


RESEARCH

Open Access



A comparison of DNA methylation in newborn blood samples from infants with and without orofacial clefts

Zongli Xu¹ , Rolv T. Lie^{3,4}, Allen J. Wilcox¹, Ola Didrik Saugstad^{5,6} and Jack A. Taylor^{1,2*}

Abstract

Background: Isolated orofacial clefts are among the most common congenital birth defects. Although the underlying biological mechanisms remain largely unknown, clefts are thought to be complex disorders influenced by genetic, environmental, and potentially epigenetic factors.

Methods: In blood samples from 2- to 3-day-old infants ($n = 747$) collected in a nationwide population-based study of orofacial clefts in Norway, we measured DNA methylation profiles for more than 450,000 CpGs and then conducted epigenome-wide association analyses (EWAS). We tested methylation profile difference at each CpG between controls ($n = 436$) and each of the cleft subtypes (92 cleft lip only, CLO; 84 cleft palate only, CPO; 132 cleft lip and palate, CLP). We also compared controls to various combinations of case groups and compared case subtypes to each other. Finally, using the EWAS results, we searched for larger differentially methylated regions (DMRs) associated with orofacial clefts.

Results: In EWAS comparing controls to individual cleft subtypes, we found no significant associations at a Bonferroni P value threshold of 10^{-7} . After pooling case groups, we found two significantly differentially methylated CpGs: cg09696939 near gene BICC1 is associated with CLO+CLP ($P = 9.58 \times 10^{-8}$); cg26985354 in gene CLASRP is associated with CPO+CLP ($P = 7.38 \times 10^{-8}$). In DMR analysis, we identified a total of 56 significant regions when comparing controls to individual cleft subtypes (10 for CLO, 6 for CPO, 41 for CLP). Only one DMR is shared among the three cleft groups. In combined case group analysis, we found 26 DMRs for CLP+CLO, 31 for CLP+CPO, and 37 when all subtypes are combined. Finally, in case-case comparisons of subtypes, we identified 10 DMRs when comparing CLP to CPO, 9 in CLP compared to CLO, and 13 in CLP compared to CPO.

Conclusions: We identified two individual CpGs and multiple DMRs that differ between controls and cleft case subtypes. Although we find some evidence for the possible role of DNA methylation in etiology of orofacial clefts, our study does not support previous reports of widespread differences in blood DNA methylation between babies with and without facial clefts.

Keywords: DNA methylation, Cleft lip, Cleft palate, Blood, Newborns

* Correspondence: taylor@niehs.nih.gov

¹Epidemiology Branch, National Institute of Environmental Health Sciences, NIH, Research Triangle Park, NC, USA

²Laboratory of Molecular Carcinogenesis, National Institute of Environmental Health Sciences, NIH, P.O. Box 12233, Mail Drop A3-05, Research Triangle Park, NC 27709, USA

Full list of author information is available at the end of the article



© The Author(s). 2019 **Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated.

Background

Orofacial clefts are among the most common congenital birth defects. Approximately 250,000 babies are born with these defects each year worldwide (<http://cleft.org.uk/>), and the prevalence varies by ethnicity, sex [1], and socioeconomic status [2]. Orofacial clefts, which include three distinct subgroups, cleft lip only (CLO), cleft lip and palate (CLP), and cleft palate only (CPO), result from fusion failures in weeks 5 (fusion of lip) and 9 (fusion of palate) from date of conception. An orofacial cleft that is left unrepaired can have a substantial adverse impact on the health, quality of life, and psychosocial well-being of the child. Full repair of clefts entails surgery, dental treatment, and speech therapy, imposing financial and emotional burden on families [3].

Approximately 70% of orofacial clefts are non-syndromic [1], and the underlying causes of these isolated cases are complex and remain largely unknown. Although familial recurrence risk estimates and segregation analysis [4–6] suggest there is strong evidence for a genetic component, environmental factors like smoking, vitamin intake, and alcohol intake may also play an important role [7–9]. To date, whole genome linkage and SNPs identified through genome-wide association studies can explain only a small percentage of the disease variation [1, 10]. Several recent studies [8, 11–13] support a multifactorial threshold model of inheritance in which small genetic risk factors may interact with environmental factors.

