Maternal red blood cell folate and infant vitamin B$_{12}$ status influence methylation of genes associated with childhood acute lymphoblastic leukaemia.

Catherine Potter[1], Anthony Vincent Moorman[2], Caroline Laura Relton[3], Dianne Ford[4], John Cummings Mathers[5], Gordon Strathdee[2], Jill Ann McKay[4,6].

[1]*Institute of Genetic Medicine, Newcastle University, UK;* [2]*Northern Institute for Cancer Research, Newcastle University, UK;* [3]*School of Social and Community Medicine, Bristol University, UK;* [4]*Faculty of Health and Life Sciences, Department of Applied Sciences, Northumbria University, UK;* [5]*Human Nutrition Research Centre, Institute of Cellular Medicine, Newcastle University, UK;* [6]*Human Nutrition Research Centre, Institute for Health & Society, Newcastle University, UK*

Corresponding author: Dr Jill A. McKay, Department of Applied Sciences, Faculty of Health and Life Sciences, Northumbria University, Newcastle upon Tyne, NE1 8ST, U.K. Tel; +44 (0)191 227 4071, Email jill.mckay@northumbria.ac.uk

Abbreviations: Developmental origins of health and disease (DOHaD); Red Blood Cell folate (RBC folate); Acute Lymphoblastic Leukaemia (ALL)

Key words; Maternal; folic acid; vitamin B$_{12}$; DNA methylation; acute lymphoblastic leukaemia; biomarkers

Abstract

Scope; Inadequate maternal folate intake is associated with increased childhood acute lymphoblastic leukaemia (ALL) risk. Folate provides methyl groups for DNA methylation, which is dramatically disrupted in ALL. We investigated if maternal folate (and related B-vitamin) intake during pregnancy may affect ALL risk via influencing DNA methylation. Methods and Results; We identified genes in which methylation changes were reported both in response to folate status and in ALL. Folate-responsive genes (n=526) were identified from mouse models of maternal folate depletion during pregnancy. Using published data, we identified 2621 genes with persistently altered methylation in ALL. Twenty-five overlapping genes were found, with the same directional methylation change in response to folate depletion and in ALL. We confirmed hypermethylation of a subset of genes (*ASCL2, KCNA1, SH3GL3, SRD5A2*) in ALL by measuring 20 patient samples using pyrosequencing. In a nested cohort of cord blood samples (n=148), *SH3GL3* methylation was inversely related to maternal RBC folate concentrations (p=0.008). Furthermore, *ASCL2* methylation was inversely related to infant vitamin B12 levels. (p=0.016). Conclusion; Findings demonstrate proof of concept for a plausible mechanism i.e. variation in DNA methylation, by which low intake of folate, and related B-vitamins during pregnancy may influence ALL risk.

Introduction

Increasingly, evidence from epidemiological studies suggests that maternal folate intake prior to and during pregnancy may influence the risk of the development of childhood acute lymphoblastic leukaemia (ALL)[1-7]. The first study to report a protective association between maternal folate and ALL risk observed an odds ratio (OR) of 0.40 (Confidence Intervals (CI) 0.21-0.73) for self-reported folate supplementation taken at any time, for any duration, during pregnancy[7]. Although this initial study was fairly small (83 cases and 166 controls), subsequent larger studies have also reported significant inverse associations between self-reported folate supplementation before or during pregnancy and ALL risk (OR = 0.4 (CI 0.3-0.6)[2] and 0.80 (0.71-0.89)[5]). Further studies found associations between higher folate intake, assessed by retrospective food frequency questionnaires, at specific times during pregnancy and altered ALL risk. Whilst Bailey *et al* observed that higher dietary folate intake during the last 6 months of pregnancy was protective[3], Singer *et al* reported that higher maternal intake of one-carbon metabolism nutrients and supplements combined in the year before pregnancy reduced offspring ALL risk[6]. Similarly, Ajrouche *et al* found that self-reported folate supplementation implemented 3 months preceding, but not during, pregnancy may reduce childhood leukaemia risk (OR 0.7 (0.5-1.0))[1]. This finding was supported by the largest meta-analysis to date, which observed that folate supplementation pre-conception reduces ALL risk (OR: 0.69, 95%CI: 0.50-0.95)[4]. Furthermore, genetic studies investigating the influence of both parental and offspring variants of folate metabolising genes on childhood ALL risk corroborate the hypothesis that lower folate status may enhance disease risk [2, 8-12]. The observational evidence that *in utero* folate exposure may modulate risk of childhood ALL development is consistent but, to date, there is limited

insight into the underlying mechanism through which folate status prior to, and during, pregnancy may be involved in disease development.

Childhood ALL arises from initial genetic abnormalities such as chromosomal rearrangement and hyperdiploidy. These abnormalities may occur *in utero* [13-17], and remain latent until a 'secondary hit' leads to overt disease. Indeed, the *ETV6-RUNX1* fusion, which occurs in approximately 25% of B-cell precursor ALL cases[18], was present in 1% of the population at birth; this is a 100-fold higher prevalence than for this disease subtype [16] suggesting that, whilst key in the disease pathway, the presence of this translocation alone is not sufficient to cause disease.

