HHS Public Access

Author manuscript

Am J Med Genet A. Author manuscript; available in PMC 2017 October 10.

Published in final edited form as:

Am J Med Genet A. 2016 November ; 170(11): 2777–2787. doi:10.1002/ajmg.a.37874.

Sequence Variation in Folate Pathway Genes and Risks of Human Cleft Lip with or without Cleft Palate

Nicholas J. Marini[1,*], **Wei Yang**[2], **Kripa Asrani**[1], **John S. Witte**[3], **Jasper Rine**[1], **Edward J. Lammer**[4], and **Gary M. Shaw**[2]

[1]Department of Molecular and Cellular Biology, California Institute for Quantitative Biosciences, University of California, Berkeley, California

[2]Department of Pediatrics, Stanford University School of Medicine, Stanford, California

[3]Department of Epidemiology and Biostatistics and Institute of Human Genetics, University of California, San Francisco, California

[4]UCSF Benioff Children's Hospital Oakland, Oakland, California

Abstract

In an effort to comprehensively interrogate genetic variation in the folate pathway for risk of cleft lip with or without cleft palate(CLP), we evaluated 504 common and rare variants in 35 folate-related genes in a panel of 330 infants with CLP and 367 non-malformed controls. Odds ratios (OR) with 95% confidence intervals were computed for common genotypes. A Case-Control Difference metric was calculated for rare variants to highlight differentially occurring alleles. Interactions between variants and a maternal folate intake variable were also evaluated. In gene-only results, significant odds ratios were observed for multiple variants in the BHMT/BHMT2/DMGDH gene cluster, particularly in Hispanic infants. Also in this cluster, rare variant analysis highlighted a substantial case-control difference in BHMT rs60340837 (synonymous Y284Y). In Hispanics, the ALDH1L1 I812V variant (rs4646750) was the most significant risk allele: OR = 3.8 (95% CI = 1.6–9.2) when heterozygous. In non-Hispanic white infants, we observed significant risk for AHCYL2 rs1095423 (homozygous OR = 3.0, 95%CI 1.1–7.8) and the 68 bp CBS insertion (c.844ins68; heterozygous OR = 2.4, 95%CI = 1.1–5.3). Rare variant analysis in this group revealed case-control differences in MTRR and several other methionine cycle genes, a process implicated previously in clefting risk. In women with low folate intake specifically, increased risks were observed for CBS rs2851391 (OR 3.6, 95%CI = 1.3–9.6) and the R259P nonsynonymous variant of TCN2 (rs1801198; OR = 2.8, 95%CI 1.2–6.3). This comprehensive study provides further direction on candidate loci to help disentangle the folate-related developmental phenomena in human clefting risk.

[*]Correspondence to: Nicholas J. Marini, Department of Molecular and Cellular Biology, California Institute for Quantitative Biosciences, University of California, Berkeley, CA 94720, nmarini@berkeley.edu.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article at the publisher's web-site.

The authors declare no conflicts of interest.

Keywords

cleft lip; cleft palate; folic acid; folate pathway; genetic association; gene-environment interaction

INTRODUCTION

Although their causes are largely unknown, orofacial clefts are suspected of being etiologically heterogeneous with both genetic and non-genetic risk factors observed [Dixon et al., 2011]. Clinical observations have long indicated that orofacial clefts should be classified into at least two distinct phenotypic groups: cleft lip with or without cleft palate (CLP), and cleft palate alone (CPO) [Fogh-Anderson, 1967]. Collectively, these are among the most common birth defects with a worldwide prevalence of approximately one in 700 live births [WHO, 2003].

Genetic associations for CLP have been observed for several genes, including *IRF6*, *FGFR2*, *FOXE1*, *MSX1*, *PVRL1*, *BMP4*, *TBX22*, and *TGFa* [Dixon et al., 2011] from either linkage analysis or candidate gene studies. In addition, genome-wide association studies have identified approximately 12 additional risk loci [Birnbaum et al., 2009; Grant et al., 2009; Beaty et al., 2010; Mangold et al., 2010, Ludwig et al., 2012] with some of these replicating in subsequent studies [e.g., Jia et al., 2015]. For most of these loci, the causative variant/gene has yet to be identified, however, large-scale sequencing efforts surrounding these loci may yield compelling candidates for further interrogation [Leslie et al., 2015]. Interestingly, there is apparently little overlap between the most significantly associated GWAS regions and genes previously implicated in CLP risk. This is probably indicative of the complex genetic etiologies underlying the CLP phenotype and the difficulty in detecting associated loci at genome-wide significance. Thus, there are likely to be additional genetic risk factors yet to be identified. Indeed, the combined efforts of all these genetic approaches have been extremely fruitful, yet still only explain a fraction of the population burden of these human birth defects.

Several non-genetic factors also appear to contribute to cleft phenotypes [Mossey et al., 2009]. Of particular relevance to the current study is that maternal use of multivitamin supplements containing folic acid in early pregnancy has been associated with decreased risk of CLP [Shaw et al., 1995; Wilcox et al., 2007]. Although some reports do not claim a benefit from maternal folate supplementation [see Little et al., 2008], a recent meta-analysis indicated a significant reduction in risk of CLP with maternal folic acid use [Butali et al., 2013]. Furthermore, using improved methods for estimating nutrient intake, Wallenstein et al. [2013] simi-] similarly observed a decreased risk for CLP with maternal folate intake as well as for other micronutrients such as niacin, riboflavin and vitamin B12. Interestingly, studies report a significant folate-risk relationship for CLP only (not CPO), which may be a hallmark distinguishing these phenotypes. It should be emphasized that the underlying mechanisms by which folic acid may reduce CLP risks are unknown, but potential candidates include available levels of the nutrient and its genetic control with downstream effects involving homocysteine accumulation, cellular methylation, and nucleotide biosynthesis [Obican et al., 2010].

In addition to epidemiological observations, the link between folate and craniofacial development has also been established in several vertebrate models. Folate deficiency and mutations in folate pathway genes cause orofacial defects [Tang et al., 2005; Lee et al., 2012; Momb et al., 2013]. In a recent study, impairment of folate pathway function causes orofacial malformations in a *Xenopus* model and, significantly, folate status affects the severity of an induced cleft phenotype in the same model [Wahl et al., 2015].

These model system results taken together with human epidemiological observations provide a compelling relationship between folate metabolism and risk for CLP. Given that there are undoubtedly additional genetic risk factors contributing to CLP occurrence, the genes of the folate pathway seem a logical place to look for such risk [Boyles et al., 2009; Blanton et al., 2011]. Thus, our objective in this study was to query in-depth the variation in folate-related pathway genes within a CLP case-control study population for disease-associated risk. To this end, we genotyped 504 variants in 35 folate pathway genes and, as such, this study represents the most comprehensive interrogation of common and low-frequency variation related to folate metabolism for CLP. We also investigated effect measure modification between variants and maternal folate intake for risk of CLP.

METHODS

Study Population

This case-control study included data on deliveries that had estimated due dates from 1995 to 1998 or 1999 to 2003. The study included live-born infants with isolated CLP (cases; N = 330 [149 from 1995 to 1998 and 181 from 1999 to 2003]) or without any structural malformation (controls; N = 367 [162 from 1995 to 1998 and 205 from 1999 to 2003]). Case information was abstracted from hospital reports and medical records following established procedures by the California Birth Defects Monitoring Program [Croen et al., 1991]. Each medical record was further reviewed by a medical geneticist (E.J.L.). Infants with trisomies were ineligible. Non-malformed controls were selected randomly to represent the population from which the cases were derived in selected counties and birth periods. This study, including the collection and use of archived newborn bloodspots, was approved by the California State Committee for the Protection of Human Subjects as well as Institutional Review Boards at Stanford University and the University of California, Berkeley. See Table I for breakdown of population demographics.