Epigenetic changes, in particular DNA methylation, can mediate both genetic and environmental factors and may advance the understanding of the molecular pathogenesis of orofacial clefts [14]. Several recent large studies [15–18] show that maternal exposure to putative cleft risk factors, such as tobacco smoking, obesity, and folate deficiency, can affect methylation profiles of blood DNA in newborns and infants. Alvizi et al. [19] performed an epigenome-wide association study of blood DNA from 67 cases and 59 controls using Illumina 450K array and reported 578 differentially methylation CpGs. In a Mendelian randomization analysis in the Avon Longitudinal Study of Parents and Children (ALSPAC) [20], Howe et al. [21] reported 21 putative CpG sites where methylation might mediate genetic allelic effects on liability to non-syndromic cleft lip/palate. Sharp et al. [22] compared DNA methylation profiles between cleft subtypes (50 samples of each subtype) and reported a total of 121 differentially methylated regions (DMRs) in pairwise comparisons using blood DNA and 15 DMRs using lip/palate DNA. Here, we report a larger epigenome-wide blood DNA methylation study from a national population-based case-control study of orofacial clefts in Norway.

Methods

Study samples

This study used blood samples from infants in our nationwide population-based Norway Facial Cleft Study (NCL), which has been previously described in detail [9]. Four hundred eighteen infants born with an orofacial cleft during the years 1996 to 2001 were included: 107 CLO, 167 CLP, and 144 CPO. A total of 480 control infants born without clefts were randomly selected from all live births in Norway during that time period. Heel-stick blood samples had been collected from all babies 2–3 days after delivery for phenylketonuria (PKU) testing. These samples were stored and later used for DNA methylation analysis. Demographic characteristics, mother's environmental exposures, and nutrition intake were collected through questionnaires approximately 3 months after delivery.

Ethical approval for this study was obtained from the Norwegian Data Inspectorate, the Regional Research Ethics Committee for Western Norway, and the Institutional Review Board of the US National Institute of Environmental Health Sciences, National Institutes of Health. Parents gave consent to data and sample collection, and all data were anonymized after collection.

DNA methylation measures

Genomic DNA was extracted from whole blood using automated equipment (Autopure LS; Gentra, Minneapolis, MN). Extracted DNA was quantified using Quant-iT™ PicoGreen dsDNA reagents (Invitrogen) and stored at -20°C . One microgram of DNA was bisulfite converted using the EZ-DNA Methylation kit (Zymo Research, Irvine, CA) following the manufacturer's protocol. Samples were randomly distributed across ten 96-well plates used for laboratory processing. The Illumina HumanMethylation450 BeadChip (San Diego, CA) was used to assess DNA methylation at 485,577 CpG sites with 99% coverage of RefSeq genes and an average of 17 CpG sites per gene distributed across the promoter, 5' untranslated region (UTR), first exon, gene body, and 3' UTR. At each CpG site on the array, the methylation level was estimated using intensity measures of probes corresponding to an unmethylated (*U*) or methylated (*M*) signal. Raw probe intensity values were extracted using R package `illuminaio` (version 0.18.0) [23].

Methylation data pre-processing

Methylation raw IDAT files were preprocessed using ENmix R package [24]. Briefly, we performed background correction using the ENmix method [24]; fluorescent dye bias between Cy3 and Cy5 was adjusted with the RELIC method [25]; inter-array quantile normalization was applied on methylation intensity values to make them comparable between samples; and Infinium I and II probe-type design

bias was adjusted using the RCP method [26]. The methylation level (β value) for each CpG site was calculated as the ratio of normalized fluorescent intensities between methylated and unmethylated alleles $\beta = M/(M + U + 100)$.

We excluded 29 samples with low-quality methylation data according to the following criteria: (1) average intensity value across Illumina's bisulfite internal control probes less than 5500, (2) more than 5% of CpG probes having low-quality data (Illumina detection P value $> 10^{-6}$, read from less than 3 beads, or outlier value for the probe in the dataset), and (3) clear outliers based on visual inspection of a density plot of total intensity ($U + M$). We excluded one sample due to sex ambiguity. We also removed poorly performing CpG probes as follows: (1) CpGs ($n = 23,264$) with more than 5% low-quality data; (2) CpGs ($n = 53,247$) with common SNPs (minor allele frequency ≥ 0.05 in Europeans based on 1000 Genomes Project data) located at the probe's target CpG site, or probes mapping to multiple genomic locations [27], or CpGs on X or Y chromosomes; (3) we further excluded 1488 CpGs with multiple mode distributions identified with ENmix R package. After sample and CpG probe exclusions, 407,513 CpGs and 868 samples (456 controls, 105 CLO, 167 CLP, and 140 CPO) remained for analysis.