DNA methylation is an important mechanism of gene regulation and methylation of CpG rich regions may inhibit binding of the regulatory machinery and lead to gene silencing[19]. As with all cancers, altered DNA methylation patterns have been widely documented in ALL [20-26]. It is therefore plausible that gene dysregulation arising from aberrant DNA methylation may contribute to the aetiology of childhood ALL, and thus be a potential 'second hit' mechanism in the causal disease pathway. Furthermore, whilst the DNA sequence is generally resistant to change in response to non-carcinogenic environmental factors, DNA methylation is more plastic and is altered by a variety of environmental and lifestyle exposures [27]. Since folate is a key source of methyl groups for synthesis of *S*-adenosyl methionine (SAM) – the universal methyl donor – it is an attractive candidate nutrient for the modulation of DNA methylation. Altered intakes of folate during pregnancy can have profound effects on DNA methylation patterns in the offspring, as has been reported in both animal models[28-33] and human studies[34, 35]. It is therefore plausible that altered DNA methylation may be an underlying mechanism through which maternal folate

intake prior to, and during, pregnancy may modulate risk of childhood ALL. Moreover, through their effects on the synthesis of SAM, related micronutrients driving one carbon metabolism, i.e. vitamins B_2, B_6 and B_{12}, may also modulate DNA methylation patterns and, hence, ALL risk. Indeed, higher maternal dietary B_{12} intake has been associated with decreased offspring ALL risk[3], and vitamin B_{12} has also been reported to influence DNA methylation patterns [36-39].

We therefore hypothesised that reduced intake of folate (and other B vitamins) during pregnancy may alter DNA methylation in genes that are also aberrantly methylated in childhood ALL. To investigate this hypothesis, we used a 'meet in the middle' approach [40] to identify genes that have altered DNA methylation in response to maternal folate depletion and are also aberrantly methylated in childhood ALL (see Figure 1). We utilised data from our mouse model of maternal folate depletion [29, 33] to determine genes which have altered methylation in response to maternal folate depletion. We then compared this list of genes with a compiled list of genes from the literature that have been reported to have altered DNA methylation in ALL. We confirmed aberrant DNA methylation in a panel of target genes in ALL patient samples, and investigated the association of maternal and infant folate and vitamin B_{12} status on methylation of those loci in a population of healthy infants at birth.

Materials and Methods

Identification of genes with altered DNA methylation in relation to maternal folate depletion AND in childhood ALL

We previously reported altered DNA methylation in response to maternal folate depletion in fetal and adult mouse models [29, 33]. Briefly, female C57Bl/6 mice were randomised to low or normal folate diets (0.4 or 2mg folic acid/kg diet) 4 weeks prior to mating. Mice were maintained on these diets throughout pregnancy, and throughout weaning (4 weeks) in the case of the adult offspring model. For the fetal model, dams were culled at 17.5 days' gestation, and foetuses removed. For the adult model of folate depletion during pregnancy, offspring were weaned on to diets containing normal levels of folate (i.e. 2mg/kg) and were maintained on this diet until 6 months of age. Low folate diets resulted in significantly lower folate status of dams compared to control fed dams (previously published data[41, 42]). Fetal (n=6/group) and adult (n=12/group) livers were excised from male mice, and DNA extracted. Promoter methylation was measured by microarray analysis following methylated DNA immunoprecipatation. Resultant data were assessed for quality, normalised and processed as described previously [29, 33]. Using data from these studies we compiled a list of 'folate-responsive' genes - i.e. those with altered DNA methylation in response to maternal folate intake in either fetal or adult offspring (n=526 genes) (See Supplementary Material).

A list of genes reported to have altered DNA methylation in childhood ALL was determined using publically-available data. A literature search was carried out in PubMed using the search terms 'DNA methylation' and 'childhood acute lymphoblastic leukaemia'. The criteria for study selection were a) methylation had been assessed at the genome-wide level

b) methylation was compared in ALL cases vs control samples c) methylation outcome data were available from the manuscript. Five studies were included for which genome-wide methylation had been assessed in childhood ALL cases compared with controls [20, 21, 23-25] (See Supplementary Material for details) and a list of 2621 gene IDs was compiled including those genes for which altered methylation was reported in at least 2 of the selected studies (See Supplementary Material).

Folate-responsive and ALL-associated gene lists were then compared using an online list comparison tool (http://jura.wi.mit.edu/bioc/tools/compare.php) with gene symbols as the common identifiers, and excluding case sensitivity.

Study population (and control samples) and ALL patient samples

DNA methylation was measured at selected loci (*ASCL2, KCNA1, SH3GL3, SRD5A2*) in 20 DNA samples taken at diagnosis from either peripheral blood (n=3) or bone marrow (n=17) of ALL patients treated on the ALL97 trial [18]. Mean age of the patients was 1.15 years. Loci were selected based on having been reported in the literature to be aberrantly methylated in ALL and having had altered DNA methylation in response to modulated maternal folate intake in mice (as above). The purpose of these measurements was to confirm that our assay detected expected (ALL-associated) changes in DNA methylation and hence was appropriate for use to measure DNA methylation at these same loci in samples for which data on human exposure to folate and other B vitamins *in utero* were available. These cases included *KMT2A* (MLL) rearrangements (n=2), high hyperdiploidy (n=7) and B-other subgroups (n=11). Participating centres obtained approval from the local ethics committee and written informed consent from patients or parents [18].

7

The samples for which data on folate exposure were available was a nested cohort study population drawn from the North Cumbria Community Genetics Project [43], which included 148 cord blood DNA samples (mean (± SD) gestation = 39.6 (1.3) weeks) in which gene specific DNA methylation was measured. This prospective, unselected, population-based cohort was collected between 1996 and 2003 at a single maternity unit in West Cumbria, UK. Mothers were recruited at their first antenatal appointment (mean (SD) gestation = 10.8 (4.7) weeks) where they completed a health and lifestyle questionnaire and provided blood samples. Cord blood from the infant and delivery details were collected at birth. Maternal and infant (cord) red blood cell (RBC) folate and serum vitamin B_{12} analyses were conducted on an aliquot of whole blood removed and stabilised prior to DNA extraction at the time of sampling - i.e. 1st trimester for maternal sample and at birth for cord blood samples. These samples were also used as the control population for comparison with DNA methylation patterns in ALL samples. Summary statistics are provided in the supplementary material. Ethical approval to undertake this study was obtained from the Newcastle and North Tyneside Local Research Ethics Committee (07/Q0906/5). Written informed consent was obtained from all participating mothers recruited during pregnancy. Consent was obtained for use of biological samples (including DNA) from each mother-child dyad for epidemiological studies.