Ascertainment of Maternal Data

Mothers who delivered in the period 1999–2003 were eligible for interview if they were the biologic mother and carried the pregnancy of the study subject, they were not incarcerated, and their primary language was English or Spanish. Maternal interviews were conducted using a standardized, computer-based questionnaire, primarily by telephone, in English or Spanish, no earlier than 6 weeks after the infant's estimated date of delivery. Information on a variety of exposures and conditions was solicited from women including age, race/ethnicity, educational level, and family history (first degree relative) of clefts. Queries specific to the periconceptional period, which was defined as 2 months before through two months after conception, included use of folic-acid containing vitamin supplements, diabetes

(gestational, Type I, and Type II), and seizure medication use. The interview also included a modified version of the National Cancer Institute's Health Habits and History Questionnaire, a well-known, semi-quantitative food frequency questionnaire with demonstrated reliability and validity [Block et al., 1990]. The food frequency questionnaire was modified to include ethnic foods appropriate to a diverse study population. This questionnaire provided information on dietary folate and energy intake. The median time between estimated date of delivery and interview completion was 11 months for cases and 9 months for controls.

Genomic DNA Isolation

Genomic DNA (gDNA) was available from newborn screening dried blood spots obtained from linkage efforts made by the California Birth Defects Monitoring Program. gDNA was extracted from dried blood spots and whole-genome amplified using QIAmp DNA Mini kits and Repli-G Midi amplification kits (Qiagen, Germantown, MD), respectively, as described previously [Marini et al., 2011]. The average yield of purified whole-genome amplified (WGA) DNA was 15 mg per individual.

SNP Selection and Genotyping

SNPs for genotyping were identified from the following sources: (i) the sequencing of 31 folate pathway genes to identify risk factors for neural tube defects [Marini et al., 2011]. From that study, we chose all common alleles (MAF > 5%), all nonsynonymous changes (including singletons), and all variants that displayed a trend toward case-control association (P value <0.15). (ii) Additional nonsynonymous SNPs in these 31 genes present in public SNP databases (NCBI). (iii) SNPs previously published as having potential functional significance in folate-related genes. (iv) Haplotype tagging SNPs for a subset of folate pathway genes to enhance coverage beyond what was seen from exon sequencing [Marini et al., 2011].

Genotyping was performed primarily using TaqMan SNP genotyping assays (ThermoFisher, Foster City, CA) on the Fluidigm EP1 platform (Fluidigm, South San Francisco, CA) using a 96×96 dynamic array. Briefly, specific target amplification in each sample was performed in pooled groups of 96 primer pairs and then subject to individual TaqMan assays taking care to not include SNP sites within 200 bp of each other. All allele calls were checked manually following automated allele calling by the Fluidigm SNP Genotyping Analysis software. For multiple SNPs in close proximity and repeat polymorphisms that could not be genotyped by TaqMan (N = 25), genotyping was performed by Sanger sequencing of amplified PCR products. As quality assurance, genotyping assays were duplicated for a 10% subset of DNA samples. All genotyping was performed blinded to case and control status. After removing QC failures and mono-allelic SNPs, the final analytical list comprised 504 variants in 35 genes which are listed, along with attributes, in Table SI.

Analytical Methods

Variants were treated as categorical variables, that is, homozygous wild-type as referent versus heterozygous or homozygous variant. Odds ratios and 95% confidence intervals were used to estimate relative risks (using SAS software version 9.4). For low frequency variants not amenable to OR calculation, we developed a simple metric ("Case-Control Difference")

which highlights potential risk alleles with the most differential occurrence between cases and controls. The Case-Control Difference (CCD) is calculated by the following formula: $|x-y| / sqrt(x+y)$, where x and y equal the normalized allele counts in cases and controls, respectively. Thus, the CCD is a measure of the differential appearance of alleles in cases and controls (numerator) with a penalty for the occurrence of that allele in both populations (denominator). For example, a low frequency variant with five occurrences in cases and zero in controls would be an interesting risk candidate and receive a higher CCD score than an allele that occurs 10 times in cases and five times in controls. Although this tool was developed for low frequency variants, CCD scores can be quite high for higher-frequency alleles if there is a considerable case-control difference, though these alleles may be better highlighted by OR calculations (Table SII). The numerator is an absolute value so the metric reveals a measure where alleles show either increased or decreased risks. It should be emphasized that the CCD merely highlights differentially behaving alleles and does not ascribe statistical significance.

Gene-folate interactions were assessed for homozygous or heterozygous variants. A dichotomized composite variable, "combined folate intake," was created by combining maternal periconceptional intake of folic acid-containing vitamin supplements with daily dietary intake of folate. Two categories were constructed corresponding to ≤ 25th percentile (to reflect low dietary folate intake) and >25th percentile, as determined from dietary folate intake levels among control mothers (312.67 μg). Combined folate intake was defined as low for women in the lowest quartile (≤ 25th percentile) who did not take supplemental folic acid in the periconceptional period. Not-low folate intake was defined as dietary folate intake above the lowest quartile (>25th percentile) regardless of maternal vitamin use. Interactions between each variant and the combined folate intake variable were investigated for effects on risk of CLP. An interaction with *P*-value <0.05 for the Wald test was considered statistically significant. Given that our study time period overlapped the introduction of folic acid fortification of grain products in the US (1998), we analyzed whether genetic results differed between the periods before and after fortification. For this analysis, we specified an index of 0 for before fortification (with estimated due dates from 1995 to 1998) and an index of 1 for after fortification (with estimated due dates from 1999 to 2003). The interaction between each SNP and this index was tested with the Wald test with *P*-value <0.05 as the statistically significant determinant.

RESULTS

As shown in Table I, case infants were not substantially different from controls with respect to race/ethnic background. The well-established male predominance for CLP was observed [Brydon et al., 2014]. Folate intake was observed to be similar in case compared to control mothers. In all, 504 variants in 35 folate-pathway genes (Table SI) were genotyped with a high success rate: 483 variant positions were called with >95% call rate with no positions below 85% call rate. All 504 variants were tested for association in the entire population, as well as within race-ethnic strata. Results for all polymorphisms are presented in Table SII.

Common Variants

Table II shows all significant associations for individual SNPs when heterozygous and homozygous (calculated separately as in Methods) within the entire population, as well as within non-Hispanic white and Hispanic white sub-groups. There were no variants statistically significantly associated (*P*-value <0.05) with CLP as both heterozygotes and homozygotes in the entire population (Table II). ALDH1L1_74483, a nonsynonymous I812V change, showed significant increased risk when heterozygous, but was too infrequently homozygous to estimate the OR. Of those alleles with frequencies high enough to estimate homozygous ORs, only GART_26477 showed the trend of increasing risk with variant allele number, but was significant only when homozygous. In addition, several variants displayed decreased risk trending with variant allele number: two tightly linked loci within the BHMT2 locus (BHMT2_8849, BHMT2_10283) significantly reduced risk when homozygous whereas MTHFD1_59019 was significant only as a heterozygote.

In non-Hispanic white infants, an intronic SNP in AHCYL2 (AHCYL2_165230) displayed significant risk when heterozygous (OR = 2.6; 95%CI 1.1–5.9) and homozygous (OR = 3.0; 95%CI 1.1–7.8). Interestingly, a second set of three tightly linked variants in the AHCYL2 gene (AHCYL2_165528, 173166, 203190) also displayed association with a significant reduction in risk seen only for homozygotes. In the CBS gene, the 68-bp insertion in intron 7 (c.844ins68, here called CBS_13779) and a second linked locus (CBS_13704) displayed significant risk when heterozygous, whereas a low-frequency intronic SNP in DHFR (DHFR_17874) displayed significantly reduced risk in heterozygotes. It should be emphasized, however, that the non-Hispanic white sub-population in the study was fairly small (Table I).

