Association of dietary folate and vitamin B-12 intake with genome-wide DNA methylation in blood: a large-scale epigenome-wide association analysis in 5841 individuals

Pooja R Mandaviya,[1,2] Roby Joehanes,[3,4,5] Jennifer Brody,[6] Juan E Castillo-Fernandez,[7] Koen F Dekkers,[8] Anh N Do,[9] Mariaelisa Graff,[10] Ismo K Hänninen,[11] Toshiko Tanaka,[12] Ester AL de Jonge,[1,13] Jessica C Kiefte-de Jong,[13,14] Devin M Absher,[15] Stella Aslibekyan,[16] Yolanda B de Rijke,[2] Myriam Fornage,[17] Dena G Hernandez,[18] Mikko A Hurme,[19] M Arfan Ikram,[13] Paul F Jacques,[20,21] Anne E Justice,[22] Douglas P Kiel,[3,4] Rozenn N Lemaitre,[6] Michael M Mendelson,[5,23] Vera Mikkilä,[24] Ann Z Moore,[12] Tess Pallister,[7] Olli T Raitakari,[25] Casper G Schalkwijk,[26] Jin Sha,[27] Eline PE Slagboom,[8] Caren E Smith,[20] Coen DA Stehouwer,[26] Pei-Chien Tsai,[7,28,29] André G Uitterlinden,[1] Carla JH van der Kallen,[26] Diana van Heemst,[30] BIOS Consortium, CHARGE Consortium Epigenetics Working Group, CHARGE Consortium Nutrition Working Group, Donna K Arnett,[31] Stefania Bandinelli,[32] Jordana T Bell,[7] Bastiaan T Heijmans,[8] Terho Lehtimäki,[11] Daniel Levy,[5] Kari E North,[10] Nona Sotoodehnia,[6] Marleen MJ van Greevenbroek,[26] Joyce BJ van Meurs,[1] and Sandra G Heil[2]

[1]Department of Internal Medicine, and [2]Department of Clinical Chemistry, Erasmus University Medical Center, Rotterdam, The Netherlands; [3]Institute for Aging Research, Hebrew SeniorLife, Department of Medicine, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA; [4]Broad Institute of MIT & Harvard, Cambridge, MA; [5]Framingham Heart Study, National Heart, Lung, and Blood Institute, NIH, Framingham, MA; [6]Department of Medicine, University of Washington, Seattle, WA; [7]Department of Twin Research and Genetic Epidemiology, King's College London, London, United Kingdom; [8]Molecular Epidemiology, Department of Biomedical Data Sciences, Leiden University Medical Center, Leiden, The Netherlands; [9]Institute for Genomics and Multiscale Biology, Icahn School of Medicine at Mount Sinai, New York, NY; [10]Department of Epidemiology, University of North Carolina, Chapel Hill, NC; [11]Department of Clinical Chemistry, Fimlab Laboratories, and Finnish Cardiovascular Research Center–Tampere, Faculty of Medicine and Life Sciences, University of Tampere, Tampere, Pirkanmaa, Finland; [12]Translational Gerontology Branch, National Institute on Aging, Baltimore, MD; [13]Department of Epidemiology, Erasmus University Medical Center, Rotterdam, The Netherlands; [14]Faculty of Governance and Global Affairs, Leiden University College, The Hague, The Netherlands; [15]HudsonAlpha Institute of Biotechnology, Huntsville, AL; [16]Department of Epidemiology, University of Alabama at Birmingham, Birmingham, AL; [17]Brown Foundation Institute of Molecular Medicine and Human Genetics Center, University of Texas Health Science Center at Houston, Houston, TX; [18]Laboratory of Neurogenetics, National Institute on Aging, Bethesda, MD; [19]Department of Microbiology and Immunology, Faculty of Medicine and Life Sciences, University of Tampere, Tampere, Pirkanmaa, Finland; [20]USDA Human Nutrition Research Center on Aging; [21]Friedman School of Nutrition Science and Policy, Tufts University, Boston, MA; [22]Biomedical and Translational Informatics, Geisinger Health, Danville, PA; [23]Department of Cardiology, Boston Children's Hospital, Harvard Medical School, Boston, MA; [24]Division of Nutrition, Department of Food and Environmental Sciences, Helsinki, Uusimaa, Finland; [25]Department of Clinical Physiology and Nuclear Medicine, Turku University Hospital, and Research Centre of Applied and Preventive Cardiovascular Medicine, University of Turku, Turku, Southwest Finland, Finland; [26]Department of Internal Medicine, Maastricht University Medical Centre and CARIM School for Cardiovascular Diseases, Maastricht University, Maastricht, The Netherlands; [27]Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA; [28]Department of Biomedical Sciences, Chang Gung University, Taoyuan, Taiwan; [29]Division of Allergy, Asthma, and Rheumatology, Department of Pediatrics, Chang Gung Memorial Hospital, Linkou, Taiwan; [30]Department of Internal Medicine, Section Gerontology and Geriatrics, Leiden University Medical Center, Leiden, The Netherlands; [31]College of Public Health, University of Kentucky, Lexington, KY; and [32]Geriatric Unit, Azienda Sanitaria di Firenze, Florence, Tuscany, Italy

ABSTRACT

Background: Folate and vitamin B-12 are essential micronutrients involved in the donation of methyl groups in cellular metabolism. However, associations between intake of these nutrients and genome-wide DNA methylation levels have not been studied comprehensively in humans.

Objective: The aim of this study was to assess whether folate and/or vitamin B-12 intake are asssociated with genome-wide changes in DNA methylation in leukocytes.

Methods: A large-scale epigenome-wide association study of folate and vitamin B-12 intake was performed on DNA from 5841 participants from 10 cohorts using Illumina 450k arrays. Folate and vitamin B-12 intakes were calculated from food-frequency questionnaires (FFQs). Continuous and categorical (low compared with high intake) linear regression mixed models were applied per cohort, controlling for confounders. A meta-analysis was performed to identify significant differentially methylated positions (DMPs) and regions (DMRs), and a pathway analysis was performed on the DMR annotated genes.

Results: The categorical model resulted in 6 DMPs, which are all negatively associated with folate intake, annotated to *FAM64A*, *WRAP73*, *FRMD8*, *CUX1*, and *LCN8* genes, which have a role in cellular processes including centrosome localization, cell proliferation, and tumorigenesis. Regional analysis showed 74 folate-associated DMRs, of which 73 were negatively associated with folate intake. The most significant folate-associated DMR was a 400-base

pair (bp) spanning region annotated to the *LGALS3BP* gene. In the categorical model, vitamin B-12 intake was associated with 29 DMRs annotated to 48 genes, of which the most significant was a 1100-bp spanning region annotated to the calcium-binding tyrosine phosphorylation-regulated gene (*CABYR*). Vitamin B-12 intake was not associated with DMPs.

Conclusions: We identified novel epigenetic loci that are associated with folate and vitamin B-12 intake. Interestingly, we found a negative association between folate and DNA methylation. Replication of these methylation loci is necessary in future studies. *Am J Clin Nutr* 2019;110:437–450.

Keywords: epigenetics, DNA methylation, diet, FFQ, Epigenome-wide Association Study, folate, vitamin B-12, genome-wide

Rotterdam Study (RS): The generation and management of the Illumina 450 K methylation array data (EWAS data) for the RS were executed by the Human Genotyping Facility of the Genetic Laboratory of the Department of Internal Medicine, Erasmus MC, the Netherlands. The EWAS data were funded by the Genetic Laboratory of the Department of Internal Medicine, Erasmus MC, and by the Netherlands Organization for Scientific Research (NWO; project number 184021007) and made available as a Rainbow Project (RP3; BIOS) of the Biobanking and Biomolecular Research Infrastructure Netherlands. We thank Michael Verbiest, Mila Jhamai, Sarah Higgins, and Marijn Verkerk for their help in creating the methylation database.

The RS is funded by the Erasmus Medical Center and Erasmus University, Rotterdam; the Netherlands Organization for Health Research and Development; the Research Institute for Diseases in the Elderly; the Ministry of Education, Culture, and Science; the Ministry for Health, Welfare, and Sports; the European Commission (DG XII); and the Municipality of Rotterdam. The authors are grateful to the study participants, the staff from the RS and the participating general practitioners and pharmacists.

Leiden Longevity Study (LLS): We thank all participants of the LLS. This study received funding from the European Union's Seventh Framework Programme (FP7/2007-2011) under grant agreement no. 259679, from the Innovation-Oriented Research Program on Genomics (SenterNovem IGE05007), the Centre for Medical Systems Biology, from the Netherlands Consortium for Healthy Ageing (grant 050-060-810), and from the Biobank-Based Integrative Omics Studies (BIOS) Consortium funded by Biobanking and Biomolecular Research Infrastructure Netherlands, a research infrastructure financed by the Dutch government (NWO 184.021.007), all in the framework of the Netherlands Genomics Initiative, the Netherlands Organization for Scientific Research (NWO).

Cohort study on Diabetes and Atherosclerosis Maastricht (CODAM): Part of this work was supported by grants of the Netherlands Organization for Scientific Research (940-35-034) and the Dutch Diabetes Research Foundation (98.901).

Invecchiare in Chianti (InCHIANTI) study: The InCHIANTI study baseline (1998–2000) was supported as a "targeted project" (ICS110.1/RF97.71) by the Italian Ministry of Health and in part by the US National Institute on Aging (Contracts: 263 MD 9164 and 263 MD 821336).

Young Finns Study (YFS): The YFS has been financially supported by the Academy of Finland: grants 286284, 134309 (Eye), 126925, 121584, 124282, 129378 (Salve), 117787 (Gendi), and 41071 (Skidi); the Social Insurance Institution of Finland; Competitive State Research Financing of the Expert Responsibility area of Kuopio, Tampere and Turku University Hospitals (grant X51001); the Juho Vainio Foundation; the Paavo Nurmi Foundation; the Finnish Foundation for Cardiovascular Research; the Finnish Cultural Foundation; the Tampere Tuberculosis Foundation; the Emil Aaltonen Foundation; the Yrjö Jahnsson Foundation; the Signe and Ane Gyllenberg Foundation; and the Diabetes Research Foundation of the Finnish Diabetes Association.

TwinsUK: TwinsUK was funded by the Wellcome Trust; the European Community's Seventh Framework Programme (FP7/2007–2013) and also receives support from the National Institute for Health Research (NIHR)-funded BioResource, Clinical Research Facility and Biomedical Research Centre based at Guy's and St Thomas' NHS Foundation Trust in partnership with King's College London. PCT, JEC-F, and JTB were supported by the European Commission and the UK Economic and Social Research Council (ES/N000404/1).

Framingham Heart Study (FHS): The Framingham Heart Study is funded by NIH contract N01-HC-25195. The laboratory work for this investigation was funded by the Division of Intramural Research, National Heart, Lung, and Blood Institute, NIH and an NIH Director's Challenge Award (DL, Principal Investigator). DPK was funded by NIH grant R01 AR041398.

Genetics of Lipid Lowering Drugs and Diet Network (GOLDN) Lipidomics Study: Phenotype/participant characteristic data were collected under NIH NHLBI grant U01 HL072524; methylation data were collected under NIH NHLBI R01 HL104135.