Bisulfite Pyrosequencing® for loci-specific DNA methylation analysis

Bisulfite conversion of DNA was performed using the EZ DNA Methylation Gold™ kit (Zymo Research) following the manufacturer's protocol. Briefly, 500ng of genomic DNA was incubated with CT conversion reagent and incubated at the following temperatures; 98°C

for 10 min, 64°C for 2.5hr, held at 4°C. DNA was then transferred to a spin column, washed, desulphonated and purified, finally eluting in a 10µl volume.

Quantitative bisulfite Pyrosequencing® was used to determine the percentage methylation at individual CpG sites within the promoter regions of the selected genes (number of CpGs: *ASCL2* n=7, *KCNA1* n=5, *SH3GL3* n=9, *SRD5A2* n=5). Briefly, 20ng of bisulfite treated DNA was added as a template in PCR reaction using 10µl Hot Star Taq mastermix and 1 µl of each forward and reverse primer (Qiagen) in a total volume 20µl. All primer sequences and annealing temperatures for each assay are given in the Supplementary material. Amplification was carried out in a Veriti® 96-Well fast thermocycler (Life Technologies) using the following protocol; 95°C 15 min, then 50 cycles of 95°C for 15 seconds, assay-specific annealing temperature for 30 seconds, 72°C for 15 seconds, followed by 72°C for 5 min, and a final holding step at 4°C. Biotin-labelled PCR products were captured with streptavidin sepharose beads (GE Healthcare), and made single stranded using sodium hydroxide denaturation with the aid of a Pyrosequencing® Vacuum Prep Tool (Qiagen). Sequencing primers were annealed to the single stranded PCR product by heating to 80°C, followed by slow cooling. Pyrosequencing® was then carried out on a Pyromark MD system. Cytosine methylation was quantified using proprietary PyroQ CpG 1.0.6 software. All PCR and Pyrosequencing® reactions were carried out in duplicate.

For each assay, 0% and 100% methylated controls (Epitect, Qiagen) were routinely run alongside sample amplification as internal controls. Assays were validated to rule out any amplification bias of primers and to assess assay reproducibility using methods described previously [44].   All primer sets were found to be unbiased and assays were reproducible.

Statistical Analysis

To determine the distribution, one sample Kolmogorov Smirov tests were applied to CpG specific and mean methylation data, with all data sets shown to be normally distributed. Pearson's correlation was used to assess correlation between percentage methylation at individual CpGs and mean percentage methylation across all CpGs for each gene. Significant correlations were observed for each individual CpG with mean methylation for all genes investigated (Supplementary Materials). To reduce false positive associations through multiple testing, initially mean methylation for each gene was examined in relation to the exposure variables maternal RBC folate, infant RBC folate, maternal vitamin B_{12} and infant vitamin B_{12} using linear regression. Estimated power to detect a co-efficient>0.2 with a sample size of 148 is 68% at p=0.05. For those genes where associations between exposure variables and mean methylation were found to be statistically significant (i.e. *p*-value<0.05), each individual CpG in the given gene was analysed in relation to the specific exposure variable of interest using linear regression. Comparisons of mean percentage methylation between 'healthy' infants and ALL patients were analysed by univariate ANOVA. Estimated power to detect a 10% methylation change assuming a conservative standard deviation estimate of 9 in each group with an n=20 is 94% at p=0.05. All statistical analyses were performed using SPSS Version 19 software (IBM Statistical Package for the Social Sciences) and a *p*-value of <0.05 was considered statistically significant.

Results

Comparison of folate and ALL-associated methylation signatures

Of the 2621 genes reported to have altered DNA methylation in ALL in at least 2 studies, 793 did not have corresponding identifiers on the mouse array used to assess methylation changes in response to maternal folate depletion, and thus could not be directly compared. The remaining potentially comparable 1828 ALL-associated genes were compared with the 526 genes observed to have altered methylation associated with maternal folate depletion, with 66 genes found to overlap (Table 1). Of these 66 genes, 34 were reported with a consistent direction of methylation change in ALL across at least two studies (i.e. 31 hypermethylated; 3 hypomethylated), and therefore considered for further analysis.

Twenty-five of the 34 genes (73.6%) that were reported to have a consistent direction of methylation change in ALL across studies displayed the same direction of methylation change in response to maternal folate depletion - i.e. hypermethylation. Of the 34 genes with consistent methylation change across studies, 4 were selected at random for DNA methylation quantification of the promoter region by pyrosequencing analysis (*ASCL2, KCNA1, SH3GL3* and *SRD5A2*). Methylation was assessed in ALL cases to confirm methylation at the region of the gene measured by these newly-designed pyrosequencing assays was detectably higher in ALL disease samples compared with a healthy population (i.e. cord blood samples). Whilst use of the cord blood samples as a comparator has limitations due to lack of age and sample type matching, it is appropriate in this instance to merely confirm the utility of the assays to capture ALL-associated hypermethylation previously reported elsewhere. As expected, mean methylation of all selected genes were statistically significantly higher in blood/bone marrow samples from ALL patients when

compared with infant cord blood samples (Figure 2). We then subsequently tested for associations between DNA methylation and measures of folate and vitamin B_{12} exposures.

Association between maternal RBC folate and vitamin B_{12} on infant gene methylation at birth

A statistically significant inverse relationship was observed between mean DNA methylation of the *SH3GL3* gene and maternal RBC folate (p=0.008; Coefficient (95%CI) 0.218 (-0.002-0.000)) Figure 3C. This pattern of altered methylation was seen at multiple CpG sites across the region (Supplementary Material). No associations were observed between mean methylation at the *ASCL2, KCNA1* and *SRD5A2* loci and maternal RBC folate status during pregnancy (Figure 3A, B & D).