In Hispanic infants, the strongest associations emanated from the ALDH1L1 gene and the BHMT/BHMT2/DMGDH locus. BHMT, BHMT2, and DMGDH lie adjacent within a 140 kb cluster on chromosome 5. Thus, several sets of variants exist that span all three genes and are in strong linkage disequilibrium effectively tagging the locus. BHMT2_13828, an intronic variant in one such set, is the only polymorphism that showed a significant OR when both heterozygous (OR = 0.6; 95%CI 0.4–0.9) and homozygous (OR = 0.5; 95%CI 0.3–0.8) with effect trending with allele number. Four other tightly linked variants from this locus (BHMT2_8849, BHMT2_10283; DMGDH_26164, DMGDH_26338) also showed significantly reduced risk when homozygous. The nonsynonymous I812 V variant of ALDH1L1 (ALDH1L1_74483) showed a very significant heterozygote risk (OR = 3.8; 95%CI 1.6–9.2) suggesting that the Hispanic sub-group drove the association seen in the overall population. Interestingly, two additional ALDH1L1 alleles, ALDH1L1_74427 (nonsynonymous, D793G) and ALDH1L1_77540 (intronic), which are tightly linked to each other but not to ALDH1L1_74483, were associated in Hispanics with significantly reduced risk when heterozygous.

Low-Frequency Variants

Low-frequency variation, particularly in coding regions, has been implicated in risk for orofacial clefting [e.g., Leslie and Murray, 2013, Leslie et al., 2015]. Our study data contained a substantial number of variants with MAF <5% (N = 267; see Table SII) which

often render OR calculations unstable. To highlight low frequency variants that may contribute to disease risk, we used a simple metric, the Case-Control Difference, which quantifies and rank orders variants with the most differential behavior in cases versus controls (see Methods). We applied this analysis to all variants with MAF <5% and the top 5% of rank-ordered lists for the entire population, as well as non-Hispanic and Hispanic sub-groups, are shown in Table III. Note that some variant alleles were frequent enough to allow OR calculations and this metric can further draw attention to variants previously observed with significant odds ratios (e.g., ALDH1L1_74483). The case-control difference also gives emphasis to low-frequency variants that narrowly miss statistical significance cut-offs. Of note in this group are DHFR_17874, significantly associated in non-Hispanic whites (Table II), but with a relatively strong case-control difference in the entire population, a new putative risk allele in BHMT (BHMT_17018; a synonymous Y284Y variant), and a variant within the 3′-UTR of SHMT1 (SHMT1_35845). In addition, the case-control difference revealed three unlinked MTHFD1 alleles (MTHFD1_12347, 32373, 54248) among variants with the most differential occurrences.

In non-Hispanic whites, the case-control difference highlighted three tightly linked rare, nonsynonymous variants within the MTRR gene with notable differential behavior [MTRR_10208 (S257T), 21088 (R415C), 23290 (P450R)]. The metric also flagged an intronic variant in MAT1A (MAT1A_5181), which displays strong differential behavior in the entire population as well, and the significantly associated DHFR_17874 allele (see Table II).

In Hispanics, the highest scoring differential allele was ALDH1L1_74483, previously identified as a significant risk allele by OR (Table II). A second, rare variant within this gene (ALDH1L1_44924) was also among the most differentially occurring alleles. In addition, this analysis flagged four unlinked variants from homocysteine trans-sulfuration genes CBS and CTH as well as the MTHFD1 allele (MTHFD1_43530), which was significantly associated in the entire population.

Gene-Folate Interaction

We explored the potential interaction between maternal folate intake and each individual SNP. Thus, the population was stratified into “low” and “not-low” folate intake groups based on the combined data of supplementation and dietary intake (see Methods). Our a priori expectation was that increased risks for a particular variant would be substantially greater in the low folate group than in the not-low folate group. As shown in Table IV, there were a few associations that fit this pattern of risk with a *P*-value <0.05 for the Wald test of gene-folate interaction. In the entire population, low-folate specific risk was seen for an intronic variant of CBS (CBS_9548; OR 3.6, 95%CI 1.3–9.6) and two linked alleles of the cobalamin (vitamin B12) transporter TCN2: TCN2_9397 (intron) and the well-studied nonsynonymous SNP R259P (TCN2_9459; rs1801198), both with OR 2.8, 95%CI 1.2–6.3). Neither the CBS allele nor the TCN2 alleles were seen in the gene-only analyses of Table II. In fact, no alleles from either gene were significantly associated in the entire population. Such folate-responsive alleles might not be expected to be observed as significant in the entire population without folate intake stratification because the low-folate intake sub-group is

relatively small (28 cases, 23 controls). The only significant odds ratio for increased risk with low-folate intake within race-ethnic sub-groups was CTH_4763 in Hispanics (OR 11.5, 95%CI 1.8–72.9), but the sample size was too small for meaningful conclusions.

We also investigated whether the gene variant analyses were modified by the introduction of folic acid fortification of grain products in the US, that is, we analyzed whether genetic results differed between the periods before and after fortification. None of the observed *P*-values indicated that this term when added to the variant models produced statistically precise results (data not shown).

DISCUSSION

In this study we estimated the association of 504 variants across 35 folate pathway genes with nonsyndromic CLP in a 697-member case-control population. This variant set was rich in low-frequency alleles(N = 267variants with MAF<5%) as well as nonsynonymous changes (N = 144) and thus represents the most comprehensive collection of common and low-frequency variation related to folate metabolism interrogated for CLP. For higher-frequency alleles, we analyzed disease association by ORs of individual genotypes (heterozygotes and homozygotes). For lower-frequency alleles where OR calculations are problematic, we used a simple but useful metric (Case-Control Difference), which merely highlights variants with relatively strong case or control skews, without ascribing any statistical significance. We further investigated whether maternal folate intake modified the effects of variants for risk of CLP.

Both OR and low-frequency variant analysis have implicated several candidate loci for CLP risk. For example, significant ORs were observed for variants in the BHMT/BHMT2/DMGDH gene cluster, particularly in Hispanic infants. Additionally, an unlinked, synonymous change in BHMT (Y284Y; BHMT_17018) was flagged by the case-control difference. Similarly, nominally significant associations were observed for alleles in MTHFD1 (Table II) while the case-control difference revealed several low-frequency MTHFD1 alleles among the most differentially occurring in cases vs. controls (Table III). In Hispanics, the ALDH1L1 gene contained variants highlighted by the case-control difference as well as multiple with significant ORs. In fact, the I812V nonsynonymous variant (ALDH1L1_74483) was the most significant risk allele in this study. In the DHFR gene, the case-control difference highlighted two unlinked alleles with relatively strong differential behavior in the entire population, with one of these (DHFR_17874) displaying significant association in non-Hispanic whites. Also in non-Hispanic white infants, it is notable that the group of genes implicated by both OR and case-control difference calculations (AHCYL2, CBS, MTRR, and MAT1A) all impinge on the methionine cycle, a process implicated previously in Cleft risk [e.g., Blanton et al., 2011].

It is worth noting that significance levels (*P* values) in this study are modest, and no findings would retain significance following any multiple comparison correction. This tends to be a hallmark of CLP/folate pathway genetic association studies [for partial review, see Bhaskar et al., 2011]. In addition, because there have been many such studies, there are numerous examples of SNP associations that fail to replicate. It has been suggested [Boyles et al.,

2009] that this conglomeration of nominal and/or conflicting results may indicate that either risks emanating from this group of genes is minimal or non-existent. On the other hand, modest risks and their attendant statistical imprecision should be expected for studies of this size since the genetics underlying CLP are probably heterogeneous involving different alterations of multiple genes. In this case, association analysis will likely only reveal multiple nominal risks rather than a small number of striking ones. Indeed, even though results may seem equivocal when looking at individual polymorphisms, there are certain recurring themes implicating genes or gene groups in CLP risk.

BHMT/BHMT2/DMGDH

One of the most consistent observations in the literature regarding folate pathway-related genes and CLP risk is the frequently observed significant association of variants within the BHMT/BHMT2/DMGDH gene cluster on chromosome 5. These three genes lie in tandem within 140 kb at chromosomal region 5q14. Because of their close proximity, there are strong blocks of linkage disequilibrium that span these genes. Thus, a positive signal from one variant could, in theory, be tagging an allele anywhere in the region. Table V summarizes previous studies (including this one) in which variants in this region have been tested for CLP association. This table lists all previously published studies querying SNPs in this region, with a brief description of study population and those SNPs displaying significant associations. Quite convincingly, out of nine studies that use a variety of analytical techniques, seven (~ 80%) report at least one significant association as determined by the investigators. In this regard, this locus stands apart from most other folate genes interrogated in CLP association studies. Consistent with our results, Blanton et al. [2011] observed association primarily in Hispanics. It is noteworthy that several studies in Table V have observed significant associations with the maternal genotype at this locus. Given the correlation between maternal and fetal genotype, it is possible that fetal associations are a consequence of the fetal genotype simply serving as a proxy for the maternal genotype. It is currently unknown whether fetal or maternal genetic variation (at this or other loci) is more relevant for CLP risk.