Atherosclerosis Risk in Communities (ARIC) Study: The ARIC Study has been funded in whole or in part with Federal funds from the National Heart, Lung, and Blood Institute, NIH, Department of Health and Human Services (contract numbers HHSN268201700001I, HHSN268201700002I, HHSN268201700003I, HHSN268201700004I, and HHSN268201700005I). We thank the staff and participants of the ARIC Study for their important contributions. Funding was also supported by 5RC2HL102419 and R01NS087541. AEJ was funded by NIH award K99/R00 HL130580.

Cardiovascular Health Study (CHS): Infrastructure for the CHARGE Consortium is supported in part by the National Heart, Lung, and Blood Institute grant R01HL105756. The CHS research was supported by NHLBI contracts HHSN268201200036C, HHSN268200800007C, N01HC55222, N01HC85079, N01HC85080, N01HC85081, N01HC85082, N01HC85083, N01HC85086; and NHLBI grants U01HL080295, K08HL116640, R01HL087652, R01HL092111, R01HL103612, R01HL105756, R01HL103612, R01HL111089, R01HL116747, and R01HL120393 with additional contribution from the National Institute of Neurological Disorders and Stroke. Additional support was provided through R01AG023629 from the National Institute on Aging, the Merck Foundation/Society of Epidemiologic Research as well as Laughlin Family, the Alpha Phi Foundation, and the Locke Charitable Foundation. A full list of principal CHS investigators and institutions can be found at CHS-NHLBI.org. The provision of genotyping data was supported in part by the National Center for Advancing Translational Sciences, CTSI grant UL1TR000124, and the National Institute of Diabetes and Digestive and Kidney Disease Diabetes Research Center grant DK063491 to the Southern California Diabetes Endocrinology Research Center. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

None of the authors report a conflict of interest related to research presented in this article.

The summary statistics data of the meta-analysis results are now provided on our website at: http://glimdna.org/publicationdata/Manuscript_AJCN-D-18-00930_MandaviyaPRetal_SummaryLevelData.zip. Additional data about the individual studies can be requested from the corresponding author.

We confirm that we will make the data (in de-identified form, if human data) used in the manuscript available to editors upon request either before or after publication for checking.

PRM, RJ, JB, JEC-F, KFD, AND, MG, IKH, and TT contributed equally as first authors.

DKA, SB, JTB, BTH, TL, DL, KEN, NS, MMJvG, JBJvM, and SGH contributed equally as last authors.

Supplemental Text 1, Supplemental Tables 1–7, and Supplemental Figures 1–4 are available from the "Supplementary data" link in the online posting of the article and from the same link in the online table of contents at https://academic.oup.com/ajcn/.

Address correspondence to SGH (e-mail: s.heil@erasmusmc.nl).

Abbreviations used: ARIC, Atherosclerosis Risk in Communities; BIOS, Biobank-Based Integrative Omics Studies; bp, base pair; BPROOF, B-Vitamins for the PRevention Of Osteoporotic Fractures; CABYR, calcium-binding tyrosine phosphorylation-regulated; CHS, Cardiovascular Health

Introduction

Folate and vitamin B-12 are essential micronutrients of the 1-carbon pathways that are involved in the donation of methyl groups to the DNA, RNA, and proteins (1). Folate is a methyl-donor itself, where its active form, 5-methyltetrahydrofolate, donates its methyl group to the remethylation of homocysteine to methionine. Vitamin B-12 is a cofactor in this reaction.

Folate and vitamin B-12 status has been connected to diverse diseases. Low concentrations of folate and vitamin B-12 during pregnancy are independently associated with the risk of neural tube defects in the child (2–4). In addition, low concentrations of folate or vitamin B-12 are associated with risk of a wide range of diseases, such as cardiovascular diseases and osteoporosis (5–8). Severe vitamin B-12 and folate deficiencies result in megaloblastic anemia, which, for vitamin B-12 deficiency, is associated with severe neurological abnormalities (9). Diagnosis of vitamin B-12 deficiency is hampered by the lack of diagnostic accuracy of available biomarkers. DNA methylation is a possible mechanism underlying the previously identified relations between folate or vitamin B-12 deficiency and disease risk, and specific alterations in methylation patterns could serve as future biomarkers for these nutrient-related disease risks. Therefore, it is important to assess the association of these nutrients with DNA methylation.

The relation between folate and/or vitamin B-12 intake and DNA methylation has been investigated mostly in studies examining individuals with a particular disease, which might confound the observed associations (10–12). A limited number of human studies were conducted in disease-free individuals (13–17). Importantly, in all previous studies, global DNA methylation levels were assessed as either total 5-methyl cytosine content, or LINE-1 or Alu repeat methylation as a proxy for global methylation status in blood leukocytes. The results from these studies were inconsistent (13–17). The relation between maternal folate intake and fetal DNA methylation was investigated in a few studies showing that maternal intake was associated with changes in DNA methylation (18–20). The previous studies related to folate and vitamin B-12 and DNA methylation in adults have been limited with respect to outcome (e.g., global methylation levels only), sample size, and participants (e.g., maternal–fetal exposures), and to date there are no large-scale epigenome-wide association studies (EWASs) in adults examining the relation between B-vitamin intake and DNA methylation.

In this study, we undertook a large-scale EWAS of folate and vitamin B-12 intake, analyzing the association with methylation at up to 485,512 CpGs assessed in whole blood measured in up to 5841 individuals across 10 cohorts from Europe and North America. Because folate and vitamin B-12 are important in the transfer of methyl groups to DNA, we hypothesize that low intake of folate or vitamin B-12 is associated with genome-wide DNA hypomethylation in the normal population.

Materials and Methods

Study populations

Data from 10 cohorts with a total of 5520 individuals of European ancestry and 321 African-American individuals were included in this meta-analysis. Written informed consent was given by all participants for genetic research. Descriptions of each cohort are provided in **Supplemental Text 1**. For all participating studies, individuals with prevalent cancer were excluded from analyses due to potential differences in dietary patterns in response to their disease (21, 22) and different methylation patterns (23, 24).

DNA methylation assessment

Genomic methylation profiling was performed on whole blood in 10 cohorts using the Infinium Illumina HumanMethylation 450k BeadChip arrays (Illumina Inc.) according to the manufacturer's protocol. The Genetics of Lipid Lowering Drugs and Diet Network (GOLDN) study used $CD4^+$ T cells from buffy coats. The Illumina array measures the methylation status of 485,512 CpG sites in the gene and nongene regions of CpG islands, shores, and shelves of the human genome (25). Poor-quality samples and probes were excluded based on cohort-specific criteria (**Supplemental Table 1**). Quantile (26), DASEN (27), Subset-quantile Within Array Normalization (28), Beta MIxture Quantile dilation (29), or Functional normalization (30) was used to correct the raw beta values that represent the methylation percentage per CpG for every sample. The normalized beta values were used for the association analysis in each cohort.

Data collection and dietary assessment

Dietary intake data were derived in each cohort from structured self-administered food-frequency questionnaires (FFQs) that contained from 78 to 389 food items (**Table 1**). Exposure variables for the nutrients folate and vitamin B-12 were calculated in μg/d from dietary intake data using national food composition tables (**Table 2**). This included foods with folic acid fortification if this was the case, but did not include B-vitamin supplement intake. Participants with missing dietary data, or those who reported very low (<500) or very high (>5000) total energy intake, were excluded.

In order to reduce the magnitude of the systematic measurement error of the FFQ, each nutrient was adjusted for total energy intake, using the residual method (31). Next, unstandardized residuals from this regression of each nutrient and total energy intake were used for association analyses. B-vitamin supplement

Study; CODAM, Cohort study on Diabetes and Atherosclerosis Maastricht; CUX1, cut like homeobox 1; DMP, differentially methylated position; DMR, differentially methylated region; EAR, estimated average requirement; EWAS, Epigenome-wide Association Study; FAM64A, family with sequence similarity 64 member A; FDR, false discovery rate; FFQ, food-frequency questionnaire; FHS, Framingham Heart Study; FRMD8, FERM domain containing 8; GOLDN, Genetics of Lipid Lowering Drugs and Diet Network; GREAT, Genomic Regions Enrichment of Annotations Tool; InCHIANTI, Invecchiare in Chianti; LCN8, lipocalin 8; LGALS3BP, galectin 3 binding protein; LLS, Leiden Longevity Study; MEGF6, multiple epidermal growth factor-like domains 6; MTHFR, methylenetetrahydrofolate reductase; RS, Rotterdam Study; WBC, white blood cell; WNT, Wnt family member; WRAP73, WD repeat containing, anti-sense to TP73; YFS, Young Finns Study.

Received August 27, 2018. Accepted for publication December 12, 2018.
First published online June 5, 2019; doi: https://doi.org/10.1093/ajcn/nqz031.

TABLE 1 Dietary assessment in participating cohorts

No.	Cohorts	Country	N^1 (5841)	Fortification[2]	No. of food items	Year(s) of dietary data collection	Reference table	Time gap[3], y
1	RS	Netherlands	900	No	389	2006–2012	The Dutch Food Composition Tables (NEVO table, 2006)	0 (46% samples had ~7-y gap)
2	LLS	Netherlands	430	No	183	2006–2007	FQ13V20061221 (Wageningen 2006), calculations based on Dutch Food Composition Database (NEVO table, 2006)	0
3	CODAM	Netherlands	154	No	178	2006–2009	NEVO Table, 2001	0
4	InCHIANTI	Italy	484	No	236	1998–2000	Italian Tables of Food Composition, 1998	0
5	YFS	Finland	155	No	128	2007	Finnish Tables of Food Composition, 2007	4
6	TwinsUK	United Kingdom	568	Yes	131	1994–2001, 2007	McCance and Widdowson's The Composition of Foods Edition 6, 2002	0
7	FHS	United States	1657	Yes	126	2005–2008	Harvard FFQ Nutrient Database, 2009	5–6
8	GOLDN	United States	983	Yes	124	2002–2004	The USDA's 1994–1996 Continuing Survey of Food Intakes by Individuals	0
9	ARIC	United States	321	No	78	1993–1995	USDA. Composition of foods: Agriculture Handbook No. 8 Series. Washington (DC): USDA, 1975–1989	0
10	CHS	United States	188	No	131	1989–1990	Harvard FFQ Nutrient Database, 1996	0

[1]Number of samples after exclusion of individuals with prevalent cancer and very low (<500) or very high (>5000) total energy intake.
[2]Cohorts that included foods with folic acid fortification.
[3]Time between dietary data collection and blood sampling for DNA-methylation analysis. Dietary data were collected before blood sampling for DNA-methylation analysis. ARIC, Atherosclerosis Risk in Communities; CHS, Cardiovascular Health Study; CODAM, Cohort study on Diabetes and Atherosclerosis Maastricht; FFQ, food-frequency questionnaire; FHS, Framingham Heart Study; GOLDN, Genetics of Lipid Lowering Drugs and Diet Network; InCHIANTI, Invecchiare in Chianti; LLS, Leiden Longevity Study; RS, Rotterdam Study; YFS, Young Finns Study.

intake was available in all cohorts for use as a covariate. Because these supplement data were recorded in different forms (B-vitamins, multivitamins, or folic acid supplements) and different units (frequency per day or per week) in each cohort, we harmonized these data across all cohorts by grouping individuals as supplement users and nonusers, regardless of specific vitamin form.