Furthermore, no associations were observed between mean methylation of any of the loci investigated and maternal vitamin B_{12} status (Figure 4).

Association between infant RBC folate and vitamin B_{12} with infant gene methylation at birth

Mean methylation at the loci investigated were not significantly associated with infant RBC folate status at birth (Figure 5).

A statistically significant inverse relationship was observed between mean methylation of the *ASCL2* gene and infant vitamin B_{12} concentration (p=0.016; Coefficient (95%CI) 0.201 (-0.002-0.000)) Figure 6A. This relationship was observed for multiple CpG sites across the

region (Supplementary Material). Mean methylation at the *KCNA1*, *SH3GL3* and *SRD5A2* loci were not significantly associated with infant vitamin B_{12} status at birth (Figure 6B-D).

Discussion

We hypothesised that, since folate is a key nutrient driving one carbon metabolism towards the production of SAM for the methylation of biological molecules, aberrant DNA methylation may be an underlying mechanism for the observed inverse associations between maternal folate intake during pregnancy and childhood ALL risk. To investigate this hypothesis we employed a 'meet in the middle' approach [40] to identify overlap between folate and ALL-associated methylation signatures to test the plausibility that folate-associated altered methylation may be mechanistically involved in the causal pathway towards childhood ALL (Figure 1). Of the genes reported to have a consistent direction of methylation change in ALL across studies (n=34), 73.5% also displayed the same direction of methylation change in response to folate depletion. Thus, in individuals harbouring ALL-initiating genetic abnormalities, maternal folate depletion may contribute to the causal disease pathway through altered DNA methylation of ALL-associated genes.

In this study we utilised previously published data from our mouse model [29, 33] to identify genes with methylation change associated with maternal folate depletion. Whilst human studies investigating the influence of maternal folate intake on offspring methylation have been conducted, there is a lack of reproducibility between these studies, with no concordant gene loci found between data sets (reviewed in [45]). Whilst this is plausibly due to differences in sample type, study population (including genetic diversity), study design, platform used to assess methylation and statistical analyses, confounding factors are likely to be major contributors to the observed differences between human studies. Since data

derived from experimental studies in inbred models are not subject to the same level of reporting bias, genetic heterogeneity and potential confounding factors as human studies due to more strict control of experimental parameters, this novel approach enabled us to include truly folate-responsive gene loci in our analysis. In this way, data from animal models may be useful in informing human studies where associations between environmental exposures and methylation patterns may be difficult to determine due to a) the difficulty in measuring or assessing certain exposures b) the large numbers of participants required to reveal small changes in methylation in association with poorly defined exposure variables and c) the large number of potential confounding factors in human studies.

It is pertinent to highlight, that the ALL-associated methylation signature was derived from studies measuring methylation in diseased cells, and therefore some of the aberrant methylation signature is likely to be a consequence of disease rather than causal. Therefore, to distinguish those methylation marks that are causal for, from those that are consequences of, disease, measurements prior to diagnosis would be required. Given the rarity of childhood ALL and lack of available pre-diagnostic samples, currently very little is known about early initiating epigenetic events in the development of childhood ALL. The utilization of historical collections of neonatal blood spot samples or collective cases from multiple large cohort studies may facilitate investigation of pre-diagnostic methylation patterns, but are unlikely to have associated exposure data or may be underpowered to detect epigenetic changes in association with exposure. Therefore, novel methodologies, such as the current meet in the middle approach, are required to ascertain the plausibility of biological mechanisms by which environment may influence childhood ALL development.

Further investigations using approaches such as one and two step Mendelian Randomisation approaches[46] will be key in providing evidence of causality.

To seek evidence for an effect of folate exposure *in utero* in humans on DNA methylation of genes found to be hypermethylated in ALL and therefore of potential relevance to disease causality, we assessed methylation of a subset of target genes (*ASCL2*, *KCNA1*, *SH3GL3* and *SRD5A2*) in healthy newborns in relation to maternal and infant folate status, and in ALL patients. Given that folate and B_{12} are key micronutrients driving one carbon metabolism, and that maternal vitamin B_{12} has been associated with offspring ALL risk [3], we also investigated the association between maternal and infant B_{12} status and methylation of these target genes. Lower maternal folate status was significantly associated with higher methylation of the *SH3GL3* locus. Therefore, since low maternal folate is associated with increased ALL risk [1-7] and *SH3GL3* hypermethylation is reported in ALL [20, 21, 24], the observed association between low folate and *SH3GL3* hypermethylation supports our hypothesis that low maternal folate intake during pregnancy may increase risk of ALL via hypermethylation of the *SH3GL3* locus. In particular, infants born to mothers with above normal (600ng/ml) RBC folate levels tended to have the lowest *SH3GL3* methylation, which may indicate that higher than normal levels of folate are protective. This theory is supported by evidence that folate supplementation is protective against ALL [1, 2, 4, 5, 7]. We also report a significant inverse association between infant vitamin B_{12} status and *ASCL2* methylation. This finding is in keeping with the hypothesis that aberrant DNA methylation may be one potential biological mechanism facilitating the previously observed association between low vitamin B_{12} status and ALL risk i.e. low vitamin B_{12} status may increase ALL risk through hypermethylation of *ASCL2* and, possibly, other genes [3].

Whilst the associations observed between *SH3GL3* and *ASCL2* methylation and maternal folate and infant B_{12} status respectively are statistically significant, the range of methylation differences that we report are fairly small. Whilst these relatively small changes in methylation in response to environmental factors are a common finding in epidemiological studies [34, 47], the functional consequences of such small effects have often been questioned. It is important to bear in mind that, whilst DNA methylation is often reported as a percentage (i.e. reports the fraction of the DNA copies in a given sample that are methylated), methylation itself is a binary phenomenon, where for any given copy of DNA, specific CpGs are either methylated or unmethylated. It is therefore the methylation pattern of a specific cell that influences gene expression patterns of that individual cell. Should such epigenetic, and thus, gene expression changes occur in individual cells harboring pre-leukaemic genetic aberrations, they may act as the "second hit" in the causal pathway, thus resulting in the transition to a leukaemic cell.