Statistical analysis

An unconditional logistic regression model was employed to perform epigenome-wide association analysis (EWAS). We tested for an association between DNA methylation M values (logit transformation of β values) at each CpG and a binary variable for each type of orofacial cleft (shared controls vs. each cleft type: CLO, CPO, CLP) or between cleft types (CLO vs CPO, CLP vs CLO, and CLP vs CPO). In all association tests, we adjusted for covariables including five surrogate variables (that explain 96% of variations in non-negative control-probe data), plate, gender, calendar year of baby's birth, gestational age, and blood cell subtype composition. The proportions of five leukocyte subtypes (T cell, B cell, granulocyte, monocyte, and natural killer cell) were estimated using a method proposed by Houseman et al. [28] with reference dataset GSE35069 [29]. To correct for multiple testing, we employed the Bonferroni method and use P value threshold of 10^{-7} as the statistical significance cutoff in all genome-wide analyses. All statistical association analyses were performed using the R software package (version 3.4). In addition to the EWAS for individual CpG, we also used software DMRcate [30] and Comb-P [31] to detect differentially methylated regions (DMRs) based on individual CpG P values from the EWAS. A DMR was defined as regions containing at least two CpGs less than 1000 base pairs apart with Sidak multiple testing-corrected P value < 0.05 .

Results

Sample characteristics

To focus our investigation on non-syndromic facial clefts, we excluded babies with other birth defects, resulting in 308 cases (92 CLO, 84 CPO, 132 CLP) and 436 controls in the analysis. Demographic characteristics of these infants and their mothers, as well as maternal exposures during pregnancy, are summarized in Table 1. Compared to controls, CLP cases are more likely to be male, and CPO cases are more likely to be female. These are consistent with the literature [32]. On average, infants with CLO or CPO weigh slightly less at birth (ANOVA test $P = 1.7 \times 10^{-2}$). Active maternal smoking during the first trimester (≥ 1 cigarette per day) is positively associated with CLO (OR = 2.5, CI = (1.6–4.0)) and CLP (OR = 1.8, CI = (1.2–2.7)), but not with CPO (OR = 1.4, CI = (0.8, 2.3)), consistent with an earlier analysis of the full dataset [7]. Mothers with less education have an elevated risk of babies with CLO (OR = 2.4, CI = (1.3–4.3)) and CLP (OR = 1.9, CI = (1.1–3.3)). Lower dietary folate during pregnancy is a risk factor for CPO (OR = 0.58, CI = (0.35–0.96)) [9]. Additional maternal characteristics were examined, such as age at delivery, body mass index (BMI), and multivitamin use, but none were associated with facial clefts.

Orofacial cleft-associated CpG sites

We used unconditional logistic regression models to test whether the DNA methylation profile in blood of 2–3-day-old infants was associated with cleft phenotype at each CpG site on the array. Considering potential etiologic differences between cleft types, we first examined the association for CLO, CPO, and CLP separately, using as a reference the same group of unaffected control babies. After Bonferroni correction for multiple testing, there were no statistically significant associations for any of the analyses (Additional file 1: Figure S1), perhaps reflecting the small sample size in each case group. Consistent with other epidemiological studies, we then combined CLO and CLP case group as the “any cleft lip” group and repeated the association analysis (Additional file 1: Figure S2). This analysis identified one significant probe (cg09696939, $P = 9.58 \times 10^{-8}$) at gene BICC1 (Table 2). We also combined CPO with the CLP group as the “any cleft palate” group (Additional file 1: Figure S2) and identified another significant probe (cg26985354, $P = 7.38 \times 10^{-8}$) in CLASRP gene (Table 2). Additional file 2: Table S1 lists 43 CpG probes with P value less than 10^{-5} in any of the comparisons. Similar to Alvizi et al. [19], we compared controls with “any cleft” group (combining all three case groups together), but did not find any significant dmCpG at genome-wide significance level of 10^{-7} . Similar to Sharp et al. [22], we also compared methylation profiles between cleft types (CLO vs CLP, CPO vs CLP, CLO vs CPO), but again did not identify any significant CpGs with P value $< 10^{-7}$.