Statistical analyses

Differentially methylated positions (DMPs): cohort-specific association analyses.

Each participating cohort used linear mixed models to investigate the associations between each nutrient and genome-wide DNA methylation. This DNA-methylation analysis was first performed in a site-specific (CpG) manner to find DMPs. Models were adjusted for age, sex, BMI (in kg/m^2), differential white blood cell (WBC) counts, smoking status, physical activity, B-vitamin supplement intake, and alcohol (g/d or drinks/wk) and coffee (g/d or servings/d) consumption as fixed effects, except for those studies for which some covariates were not present (Table 2). Technical covariates such as array number and position on array were also adjusted for and were treated as random effects. Differential WBC counts were either represented as a percentage of measured cell counts or imputed using the Houseman method (32). We first used a continuous model (Model 1), where the unstandardized residuals of each nutrient were used as continuous variables to estimate their association with genome-wide DNA methylation. To determine whether large differences in nutrient intake affect DNA methylation, we also performed a categorical analysis where the individuals of each residual nutrient were divided into tertiles, and the first and third tertiles were used to define low and high nutrient intakes, respectively (Model 2). Food intake measured by FFQ is prone to measurement error, but is adequate for ranking individuals based on their intake (33). Thus, a model based on categorical ranks rather than absolute intake is the recommended analytical approach. In addition, because the use of supplements may be a confounder to the analysis, we performed a sensitivity analysis of both folate and vitamin B-12 intake in nonusers of B-vitamin, multivitamin, or folic acid supplements (Model 3).

DMPs: meta-analyses.

Figure 1 shows the stepwise study design that was followed. Using the software GWAMA (34), fixed effect meta-analyses as weighted by inverse variance were performed from the summary statistics of each participating cohort on the continuous

TABLE 2 Demographic and lifestyle characteristics of participating cohorts (measured cell counts in 10^9/L)[1]

No.	Cohorts	Age (mean ± SD)	Women (%)	BMI (mean ± SD)	White blood cell			Current smokers (%)	Alcohol intake (mean ± SD)	Coffee intake (mean ± SD)	Physical activity (mean ± SD) or active (%)	Supplement users (%)	Folate, μg/d (median [range])	Vitamin B-12, μg/d (median [range])
					Granulocytes (mean ± SD)	Lymphocytes (mean ± SD)	Monocytes (mean ± SD)							
1	RS	64.5 ± 9.0	57.4	27.6 ± 4.1	0.4 ± 0.1	0.2 ± 0.1[5]	0.1 ± 0.0[8]	7.4	12.3 ± 15.7[9]	385.8 ± 240.0[10]	63.4 ± 64.2[13]	32.1	415.2 (53.9–1173.9)	4.6 (0.6–14.1)
2	LLS	58.6 ± 6.3	51.3	25.3 ± 3.8	4.3 ± 1.3	—	—	12.3	16.7 ± 15.1	NA	NA	26.3	192.7 (68.0–415.4)	4.0 (0.9–11.0)
3	CODAM	65.5 ± 6.9	44.2	28.6 ± 4.2	0.3 ± 0.1	0.3 ± 0.1[5]	0.1 ± 0.0[8]	15.6	12.5 ± 14.2[9]	491.9 ± 276.8[10]	7150.4 ± 4955.3[14]	3.2	200.3 (96.2–364.7)	4.2 (0.8–10.8)
4	InCHIANTI	62.6 ± 15.7	54.1	27.1 ± 3.9	0.031 ± 0.02[2] 0.59 ± 0.08[3]	0.32 ± 0.08	—	18.6	16.5 ± 21.7	2.4 ± 1.4[11]	51	NA	273.9 (88.9–645.9)	NA
5	YFS	40.2 ± 3.3	61.9	25.6 ± 4.4	0.5 ± 0.1	0.2 ± 0.1[5]	0.1 ± 0.0[7]	24.5	6.9 ± 7.3	409.9 ± 265.8[10]	77.4	39.4	332.8 (159.1–732.0)	7.4 (3.0–16.9)
6	TwinsUK	58.9 ± 9.9	100.0	26.4 ± 4.8	0.5 ± 0.1	0.2 ± 0.1[5]	0.1 ± 0.0[8]	10.9	7.4 ± 11.2	NA	NA	53.3	376.7 (122.6–976.3)	6.0 (0.5–16.0)
7	FHS	64.8 ± 8.6	57.1	28.2 ± 5.4	0.5 ± 0.1	0.2 ± 0.1[5]	0.04 ± 0.0[7]	10.3	10.1 ± 15.4	157.2 ± 123.8[12]	88.3	57.7	405.9 (99.2–1141.0)	5.5 (0.0–18.2)
8	GOLDN	49.0 ± 16.0	53.0	28.3 ± 5.7	NAp	—	—	7.0	6.0 ± 20.0	277.5 ± 427.1	2.97 ± 2.47[15]	58.0	233.6 (0.0–1143.6)	4.5 (0.0–31.2)
9	ARIC	59.5 ± 5.8	67.6	30.5 ± 6.4	1.7 ± 2.2[2]	54.1 ± 10.2[3] NA	—	23.4	2.7 ± 10.0	8.9 ± 12.1	66.8 ± 52.0[16,17]	24.3	248.1 (36.0–811.2)	6.3 (0.7–26.3)
10	CHS	77.3 ± 4.9	66.5	27.5 ± 4.8	0.5 ± 0.15	— 0.1 ± 0.1[4]	0.2 ± 0.1[5] 0.1 ± 0.1[6]	0.1 ± 0.0[7]	0.1 ± 0.04[8]	7.98	2.34 ± 6.6	0.73 ± 1.1[11]	763.9 ± 1188.1	33.0	360.0 (93.7–886.7)	5.5 (0.8–28.7)

[1] ARIC, Atherosclerosis Risk in Communities; CHS, Cardiovascular Health Study; CODAM, Cohort study on Diabetes and Atherosclerosis Maastricht; FHS, Framingham Heart Study; GOLDN, Genetics of Lipid Lowering Drugs and Diet Network; InCHIANTI, Invecchiare in Chianti; LLS, Leiden Longevity Study; NA, not available; NAp, not applicable; RS, Rotterdam Study; YFS, Young Finns Study.

[3] [2] Granulocytes: eosinophils and neutrophils (measured).

[7] [6] [5] [4] Lymphocytes; CD8T, CD4T, NK, B cell (Houseman-imputed).

[8] Monocytes: Houseman-imputed.

[9] Alcohol intake: g/d.

[12] [11] [10] Coffee intake: g/d or servings/d; FHS used caffeine intake in g/d.

[16] [15] [14] [13] Physical activity: Total metabolic equivalent (MET) h/wk; total score of all activities; h/wk moderate activity; total metabolic equivalent (MET) min/wk.

[17] Not included in analyses (only available to $N = 174$ individuals).

(Model 1), categorical (Model 2) and supplement nonusers sensitivity models (Model 2a) of folate and vitamin B-12 intake EWASs. An additional sensitivity meta-analysis was also performed with limitation to studies that had FFQ data at the same time point as DNA-methylation measurements (Model 2b). Heterogeneity (I^2) was used to evaluate differences between cohorts in the fixed effect meta-analysis. The probes with SNPs at single base extension and probes with improper binding (35) were excluded to avoid spurious signals and cohybridization with alternate homologous sequences. All participating cohorts had different probe exclusions for quality control, and therefore as part of the meta-analysis stage, we removed probes if they were not present in at least 5 cohorts. The significance was defined by the Benjamini–Hochberg method (36) of the false discovery rate (FDR) < 0.05. The gene annotations for the DMPs we identified were performed using the Genomic Regions Enrichment of Annotations Tool (GREAT) (37) with the University of California, Santa Cruz (UCSC) (38) where they assigned a regulatory domain consisting of a basal domain to each gene, that extends up to 5 kb upstream and 1 kb downstream from its transcription start site. The DMP is annotated with a gene if it overlaps with its basal domain. In addition, DMP is annotated to a gene if an extension is reached up to the basal regulatory domain of the nearest upstream and downstream genes within 1 Mb. The genomic inflation factors (39) were computed to estimate the rate of false positives due to population structure. Although such a measure has been successfully applied in genome-wide association studies, its application to EWAS may not confer the same benefit due to the inherent nature of the correlation structure in CpG sites of interdependent pathways and environment exposure (40).

Differentially methylated regions (DMRs).

Because adjacent CpG sites across the genome are often spatially correlated and known to regulate in longer genomic regions (41), we extended our analysis to find DMRs, which in general are genomic regions of 500–1000 base pairs (bp) with different methylation status. We identified DMRs for each nutrient using the software comb-p (42), through analysis of the nominal P values of DMPs generated from the meta-analysis (Figure 1, Model 3). Nominal P values of DMPs were adjusted according to their weighted correlation with adjacent P values using the Stouffer–Liptak–Kechris method, in a sliding window of 500 bp with varying lags of 50 bp (42). Regions were identified by the peak finding algorithm on the adjusted P values that qualified according to the Benjamini–Hochberg FDR threshold < 0.05. The identified regions were then given new P values and corrected for multiple testing using the Sidak correction (42). The significance for DMRs was defined as Sidak < 0.05 for each nutrient. The gene annotations for the identified DMRs were performed using the GREAT tool (37) with the same extensions as applied for the DMP gene annotations.

Pathway analysis was performed for genes annotated to the DMRs related to each nutrient, using the WEB-based GEne SeT AnaLysis Toolkit (43). Overrepresentation enrichment analysis was performed for the GREAT annotated genes of the folate and vitamin B-12-associated DMRs, respectively, and compared against the reference genome, in the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways. The multiple testing