Whilst these findings support our hypothesis that altered DNA methylation may be a "second hit" that is causal for the development of ALL, it is pertinent to note we observed a significant association between folate status and methylation for only one of the four selected loci investigated. There are several potential explanations for this finding. It is plausible that the genes selected for investigation may not be folate-responsive in humans, or that methylation is altered at these loci only in response to more severe levels of depletion than were observed in this study (the majority (75%) of participants had folate levels within the normal range i.e. 200-600ng/ml; with 7.4% and 17.6% being below and above these levels respectively). Additionally, given the utilisation of SAM across a range of biological pathways, the relationship between folate status, SAM and DNA methylation is

complex and therefore not always linear. Moreover, these relationships are highly tissue-specific, therefore changes in gene-specific methylation observed in the murine liver in response to maternal folate intake may not be similar to folate responsive genes in human blood samples. Furthermore, whilst pyrosequencing is the gold standard measure of DNA methylation at specific CpG loci, this technique measures only a limited number of CpG's within a small genomic region (approximately 50-80bp). It is therefore possible, that methylation at other CpG sites not assayed here may be altered in response to low folate intake. Finally, the small population size (n=148 mother-child dyads) may have limited our power to detect significant relationships between gene-specific methylation and B-vitamin exposure.

To our knowledge, this is the first study to address the potential for a plausible biological epigenetic mechanism through which low maternal folate status may adversely influence the causal pathway towards childhood ALL development. Furthermore, we provide proof of principle that the meet in the middle approach used to test our hypothesis (Figure 1) is effective in its triangulation of evidence for complementary sources, using data derived from animal studies to inform candidate markers of exposure in human studies. Whilst further investigation of the role of DNA methylation in this disease pathway is warranted, this novel approach will allow further investigation of the role of folate (and other environmental factors) in ALL development. Improved understanding of the causes and underlying mechanisms that promote disease development, may lead to preventative (e.g. public health campaigns for improved nutrition during pregnancy) and predictive (e.g. biomarkers for high risk individuals such as those with Downs syndrome) strategies, to reduce the burden of childhood ALL.

Authors' Contributions

JAM designed the study and drafted the manuscript. JAM compiled and determined the altered- methylation in ALL gene list from the literature, carried out laboratory analysis (with the exception of pyrosequencing analysis of the *ASCL2* gene) and statistical analysis. CH carried out data analysis and interpretation. GS contributed to study design. AVM and CR provided samples for analysis. DF and JCM give permission to access and utilise genome-wide methylation data from mouse model of folate depletion. All authors contributed to revisions of, and approved, the final manuscript.

Conflict of interest statement

The authors report no conflicts of interests**.**

References

1.   Ajrouche, R., J. Rudant, L. Orsi, A. Petit, A. Baruchel, B. Nelken, M. Pasquet, G. Michel, C. Bergeron, S. Ducassou, V. Gandemer, P. Lutz, L. Saumet, X. Rialland, D. Hemon, and J. Clavel, *Maternal reproductive history, fertility treatments and folic acid supplementation in the risk of childhood acute leukemia: the ESTELLE study.* Cancer Causes Control, 2014. **25**(10): p. 1283-93.

2.   Amigou, A., J. Rudant, L. Orsi, S. Goujon-Bellec, G. Leverger, A. Baruchel, Y. Bertrand, B. Nelken, G. Plat, G. Michel, S. Haouy, P. Chastagner, S. Ducassou, X. Rialland, D. Hemon, and J. Clavel, *Folic acid supplementation, MTHFR and MTRR polymorphisms, and the risk of childhood leukemia: the ESCALE study (SFCE).* Cancer Causes Control, 2012. **23**(8): p. 1265-77.

3.   Bailey, H.D., M. Miller, A. Langridge, N.H. de Klerk, F.M. van Bockxmeer, J. Attia, R.J. Scott, B.K. Armstrong, and E. Milne, *Maternal dietary intake of folate and vitamins B6 and B12 during pregnancy and the risk of childhood acute lymphoblastic leukemia.* Nutr Cancer, 2012. **64**(7): p. 1122-30.

4.   Dessypris, N., M.A. Karalexi, E. Ntouvelis, A.A. Diamantaras, V. Papadakis, M. Baka, E. Hatzipantelis, M. Kourti, M. Moschovi, S. Polychronopoulou, V. Sidi, E. Stiakaki, and E.T. Petridou, *Association of maternal and index child's diet with subsequent leukemia risk: A systematic review and meta analysis.* Cancer Epidemiol, 2017. **47**: p. 64-75.

5.   Metayer, C., E. Milne, J.D. Dockerty, J. Clavel, M.S. Pombo-de-Oliveira, C. Wesseling, L.G. Spector, J. Schuz, E. Petridou, S. Ezzat, B.K. Armstrong, J. Rudant, S. Koifman, P. Kaatsch, M. Moschovi, W.M. Rashed, S. Selvin, K. McCauley, R.J. Hung, A.Y. Kang, and C. Infante-Rivard, *Maternal supplementation with folic acid and other vitamins and risk of leukemia in*

offspring: a Childhood Leukemia International Consortium study. Epidemiology, 2014. **25**(6): p. 811-22.

6.   Singer, A.W., S. Selvin, G. Block, C. Golden, S.L. Carmichael, and C. Metayer, *Maternal prenatal intake of one-carbon metabolism nutrients and risk of childhood leukemia.* Cancer Causes Control, 2016. **27**(7): p. 929-40.