Table 1 Characteristics of infants and mothers

	Control	CLO	CPO	CLP
Infants	436	92	84	132
Sex, <i>n</i> (%)				
Female	198 (45)	41 (45)	47 (56)	37 (28)
Male	238 (55)	51 (55)	37 (44)	95 (72)
Birth weight, kg	3.6 ± 0.6	3.4 ± 0.7	3.5 ± 0.6	3.6 ± 0.6
Gestational age, days	277.7 ± 11.6	274.8 ± 14.6	276.4 ± 13.8	278.6 ± 11.7
Mothers				
Age at delivery, years	29.1 ± 4.8	28.9 ± 4.6	29.2 ± 5.0	29.4 ± 5.4
Active smoker ^a , <i>n</i> (%)				
No	316 (72)	47 (51)	55 (65)	78 (59)
Yes	120 (28)	45 (49)	29 (35)	54 (41)
Alcohol use ^b , <i>n</i> (%)				
0	292 (66)	52 (56)	56 (66)	80 (61)
1–3	73 (17)	19 (21)	14 (17)	28 (21)
4–6	26 (6)	6 (7)	8 (10)	8 (6)
≥ 7	42 (10)	12 (13)	6 (7)	14 (11)
Missing	3 (1)	3 (3)	0 (0)	2 (1)
BMI (kg/m ²)	23.5 ± 3.9	23.5 ± 3.7	23.7 ± 4.8	23.3 ± 4.1
Folic acid supplement, µg ^{c,d} , <i>n</i> (%)				
0	261 (60)	59 (65)	47 (56)	94 (71)
1–399	90 (21)	17 (18)	19 (23)	24 (18)
≥ 400	85 (19)	16 (17)	18 (21)	14 (11)
Dietary folate, µg, <i>n</i> (%)				
0–171	110 (25)	28 (31)	31 (37)	41 (31)
172–214	96 (22)	21 (23)	21 (25)	36 (27)
215–264	112 (26)	16 (17)	9 (11)	21 (16)
≥ 265	96 (22)	16 (17)	20 (24)	25 (19)
Missing	22 (5)	11 (12)	3 (3)	9 (7)
Multivitamins ^d , <i>n</i> (%)				
No	277 (64)	65 (71)	50 (60)	96 (73)
Yes	159 (36)	27 (29)	34 (40)	36 (27)
Education, <i>n</i> (%)				
Less than high school	45 (10)	20 (22)	9 (11)	24 (18)
High school and above	391 (90)	72 (78)	75 (89)	108 (82)

^aDuring the first trimester

^bTotal number of drinks during the first trimester

^cAny intake of folic acid supplements (either alone or with multivitamins)

^dDuring month prior to pregnancy and first 2 months of pregnancy

Table 2 CpGs associated with non-syndromic clefts

CpG	Chr	Position	Gene	Methylation	Coef.	<i>P</i>	Phenotype
cg09696939	10	60,272,079	BICC1	0.019	2.304	9.6 × 10 ⁻⁸	Any cleft lip
cg26985354	19	45,567,180	CLASRP	0.889	-1.45	7.4 × 10 ⁻⁸	Any cleft palate

between epigenetic age and chronological age) could be used to indicate development rate in newborns. The blood samples in our study were collected from all babies 2–3 days after delivery, and thus, the chronological age is essentially 0 year old. We found that methylation age estimated using the Horvath method [34] was in general younger (~0.4 years) than the chronological age for our study participants. The means and standard deviations of epigenetic ages are -0.39 ± 0.10 , -0.40 ± 0.11 , -0.41 ± 0.09 , and -0.40 ± 0.10 , respectively, for controls, CLO, CPO, and CLP. Epigenetic age estimates for any of the cleft groups are slightly younger than those of the controls, but none of the pairwise comparisons are statistically significant (Tukey honest significant differences test, $P > 0.1$).

Discussion

We conducted an epigenome-wide association study investigating 2–3-day-old infant blood DNA methylation and isolated orofacial clefts. We did not find individual dmCpG when we analyzed each cleft type separately. When we combined the cleft palate only with the cleft lip and palate, we identified one significantly methylated probe cg09696939 near gene BICC1. Similarly, when we combined the cleft lip only with the cleft lip and palate, we identified another significant dmCpG cg26985354 in gene CLASRP. Based on EWAS P values for individual CpGs and their chromosome location, we identified multiple differentially methylated regions (DMRs) between controls and case subtypes, and in case-case analysis of different subtypes. We confirmed one of the DMRs (17:80,541,737–80,542,007 in gene FOXK2) previously reported in case-case analysis by Sharp et al. [22].