While significant associations of common variants within this gene cluster are often observed, the risk attributed to the queried variants may be increased (odds ratios >1) or decreased (odds ratios <1) as reported in this study (see Table II and references in Table V). While the reason for this is not clear, this may reflect the contribution of underlying rare variants at this locus (which may occur on the haplotype of the major or minor allele of a nearby common variant) as has been previously suggested for clefts [Leslie and Murray, 2013] and as has been seen for regions highlighted by GWAS studies [Leslie et al., 2015].

There are further lines of evidence that lend support to this region as a genuine risk locus for CLP. BHMT/BHMT2 (Betaine-homocysteine methyltransferase) and DMGDH (Dimethyl-glycine dehydrogenase) function sequentially in a multi-step process converting choline to glycine. Measuring the relevant metabolites choline and betaine in maternal mid-pregnancy sera, Shaw et al. [2009] found altered metabolite levels in CLP case mothers, suggesting metabolic imbalances. Furthermore, increased intake of choline in the maternal diet has been associated with decreased CLP risk [Shaw et al., 2006]. In mouse models, BHMT knock-

outs are not reported to increase the risk of clefts in offspring, though choline and one-carbon metabolism are perturbed [Teng et al., 2011]. Thus genetic studies, metabolite analysis and epidemiology converge at this locus for CLP risk.

DHFR

We identified two low-frequency intronic variants within the DHFR gene with relatively strong differences in cases and controls (DHFR_17874, DHFR_26717). DHFR_17874 had a significant odds ratio in non-Hispanic whites (OR 0.2; 95%CI 0.1–0.8). Martinelli et al. [2014] also observed significant associations with reduced risk for CLP within DHFR in an Italian population, though one of these SNPs (DHFR_15483 here) did not show a significant association in the current study. Mechanistically, DHFR has been implicated in orofacial clefting from model system studies. Wahl et al. [2015] reported that methotrexate, a Dihydrofolate reductase inhibitor, causes craniofacial malformations and exacerbates clefting phenotypes in a *Xenopus* model. In addition, teratology studies in mice and rabbits show that maternal methotrexate exposure during early pregnancy results in orofacial clefting outcomes in offspring [Hyoun et al., 2012].

Methionine Cycle

A subset of folate-dependent reactions plays a critical role in the synthesis of S-adenosylmethionine (AdoMet), which serves as the methyl group donor in DNA and protein methylation reactions. Homocysteine, which is the precursor of methionine (and, therefore, AdoMet) is also subsequently regenerated following AdoMet methyl transfer in a cyclic series of reactions referred to as the “methionine cycle.” The genes/enzymes of the methionine cycle have garnered interest in identifying risk for orofacial clefts because of some studies that report high plasma homocysteine levels in cleft case mothers [Wong et al., 1999; Kumari et al., 2013], though this does not appear to be the case in all populations queried [Shaw et al., 2009; Munger et al., 2011]. Nevertheless, it is noteworthy that the BHMT/BHMT2 reactions (whose genes are often implicated in cleft risk as discussed above) convert homocysteine to methionine, and thus can impact the methionine cycle. Furthermore, we have seen a strong representation of genes that impinge on the methionine cycle in non-Hispanic white cases: ACHYL2_165230 was the most significant risk allele in this group. A significant risk signal also emanated from the CBS (cystathionine-beta-synthase) 68bp insertion (c.844ins68; here CBS_13779) and a tightly linked allele. The case-control difference metric highlighted alleles in MTRR (Methionine Synthase Reductase; necessary for the folate-dependent conversion of homocysteine to methionine) and MAT1A (necessary for AdoMet synthesis). The 844ins68 allele of CBS has been tested in other populations with typically conflicting results. For example, Rubini et al.[2005] reported evidence of transmission distortion and imprinting for this allele in an Italian sample of 134 CLP triads, whereas Martinelli et al. [2011] found no association in a different Italian population. However, evidence of association for other alleles of CBS has been reported in multiple studies [Boyles et al., 2008, 2009; Blanton et al., 2011; Martinelli et al., 2011]. Finally, it is worth noting that the alleles/genes highlighted as conferring risk specifically with low maternal folate intake (Table IV) participate in the methionine cycle: CBS (CBS_9548) and the TCN2 cobalamin (vitamin B12) transporter (TCN2_9397, TCN2_9450). Vitamin B12 is essential for the Methionine synthase (MTR) reaction

converting homocysteine to methionine and, thus, TCN2 activity impinges on homocysteine levels [Hsu et al., 2011].

Thus, these results re-focus emphasis on homocysteine clearance and the methionine cycle, yet the underlying mechanism is still unclear. Elevated homocysteine levels may be a specific marker for DNA methylation cycle aberration or could merely be indicative of altered folate metabolism in general. Wahl et al. [2015] have suggested that deficiencies in the folate-dependent synthesis of nucleotide precursors, rather than cellular methylation changes, underlie clefts in a *Xenopus* model. These defects lead to impairment of DNA synthesis, genomic instability and, ultimately, apoptosis.

MTHFD1

We have identified alleles of MTHFD1 that are significantly associated with CLP in the entire population (MTHFD1_43530, 59019; Table II) as well as three additional, unlinked alleles highlighted by the case-control difference as some of the most differentially occurring in cases and controls (Table III). Similarly, alleles of MTHFD1 were also flagged by odds ratio and the case-control difference in Hispanics. To our knowledge, these variants have not been tested in previous studies, though other MTHFD1 variants have been tested for CLP risk. [Mostowska et al., 2006; Mills et al., 2008; Palmieri et al., 2008; Boyles et al., 2009; Blanton et al., 2011]. Mills et al. [2008] reported a significant association with the common, nonsynonymous R653Q variant and Blanton et al. [2011] reported significant associations in non-Hispanic white individuals for multiple MTHFD1 variants. It is noteworthy that, whereas homozygous deletions of MTHFD1 are embryonic lethal in mice, heterozygous mutations have negative impacts on purine synthesis (a folate-dependent process) and increased incidence of developmental defects in embryos including craniofacial malformations [Christensen et al., 2013]. MTHFD1 may therefore be considered a candidate risk gene.

The most significant risk allele in our study was ALDH1L1_74483, a nonsynonymous I812V alteration strongly associated in Hispanics. To the best of our knowledge, this polymorphism has not been previously tested for CLP association. Boyles et al. [2009] tested several ALDH1L1 variants, but none were significantly associated with CLP in their study. There is little phenotypic information surrounding ALDH1L1 mutation to implicate it in craniofacial development, thus the importance of ALDH1L1 variation in cleft risk remains to be determined.

This study has implicated several loci related to folate metabolism as important in conferring risk for isolated clefts. Although these findings are consistent with previous studies and with biological mechanisms, they were based on relatively small sample sizes. In particular, the power to detect gene-nutrient interaction effects was low. Therefore, replication of these findings in additional populations is warranted.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank the California Department of Public Health Maternal Child and Adolescent Health Division for providing data. This project was partially supported by NIH R01 HD074695 (N.J.M), R01 GM072859 (J.D.R.), CDC 6U01-DD-000489 (G.M.S.), and 1U01-DD-0006982 (G.M.S.). The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the California Department of Public Health. The authors wish to give special acknowledgement to co-author Edward Lammer, who passed away as this manuscript was nearing completion. Ed was instrumental to this research and, more broadly, to the understanding of birth defects risk and causation. Ed was an outstanding colleague and we will surely miss him.

Grant sponsor: NIH; Grant numbers: R01 HD074695, R01 GM072859, CDC 6U01-DD-000489, 1U01-DD-0006982.