7.   Thompson, J.R., P.F. Gerald, M.L. Willoughby, and B.K. Armstrong, *Maternal folate supplementation in pregnancy and protection against acute lymphoblastic leukaemia in childhood: a case-control study.* Lancet, 2001. **358**(9297): p. 1935-40.

8.   Lightfoot, T.J., W.T. Johnston, D. Painter, J. Simpson, E. Roman, C.F. Skibola, M.T. Smith, J.M. Allan, and G.M. Taylor, *Genetic variation in the folate metabolic pathway and risk of childhood leukemia.* Blood, 2010. **115**(19): p. 3923-9.

9.   Lupo, P.J., D. Nousome, K.Y. Kamdar, M.F. Okcu, and M.E. Scheurer, *A case-parent triad assessment of folate metabolic genes and the risk of childhood acute lymphoblastic leukemia.* Cancer Causes Control, 2012. **23**(11): p. 1797-803.

10.   Metayer, C., G. Scelo, A.P. Chokkalingam, L.F. Barcellos, M.C. Aldrich, J.S. Chang, N. Guha, K.Y. Urayama, H.M. Hansen, G. Block, V. Kiley, J.K. Wiencke, J.L. Wiemels, and P.A. Buffler, *Genetic variants in the folate pathway and risk of childhood acute lymphoblastic leukemia.* Cancer Causes Control, 2011. **22**(9): p. 1243-58.

11.   Milne, E., K.R. Greenop, R.J. Scott, M. Haber, M.D. Norris, J. Attia, S.E. Jamieson, M. Miller, C. Bower, H.D. Bailey, S. Dawson, G.B. McCowage, N.H. de Klerk, F.M. van Bockxmeer, and B.K. Armstrong, *Folate pathway gene polymorphisms, maternal folic acid use, and risk of childhood acute lymphoblastic leukemia.* Cancer Epidemiol Biomarkers Prev, 2015. **24**(1): p. 48-56.

12.   Reddy, H. and K. Jamil, *Polymorphisms in the MTHFR gene and their possible association with susceptibility to childhood acute lymphocytic leukemia in an Indian population.* Leuk Lymphoma, 2006. **47**(7): p. 1333-9.

13.   Greaves, M., *In utero origins of childhood leukaemia.* Early Hum Dev, 2005. **81**(1): p. 123-9.

14.   Greaves, M.F., A.T. Maia, J.L. Wiemels, and A.M. Ford, *Leukemia in twins: lessons in natural history.* Blood, 2003. **102**(7): p. 2321-33.

15.   Greaves, M.F. and J. Wiemels, *Origins of chromosome translocations in childhood leukaemia.* Nat Rev Cancer, 2003. **3**(9): p. 639-49.

16.   Mori, H., S.M. Colman, Z. Xiao, A.M. Ford, L.E. Healy, C. Donaldson, J.M. Hows, C. Navarrete, and M. Greaves, *Chromosome translocations and covert leukemic clones are generated during normal fetal development.* Proc Natl Acad Sci U S A, 2002. **99**(12): p. 8242-7.

17.   Smith, M.T., C.M. McHale, J.L. Wiemels, L. Zhang, J.K. Wiencke, S. Zheng, L. Gunn, C.F. Skibola, X. Ma, and P.A. Buffler, *Molecular biomarkers for the study of childhood leukemia.* Toxicol Appl Pharmacol, 2005. **206**(2): p. 237-45.

18.   Moorman, A.V., H.M. Ensor, S.M. Richards, L. Chilton, C. Schwab, S.E. Kinsey, A. Vora, C.D. Mitchell, and C.J. Harrison, *Prognostic effect of chromosomal abnormalities in childhood B-cell precursor acute lymphoblastic leukaemia: results from the UK Medical Research Council ALL97/99 randomised trial.* Lancet Oncol, 2010. **11**(5): p. 429-38.

19.   Mathers, J.C. and J.A. McKay, *Epigenetics - potential contribution to fetal programming.* Adv Exp Med Biol, 2009. **646:119-23.**(doi): p. 10.1007/978-1-4020-9173-5_13.

20.   Almamun, M., B.T. Levinson, A.C. van Swaay, N.T. Johnson, S.D. McKay, G.L. Arthur, J.W. Davis, and K.H. Taylor, *Integrated methylome and transcriptome analysis reveals novel*

regulatory elements in pediatric acute lymphoblastic leukemia. Epigenetics, 2015. **10**(9): p. 882-90.

21.   Chatterton, Z., L. Morenos, F. Mechinaud, D.M. Ashley, J.M. Craig, A. Sexton-Oates, M.S. Halemba, M. Parkinson-Bates, J. Ng, D. Morrison, W.L. Carroll, R. Saffery, and N.C. Wong, *Epigenetic deregulation in pediatric acute lymphoblastic leukemia.* Epigenetics, 2014. **9**(3): p. 459-67.

22.   Davidsson, J., H. Lilljebjorn, A. Andersson, S. Veerla, J. Heldrup, M. Behrendtz, T. Fioretos, and B. Johansson, *The DNA methylome of pediatric acute lymphoblastic leukemia.* Hum Mol Genet, 2009. **18**(21): p. 4054-65.

23.   Dunwell, T., L. Hesson, T.A. Rauch, L. Wang, R.E. Clark, A. Dallol, D. Gentle, D. Catchpoole, E.R. Maher, G.P. Pfeifer, and F. Latif, *A genome-wide screen identifies frequently methylated genes in haematological and epithelial cancers.* Mol Cancer, 2010. **9**: p. 44.