CpG cg09696939 is located on chromosome 10, 824 base pairs upstream of gene BICC1. It is located in a CpG shore region and hypomethylated (averaged methylation level is 0.02). BICC1 encodes an RNA-binding protein, acts as a negative regulator of Wnt signaling, and is active in regulating gene expression by modulating protein translation during embryonic development. The gene is reported to be associated with cystic and renal dysplasia [35], myopia [36], and Alzheimer disease [37].

CpG cg26985354 is located on chromosome 19, in an intron of the gene CLASRP (CLK4-associating serine/arginine-rich protein). It is in a CpG island region and is hypermethylated (averaged methylation level is 0.89). CLASRP is a protein-coding gene, which functions as an alternative splicing regulator and may regulate members of the CLK kinase family [38].

Although we did not find any differentially methylated genomic regions using the DMRcate method [30], we found multiple DMRs using the Comp-P method [30], perhaps reflecting increased sensitivity of Comp-P when effect sizes are small. Compared to controls, CLP had

the most DMRs (41 DMRs), followed by CLO (10 DMRs) and CPO (6 DMRs); only one DMR, in an intra-genic region on chromosome 6, was shared among the three case subtypes. The number of DMRs found in each subtype could reflect sample size and resulting power ($n = 132$ CLP samples, 92 CLO samples, and 84 CPO samples). Subtype-specific DMRs might arise because of etiologic heterogeneity [22], but we cannot dismiss the possibility that they arise by chance.

In a study of a general population cohort without clefts, Howe et al. [21] found that cleft-associated polymorphisms predicted differential methylation of 21 CpGs. Although we found weak statistical support for three of these CpGs (threshold $P < 0.05$), the directions of effect for two were opposite from Howe's prediction [21] and the third had an opposite direction of effect in an independent EWAS analysis of clefts [19, 21]. A study of blood DNA from 5-year-old children by Alvizi et al. [19] compared 67 combined cleft cases to 59 controls and reported 578 dmCpGs. But in our study, only 3 of these had the same direction of effect and met a threshold of unadjusted $P < 0.05$ —fewer than expected by chance. A case-only study by Sharp et al. [22] of both blood and cleft tissue (with 50 samples of each cleft subtype) reported 335 dmCpGs when comparing CLO to CPO. However, we did not find any significant CpGs when performing similar case-case analysis in our NCL dataset. Sharp et al. also reported 121 DMRs, and we confirm one of the DMRs they report in the gene FOXK2. But while they found this DMR in a comparison of CLO vs CPO, we found it in CLO vs CPO, CLP vs CPO, and control vs CPO. Similar to Sharp et al., we also found that methylation age estimates for CLP were slightly higher than those for CLO and CPO although in our analysis none of these differences were statistically significant. The low replicate rate could be due to small sample size, age, and study design differences.

Unlike genotype, which remains constant, epigenetic profiles may change over time and in response to exposure. One of the most important strengths of our study is that blood samples for cases and controls used in analysis were uniformly collected and stored soon after delivery as part of a standardized PKU screening program across all regions of Norway, thus minimizing case-control differences in age, collection, handling, and storage. The study includes nationwide enrollment of all newborns referred for cleft surgery in Norway, with controls selected by random sampling from among all live births in Norway during the same period.

Our study also has limitations. Although larger than prior cleft-methylation studies, our sample size is still relatively small, particularly when considering specific cleft subtypes. Orofacial clefts result from fusion failures in very early pregnancy. We compared DNA methylation profiles between cleft cases and controls from blood samples at birth, and those results may not be an accurate surrogate for

methylation differences in the relevant target tissues during early fetal life. Methylation differences in the blood of newborns could be caused by cleft-related differences in genetic or environmental exposures, but could also be the consequence of other unmeasured factors or due to chance.

Conclusions

Using newborn blood DNA, we identified two CpGs and multiple genomic regions that are differentially methylated between controls and cleft cases, and additional differences between cleft subtypes. These findings lend some support to the hypothesis that aberrant DNA methylation profiles may be related to orofacial clefts.