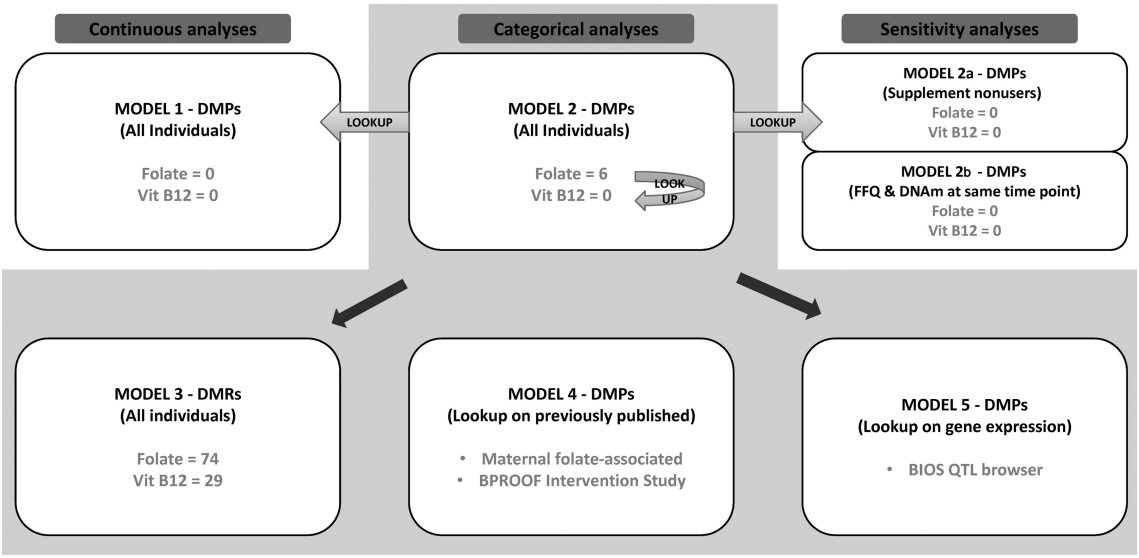

FIGURE 1 Analysis flow scheme. *1*) Meta-analysis of continuous model EWAS of folate and vitamin B-12 intake as well as *2*) meta-analysis of categorical model EWAS were performed. Significant (FDR < 0.05) associations of the folate-intake categorical meta-analysis were looked up in the folate intake (as a continuous variable) meta-analysis and vitamin B-12 intake (as a categorical variable) meta-analysis. Meta-analyses of categorical EWAS were followed by *2a*) sensitivity meta-analyses of B-vitamin supplement nonusers that included lookup of the folate-intake-associated CpGs of the main model, *2b*) sensitivity meta-analysis of studies with FFQ collection at the same time point as the DNA-methylation measurements, *3*) DMR analyses to find regions of association, and lastly, *4*) comparison to previously identified DMPs associated with maternal folate intake. *5*) Lookup of the folate-intake-associated DMPs on the previously published methylation–expression associations in the BIOS QTL browser. BIOS, Biobank-Based Integrative Omics Studies; BPROOF, B-Vitamins for the PRevention Of Osteoporotic Fractures; DMP, differentially methylated position; DMR, differentially methylated region; DNAm, DNA methylation; EWAS, Epigenome-wide Association Study; FDR, false discovery rate; FFQ, food-frequency questionnaire; Vit B12, vitamin B-12.

was set to FDR < 0.05, and the minimum number of genes in a functional gene set category was set to 2.

Previously published DMPs using Illumina 450k arrays.

We also investigated whether previously published DMPs related to maternal folate (19, 20) and DMPs related to folate and vitamin B-12 intervention (44) were associated with folate and vitamin B-12 intake in our population-based meta-analysis in adults. Gonseth et al. observed 4 DMPs (cg22664307, cg21039708, cg15219145, and cg13499966) in neonatal blood that were associated with maternal folate intake ($n = 167$) (19). In a second study examining cord blood methylation with maternal plasma folate ($n = 1988$), 443 DMPs in cord blood were identified (20). In a 2-y folic acid and vitamin B-12 intervention study, 1 DMP (cg19380919) had the greatest change in methylation in the treatment group compared to the controls with marginal significance (44). We compared the nominal P values and beta coefficients to examine the direction of effects. In addition, we also checked for the enrichment of these previously found DMPs in our meta-analysis by comparing these CpGs to the same number of randomly selected CpGs from the array using 100 permutations. From both the CpG sets, the number of CpGs that had nominal $P < 0.05$ were determined and compared. Using Fisher's exact test, the significant enrichment was then determined.

Effect of DMPs and DMRs on gene expression.

Finally, in the Biobank-Based Integrative Omics Studies (BIOS) QTL browser (35), we assessed whether the folate and vitamin B-12 intake-associated DMPs were associated with expression levels of the nearby genes.

Results

Cohort characteristics

Among the European cohorts, the median intakes of dietary folate ranged from 193 μg/d in the Leiden Longevity Study (LLS) to 415 μg/d in the Rotterdam Study (RS). For the LLS and Cohort study on Diabetes and Atherosclerosis Maastricht (200 μg/d) cohorts, this is slightly lower than the European estimated average requirement (EAR) of 250 μg/d (45). The RS had the highest median folate intake at 415 μg/d. Among the American cohorts, the median intake of dietary folate was 234 μg/d in the GOLDN study, which is lower than the American EAR of 320 μg/d (46, 47). The Cardiovascular Health Study, Framingham Heart Study (FHS) and Atherosclerosis Risk in Communities (ARIC) fulfilled the American EAR for their median dietary folate intakes at 360, 406, and 448 μg/d, respectively.

Among the European cohorts, the median intake of dietary vitamin B-12 ranged from 4.0 μg/d in the LLS to 7.4 μg/d in the Young Finns Study (YFS). For the American cohorts, the median intake of dietary vitamin B-12 ranged from 4.5 μg/d in the GOLDN study to 6.3 μg/d in the ARIC study. Characteristics of each participating study are given in Table 1.

Study participants were adults (44–100% women) with the mean age ranging from 40 to 77 y. There were 7–25% current smokers and 3–58% B-vitamin, multivitamin, or folic acid

supplement users. Further characteristics such as BMI, WBC counts, physical activity, and alcohol and coffee intake of each cohort are provided in Table 2.

Association analysis between B-vitamin intake and methylation

Model 1: Continuous (DMPs).

A meta-analysis of the 5815 individuals from the EWAS analysis of the folate-intake continuous model showed no significant DMP associations (FDR < 0.05). A meta-analysis of the 5302 individuals of the EWAS analysis of the vitamin B-12-intake continuous model also showed no significant DMP associations (FDR < 0.05) (**Supplemental Figure 1**).

Model 2: Categorical (DMPs).

Meta-analyses of 3894 individuals evaluated categorically from the EWASs of folate intake showed 6 significant DMP associations (FDR < 0.05) (**Table 3**, **Figure 2**A–G, and **Supplemental Figure 2**). These 6 DMPs showed consistent association with nominal significance ($P < 0.05$) in the continuous model (Table 3). The most significant DMP was at cg23465990 ($P = 3.87 \times 10^{-8}$, FDR = 0.018) on chromosome 17, annotated to the nearest gene *FAM64A* (605 bp upstream), and showed a 0.12% decrease in methylation per microgram per day increase in residual folate intake. Other significant DMPs at cg11832534, cg03249011, cg14398883, cg00826902, and cg14145338 were annotated to the nearest genes *WRAP73* (2648 bp downstream), *FRMD8* (41,931 bp downstream), *CUX1* (62,673 bp upstream), *WRAP73* (2692 bp downstream), and *LCN8* (3667 bp downstream), respectively, of chromosomes 1, 11, 7, and 9. All identified DMPs were negative associations between folate intake and methylation levels with a 0.12–0.79% decrease in methylation per microgram per day increase in residual folate intake.

EWAS results of the vitamin B-12-intake categorical model did not show any significant DMPs in the meta-analysis of 3566 individuals (**Figure 3** and Supplemental Figure 2). Of the 6 significant DMPs from the folate-intake analysis, 2 folate-intake-associated DMPs (cg23465990 and cg14398883) showed borderline nominal significance in the same direction as folate intake in the categorical model of vitamin B-12 intake ($P = 0.02$ and 0.04, respectively) (**Supplemental Table 2**).

Model 2a: Sensitivity model in supplement nonusers.

A sensitivity analysis was further performed to reduce heterogeneity in the models caused by the use of B-vitamin supplements. This may potentially help with finding new true positive DMPs. We therefore restricted the analysis to supplement nonusers, but no significant DMPs were found in the meta-analysis of 2183 individuals for the categorical model of folate intake (**Supplemental Figure 3**). The 6 significant DMPs identified in the full categorical model were nominally associated ($P < 0.05$) in the same direction with folate intake in the nonusers (Table 3). The effect sizes were similar to the full categorical model with a 0.12–0.74% decrease in methylation per micrograms per day increase in folate intake. Furthermore,

TABLE 3 Differentially methylated positions significantly associated with folate intake at the epigenome-wide level in the meta-analysis

	CpG	Nearest gene	Chromosome	bp	Categorical model: all individuals (Model 2)						Categorical model: sensitivity analysis of supplement nonusers LOOKUP (Model 2a)			Categorical model: sensitivity analysis of same time point FFQ and DNAm LOOKUP (Model 2b)			Continuous model: all individuals LOOKUP (Model 1)		
					N	Effect	SE	*P*	FDR	I^2	*N*	Effect	SE	*P*	*N*	Effect	SE	*P*	*N*	Effect	SE	*P*
1	cg23465990	*FAM64A* (−605)	17	6,347,153	3791	−0.0012	2.2×10^{4}	3.9×10^{8}	0.018	0.02	2120	−0.0012	2.6×10^{4}	5.3×10^{6}	2,090	−0.0010	2.9×10^{4}	9.2×10^{4}	5.661	−0.000004	1.0×10^{6}	2.4×10^{5}
2	cg11832534	*WRAP73* (+2,648), *TPRG1L* (+22,467)	1	3,563,998	3894	−0.0041	7.7×10^{4}	9.0×10^{8}	0.021	0	2183	−0.0032	9.6×10^{4}	7.9×10^{4}	2,193	−0.0041	9.7×10^{4}	2.2×10^{5}	5.815	−0.000015	3.0×10^{6}	8.9×10^{6}
3	cg03249011[1]	*SCYL1* (−96,576), *FRMD8* (+41,931)	11	65,195,995	3894	−0.0075	1.4×10^{3}	1.8×10^{7}	0.029	0	2183	−0.0074	1.8×10^{3}	4.2×10^{5}	2,193	−0.0077	1.9×10^{3}	3.0×10^{5}	5.815	−0.000016	6.0×10^{6}	7.5×10^{3}
4	cg14398883	*MYLI0* (−125,633), *CUX1* (−62,673)	7	101,398,184	3894	−0.0079	1.6×10^{3}	5.1×10^{7}	0.049	0	2183	−0.0048	1.4×10^{3}	5.2×10^{4}	2,193	−0.0066	2.2×10^{3}	2.9×10^{3}	5.815	−0.000028	6.0×10^{6}	1.0×10^{6}
5	cg00826902	*WRAP73* (+2,692), *TPRG1L* (+22,423)	1	3,563,954	3894	−0.0049	9.8×10^{4}	5.9×10^{7}	0.049	0.49	2183	−0.0053	1.2×10^{3}	1.5×10^{5}	2,193	−0.0026	1.2×10^{3}	3.4×10^{2}	5.815	−0.000019	4.0×10^{6}	3.2×10^{6}
6	cg14145338	*LCN6* (−6,084), *LCN8* (+3,667)	9	139,649,039	3894	−0.0061	1.2×10^{3}	6.2×10^{7}	0.049	0.05	2183	−0.0061	1.4×10^{3}	2.1×10^{5}	2,193	−0.0042	1.5×10^{3}	6.2×10^{3}	5.815	−0.000022	5.0×10^{6}	2.0×10^{5}

[1]Location = Enhancer, based on Illumina annotation, derived from the University of California, Santa Cruz (UCSC). FDR threshold = 0.05; statistical test: linear mixed models, adjusted for BMI, white blood cell counts, smoking status, physical activity, B-vitamin supplement intake, alcohol and coffee consumption, and batch effects. bp, base pair location based on Illumina annotation; Effect, beta coefficients based on unstandardized residuals of folate intake (1 μg/d), adjusted for total energy intake; FDR, false discovery rate; FFQ, food-frequency questionnaire; I^2, heterogeneity I^2 parameter.