24.   Nordlund, J., C.L. Backlin, P. Wahlberg, S. Busche, E.C. Berglund, M.L. Eloranta, T. Flaegstad, E. Forestier, B.M. Frost, A. Harila-Saari, M. Heyman, O.G. Jonsson, R. Larsson, J. Palle, L. Ronnblom, K. Schmiegelow, D. Sinnett, S. Soderhall, T. Pastinen, M.G. Gustafsson, G. Lonnerholm, and A.C. Syvanen, *Genome-wide signatures of differential DNA methylation in pediatric acute lymphoblastic leukemia.* Genome Biol, 2013. **14**(9): p. r105.

25.   Sandoval, J., H. Heyn, J. Mendez-Gonzalez, A. Gomez, S. Moran, M. Baiget, M. Melo, I. Badell, J.F. Nomdedeu, and M. Esteller, *Genome-wide DNA methylation profiling predicts relapse in childhood B-cell acute lymphoblastic leukaemia.* Br J Haematol, 2013. **160**(3): p. 406-9.

26.   Schafer, E., R. Irizarry, S. Negi, E. McIntyre, D. Small, M.E. Figueroa, A. Melnick, and P. Brown, *Promoter hypermethylation in MLL-r infant acute lymphoblastic leukemia: biology and therapeutic targeting.* Blood, 2010. **115**(23): p. 4798-809.

27.   Mathers, J.C., G. Strathdee, and C.L. Relton, *Induction of epigenetic alterations by dietary and other environmental factors.* Adv Genet, 2010. **71**: p. 3-39.

28.   Kim, J.M., K. Hong, J.H. Lee, S. Lee, and N. Chang, *Effect of folate deficiency on placental DNA methylation in hyperhomocysteinemic rats.* J Nutr Biochem, 2009. **20**(3): p. 172-6.

29.   McKay, J.A., M. Adriaens, C.T. Evelo, D. Ford, and J.C. Mathers, *Gene promoter DNA methylation patterns have a limited role in orchestrating transcriptional changes in the fetal liver in response to maternal folate depletion during pregnancy.* Mol Nutr Food Res, 2016. **60**(9): p. 2031-42.

30.   McKay, J.A., K.J. Waltham, E.A. Williams, and J.C. Mathers, *Folate depletion during pregnancy and lactation reduces genomic DNA methylation in murine adult offspring.* Genes Nutr, 2011. **6**(2): p. 189-96.

31.   McKay, J.A., E.A. Williams, and J.C. Mathers, *Effect of maternal and post-weaning folate supply on gene-specific DNA methylation in the small intestine of weaning and adult apc and wild type mice.* Front Genet, 2011. **2**: p. 23.

32.   McKay, J.A., Y.K. Wong, C.L. Relton, D. Ford, and J.C. Mathers, *Maternal folate supply and sex influence gene-specific DNA methylation in the fetal gut.* Mol Nutr Food Res, 2011. **55**(11): p. 1717-23.

33.   McKay, J.A., L. Xie, M. Adriaens, C.T. Evelo, D. Ford, and J.C. Mathers, *Maternal folate depletion during early development and high fat feeding from weaning elicit similar changes in gene expression, but not in DNA methylation, in adult offspring.* Mol Nutr Food Res, 2017. **61**(4).

34.   Joubert, B.R., H.T. den Dekker, J.F. Felix, J. Bohlin, S. Ligthart, E. Beckett, H. Tiemeier, J.B. van Meurs, A.G. Uitterlinden, A. Hofman, S.E. Haberg, S.E. Reese, M.J. Peters, B.K. Andreassen,

E.A. Steegers, R.M. Nilsen, S.E. Vollset, O. Midttun, P.M. Ueland, O.H. Franco, A. Dehghan, J.C. de Jongste, M.C. Wu, T. Wang, S.D. Peddada, V.W. Jaddoe, W. Nystad, L. Duijts, and S.J. London, *Maternal plasma folate impacts differential DNA methylation in an epigenome-wide meta-analysis of newborns.* Nat Commun, 2016. **7**: p. 10577.

35.   Steegers-Theunissen, R.P., S.A. Obermann-Borst, D. Kremer, J. Lindemans, C. Siebel, E.A. Steegers, P.E. Slagboom, and B.T. Heijmans, *Periconceptional maternal folic acid use of 400 microg per day is related to increased methylation of the IGF2 gene in the very young child.* PLoS One, 2009. **4**(11): p. e7845.

36.   Kok, D.E., R.A. Dhonukshe-Rutten, C. Lute, S.G. Heil, A.G. Uitterlinden, N. van der Velde, J.B. van Meurs, N.M. van Schoor, G.J. Hooiveld, L.C. de Groot, E. Kampman, and W.T. Steegenga, *The effects of long-term daily folic acid and vitamin B12 supplementation on genome-wide DNA methylation in elderly subjects.* Clin Epigenetics, 2015. **7**: p. 121.

37.   McCullough, L.E., E.E. Miller, M.A. Mendez, A.P. Murtha, S.K. Murphy, and C. Hoyo, *Maternal B vitamins: effects on offspring weight and DNA methylation at genomically imprinted domains.* Clin Epigenetics, 2016. **8**: p. 8.

38.   McKay, J.A., A. Groom, C. Potter, L.J. Coneyworth, D. Ford, J.C. Mathers, and C.L. Relton, *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**(3): p. e33290.

39.   Forster, V.J., A. McDonnell, R. Theobald, and J.A. McKay, *Effect of methotrexate/vitamin B12 on DNA methylation as a potential factor in leukemia treatment-related neurotoxicity.* Epigenomics, 2017. **9**(9): p. 1205-1218.

40.   Vineis, P. and F. Perera, *Molecular epidemiology and biomarkers in etiologic cancer research: the new in light of the old.* Cancer Epidemiol Biomarkers Prev, 2007. **16**(10): p. 1954-65.

41.   McKay, J.A., L. Xie, S. Harris, Y.K. Wong, D. Ford, and J.C. Mathers, *Blood as a surrogate marker for tissue-specific DNA methylation and changes due to folate depletion in post-partum female mice.* Mol Nutr Food Res., 2011. **55**(7): p. 1026-35. doi: 10.1002/mnfr.201100008. Epub 2011 Mar 24.