Additional files

Additional file 1: Figure S1. EWAS test results for comparisons between shared controls and cleft subtypes (CLO, CPO and CLP). Shown are the $-\log_{10}(P)$ multiplied by sign of methylation coefficient estimated using unconditional logistic regression. **Figure S2.** EWAS test results for comparisons between shared controls and combined cleft subtypes (CLP +CPO and CLP+CLO). Shown are the $-\log_{10}(P)$ multiplied by sign of methylation coefficient estimated using unconditional logistic regression. (DOCX 463 kb)

Additional file 2 Table S1. CpG probes with association P value less than 10^{-5} in any one of the cleft subtype analysis. Unconditional logistic regression model was used to test for association between DNA methylation M value and cleft status (shared controls vs. each cleft type). Covariables adjusted in each test are control surrogate variables, plate, gender, calendar year of baby's birth, gestational age, and blood cell subtype composition. **Table S2.** Sensitivity analysis results by further adjusting for mother's smoking status, BMI, drinking, folate intake and gestational age. Shown are CpG probes with association P value less than 10^{-5} in any one of the cleft subtype analysis. Unconditional logistic regression model was used to test for association between DNA methylation M value and cleft status (shared controls vs. each cleft type). Total covariables adjusted in each test are control surrogate variables, plate, gender, calendar year of baby's birth, gestational age, blood cell subtype composition, smoking, BMI, drinking, folate intake and gestational age. **Table S3.** CpGs reported to be associated with cleft using Mendelian randomization analysis by Howe et al. 2018. Shown are the analysis results in NCL dataset: mean methylation level in control groups, all cleft case combined group(case), and case subgroups; case-control logistic regression estimated coefficient and P values for the combined case group and case subgroups. **Table S4.** Methylation means and association P values (controls compared to combined cleft cases) in Alvizi et al. (Scientific Reports, 2016) study and NCL study. Shown are the overlapped CpGs between NCL study and 578 CpGs reported by Alvizi et al. **Table S5.** Candidate gene results. Shown are CpGs around literature reported cleft-associated genes with methylation association P value less than 0.01 in any one of the cleft subtype analysis. **Table S6.** Differentially methylated regions (DMR) with Sidak multiple testing corrected P value < 0.05 . **Table S7.** CpGs and association results within differentially methylated regions. (ZIP 229 kb)

Abbreviations

ALSPAC: Avon Longitudinal Study of Parents and Children; BMI: Body mass index; CLO: Cleft lip only; CLP: Cleft lip and palate; CpG: Cytosine-guanine dinucleotide; CPO: Cleft palate only; dmCpG: Differentially methylated CpG; DMR: Differentially methylated region; ENmix: Exponential-Normal mixture signal intensity background correction; EWAS: Epigenome-wide association studies; GWAS: Genome-wide association studies; NCL: Norway Facial Clefts Study; PKU: Phenylketonuria; RCP: Regression on Correlated Probes; RELIC: REgression on Logarithm of Internal Control probes; UTR: Untranslated region

Acknowledgements

We are very grateful to all the participants of this study.

Funding

This research was supported by the Intramural Research Program of the NIH, National Institute of Environmental Health Sciences (Z01 ES049033, and Z01 ES049032).

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Authors' contributions

ZX contributed to the design of the study, analyzed and interpreted the data, and wrote the original draft. JAT contributed the design and conceptualization of the study, interpreted the data, and revised the draft. AJW contributed to the design and conceptualization of the study, interpreted the data, and reviewed the final draft. RTL contributed to the design and conceptualization of the study and data collection and reviewed the final draft. OS contributed to the data collection and reviewed the final draft. All authors critically read and approved the final manuscript.

Ethics approval and consent to participate

Ethical approval for this study was obtained from the Norwegian Data Inspectorate, the Regional Research Ethics Committee for Western Norway, and the Institutional Review Board of the US National Institute of Environmental Health Sciences, National Institutes of Health. Parents gave consent to data and sample collection and all data were anonymized after collection.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

¹Epidemiology Branch, National Institute of Environmental Health Sciences, NIH, Research Triangle Park, NC, USA. ²Laboratory of Molecular Carcinogenesis, National Institute of Environmental Health Sciences, NIH, P.O. Box 12233, Mail Drop A3-05, Research Triangle Park, NC 27709, USA. ³Department of Global Public Health and Primary Care, University of Bergen, Bergen, Norway. ⁴Centre for Fertility and Health, Norwegian Institute of Public Health, Oslo, Norway. ⁵Department of Pediatric Research, Oslo University Hospital, Oslo, Norway. ⁶Department of Clinical Medicine, University of Oslo, Oslo, Norway.