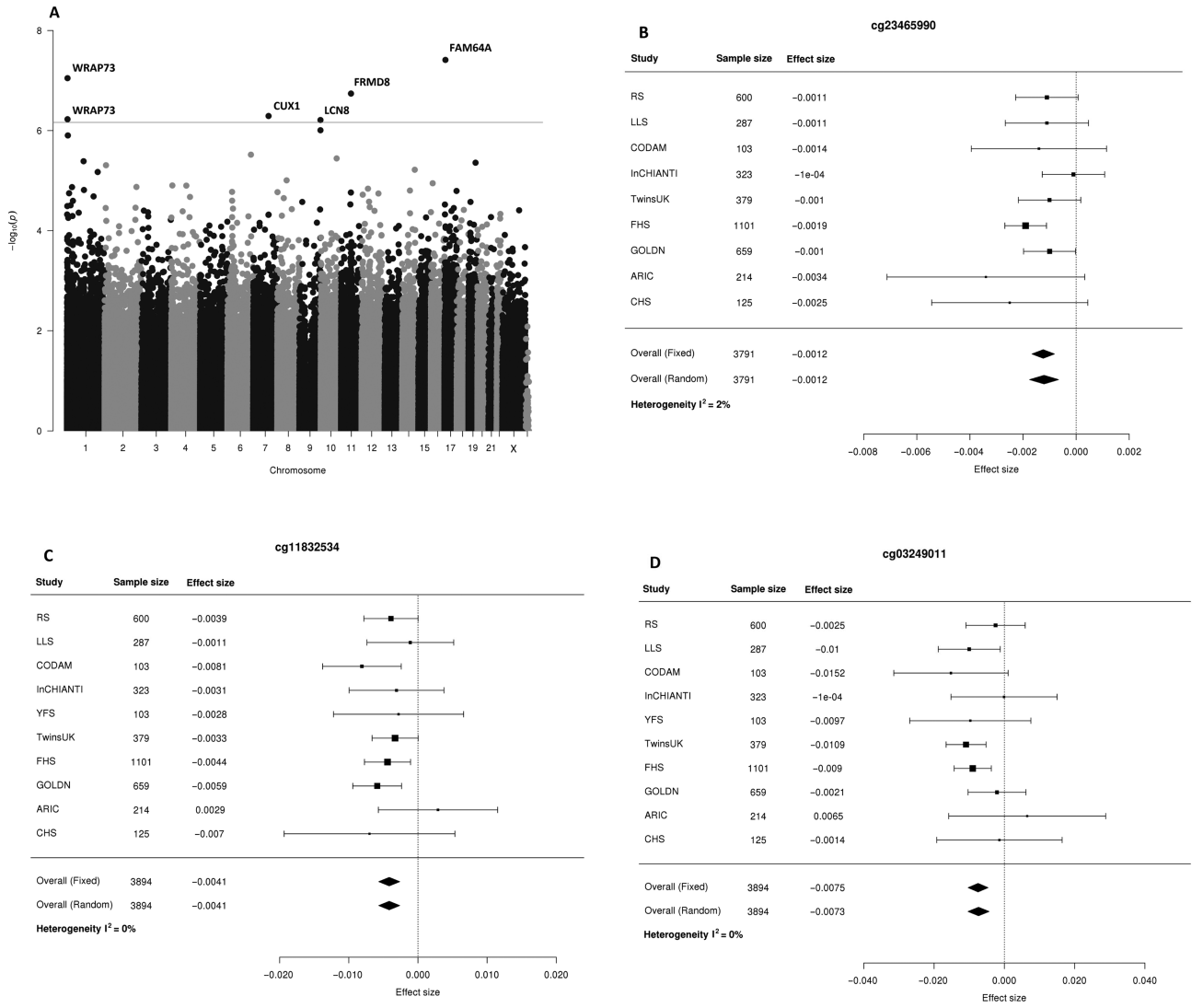

FIGURE 2 (A) Manhattan plot of the folate categorical model, adjusted for BMI, WBC counts, smoking status, physical activity, B-vitamin supplement intake, alcohol and coffee consumption, and batch effects, showing the association between folate intake and genome-wide DNA methylation (Model 2), with 6 significant DMPs at FDR < 0.05 (red line), in 3894 individuals. Nearest genes for these 6 DMPs are reported. (B) Forest plot of the categorical folate-intake model 2, showing the association between folate intake and the most significant DMP cg23465990 (*FAM64A*) across all studies and in a meta-analysis of 3791 individuals. (C) Forest plot of the categorical folate-intake model 2, showing the association between folate intake and the DMP cg11832534 (*WRAP73*) across all studies and in a meta-analysis of 3894 individuals. (D) Forest plot of the categorical folate-intake model 2, showing the association between folate intake and the DMP cg03249011 (*FRMD8*) across all studies and in a meta-analysis of 3894 individuals. (E) Forest plot of the categorical folate-intake model 2, showing the association between folate intake and the DMP cg14398883 (*CUX1*) across all studies and in a meta-analysis of 3894 individuals. (F) Forest plot of the categorical folate-intake model 2, showing the association between folate intake and the DMP cg00826902 (*WRAP73*) across all studies and in a meta-analysis of 3894 individuals. (G) Forest plot of the categorical folate-intake model 2, showing the association between folate intake and the DMP cg14145338 (*LCN8*) across all studies and in a meta-analysis of 3894 individuals. ARIC, Atherosclerosis Risk in Communities; CHS, Cardiovascular Health Study; CODAM, Cohort study on Diabetes and Atherosclerosis Maastricht; DMP, differentially methylated position; FDR, false discovery rate; FHS, Framingham Heart Study; GOLDN, Genetics of Lipid Lowering Drugs and Diet Network; InCHIANTI, Invecchiare in Chianti; LLS, Leiden Longevity Study; RS, Rotterdam Study; WBC; white blood cell; YFS, Young Finns Study.

no significant DMPs were identified either in the individual cohort results or in the meta-analysis of 1855 individuals for the categorical model of vitamin B-12 intake in supplement nonusers (Supplemental Figure 3).

Model 2b: Sensitivity model time of measurements.

To determine whether there is a confounding effect by studies with a time lag between DNA methylation and folate or vitamin B-12 intake assessment, a second sensitivity meta-analysis was performed including only studies where DNA-methylation measurements and FFQ collection were assessed at the same time. The FHS, YFS, and a subset of the RS cohort had an FFQ collection 4–7 y away from DNA collection, and so these 3 cohorts were omitted (Model 2b, Figure 1). A meta-analysis of the 2183 individuals for the categorical model of folate intake did not result in any significant DMPs (**Supplemental Figure 4**). The 6 significant DMPs associated with folate intake in the

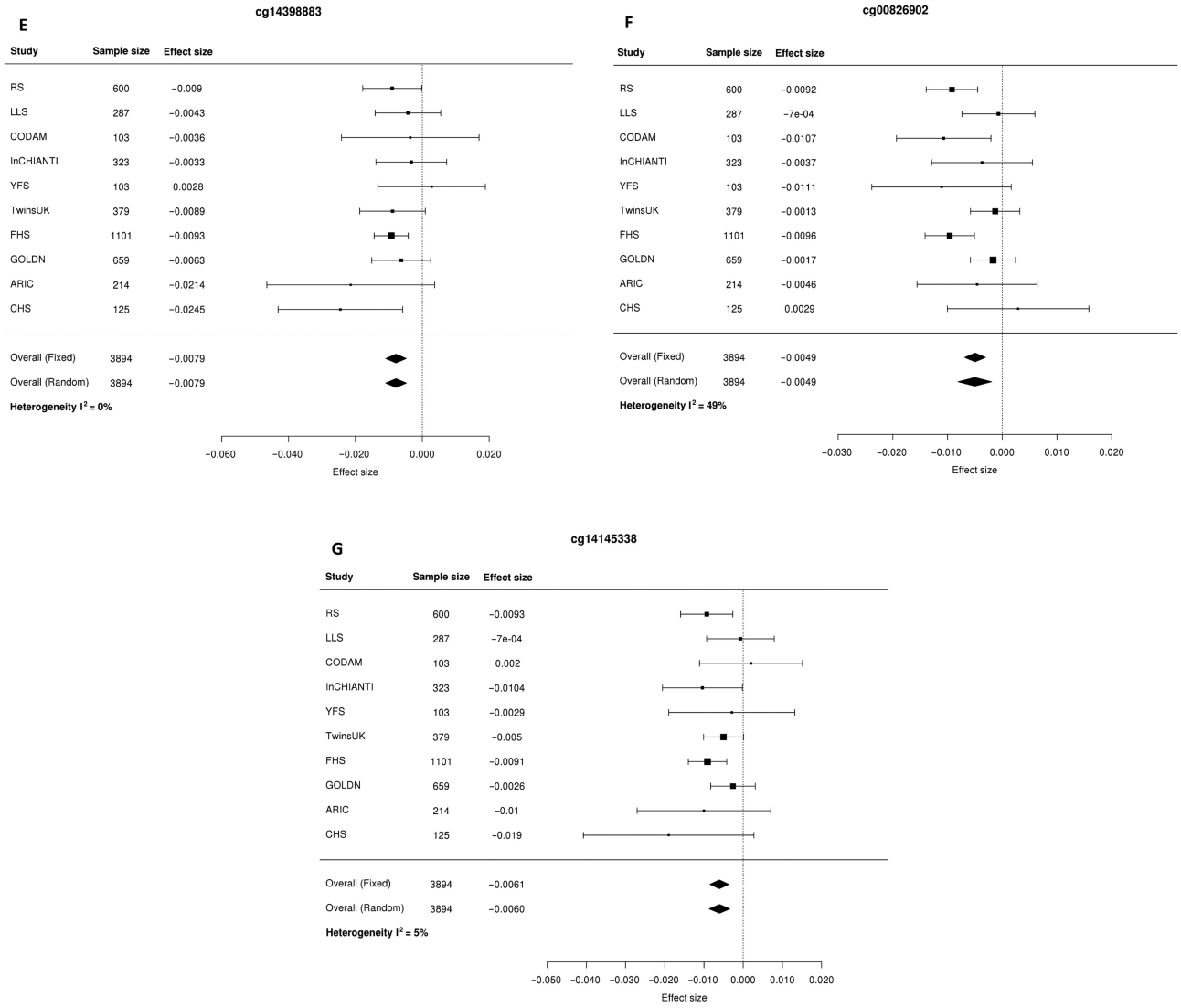

FIGURE 2 continued

full categorical model had consistent associations with nominal significance (Table 3). Furthermore, sensitivity meta-analysis of the 1865 individuals for the categorical model of the vitamin B-12 intake also did not identify any significant DMPs (Supplemental Figure 4).

Model 3: Categorical (DMRs).

Folate intake. We additionally performed DMR analysis using *P* values of the folate and vitamin B-12-intake meta-analyses of the categorical model (Model 2). By investigation of significant regions (Sidak $P < 0.05$) using the software comb-p, we observed 74 significant DMRs associated with folate intake (**Supplemental Table 3**, **Figure 4**). At most DMRs (73/74), methylation was negatively associated with folate intake. The most significant DMR associated with folate intake was the chr17:76,975,944–76,976,358 ($P = 1.47 \times 10^{14}$) containing a 414-bp region of 8 DMPs, which annotates to the *LGALS3BP* gene. Overrepresentation enrichment analysis for the 117 genes of the folate-associated 74 DMRs showed 1 significant pathway: signaling pathways regulating pluripotency of stem cells ($P = 1.61 \times 10^{4}$).

