Boeke CE, Baccarelli A, Kleinman KP, Burris HH, Litonjua AA, Rifas-Shiman SL, Tarantini L, Gillman MW. Gestational intake of methyl donors and global LINE-1 DNA methylation in maternal and cord blood. Epigenetics 2012; 7:253-60; PMID:22430801; http://dx.doi.org/10.4161/epi.7.3.19082
54. Aarabi M, Gabriel MCS, Chan D, Behan NA, Caron M, Pastinen T, Bourque G, MacFarlane AJ, Zini A, Trasler J. High-dose folic acid supplementation alters the human sperm methylome and is influenced by the MTHFR C677T polymorphism. Hum Mol Genet 2015; 24(22):6301-13; PMID:26307085; http://dx.doi.org/10.1093/hmg/ddv338
55. Branum AM, Bailey R, Singer BJ. Dietary Supplement Use and Folate Status during Pregnancy in the United States. J Nutr 2013; 143:486-92; PMID:23365107; http://dx.doi.org/10.3945/jn.112.169987
56. Sobell J, Block G, Koslowe P, Tobin J, Andres R. Validation of a retrospective questionnaire assessing diet 10-15 years ago. Am J Epidemiol 1989; 130:173-87; PMID:2741904
57. Teschendorff AE, Zhuang J, Widschwendter M. Independent surrogate variable analysis to deconvolve confounding factors in large-scale microarray profiling studies. Bioinformatics 2011; 27:1496-505; PMID:21471010; http://dx.doi.org/10.1093/bioinformatics/btr171
58. Leek JT, Storey JD. Capturing heterogeneity in gene expression studies by surrogate variable analysis. PLoS Genet 2007; 3:1724-35; PMID:17907809; http://dx.doi.org/10.1371/journal.pgen.0030161
59. Knight RD, Nair S, Nelson SS, Afshar A, Javidan Y, Geisler R, Rauch GJ, Schilling TF. lockjaw encodes a zebrafish tfap2a required for early neural crest. Development 2003; 130:5755-68; PMID:14534133; http://dx.doi.org/10.1242/dev.00575
60. Milne E, Greenop KR, Bower C, Miller M, Bockxmeer FM, van Scott RJ, Klerk NH, de Ashton LJ, Gottardo NG, Armstrong BK. Maternal Use of Folic Acid and Other Supplements and Risk of Childhood Brain Tumors. Cancer Epidemiol Biomarkers Prev 2012; 21:1933-41; PMID:22941336; http://dx.doi.org/10.1158/1055-9965.EPI-12-0803
61. Zhang X, Inukai T, Hirose K, Akahane K, Kuroda I, Honna-Oshiro H, Kagami K, Goi K, Nakamura K, Kobayashi M, et al. Oncogenic fusion E2A-HLF sensitizes t(17;19)-positive acute lymphoblastic leukemia to TRAIL-mediated apoptosis by upregulating the expression of death receptors. Leukemia 2012; 26:2483-93; PMID:22743623; http://dx.doi.org/10.1038/leu.2012.139
62. SP1 Sp1 transcription factor [Homo sapiens (human)] - Gene - NCBI
63. LEF1 lymphoid enhancer-binding factor 1 [Homo sapiens (human)] - Gene - NCBI
64. Jensen CD, Block G, Buffler P, Ma X, Selvin S, Month S. Maternal dietary risk factors in childhood acute lymphoblastic leukemia (United States). Cancer Causes Control CCC 2004; 15:559-70; PMID:15280635; http://dx.doi.org/10.1023/B:CACO.0000036161.98734.17
65. Fortin JP, Labbe A, Lemire M, Zanke BW, Hudson TJ, Fertig EJ, Greenwood CM, Hansen KD. Functional normalization of 450k methylation array data improves replication in large cancer studies. Genome Biol 2014; 15(12):503; PMID:25599564; http://dx.doi.org/10.1186/s13059-014-0503-2
66. Joo JE, Wong EM, Baglietto L, Jung CH, Tsimiklis H, Park DJ, Wong NC, English DR, Hopper JL, Severi G, et al. The use of DNA from archival dried blood spots with the Infinium HumanMethylation450 array. BMC Biotechnol 2013; 13:23; PMID:23497093; http://dx.doi.org/10.1186/1472-6750-13-23
67. Houseman EA, Accomando WP, Koestler DC, Christensen BC, Marsit CJ, Nelson HH, Wiencke JK, Kelsey KT. DNA methylation arrays as surrogate measures of cell mixture distribution. BMC Bioinformatics 2012; 13:86; PMID:22568884; http://dx.doi.org/10.1186/1471-2105-13-86
68. Devlin B, Roeder K. Genomic Control for Association Studies. Biometrics 1999; 55:997-1004; PMID:11315092; http://dx.doi.org/10.1111/j.0006-341X.1999.00997.x
69. R Core Team, R: A Language and Environment for Statistical Computing 2013
70. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res 2015; 43(7):e47; PMID:25605792; http://dx.doi.org/10.1093/nar/gkv007
71. Hahne F, Durinck S, Ivanek R, Mueller A, Lianoglou S, Tan G, Gviz LP. Plotting data and annotation information along genomic coordinates
72. Gjessing HK. Haplin: Analyzing Case-Parent Triad and/or Case-Control Data with SNP Haplotypes 2015
73. Institute of Medicine (US) Standing Committee on the Scientific Evaluation of Dietary Reference Intakes and its Panel on Folate, Other B Vitamins, and Choline. Dietary Reference Intakes for Thiamin, Riboflavin, Niacin, Vitamin B6, Folate, Vitamin B12, Pantothenic Acid, Biotin, and Choline National Academies Press (US), Washington (DC) 1998
74. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci USA 2005; 102:15545-50; PMID:16199517; http://dx.doi.org/10.1073/pnas.0506580102
75. Mootha VK, Lindgren CM, Eriksson KF, Subramanian A, Sihag S, Lehar J, Puigserver P, Carlsson E, Ridderstråle M, Laurila E, et al. PGC-1α-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. Nat Genet 2003; 34:267-73; PMID:12808457; http://dx.doi.org/10.1038/ng1180