Received: 29 October 2018 Accepted: 20 February 2019

Published online: 04 March 2019

References

- Dixon MJ, Marazita ML, Beaty TH, Murray JC. Cleft lip and palate: understanding genetic and environmental influences. *Nat Rev Genet.* 2011; 12:167–78.
- Murray JC, Daack-Hirsch S, Buetow KH, Munger R, Espina L, Paglinawan N, Villanueva E, Rary J, Magee K, Magee W. Clinical and epidemiologic studies of cleft lip and palate in the Philippines. *Cleft Palate Craniofac J.* 1997;34:7–10.
- Wehby GL, Cassell CH. The impact of orofacial clefts on quality of life and healthcare use and costs. *Oral Dis.* 2010;16:3–10.
- Mitchell LE, Risch N. Mode of inheritance of nonsyndromic cleft lip with or without cleft palate: a reanalysis. *Am J Hum Genet.* 1992;51:323–32.
- Sivertsen A, Wilcox AJ, Skjaerven R, Vindenes HA, Abyholm F, Harville E, Lie RT. Familial risk of oral clefts by morphological type and severity: population based cohort study of first degree relatives. *Bmj.* 2008;336:432–4.
- Vieira AR, Romitti PA, Orioli IM, Castilla EE. Inheritance of cleft palate in South America: evidence for a major locus recessive. *Orthod Craniofac Res.* 2003;6:83–7.