References

Beaty TH, Murray JC, Marazita ML, Munger RG, Ruczinski I, Hetmanski JB, Liang KY, Wu T, Murray T, Fallin MD, Redett RA, Raymond G, Schwender H, Jin SC, Cooper ME, Dunnwald M, Mansilla MA, Leslie E, Bullard S, Lidral AC, Moreno LM, Menezes R, Vieira AR, Petrin A, Wilcox AJ, Lie RT, Jabs EW, Wu-Chou YH, Chen PK, Wang H, Ye X, Huang S, Yeow V, Chong SS, Jee SH, Shi B, Christensen K, Melbye M, Doheny KF, Pugh EW, Ling H, Castilla EE, Czeizel AE, Ma L, Field LL, Brody L, Pangilinan F, Mills JL, Molloy AM, Kirke PN, Scott JM, Arcos-Burgos M, Scott AF. A genome-wide association study of cleft lip with and without cleft palate identifies risk variants near MAFB and ABCA4. Nat Genet. 2010; 42:525–529. [PubMed: 20436469]

Bhaskar LV, Murthy J, Venkatesh Babu G. Polymorphisms in genes involved in folate metabolism and orofacial clefts. Arch Oral Biol. 2011; 56:723–737. [PubMed: 21310392]

Birnbaum S, Ludwig KU, Reutter H, Herms S, Steffens M, Rubini M, Baluardo C, Ferrian M, Almeida de Assis N, Alblas MA, Barth S, Freuden-berg J, Lauster C, Schmidt G, Scheer M, Braumann B, Bergé SJ, Reich RH, Schiefke F, Hemprich A, Pötzsch S, Steegers-Theunissen RP, Pötzsch B, Moebus S, Horsthemke B, Kramer FJ, Wienker TF, Mossey PA, Propping P, Cichon S, Hoffmann P, Knapp M, Nöthen MM, Mangold E. Key susceptibility locus for non-syndromic cleft lip with or without cleft palate on chromosome 8q24. Nat Genet. 2009; 41:473–477. [PubMed: 19270707]

Blanton SH, Henry RR, Yuan Q, Mulliken JB, Stal S, Finnell RH, Hecht JT. Folate pathway and nonsyndromic cleft lip and palate. Birth Defects Res A Clin Mol Teratol. 2011; 91:50–60. [PubMed: 21254359]

Block G, Woods M, Potosky A, Clifford C. Validation of a self-administered diet history questionnaire using multiple diet records. J Clin Epidemiol. 1990; 43:1327–1335. [PubMed: 2254769]

Boyles AL, Wilcox AJ, Taylor JA, Meyer K, Fredriksen A, Ueland PM, Drevon CA, Vollset SE, Lie RT. Folate and one-carbon metabolism gene polymorphisms and their associations with oral facial clefts. Am J Med Genet Part A. 2008; 146A:440–449. [PubMed: 18203168]

Boyles AL, Wilcox AJ, Taylor JA, Shi M, Weinberg CR, Meyer K, Fredriksen A, Ueland PM, Johansen AM, Drevon CA, Jugessur A, Trung TN, Gjessing HK, Vollset SE, Murray JC, Christensen K, Lie RT. Oral facial clefts and gene polymorphisms in metabolism of folate/one-carbon and vitamin A: A pathway-wide association study. Genet Epidemiol. 2009; 33:247–255. [PubMed: 19048631]

Brydon CA, Conway J, Kling R, Mehta L, Jabs EW, Taub PJ. Cleft lip and/or palate: One organization's experience with more than a quarter million surgeries during the past decade. J Craniofac Surg. 2014; 25:1601–1609. [PubMed: 25072973]

Butali A, Little J, Chevrier C, Cordier S, Steegers-Theunissen R, Jugessur A, Oladugba B, Mossey PA. Folic acid supplementation use and the MTHFR C677T polymorphism in orofacial clefts etiology: An individual participant data pooled-analysis. Birth Defects Res A Clin Mol Teratol. 2013; 97:509–514. [PubMed: 23670871]

Christensen KE, Deng L, Leung KY, Arning E, Bottiglieri T, Malysheva OV, Caudill MA, Krupenko NI, Greene ND, Jerome-Majewska L, MacKenzie RE, Rozen R. A novel mouse model for genetic variation in 10-formyltetrahydrofolate synthetase exhibits disturbed purine synthesis with impacts

on pregnancy and embryonic development. Hum Mol Genet. 2013; 22:3705–3719. [PubMed: 23704330]

Croen LA, Shaw GM, Jensvold NG, Harris JA. Birth defects monitoring in California: A resource for epidemiological research. Paediatr Perinat Epidemiol. 1991; 5:423–427. [PubMed: 1754501]

Dixon MJ, Marazita ML, Beaty TH, Murray JC. Cleft lip and palate: Understanding genetic and environmental influences. Nat Rev Genet. 2011; 12:167–178. [PubMed: 21331089]

Fogh-Anderson P. Genetic and non-genetic factors in the etiology of facial clefts. Scand J Plastic Reconstr Surg. 1967; 1:22–29.

Grant SF, Wang K, Zhang H, Glaberson W, Annaiah K, Kim CE, Bradfield JP, Glessner JT, Thomas Garris KAM, Frackelton EC, Otieno FG, Chiavacci RM, Nah HD, Kirschner RE, Hakonarson H. A genome-wide association study identifies a locus for nonsyndromic cleft lip with or without cleft palate on 8q24. J Pediatr. 2009; 155:909–913. [PubMed: 19656524]

Hsu FC, Sides EG, Mychaleckyj JC, Worrall BB, Elias GA, Liu Y, Chen WM, Coull BM, Toole JF, Rich SS, Furie KL, Sale MM. Transcobalamin 2 variant associated with poststroke homocysteine modifies recurrent stroke risk. Neurology. 2011; 77:1543–1550. [PubMed: 21975197]

Hyoun SC, Običan SG, Scialli AR. Teratogen update: Methotrexate. Birth Defects Res A Clin Mol Teratol. 2012; 94:187–207. [PubMed: 22434686]

Jia Z, Leslie EJ, Cooper ME, Butali A, Standley J, Rigdon J, Suzuki S, Gongorjav A, Shonkhuuz TE, Natsume N, Shi B, Marazita ML, Murray JC. Replication of 13q31.1 association in nonsyndromic cleft lip with cleft palate in Europeans. Am J Med Genet Part A. 2015; 167A:1054–1060. [PubMed: 25786657]

Jin LL, Chen EJ, Hou W, Liu XH, Hu Y. The association between folate pathway genes and cleft lip with or without cleft palate in a chinese population. Biomed Environ Sci. 2015; 28:136–139. [PubMed: 25716564]

Kumari P, Ali A, Sukla KK, Singh SK, Raman R. Lower incidence of nonsyndromic cleft lip with or without cleft palate in females: Is homocysteine a factor? J Biosci. 2013; 38:21–26. [PubMed: 23385809]

Lee MS, Bonner JR, Bernard DJ, Sanchez EL, Sause ET, Prentice RR, Burgess SM, Brody LC. Disruption of the folate pathway in zebrafish causes developmental defects. BMC Dev Biol. 2012; 12:12. [PubMed: 22480165]

Leslie EJ, Murray JC. Evaluating rare coding variants as contributing causes to non-syndromic cleft lip and palate. Clin Genet. 2013; 84:496–500. [PubMed: 22978696]

Leslie EJ, Taub MA, Liu H, Steinberg KM, Koboldt DC, Zhang Q, Carlson JC, Hetmanski JB, Wang H, Larson DE, Fulton RS, Kousa YA, Fakhouri WD, Naji A, Ruczinski I, Begum F, Parker MM, Busch T, Standley J, Rigdon J, Hecht JT, Scott AF, Wehby GL, Christensen K, Czeizel AE, Deleyiannis FW, Schutte BC, Wilson RK, Cornell RA, Lidral AC, Weinstock GM, Beaty TH, Marazita ML, Murray JC. Identification of functional variants for cleft lip with or without cleft palate in or near PAX7, FGFR2, and NOG by targeted sequencing of GWAS loci. Am J Hum Genet. 2015; 96:397–411. [PubMed: 25704602]