Vitamin B-12 intake. A regional analysis using comb-p found 29 significant DMRs with Sidak $P < 0.05$ associated with vitamin B-12 intake (**Supplemental Table 4**, **Figure 5**), of which 15 showed a negative direction of effects. The most significant DMR associated with vitamin B-12 intake was the chr18:21,718,458–21,719,569 ($P = 1.09 \times 10^{13}$) containing a 1111-bp region of 18 DMPs, which annotates to the promoter region of the calcium-binding tyrosine phosphorylation-regulated gene (*CABYR*). The vitamin B-12-associated DMRs did not overlap with the folate intake-associated DMRs. An enrichment analysis for the 48 genes of the vitamin B-12-associated 29 DMRs did not reveal any significant pathways.

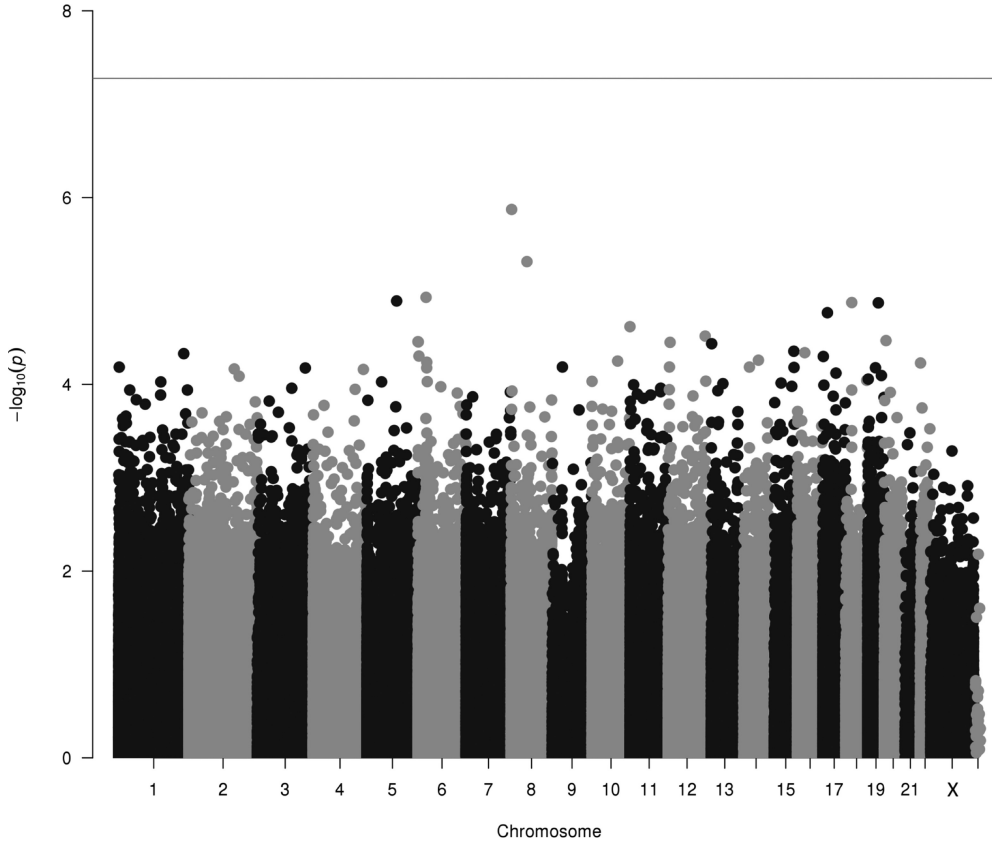

FIGURE 3 Manhattan plot of the vitamin B-12 categorical model, adjusted for BMI, WBC counts, smoking status, physical activity, B-vitamin supplement intake, alcohol and coffee consumption, and batch effects, showing the association between vitamin B-12 intake and genome-wide DNA methylation (Model 2), with no significant DMPs at FDR < 0.05 (red line), in 3566 individuals. DMP, differentially methylated position; FDR, false discovery rate; WBC, white blood cell.

Previously published DMPs using Illumina 450k arrays.

Maternal folate exposure and cord blood DMPs. We analyzed previously identified DMPs that have previously been reported to be associated in neonates with maternal folate intake. In the current folate-intake meta-analysis (Categorical Model 2), 1 (cg15219145) of the 4 previously identified DMPs from the Gonseth et al. study (19) showed a nominal significance ($P < 0.05$) with a similar direction. None of these 4 CpGs were significant with vitamin B-12 intake (**Supplemental Table 5**). Next, 28 and 27 of the 443 previously identified DMPs in newborns from the Joubert et al. study (20) were associated with folate and vitamin B-12 intake, respectively, in the same direction with nominal significance ($P < 0.05$) in our study (**Supplemental Table 6**). However, no enrichment of significant P values was found in these 443 DMPs in either folate ($P = 0.60$) or vitamin B-12 ($P = 0.22$) intake models in our study.

Folic acid and vitamin B-12 intervention study. The previously identified DMP (cg19380919) of a 2-y folic acid and vitamin B-12 intervention study (44) was not associated with either folate intake ($P = 0.78$) or vitamin B-12 intake ($P = 0.43$) in the current study (**Supplemental Table 7**).

Effect of DMPs on gene expression.

Of the 6 folate-intake-associated DMPs, cg00826902 showed a positive association with expression of the *MEGF6* gene (mean effect $= 0.07$, $P = 8.52 \times 10^5$, FDR $= 0.02$) (35).

Discussion

We conducted the first large-scale EWAS of the association between dietary folate and vitamin B-12 intake and genome-wide DNA methylation in humans. We identified 6 novel DMPs and 74 DMRs significantly associated with folate intake, and 29 DMRs significantly associated with vitamin B-12 intake. These novel epigenetic loci might be mechanistic indicators of low folate and vitamin B-12 intake.

The most significant DMP at cg23465990 is annotated to the gene family with sequence similarity 64 member A (*FAM64A*) on chromosome 17, which is identified as a marker for cell proliferation control (48, 49). The DMPs at cg11832534 and cg00826902 on chromosome 1 are annotated to the WD repeat containing protein coding gene (*WRAP73*), which ensures spindle morphology (50) and functions in ciliogenesis (51). The DMP at cg14398883 on chromosome 7 is annotated to cut like homeobox 1 (*CUX1*), which is a tumor suppressor (52). The DMP at cg03249011 is annotated to the FERM domain containing 8 (*FRMD8*) on chromosome 11. *FRMD8* is

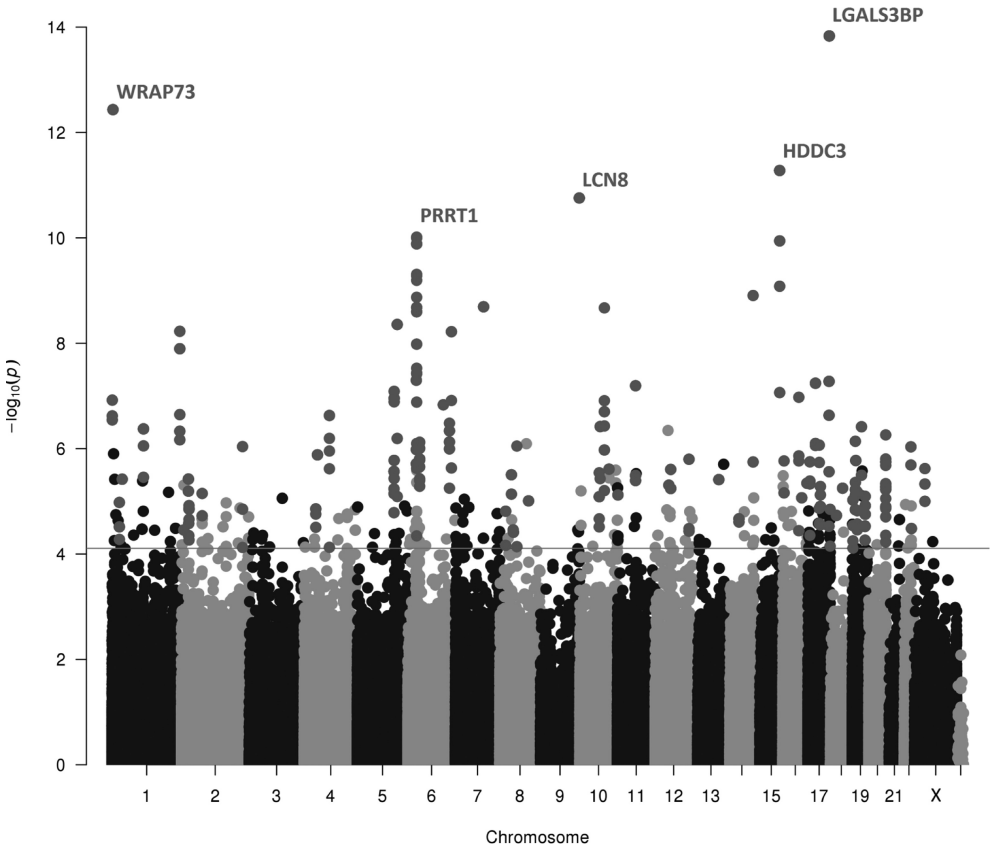

FIGURE 4 Manhattan plot of the folate-intake categorical model 3, showing the association between folate intake and genome-wide DNA methylation, with 74 significant DMRs adjusted for multiple comparisons using Sidak < 0.05 in 3894 samples. The red line represents the threshold of FDR < 0.05 of the correlation-adjusted *P* values of the CpGs. Green circles denote CpGs that are present within the DMRs, whereas black/gray circles denote the single CpGs that do not represent DMRs. The nearest genes for the top 5 DMRs are reported. DMR, differentially methylated region; FDR, false discovery rate.

associated with survival rate in lung adenocarcinoma patients (53). Lastly, the DMP at cg14145338 is annotated to lipocalin 8 (*LCN8*) on chromosome 9. *LCN8* expresses in epididymis and is suggested to be involved in male fertility (54). Further studies are needed to understand how these CpGs with seemingly no obvious link with B-vitamin homeostasis are linked to folate intake.

In contrast to the relatively small number of DMPs for folate intake, we observed 74 significant DMRs in association with folate intake, with several of potential relevance to immune function and stem cell function. The most significant DMR was the chromosome 17 locus spanning a 414-bp region including 8 CpGs annotating to the galectin 3 binding protein (*LGALS3BP*) gene, which is implicated in immune response. Signaling pathways regulating pluripotency of stem cells are overrepresented among the genes annotated near the folate-associated DMR CpGs, comprising AKT serine/threonine kinase 3, inhibitor of DNA binding 2, HLH protein, Wnt family member 6 (*WNT6*), *WNT9B*, and *WNT10A* genes.

In contrast to the null findings for vitamin B-12 and DMP, a regional analysis showed 29 significant DMRs associated with vitamin B-12 intake. The most significant DMR was the chromosome 18 locus spanning 1111 bp containing 18 CpGs and a promoter of *CABYR*. Surprisingly, there was no overlap of DMRs between the folate and vitamin B-12 intake despite the correlation between folate and vitamin B-12 intake, thus suggesting that their roles may be specific to different genes and pathways.