42.   McKay, J.A., L. Xie, C. Manus, S.A. Langie, R.J. Maxwell, D. Ford, and J.C. Mathers, *Metabolic effects of a high-fat diet post-weaning after low maternal dietary folate during pregnancy and lactation.* Mol Nutr Food Res., 2014. **58**(5): p. 1087-97. doi: 10.1002/mnfr.201300615. Epub 2014 Jan 15.

43.   Chase, D.S., E.J. Tawn, L. Parker, P. Jonas, C.O. Parker, and J. Burn, *The North Cumbria Community Genetics Project.* J Med Genet, 1998. **35**(5): p. 413-6.

44.   White, H.E., V.J. Durston, J.F. Harvey, and N.C. Cross, *Quantitative analysis of SNRPN(correction of SRNPN) gene methylation by pyrosequencing as a diagnostic test for Prader-Willi syndrome and Angelman syndrome.* Clin Chem, 2006. **52**(6): p. 1005-13.

45.   Richmond, R.C. and B.R. Joubert, *Contrasting the effects of intra-uterine smoking and one-carbon micronutrient exposures on offspring DNA methylation.* Epigenomics, 2017. **9**(3): p. 351-367.

46.   Relton, C.L. and G. Davey Smith, *Two-step epigenetic Mendelian randomization: a strategy for establishing the causal role of epigenetic processes in pathways to disease.* Int J Epidemiol., 2012. **41**(1): p. 161-76. doi: 10.1093/ije/dyr233.

47.   Markunas, C.A., Z. Xu, S. Harlid, P.A. Wade, R.T. Lie, J.A. Taylor, and A.J. Wilcox, *Identification of DNA methylation changes in newborns related to maternal smoking during pregnancy.* Environ Health Perspect, 2014. **122**(10): p. 1147-53.

Table 1. Genes with altered methylation in childhood ALL and in response to maternal folate depletion, showing the direction of methylation change. Where both Hyper/Hypo is stated, differential methylation change was observed either between studies or within a given study. *Denotes where reported direction of methylation change was consistent BETWEEN studies, but both hyper and hypomethylation were reported for that gene WITHIN ONE study.

Gene ID	Hyper/ Hypomethylated in childhood ALL	Hyper/ Hypomethylated in response to maternal folate depletion	Gene ID	Hyper/ Hypomethylated in childhood ALL	Hyper/ Hypomethylated in response to maternal folate depletion
Abca4	Hyper/Hypo	Hyper	*Limch1*	Hyper	Hyper
Adrb3	Hyper	Hyper	*Mcoln2*	Hyper/Hypo	Hyper
Aldh1l1	Hyper	Hyper	*Meox2*	Hyper	Hyper
Ank3	Hyper	Hypo	*Mocs1*	Hyper	Hyper
Ankk1	Hyper/Hypo*	Hypo	*Mrps27*	Hyper/Hypo	Hypo
Ascl2	Hyper	Hyper	*Nde1*	Hyper/Hypo	Hypo
Btbd11	Hyper/Hypo*	Hyper	*Npy5r*	Hyper	Hyper
Ccdc37	Hyper	Hyper	*P4ha3*	Hyper/Hypo*	Hyper
Cdh22	Hyper	Hyper	*Palm2-akap2*	Hyper/Hypo	Hyper
Chst15	Hypo	Hyper	*Pcdha1*	Hyper	Hyper
Clip4	Hyper/Hypo	Hyper	*Pcdha11*	Hyper	Hyper

Cxadr	Hyper	Hyper	*Pcdhb6*	Hyper	Hyper
Dapk2	Hyper/Hypo	Hyper	*Pdlim3*	Hyper	Hyper
Dntt	Hypo	Hyper	*Phactr3*	Hyper/Hypo*	Hyper
Dupd1	Hyper	Hyper	*Prdm14*	Hyper	Hypo
Dusp22	Hyper/Hypo*	Hyper	*Rgs10*	Hyper/Hypo	Hyper
Edn3	Hyper	Hyper	*Rin2*	Hyper/Hypo	Hypo
Eif2c2	Hyper/Hypo	Hyper	*Rspo4*	Hyper	Hyper
Eif5a2	Hyper	Hypo	*Sh3gl3*	Hyper	Hyper
Ern1	Hyper/Hypo	Hyper	*Sh3rf3*	Hyper/Hypo	Hyper
Gabra5	Hyper/Hypo*	Hyper	*Slc12a7*	Hyper/Hypo*	Hyper
Galnt2	Hyper/Hypo*	Hyper	*Slc46a1*	Hyper	Hyper
Gas2	Hyper	Hypo	*Slc6a1*	Hyper/Hypo*	Hyper
Gnao1	Hyper/Hypo*	Hypo	*Sox30*	Hyper	Hypo
Gpr157	Hyper/Hypo	Hyper	*Spag17*	Hyper	Hyper
Hipk1	Hyper/Hypo	Hyper	*Srd5a2*	Hyper	Hypo
Hdac11	Hyper/Hypo	Hyper	*Stx2*	Hyper/Hypo	Hyper
Jakmip1	Hyper/Hypo*	Hyper	*Tmc5*	Hyper/Hypo	Hyper
Kcna1	Hyper	Hyper	*Tmem132c*	Hyper	Hyper
Kcnip1	Hyper/Hypo	Hyper	*Traf3*	Hyper/Hypo*	Hyper
Kcnk2	Hyper	Hyper	*Trpm3*	Hypo	Hyper
Kdr	Hyper	Hyper	*Usp10*	Hyper/Hypo	Hyper
Kif1b	Hyper/Hypo	Hyper	*Vsx1*	Hyper	Hyper