Little J, Gilmour M, Mossey PA, Fitzpatrick D, Cardy A, Clayton-Smith J, Fryer AE, ITS MAGIC collaboration. Folate and clefts of the lip and palate-a U.K.-based case-control study: Part I: Dietary and supplemental folate. Cleft Palate Craniofac J. 2008; 45:420–427. [PubMed: 18616361]

Ludwig KU, Mangold E, Herms S, Nowak S, Reutter H, Paul A, Becker J, Herberz R, AlChawa T, Nasser E, Böhmer AC, Mattheisen M, Alblas MA, Barth S, Kluck N, Lauster C, Braumann B, Reich RH, Hemprich A, Pötzsch S, Blaumeiser B, Daratsianos N, Kreusch T, Murray JC, Marazita ML, Ruczinski I, Scott AF, Beaty TH, Kramer FJ, Wienker TF, Steegers-Theunissen RP, Rubini M, Mossey PA, Hoffmann P, Lange C, Cichon S, Propping P, Knapp M, Nöthen MM. Genome-wide meta-analyses of nonsyndromic cleft lip with or without cleft palate identify six new risk loci. Nat Genet. 2012; 44:968–971. [PubMed: 22863734]

Mangold E, Ludwig KU, Birnbaum S, Baluardo C, Ferrian M, Herms S, Reutter H, de Assis NA, Chawa TA, Mattheisen M, Steffens M, Barth S, Kluck N, Paul A, Becker J, Lauster C, Schmidt G, Braumann B, Scheer M, Reich RH, Hemprich A, Pötzsch S, Blaumeiser B, Moebus S, Krawczak M, Schreiber S, Meitinger T, Wichmann HE, Steegers-Theunissen RP, Kramer FJ, Cichon S, Propping P, Wienker TF, Knapp M, Rubini M, Mossey PA, Hoffmann P, Nöthen MM. Genome-

wide association study identifies two susceptibility loci for nonsyndromic cleft lip with or without cleft palate. Nat Genet. 2010; 42:24–26. [PubMed: 20023658]

Marini NJ, Hoffmann TJ, Lammer EJ, Hardin J, Lazaruk K, Stein JB, Gilbert DA, Wright C, Lipzen A, Pennacchio LA, Carmichael SL, Witte JS, Shaw GM, Rine J. A genetic signature of spina bifida risk from pathway-informed comprehensive gene-variant analysis. PLoS ONE. 2011; 6:e28408. [PubMed: 22140583]

Martinelli M, Masiero E, Carinci F, Morselli PG, Pezzetti F, Scapoli L. New evidence for the role of cystathionine beta-synthase in non-syndromic cleft lip with or without cleft palate. Eur J Oral Sci. 2011; 119:193–197.

Martinelli M, Girardi A, Cura F, Carinci F, Morselli PG, Scapoli L. Evidence of the involvement of the DHFR gene in nonsyndromic cleft lip with or without cleft palate. Eur J Med Genet. 2014; 57:1–4. [PubMed: 24361572]

Mills JL, Molloy AM, Parle-McDermott A, Troendle JF, Brody LC, Conley MR, Cox C, Pangilinan F, Orr DJ, Earley M, McKiernan E, Lynn EC, Doyle A, Scott JM, Kirke PN. Folate-related gene polymorphisms as risk factors for cleft lip and cleft palate. Birth Defects Res A Clin Mol Teratol. 2008; 82:636–643. [PubMed: 18661527]

Momb J, Lewandowski JP, Bryant JD, Fitch R, Surman DR, Vokes SA, Appling DR. Deletion of Mthfd1 l causes embryonic lethality and neural tube and craniofacial defects in mice. Proc Natl Acad Sci USA. 2013; 110:549–564. [PubMed: 23267094]

Mossey PA, Little J, Munger RG, Dixon MJ, Shaw WC. Cleft lip and palate. Lancet. 2009; 374:1773–1785. [PubMed: 19747722]

Mostowska A, Hozyasz KK, Jagodzinski PP. Maternal MTR genotype contributes to the risk of non-syndromic cleft lip and palate in the Polish population. Clin Genet. 2006; 69:512–517. [PubMed: 16712703]

Mostowska A, Hozyasz KK, Biedziak B, Misiak J, Jagodzinski PP. Polymorphisms located in the region containing BHMT and BHMT2 genes as maternal protection factors for orofacial clefts. Eur J Oral Sci. 2010; 118:325–332. [PubMed: 20662904]

Mostowska A, Hozyasz KK, Wojcicki P, Dziegelewska M, Jagodzinski PP. Associations of folate and choline metabolism gene polymorphisms with orofacial clefts. J Med Genet. 2010b; 47:809–815. [PubMed: 19737740]

Munger RG, Tamura T, Johnston KE, Feldkamp ML, Pfister R, Cutler R, Murtaugh MA, Carey JC. Oral clefts and maternal biomarkers of folate-dependent one-carbon metabolism in Utah. Birth Defects Res A Clin Mol Teratol. 2011; 91:153–161. [PubMed: 21290562]

Obican SG, Finnell RH, Mills JL, Shaw GM, Scialli AR. Folic acid in early pregnancy: A public health success story. FASEB J. 2010; 11:4167–4174.

Palmieri A, Masiero E, Martinelli M, Scapoli L, Pezzetti F, Caramelli E, Guidotti L, Carinci F. The MTHFD1 gene is not involved in cleft lip with or without palate onset among the Italian population. Ann Hum Genet. 2008; 72:297–299. [PubMed: 18261183]

Rubini M, Brusati R, Garattini G, Magnani C, Liviero F, Bianchi F, Tarantino E, Massei A, Pollastri S, Carturan S, Amadori A, Bertagnin E, Cavallaro A, Fabiano A, Franchella A, Calzolari E. Cystathionine beta-synthase c.844ins68 gene variant and non-syndromic cleft lip and palate. Am J Med Genet Part A. 2005; 136A:368–372. [PubMed: 16007597]

Shaw GM, Lammer EJ, Wasserman CR, O'Malley CD, Tolarova MM. Risks of orofacial clefts in children born to women using multivitamins containing folic acid periconceptionally. Lancet. 1995; 346:393–396. [PubMed: 7623568]

Shaw GM, Carmichael SL, Laurent C, Rasmussen SA. Maternal nutrient intakes and risk of orofacial clefts. Epidemiology. 2006; 17:285–291. [PubMed: 16570024]

Shaw GM, Vollset SE, Carmichael SL, Yang W, Finnell RH, Blom H, Ueland PM. Nested case-control study of one-carbon metabolites in mid-pregnancy and risks of cleft lip with and without cleft palate. Pediatr Res. 2009; 66:501–506. [PubMed: 19668105]

Tang LS, Santillano DR, Wlodarczyk BJ, Miranda RC, Finnell RH. Role of Folbp1 in the regional regulation of apoptosis and cell proliferation in the developing neural tube and craniofacies. Am J Med Genet Part C Semin Med Genet. 2005; 135C:48–58. [PubMed: 15800851]

Teng YW, Mehedint MG, Garrow TA, Zeisel SH. Deletion of betaine-homocysteine S-methyltransferase in mice perturbs choline and 1-carbon metabolism, resulting in fatty liver and hepatocellular carcinomas. J Biol Chem. 2011; 286:36258–36267. [PubMed: 21878621]

Wahl SE, Kennedy AE, Wyatt BH, Moore AD, Pridgen DE, Cherry AM, Mavila CB, Dickinson AJ. The role of folate metabolism in orofacial development and clefting. Dev Biol. 2015; 405:108–122. [PubMed: 26144049]

Wallenstein MB, Shaw GM, Yang W, Carmichael SL. Periconceptional nutrient intakes and risks of orofacial clefts in California. Pediatr Res. 2013; 74:457–465. [PubMed: 23823175]

WHO. Global registry and database on craniofacial anomalies. In: Mossey, P., Catilla, E., editors. Report of a WHO Registry Meeting on Craniofacial Anomalies. 2003.