We examined previously identified folate-associated DMRs in mother–offspring pairs (19, 20, 44), but were not able to replicate the findings. This suggests that the folate-related DMPs in newborns and adults are not similar and might differ across the life course. We also tried to replicate previously identified CpGs in the B-Vitamins for the PRevention Of Osteoporotic Fractures (BPROOF) intervention study, where folate/vitamin B-12 supplementation was performed in elderly individuals. Again, we were unable to replicate those findings. A possible explanation for this latter finding is that the BPROOF study is an intervention study in which the effects of folate and vitamin B-12 on DNA methylation were studied over a period of 2 y. In addition, only individuals with elevated homocysteine (>15 μmol/L) were included in the BPROOF study (44). In the studies reported in this paper, we included generally healthy individuals with no limitation on homocysteine concentrations.

We hypothesized that low levels of folate give rise to DNA hypomethylation (13, 16, 17). However, all 6 folate-associated DMPs and most (73 of 74) DMRs were negatively associated with folate intake. The occurrence of general hypomethylation with higher folate intake is contrary to our hypothesis, where a higher

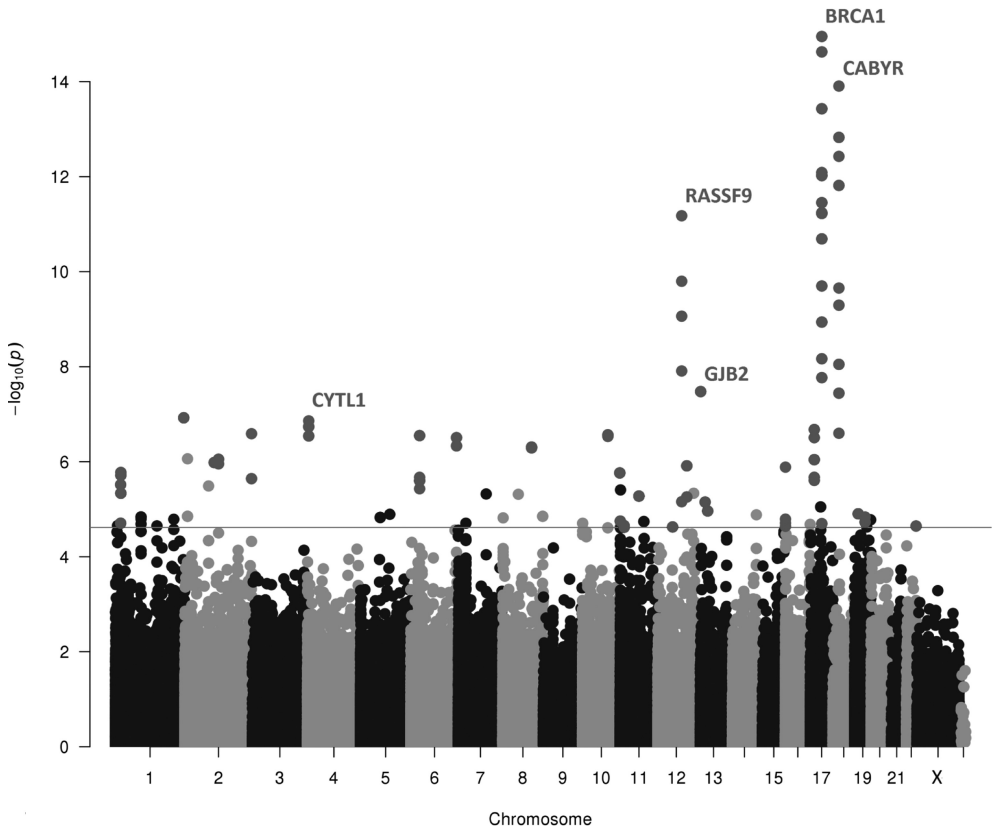

FIGURE 5 Manhattan plot of the vitamin B-12-intake categorical model 3, showing the association between vitamin B-12 intake and genome-wide DNA methylation, with 29 significant DMRs adjusted for multiple comparisons using Sidak < 0.05 in 3565 samples. The red line represents the threshold of FDR < 0.05 of the correlation-adjusted *P* values of the CpGs. Green circles denote the CpGs that are present within the DMRs, whereas black/gray circles denote the single CpGs that do not represent DMRs. The nearest genes for the top 5 DMRs are reported. DMR, differentially methylated region; FDR, false discovery rate.

folate intake would donate methyl groups and result in relative genome-wide hypermethylation. However, our results are in line with the study from Ono et al., who showed that a higher folate intake was associated with lower global methylation (15). Our results are also in line with the studies with maternal folate intake that showed a majority of identified DMPs in newborns to be negatively associated with folate intake (19, 20). Furthermore, an intervention study of folic acid supplementation in mice demonstrated that folic acid inhibited methylenetetrahydrofolate reductase (MTHFR) activity and reduced *S*-adenosylmethionine and the *S*-adenosylmethionine:*S*-adenosylhomocysteine ratio (55). The inhibition of MTHFR due to higher folic acid intake may explain the relatively lower DNA methylation observed in our study.

The strengths of our study are that this was conducted in a large sample size of 5,841 individuals from 10 well-characterized cohorts, using the Illumina 450k methylation data. The nutrient data were harmonized across studies, and all studies ran similar models with the same covariates. Although our study has yielded a number of interesting findings, among the most important insights is that folate or vitamin B-12 intake, a major determinant of B-vitamin status, is not related to large-scale differences in genome-wide methylation DMPs. The weaknesses of our study are that we were not able to replicate our findings, as we included all possible cohorts in the discovery. In addition, our study results show DMPs that approach borderline significance and have modest effect sizes. In such cases, the comb-p package for DMR analysis may return false-positive DMRs (56). Therefore, replication of our findings would be necessary before definite conclusions can be drawn. However, despite few identified methylation loci, 1 of the 6 folate-intake-associated DMPs, cg00826902, showed a positive association with expression of the *MEGF6* gene. This means that high folate intake is associated with hypomethylation of cg00826902 in chromosome 1, which is in turn associated with lower expression of the *MEGF6* gene (35). For the CpGs that were not associated with expression levels, their biological role needs to be validated further. Furthermore, dietary data are semiquantitative and prone to measurement errors, which can lead to misclassification and can compromise our ability to detect statistically significant associations. However, we addressed these limitations by using categorical models. Although this analysis reduces statistical power due to the one-third reduction in sample size, greater effects were observed because the comparison is made between extreme tertile groups of the nutrient intakes. Moreover, none of the cohorts had a median intake of vitamin B-12 that was lower than the European or American EAR of 4 μg/d (57) or 2 μg/d (47), respectively. This could mean that the vitamin B-12 intake was consistently high enough, and our population was relatively healthy enough to prevent any measurable effects on

methylation. Furthermore, the human genome contains more than 28 million CpGs (58), of which 1.7% is represented in Illumina 450k arrays. Therefore, other regions that could be associated with folate or vitamin B-12 intake might be missed. In addition, we used whole blood leukocytes for our study and acknowledge the possibility that these cells may not be the ideal tissue for evaluating the association between B-vitamin intake and methylation, and that larger tissue-specific effects may be present but remain undetected in our study. Lastly, to validate the results, measurement of blood levels of folate and vitamin B-12 would be of interest for the future.

In conclusion, we observed 6 DMPs associated with dietary folate intake. Regional associations showed 74 DMRs for dietary folate intake and 29 DMRs for dietary vitamin B-12 intake. Our meta-analysis identified several novel differentially methylated loci that could serve as mechanistic indicators of low folate and vitamin B-12 intake. Further studies are necessary to replicate findings.

The authors' responsibilities were as follows—PRM: contributed to the analysis plan and performed the statistical data analysis, interpretation of the results, and writing of the manuscript; RJ: contributed to the statistical data analysis, writing of the scripts, and critical review of the manuscript; JB, JEC-F, KFD, AND, MG, IKH, and TT: contributed to the statistical data analysis; EALdJ and JCKdJ: contributed to the analysis plan, result interpretation, and critical review of the manuscript; YBdR, MF, MAI, PFJ, AEJ, DPK, MMM, CES, SA, DvH, KFD, TT, MG, and KEN: contributed to the critical review of the manuscript; DMA, DGH, MAH, RNL, VM, AZM, TP, OTR, CGS, JS, EPES, CDAS, P-CT, and CJHvdK: contributed to the interpretation of the results; AGU: conducted the supervision and contributed to the interpretation of the results and critical review of manuscript; DKA, SB, JTB, BTH, TL, DL, KEN, NS, and MMJvG: were the principal investigators of the respective participating cohorts and contributed to the interpretation of the results; JBJvM and SGH: were the principal investigators of this project, conducted the supervision, and contributed to the analysis plan, interpretation of the results, and writing of the manuscript; and all authors: read and approved the final manuscript.