7. Lie RT, Wilcox AJ, Taylor J, Gjessing HK, Saugstad OD, Abyholm F, Vindenes H. Maternal smoking and oral clefts: the role of detoxification pathway genes. *Epidemiology*. 2008;19:606–15.
8. Boyles AL, DeRoo LA, Lie RT, Taylor JA, Jugessur A, Murray JC, Wilcox AJ. Maternal alcohol consumption, alcohol metabolism genes, and the risk of oral clefts: a population-based case-control study in Norway, 1996-2001. *Am J Epidemiol*. 2010;172:924–31.
9. Wilcox AJ, Lie RT, Solvoll K, Taylor J, McConaughy 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.
10. Ludwig KU, Bohmer AC, Bowes J, Nikolic M, Ishorst N, Wyatt N, Hammond NL, Golz L, Thieme F, Barth S, et al. Imputation of orofacial clefting data identifies novel risk loci and sheds light on the genetic background of cleft lip 6 cleft palate and cleft palate only. *Hum Mol Genet*. 2017;26:829–42.
11. Beaty TH, Ruczinski I, Murray JC, Marazita ML, Munger RG, Hetmanski JB, Murray T, Redett RJ, Fallin MD, Liang KY, et al. Evidence for gene-environment interaction in a genome wide study of nonsyndromic cleft palate. *Genet Epidemiol*. 2011;35:469–78.
12. Shi M, Christensen K, Weinberg CR, Romitti P, Bathum L, Lozada A, Morris RW, Lovett M, Murray JC. Orofacial cleft risk is increased with maternal smoking and specific detoxification-gene variants. *Am J Hum Genet*. 2007;80:76–90.
13. Wu T, Liang KY, Hetmanski JB, Ruczinski I, Fallin MD, Ingersoll RG, Wang H, Huang S, Ye X, Wu-Chou Y-H, et al. Evidence of gene-environment interaction for the IRF6 gene and maternal multivitamin supplementation in controlling the risk of cleft lip with/without cleft palate. *Hum Genet*. 2010;128:401–10.
14. Sharp GC, Stergiakouli E, Sandy J, Relton C. Epigenetics and orofacial clefts: a brief introduction. *Cleft Palate Craniofac J*. 2018;55:795–7.
15. Markunas CA, Xu Z, Harlid S, Wade PA, Lie RT, Taylor JA, Wilcox AJ. Identification of DNA methylation changes in newborns related to maternal smoking during pregnancy. *Environ Health Perspect*. 2014;122:1147–53.
16. Sharp GC, Lawlor DA, Richmond RC, Fraser A, Simpkin A, Suderman M, Shihab HA, Lyttleton O, McArdle W, Ring SM, et al. Maternal pre-pregnancy BMI and gestational weight gain, offspring DNA methylation and later offspring adiposity: findings from the Avon Longitudinal Study of Parents and Children. *Int J Epidemiol*. 2015;44:1288–304.
17. Joubert BR, den Dekker HT, Felix JF, Bohlén 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.
18. Joubert BR, Felix JF, Yousefi P, Bakulski KM, Just AC, Breton C, Reese SE, Markunas CA, Richmond RC, Xu CJ, et al. DNA methylation in newborns and maternal smoking in pregnancy: genome-wide consortium meta-analysis. *Am J Hum Genet*. 2016;98:680–96.
19. Alvizi L, Ke XY, Brito LA, Seselgyte R, Moore GE, Stanier P, Passos-Bueno MR. Differential methylation is associated with non-syndromic cleft lip and palate and contributes to penetrance effects. *Sci Rep-Uk*. 2017;7:2441.
20. Golding J, Pembrey M, Jones R, Team AS. ALSPAC-the Avon Longitudinal Study of Parents and Children - I. Study methodology. *Paediatr Perinat Ep*. 2001;15:74–87.
21. Howe LJ, Richardson TG, Arathimos R, Alvizi L, Passos-Bueno MR, Stanier P, Nohr E, Ludwig KU, Mangold E, Knapp M, et al. Evidence for DNA methylation mediating genetic liability to non-syndromic cleft lip/palate. *Epigenomics*. 2019;11:133–45.
22. Sharp GC, Ho K, Davies A, Stergiakouli E, Humphries K, McArdle W, Sandy J, Davey Smith G, Lewis SJ, Relton CL. Distinct DNA methylation profiles in subtypes of orofacial cleft. *Clin Epigenetics*. 2017;9:63.
23. Smith ML, Baggerly KA, Bengtsson H, Ritchie ME, Hansen KD. illuminaio: an open source IDAT parsing tool for Illumina microarrays. *F1000Research*. 2013;2:264.
24. Xu Z, Niu L, Li L, Taylor JA. ENmix: a novel background correction method for Illumina HumanMethylation450 BeadChip. *Nucleic Acids Res*. 2016;44:e20.
25. Xu Z, Langie SA, De Boever P, Taylor JA, Niu L. RELIC: a novel dye-bias correction method for Illumina Methylation BeadChip. *BMC Genomics*. 2017;18:4.
26. Niu L, Xu Z, Taylor JA. RCP: a novel probe design bias correction method for Illumina Methylation BeadChip. *Bioinformatics*. 2016;32:2659–63.
27. Price ME, Cotton AM, Lam LL, Farre P, Emberly E, Brown CJ, Robinson WP, Kobor MS. Additional annotation enhances potential for biologically-relevant analysis of the Illumina Infinium HumanMethylation450 BeadChip array. *Epigenetics Chromatin*. 2013;6:4.
28. 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.
29. Reinius LE, Acevedo N, Joerink M, Pershagen G, Dahlen S-E, Greco D, Soderhall C, Scheynius A, Kere J. Differential DNA methylation in purified human blood cells: implications for cell lineage and studies on disease susceptibility. *PLoS One*. 2012;7:e41361.
30. Peters TJ, Buckley MJ, Statham AL, Pidsley R, Samaras K, Lord RV, Clark SJ, Molloy PL. De novo identification of differentially methylated regions in the human genome. *Epigenetics Chromatin*. 2015;8:6.
31. 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.
32. Tolarova M. A study of the incidence, sex-ratio, laterality and clinical severity in 3,660 probands with facial clefts in Czechoslovakia. *Acta Chir Plast*. 1987;29:77–87.
33. Marioni RE, Shah S, McRae AF, Chen BH, Colicino E, Harris SE, Gibson J, Henders AK, Redmond P, Cox SR, et al. DNA methylation age of blood predicts all-cause mortality in later life. *Genome Biol*. 2015;16:25.
34. Horvath S. DNA methylation age of human tissues and cell types. *Genome Biol*. 2013;14:R115.
35. Kraus MR, Clauin S, Pfister Y, Di Maio M, Ulinski T, Constam D, Bellanne-Chantelot C, Grapin-Botton A. Two mutations in human BICC1 resulting in Wnt pathway hyperactivity associated with cystic renal dysplasia. *Hum Mutat*. 2012;33:86–90.
36. Verhoeven VJ, Hysi PG, Wojciechowski R, Fan Q, Guggenheim JA, Hohn R, MacGregor S, Hewitt AW, Nag A, Cheng CY, et al. Genome-wide meta-analyses of multiethnic cohorts identify multiple new susceptibility loci for refractive error and myopia. *Nat Genet*. 2013;45:314–8.
37. Grupe A, Li Y, Rowland C, Nowotny P, Hinrichs AL, Smemo S, Kauwe JS, Maxwell TJ, Cherry S, Doil L, et al. A scan of chromosome 10 identifies a novel locus showing strong association with late-onset Alzheimer disease. *Am J Hum Genet*. 2006;78:78–88.
38. Katsu R, Onogi H, Wada K, Kawaguchi Y, Hagiwara M. Novel SR-rich-related protein clasp specifically interacts with inactivated Clk4 and induces the exon EB inclusion of Clk. *J Biol Chem*. 2002;277:44220–8.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