Wilcox AJ, Lie RT, Solvoll K, Taylor J, McConnaughey DR, Abyholm F, Vindenes H, Vollset SE, Drevon CA. Folic acid supplements and risk of facial clefts: National population based case-control study. BMJ. 2007; 334:464. [PubMed: 17259187]

Wong WY, Eskes TK, Kuijpers-Jagtman AM, Spauwen PH, Steegers EA, Thomas CM, Hamel BC, Blom HJ, Steegers-Theunissen RP. Nonsyndromic orofacial clefts: Association with maternal hyperhomo-cysteinemia. Teratology. 1999; 60:253–257. [PubMed: 10525201]

Zhu H, Curry S, Wen S, Wicker NJ, Shaw GM, Lammer EJ, Yang W, Jafarov T, Finnell RH. Are the betaine-homocysteine methyltransferase (BHMT and BHMT2) genes risk factors for spina bifida and orofacial clefts? Am J Med Genet Part A. 2005; 135A:274–277.

TABLE I

Population Demographics

Variables	Cases (collective percentage)	Controls (collective percentage)
Race/ethnicity		
Asian	37 (11.2)	32 (8.7)
Black	13 (3.9)	15 (4.1)
White, Hispanic	193 (58.5)	226 (61.6)
White, Non-Hispanic	79 (23.9)	82 (22.3)
Other	8 (2.4)	10 (2.7)
No race reported	0 (0)	2 (0.5)
Total	**330**	**367**
Infant sex		
Female	120 (36.4)	182 (49.6)
Male	210 (63.6)	185 (50.4)
Total	**330**	**367**
Combined maternal folate intake *		
Low	28 (15.5)	23 (11.2)
Not-low	152 (84.0)	177 (86.3)
Missing	1 (0.6)	5 (2.4)
Total	**181**	**205**

* Maternal folate intake during the periconceptional period from supplements and diet was ascertained for a subset of the sample population from interview data. Folate intake categories are defined in Methods.

TABLE II

Variants With Significant Associations in Population Sub-Groups

SNP ID	Sub-group	Case allele counts*	Control allele counts*	Heterozygote odds ratio (95%CI)**	Homozygote odds ratio (95%CI)**
ALDH1L1_74483	All	295/34/1	344/21/2	**1.9 (1.1–3.3)**	n/d
ATIC_22288		295/34/0	307/60/0	**0.6 (0.4–0.9)**	n/d
ATIC_31585		290/40/0	297/70/0	**0.6 (0.4–0.9)**	n/d
BHMT_28849		122/159/49	114/170/82	0.9 (0.6–1.2)	**0.6 (0.4–0.9)**
BHMT2_10283		121/158/51	114/172/79	0.9 (0.6–1.2)	**0.6 (0.4–0.9)**
FOLR2_3389		170/117/43	164/164/38	**0.7 (0.5–0.9)**	1.1 (0.7–1.8)
GART_22073		197/121/12	261/91/9	**1.8 (1.3–2.4)**	1.8 (0.7–4.3)
GART_26477		87/174/69	121/191/54	1.3 (0.9–1.8)	**1.8 (1.1–2.8)**
MAT2A_4688		176/140/14	224/120/23	**1.5 (1.1–2.0)**	0.8 (0.4–1.5)
MTHFD1_43530		298/31/1	305/58/4	**0.5 (0.3–0.9)**	n/d
MTHFD1_59019		287/40/3	285/74/8	**0.5 (0.4–0.8)**	0.4 (0.1–1.4)
MTHFR_12220		257/69/4	307/50/4	**1.6 (1.1–2.5)**	1.2 (0.3–4.8)
AHCYL2_165230	White, non–Hispanic	11/47/21	25/41/16	**2.6 (1.1–5.9)**	**3.0 (1.1–7.8)**
AHCYL2_165528		35/40/4	31/36/15	1.0 (0.5–1.9)	**0.2 (0.1–0.8)**
AHCYL2_173166		35/39/4	32/35/14	1.0 (0.5–2.0)	**0.3 (0.1–0.9)**
AHCYL2_203190		36/39/4	32/35/15	1.0 (0.5–1.9)	**0.2 (0.1–0.8)**
CBS_13704		57/21/1	70/10/0	**2.6 (1.1–5.9)**	n/d
CBS_13779		57/21/1	71/11/0	**2.4 (1.1–5.3)**	n/d
DHFR_17874		76/3/0	69/12/0	**0.2 (0.1–0.8)**	n/d
AHCY_14358	Hispanic	165/26/2	174/50/1	**0.5 (0.3–0.9)**	n/d
ALDH1L1_74427		157/33/3	158/59/8	**0.6 (0.3–0.9)**	0.4 (0.1–1.4)
ALDH1L1_74483		171/21/1	218/7/1	**3.8 (1.6–9.2)**	n/d
ALDH1L1_77540		163/26/4	160/56/9	**0.5 (0.3–0.8)**	0.4 (0.1–1.4)
ALDH1L2_38374		63/100/30	97/96/33	**1.6 (1.1–2.4)**	1.4 (0.8–2.5)
ATIC_22288		175/18/0	187/39/0	**0.5 (0.3–0.9)**	n/d
ATIC_31585		172/21/0	182/44/0	**0.5 (0.3–0.9)**	n/d
BHMT2_8849		82/86/25	68/107/50	0.7 (0.4–1.0)	**0.4 (0.2–0.7)**
BHMT2_10283		80/87/26	67/111/47	0.7 (0.4–1.0)	**0.5 (0.3–0.8)**

Am J Med Genet A. Author manuscript; available in PMC 2017 October 10.

SNP ID	Sub-group	Case allele counts*	Control allele counts*	Heterozygote odds ratio (95%CI)**	Homozygote odds ratio (95%CI)**
BHMT2_13828		82/85/26	66/112/46	**0.6 (0.4–0.9)**	**0.5 (0.3–0.8)**
CTH_2898		141/49/3	188/37/1	**1.8 (1.1–2.9)**	n/d
DMGDH_26164		77/88/28	64/111/51	0.7 (0.4–1.0)	**0.5 (0.3–0.8)**
DMGDH_26388		77/88/28	64/110/51	0.7 (0.4–1.0)	**0.5 (0.3–0.8)**
FOLR2_3389		120/54/19	117/96/12	**0.5 (0.4–0.8)**	1.5 (0.7–3.3)
GART_22073		113/76/4	160/58/5	**1.9 (1.2–2.8)**	1.1 (0.3–4.3)
MAT1A_11656		49/106/38	77/97/52	**1.7 (1.1–2.7)**	1.1 (0.7–2.0)
MAT2A_4688		106/79/8	147/68/11	**1.6 (1.1–2.4)**	1.0 (0.4–2.6)
MTHFD1_3004		55/91/47	71/123/31	1.0 (0.6–1.5)	**2.0 (1.1–3.5)**

*Number of individuals with genotypes: major homozygote/heterozygote/minor homozygote.

**Significant odds ratios (*P*-value <0.05) are boldface; not determined (n/d) for <3 occurrences in cases or controls.

Am J Med Genet A. Author manuscript; available in PMC 2017 October 10.