References

1. Friso S, Udali S, De Santis D, Choi SW. One-carbon metabolism and epigenetics. Mol Aspects Med 2017;54:28–36.
2. Kirke PN, Molloy AM, Daly LE, Burke H, Weir DG, Scott JM. Maternal plasma folate and vitamin B12 are independent risk factors for neural tube defects. Q J Med 1993;86:703–8.
3. Molloy AM, Mills JL, Kirke PN, Weir DG, Scott JM. Folate status and neural tube defects. Biofactors 1999;10:291–4.
4. Molloy AM, Kirke PN, Troendle JF, Burke H, Sutton M, Brody LC, Scott JM, Mills JL. Maternal vitamin B12 status and risk of neural tube defects in a population with high neural tube defect prevalence and no folic acid fortification. Pediatrics 2009;123:917–23.
5. Ma Y, Peng D, Liu C, Huang C, Luo J. Serum high concentrations of homocysteine and low levels of folic acid and vitamin B12 are significantly correlated with the categories of coronary artery diseases. BMC Cardiovasc Disord 2017;17:37.
6. Peng Y, Dong B, Wang Z. Serum folate concentrations and all-cause, cardiovascular disease and cancer mortality: A cohort study based on 1999–2010 National Health and Nutrition Examination Survey (NHANES). Int J Cardiol 2016;219:136–42.
7. Ebesunun MO, Umahoin KO, Alonge TO, Adebusoye LA. Plasma homocysteine, B vitamins and bone mineral density in osteoporosis: A possible risk for bone fracture. Afr J Med Med Sci 2014;43:41–7.
8. Herrmann M, Peter Schmidt J, Umanskaya N, Wagner A, Taban-Shomal O, Widmann T, Colaianni G, Wildemann B, Herrmann W. The role of hyperhomocysteinemia as well as folate, vitamin B(6) and B(12) deficiencies in osteoporosis: A systematic review. Clin Chem Lab Med 2007;45:1621–32.
9. Castellanos-Sinco HB, Ramos-Peñafiel CO, Santoyo-Sánchez A, Collazo-Jaloma J, Martínez-Murillo C, Montaño-Figueroa E, Sinco-Ángeles A. Megaloblastic anaemia: Folic acid and vitamin B12 metabolism. Revista Médica Del Hospital General De México 2015;78:135–43.
10. Jung AY, Botma A, Lute C, Blom HJ, Ueland PM, Kvalheim G, Midttun Ø, Nagengast F, Steegenga W, Kampman E. Plasma B vitamins and LINE-1 DNA methylation in leukocytes of patients with a history of colorectal adenomas. Mol Nutr Food Res 2013;57:698–708.
11. Schernhammer ES, Giovannucci E, Kawasaki T, Rosner B, Fuchs CS, Ogino S. Dietary folate, alcohol and B vitamins in relation to LINE-1 hypomethylation in colon cancer. Gut 2010;59:794–9.
12. Piyathilake CJ, Johanning GL, Macaluso M, Whiteside M, Oelschlager DK, Heimburger DC, Grizzle WE. Localized folate and vitamin B-12 deficiency in squamous cell lung cancer is associated with global DNA hypomethylation. Nutr Cancer 2000;37:99–107.
13. Zhang FF, Morabia A, Carroll J, Gonzalez K, Fulda K, Kaur M, Vishwanatha JK, Santella RM, Cardarelli R. Dietary patterns are associated with levels of global genomic DNA methylation in a cancer-free population. J Nutr 2011;141:1165–71.
14. Perng W, Villamor E, Shroff MR, Nettleton JA, Pilsner JR, Liu Y, Diez-Roux AV. Dietary intake, plasma homocysteine, and repetitive element DNA methylation in the Multi-Ethnic Study of Atherosclerosis (MESA). Nutr Metab Cardiovasc Dis 2014;24:614–22.
15. Ono H, Iwasaki M, Kuchiba A, Kasuga Y, Yokoyama S, Onuma H, Nishimura H, Kusama R, Ohnami S, Sakamoto H, et al. Association of dietary and genetic factors related to one-carbon metabolism with global methylation level of leukocyte DNA. Cancer Sci 2012;103:2159–64.
16. Zhang FF, Santella RM, Wolff M, Kappil MA, Markowitz SB, Morabia A. White blood cell global methylation and IL-6 promoter methylation in association with diet and lifestyle risk factors in a cancer-free population. Epigenetics 2012;7:606–14.
17. Agodi A, Barchitta M, Quattrocchi A, Maugeri A, Canto C, Marchese AE, Vinciguerra M. Low fruit consumption and folate deficiency are associated with LINE-1 hypomethylation in women of a cancer-free population. Genes Nutr 2015;10:480.
18. 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;28:4068–76.
19. Gonseth S, Roy R, Houseman EA, de Smith AJ, Zhou M, Lee ST, Nusslé S, Singer AW, Wrensch MR, Metayer C, et al. Periconceptional folate consumption is associated with neonatal DNA methylation modifications in neural crest regulatory and cancer development genes. Epigenetics 2015;10:1166–76.
20. Joubert BR, den Dekker HT, Felix JF, Bohlin J, Ligthart S, Beckett E, Tiemeier H, van Meurs JB, Uitterlinden AG, Hofman A, et al. Maternal plasma folate impacts differential DNA methylation in an epigenome-wide meta-analysis of newborns. Nat Commun 2016;7:10577.
21. Supportive PDQ, Palliative Care Editorial B. Nutrition in Cancer Care (PDQ(R)): Health Professional Version. 2002.
22. Hebuterne X, Lemarie E, Michallet M, de Montreuil CB, Schneider SM, Goldwasser F. Prevalence of malnutrition and current use of nutrition support in patients with cancer. JPEN J Parenter Enteral Nutr 2014;38:196–204.
23. Hanley MP, Hahn MA, Li AX, Wu X, Lin J, Wang J, Choi AH, Ouyang Z, Fong Y, Pfeifer GP, et al. Genome-wide DNA methylation profiling reveals cancer-associated changes within early colonic neoplasia. Oncogene 2017;36:5035–44.
24. Ye D, Jiang D, Li Y, Jin M, Chen K. The role of LINE-1 methylation in predicting survival among colorectal cancer patients: A meta-analysis. Int J Clin Oncol 2017;22:749–57.
25. Bibikova M, Barnes B, Tsan C, Ho V, Klotzle B, Le JM, Delano D, Zhang L, Schroth GP, Gunderson KL, et al. High density DNA methylation array with single CpG site resolution. Genomics 2011;98:288–95.
26. Lehne B, Drong AW, Loh M, Zhang W, Scott WR, Tan ST, Afzal U, Scott J, Jarvelin MR, Elliott P, et al. A coherent approach for analysis of the Illumina HumanMethylation450 BeadChip improves data quality

and performance in epigenome-wide association studies. Genome Biol 2015;16:37.
27. Pidsley R, Y Wong CC, Volta M, Lunnon K, Mill J, Schalkwyk LC. A data-driven approach to preprocessing Illumina 450 K methylation array data. BMC Genomics 2013;14:293.
28. Maksimovic J, Gordon L, Oshlack A. SWAN: Subset-quantile within array normalization for illumina infinium HumanMethylation450 BeadChips. Genome Biol 2012;13:R44.
29. Teschendorff AE, Marabita F, Lechner M, Bartlett T, Tegner J, Gomez-Cabrero D, Beck S. A beta-mixture quantile normalization method for correcting probe design bias in Illumina Infinium 450 k DNA methylation data. Bioinformatics 2013;29:189–96.
30. 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;15:503.
31. Willett WC, Howe GR, Kushi LH. Adjustment for total energy intake in epidemiologic studies. Am J Clin Nutr 1997;65:1220S–8S; discussion 1229S-1231S.
32. 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.
33. Preacher KJ, Rucker DD, MacCallum RC, Nicewander WA. Use of the extreme groups approach: A critical reexamination and new recommendations. Psychol Methods 2005;10:178–92.
34. Magi R, Morris AP. GWAMA: Software for genome-wide association meta-analysis. BMC Bioinformatics 2010;11:288.
35. Bonder MJ, Luijk R, Zhernakova DV, Moed M, Deelen P, Vermaat M, van Iterson M, van Dijk F, van Galen M, Bot J, et al. Disease variants alter transcription factor levels and methylation of their binding sites. Nat Genet 2017;49(1):131–8.
36. Benjamini Y, Hochberg Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. J Roy Stat Soc 1995;57:289–300.
37. McLean CY, Bristor D, Hiller M, Clarke SL, Schaar BT, Lowe CB, Wenger AM, Bejerano G. GREAT improves functional interpretation of cis-regulatory regions. Nat Biotechnol 2010;28: 495–501.
38. Speir ML, Zweig AS, Rosenbloom KR, Raney BJ, Paten B, Nejad P, Lee BT, Learned K, Karolchik D, Hinrichs AS, et al. The UCSC Genome Browser database: 2016 update. Nucleic Acids Res 2016;44:D717–25.
39. Devlin B, Roeder K. Genomic control for association studies. Biometrics 1999;55:997–1004.
40. Liu L, Li Y, Tollefsbol TO. Gene–environment interactions and epigenetic basis of human diseases. Curr Issues Mol Biol 2008;10:25–36.
41. Lienert F, Wirbelauer C, Som I, Dean A, Mohn F, Schubeler D. Identification of genetic elements that autonomously determine DNA methylation states. Nat Genet 2011;43: 1091–7.
42. Pedersen BS, Schwartz DA, Yang IV, Kechris KJ. Comb-p: Software for combining, analyzing, grouping and correcting spatially correlated P-values. Bioinformatics 2012;28:2986–8.
43. Wang J, Duncan D, Shi Z, Zhang B. WEB-based GEne SeT AnaLysis Toolkit (WebGestalt): Update 2013. Nucleic Acids Res 2013;41: W77–83.
44. Kok DE, Dhonukshe-Rutten RA, Lute C, Heil SG, Uitterlinden AG, van der Velde N, van Meurs JB, van Schoor NM, Hooiveld GJ, de Groot LC, et al. The effects of long-term daily folic acid and vitamin B12 supplementation on genome-wide DNA methylation in elderly subjects. Clin Epigenetics 2015;7:121.
45. EFSA Panel on Dietetic Products NaAN. Scientific Opinion on Dietary Reference Values for folate. European Food Safety Authority (EFSA) 2015.
46. Institute of Medicine, Food and Nutrition Board. A Report of the Standing Committee on the Scientific Evaluation of Dietary Reference Intakes and its Panel on Folate, Other B Vitamins, and Choline and Subcommittee on Upper Reference Levels of Nutrients. Dietary reference intakes for thiamin, riboflavin, niacin, vitamin b6, folate, vitamin b12, pantothenic acid, biotin, and choline. (Washington (DC): National Academies Press (US); 1998.
47. Medicine Io. Dietary reference intakes: The essential guide to nutrient requirements. Washington (DC): The National Academies Press; 2006.
48. Archangelo LF, Glasner J, Krause A, Bohlander SK. The novel CALM interactor CATS influences the subcellular localization of the leukemogenic fusion protein CALM/AF10. Oncogene 2006;25: 4099–109.
49. Archangelo LF, Greif PA, Holzel M, Harasim T, Kremmer E, Przemeck GK, Eick D, Deshpande AJ, Buske C, de Angelis MH, et al. The CALM and CALM/AF10 interactor CATS is a marker for proliferation. Mol Oncol 2008;2:356–67.
50. Hori A, Morand A, Ikebe C, Frith D, Snijders AP, Toda T. The conserved Wdr8-hMsd1/SSX2IP complex localises to the centrosome and ensures proper spindle length and orientation. Biochem Biophys Res Commun 2015;468:39–45.
51. Kurtulmus B, Wang W, Ruppert T, Neuner A, Cerikan B, Viol L, Dueñas-Sánchez R, Gruss OJ, Pereira G. WDR8 is a centriolar satellite and centriole-associated protein that promotes ciliary vesicle docking during ciliogenesis. J Cell Sci 2016;129:621–36.
52. Wong CC, Martincorena I, Rust AG, Rashid M, Alifrangis C, Alexandrov LB, Tiffen JC, Kober C, Chronic Myeloid Disorders Working Group of the International Cancer Genome Consortium, Green AR, et al. Inactivating CUX1 mutations promote tumorigenesis. Nat Genet 2014;46:33–8.
53. Galvan A, Frullanti E, Anderlini M, Manenti G, Noci S, Dugo M, Ambrogi F, De Cecco L, Spinelli R, Piazza R, et al. Gene expression signature of non-involved lung tissue associated with survival in lung adenocarcinoma patients. Carcinogenesis 2013;34:2767–73.
54. Suzuki K, Lareyre JJ, Sanchez D, Gutierrez G, Araki Y, Matusik RJ, Orgebin-Crist MC. Molecular evolution of epididymal lipocalin genes localized on mouse chromosome 2. Gene 2004;339:49–59.
55. Christensen KE, Mikael LG, Leung KY, Levesque N, Deng L, Wu Q, Malysheva OV, Best A, Caudill MA, Greene ND, et al. High folic acid consumption leads to pseudo-MTHFR deficiency, altered lipid metabolism, and liver injury in mice. Am J Clin Nutr 2015;101: 646–58.
56. Kolde R, Martens K, Lokk K, Laur S, Vilo J. seqlm: An MDL based method for identifying differentially methylated regions in high density methylation array data. Bioinformatics 2016;32:2604–10.
57. EFSA Panel on Dietetic Products, Nutrition, and Allergies (NDA). Scientific opinion on dietary reference values for cobalamin (vitamin B12). European Food Safety Authority (EFSA) 2015.
58. Stirzaker C, Taberlay PC, Statham AL, Clark SJ. Mining cancer methylomes: Prospects and challenges. Trends Genet 2014;30:75–84.