TABLE III

Low-Frequency Variants With Strongest Case-Control Differences in Population Sub-Groups

SNP ID	Sub-group	Case allele counts*	Control allele counts*	Heterozygote odds ratio (95%CI)**	Homozygote odds ratio (95%CI)**	Case-control difference
MAT1A_5181	All	329/1/0	358/9/0	n/d	n/d	2.5
MTHFD2_1258		311/18/0	325/32/2	0.6 (0.3–1.1)	n/d	2.2
MTHFD1_54248		306/4/0	341/0/0	n/d	n/d	2.1
FPGS_2386		319/10/0	341/19/2	0.6 (0.3–1.2)	n/d	2.1
AMT_4255		327/3/0	355/11/0	0.3 (0.1–1.1)	n/d	2.0
MTHFD1_12347		318/12/0	344/21/2	0.6 (0.3–1.3)	n/d	1.9
SHMT1_35845		317/13/0	360/6/0	2.5 (0.9–6.5)	n/d	1.9
ALDH1L1_74483		295/34/1	344/21/2	**1.9 (1.1–3.3)**	n/d	1.9
BHMT_17018		295/33/2	341/26/0	1.5 (0.9–2.5)	n/d	1.8
DHFR_17874		317/13/0	333/25/0	0.5 (0.3–1.1)	n/d	1.7
ATIC_8754		330/0/0	364/3/0	n/d	n/d	1.6
MTR_96924		330/0/0	364/3/0	n/d	n/d	1.6
MTHFD1_32373		330/0/0	363/3/0	n/d	n/d	1.6
DHFR_26717		321/9/0	362/4/0	2.5 (0.8–8.3)	n/d	1.6
MTRR_23290	White, non–Hispanic	77/2/0	70/12/0	n/d	n/d	2.6
MTRR_21088		77/2/0	71/11/0	n/d	n/d	2.5
DHFR_17874		76/3/0	69/12/0	**0.2 (0.1–0.8)**	n/d	2.3
MAT1A_5181		79/0/0	77/5/0	n/d	n/d	2.2
MTRR_10208		77/2/0	73/9/0	n/d	n/d	2.1
SHMT1_35845		75/4/0	82/0/0	n/d	n/d	2.0
AHCY_8816		77/2/0	74/8/0	n/d	n/d	1.9
AHCYL2_1458		76/3/0	82/0/0	n/d	n/d	1.8
ALDH1L1_74483	Hispanic	171/21/1	218/7/1	**3.8 (1.6–9.2)**	n/d	3.0
CTH_12147		186/7/0	207/19/0	0.4 (0.2–1.0)	n/d	2.1
CBS_4605		190/3/0	225/0/0	n/d	n/d	1.9
CBS_14399		193/0/0	221/4/0	n/d	n/d	1.9
MTHFD1_43530		179/14/0	198/28/0	0.6 (0.3–1.1)	n/d	1.7
MTHFR_16366		187/6/0	169/1/0	n/d	n/d	1.7

SNP ID	Sub-group	Case allele counts*	Control allele counts*	Heterozygote odds ratio (95%CI)**	Homozygote odds ratio (95%CI)**	Case-control difference
SHMT2_4031		180/13/0	213/7/0	2.2 (0.9–5.6)	n/d	1.7
MAT1A_1103		187/6/0	224/2/0	n/d	n/d	1.7
CTH_5716		192/0/0	218/3/0	n/d	n/d	1.6
DMGDH_26227		193/0/0	223/3/0	n/d	n/d	1.6
ALDH1L1_44924		192/1/0	220/5/0	n/d	n/d	1.5
SHMT2_4974		189/4/0	224/1/0	n/d	n/d	1.5

*Allele counts are individuals with genotypes: major homozygote/heterozygote/minor homozygote.

**Significant odds ratios (*P*-value <0.05) are boldface; not determined (n/d) for <3 occurrences in cases or controls.

TABLE IV

Gene-Folate Interactions in Population Sub-Groups

		Not-low folate intake*			Low folate intake*		
SNP ID	**Sub-group**	**Case/control**	**Odds ratio (95%CI)**	***P*-value**	**Case/control**	**Odds ratio (95%CI)**	***P*-value**
ALDH1L1_28193	All	152/176	1.2 (0.5–2.5)	0.69	28/23	0.2 (0.03–0.9)	0.04
CBS_10681		152/177	0.8 (0.5–1.3)	0.31	28/23	0.1 (0.01–0.6)	0.01
CBS_15061		152/176	0.8 (0.5–1.3)	0.36	28/23	0.1 (0.01–0.6)	0.01
CBS_9548		152/177	1.0 (0.7–1.4)	0.97	28/23	3.6 (1.3–9.6)	0.01
TCN2_16657		152/177	1.1 (0.8–1.6)	0.48	28/23	0.4 (0.2–0.9)	0.02
TCN2_16224		152/176	1.0 (0.7–1.5)	0.99	28/23	0.2 (0.1–0.7)	0.01
TCN2_8396		152/176	1.1 (0.8–1.6)	0.51	28/23	0.4 (0.2–0.9)	0.02
TCN2_9392		152/176	1.0 (0.7–1.4)	0.99	28/23	2.8 (1.2–6.3)	0.02
TCN2_9450		152/177	1.0 (0.7–1.3)	0.85	28/23	2.8 (1.2–6.3)	0.02
CTH_4263	Hispanic	92/112	0.8 (0.5–1.3)	0.45	16/11	11.5 (1.8–72.9)	0.01
MTRR_1019		92/113	0.9 (0.6–1.4)	0.66	16/11	0.1 (0.03–0.7)	0.02
MTRR_17310		92/113	0.9 (0.6–1.3)	0.51	16/11	0.1 (0.03–0.7)	0.02
MTRR_17743		92/113	0.8 (0.5–1.2)	0.30	16/11	0.1 (0.03–0.7)	0.02
MTRR_28975		92/113	0.9 (0.6–1.3)	0.50	16/11	0.1 (0.03–0.7)	0.02
SLC46A1_8485		92/112	1.1 (0.7–1.6)	0.73	16/11	0.3 (0.1–1.0)	0.05

*Folate intake was defined on the basis of combining maternal periconceptional intake of folic acid-containing vitamin supplements with daily dietary intake of folate as described in Methods.

TABLE V

Previous Reports of Significant CLP Associations Within the BHMT/BHMT2/DMGDH Locus

				Variants significantly associated*		
Reference	**Study**	**Population**	**Size**	**Gene**	**SNP rs#**	**Sub-group**
Zhu et al. [2005]	Case-control	California	563 CLP + CPO cases; 201 controls	NONE		
Boyles et al. [2008]	Family based	Norway	362 CLP families; 191 CPO families	BHMT	rs3733890	Maternal
Boyles et al. [2009]	Family based	Norway	362 CLP families; 191 CPO families	BHMT	rs3733890	Maternal—low folate
				BHMT2	rs10944	Maternal—low folate
				DMGDH	rs1805074	Maternal—low folate
				DMGDH	rs250513	Fetal—low folate
				DMGDH	rs479405	Maternal—low folate
				DMGDH	rs532964	Maternal—low folate
Blanton et al. [2011]	Family based	Massachusetts/Texas	120 multiplex; CLP families; 325 simplex trios	BHMT	rs3733890	Hispanic, fetal
				BHMT	rs645112	Hispanic, fetal
				BHMT2	rs1422086	Hispanic, fetal
				BHMT2	rs2253262	Hispanic, White, Fetal
				BHMT2	rs476620	Hispanic, fetal
				BHMT2	rs626105	Hispanic, fetal
				BHMT2	rs682985	Hispanic, White, fetal
Mostowska et al. [2010]	Case-control mothers	Poland	180 CLP mothers; 162 controls	BHMT	rs600473	Maternal
				BHMT	rs7356530	Maternal
				BHMT2	rs526264	Maternal
				BHMT2	rs625879	Maternal
Mostowska et al. [2010b]	Case-control	Poland	174 CLP cases; 176 controls	BHMT	rs3733890	Fetal
				BHMT	rs585800	Fetal
Martinelli et al. [2011]	Family based	Italy	238 CLP triads	NONE		
Jin et al. [2015]	Case-control	China	481 CLP cases; 558 controls	BHMT	rs3733890	Fetal
This study	Case-control	California	330 CLP cases; 367 controls	BHMT2	rsl422086	Hispanic, fetal
				BHMT2	rs526264	Hispanic, fetal
				BHMT2	rs682985	Hispanic, fetal
				DMGDH	rs532964	Hispanic, fetal

Reference	Study	Population	Size	Variants significantly associated*		
				Gene	SNP rs#	Sub-group
				DMGDH	rs248384	Hispanic, fetal

*SNPs listed are described in the original publications as significantly associated with CLP according to the analytical methods employed in that particular study; “NONE” means that SNPs in this region were tested, but none were observed to be significantly associated with CLP. In studies where cleft palate only [CPO] samples were also queried, the results shown are for CLP associations only.