Received: 14 July 2023 | Revised: 22 February 2024 | Accepted: 27 March 2024

DOI: 10.1002/gepi.22564

RESEARCH ARTICLE

Genetic Epidemiology

OFFICIAL JOURNAL INTERNATIONAL GENETIC EPIDEMIOLOGY SOCIETY www.geneticepi.org

WILEY

Shared genetic risk between major orofacial cleft phenotypes in an African population

Azeez Alade[1,2] | **Tabitha Peter**[1] | **Tamara Busch**[1] | **Waheed Awotoye**[3] | **Deepti Anand**[4] | **Oladayo Abimbola**[1] | **Emmanuel Aladenika**[1] | **Mojisola Olujitan**[1] | **Oscar Rysavy**[1] | **Phuong Fawng Nguyen**[1] | **Thirona Naicker**[5] | **Peter A. Mossey**[6] | **Lord J. J. Gowans**[7] | **Mekonen A. Eshete**[8] | **Wasiu L. Adeyemo**[9] | **Erliang Zeng**[1] | **Eric Van Otterloo**[1,10] | **Michael O'Rorke**[2] | **Adebowale Adeyemo**[11] | **Jeffrey C. Murray**[12] | **Salil A. Lachke**[4,13] | **Paul A. Romitti**[2] | **Azeez Butali**[1,14]

[1]Iowa Institute of Oral Health Research, University of Iowa, Iowa City, Iowa, USA

[2]Department of Epidemiology, College of Public Health, University of Iowa, Iowa City, Iowa, USA

[3]Department of Orthodontics, College of Dentistry, University of Iowa, Iowa City, Iowa, USA

[4]Department of Biological Sciences, University of Delaware, Newark, Delaware, USA

[5]Department of Paediatrics, Clinical Genetics, University of KwaZulu-Natal and Inkosi Albert Luthuli Central Hospital, Durban, South Africa

[6]Department of Orthodontics, University of Dundee, Dundee, UK

[7]Komfo Anokye Teaching Hospital and Kwame Nkrumah University of Science and Technology, Kumasi, Ghana

[8]Department of Surgery, Addis Ababa University, School of Medicine, Addis Ababa, Ethiopia

[9]Department of Oral and Maxillofacial Surgery, University of Lagos, Lagos, Nigeria

[10]Department of Periodontics, College of Dentistry, University of Iowa, Iowa City, Iowa, USA

[11]National Human Genomic Research Institute, Bethesda, Maryland, USA

[12]Department of Pediatrics, University of Iowa, Iowa City, Iowa, USA

[13]Center for Bioinformatics and Computational Biology, University of Delaware, Newark, Delaware, USA

[14]Department of Oral Pathology, Radiology and Medicine, College of Dentistry, University of Iowa, Iowa City, Iowa, USA

Correspondence
Azeez Alade, Iowa Institute of Oral Health Research, University of Iowa, Iowa City, IA, USA.
Email: azeez-alade@uiowa.edu

Funding information
National Institute of Dental and Craniofacial Research; International Association for Dental Research; IADR/Smile Train; National Institutes of Health/National Institute of Dental and Craniofacial Research,

Abstract

Nonsyndromic orofacial clefts (NSOFCs) represent a large proportion (70%–80%) of all OFCs. They can be broadly categorized into nonsyndromic cleft lip with or without cleft palate (NSCL/P) and nonsyndromic cleft palate only (NSCPO). Although NSCL/P and NSCPO are considered etiologically distinct, recent evidence suggests the presence of shared genetic risks. Thus, we investigated the genetic overlap between NSCL/P and NSCPO using African genome-wide association study (GWAS) data on NSOFCs. These data consist of 814 NSCL/P, 205 NSCPO cases, and 2159 unrelated controls. We generated common single-nucleotide variants (SNVs) association summary

Grant/Award Numbers: DE022378, DE28300, DE024776

statistics separately for each phenotype (NSCL/P and NSCPO) under an additive genetic model. Subsequently, we employed the pleiotropic analysis under the composite null (PLACO) method to test for genetic overlap. Our analysis identified two loci with genome-wide significance (rs181737795 [$p = 2.58E{-}08$] and rs2221169 [$p = 4.5E{-}08$]) and one locus with marginal significance (rs187523265 [$p = 5.22E{-}08$]). Using mouse transcriptomics data and information from genetic phenotype databases, we identified *MDN1, MAP3k7, KMT2A, ARCN1*, and *VADC2* as top candidate genes for the associated SNVs. These findings enhance our understanding of genetic variants associated with NSOFCs and identify potential candidate genes for further exploration.

KEYWORDS

craniofacial, genetics, nonsyndromic, orofacial clefts, pleiotropy, single-nucleotide variations, transcriptomics

1 | INTRODUCTION

Orofacial clefts (OFC) are one of the most common birth defects affecting one in 700 live births worldwide (Rahimov et al., 2012). Anatomically, OFC phenotypes include cleft lip only (CLO), cleft lip and palate (CLP), and cleft palate only (CPO) (Carinci et al., 2007). CLO and CLP are often grouped as cleft lip with or without palate (CL/P) due to a common embryological and epidemiological pattern (Sperber, 2002). OFCs can also be grouped as syndromic and nonsyndromic (NS), depending upon the presence or absence of other birth defects (Calzolari et al., 2007). NSOFCs occur more frequently than syndromic OFCs, and among NSOFCs, the estimated prevalence of NSCL/P is more than double that of NSCPO (70% vs. 30%) (Marazita et al., 2002).

NSOFCs are thought to be caused by multiple genes acting alone or in combination with environmental factors. NSCL/P and NSCPO are considered etiologically distinct and genetic studies for each phenotype have identified over 50 loci for NSCL/P and 13 loci for NSCPO (Butali et al., 2019; Leslie et al., 2017; Slavec et al., 2022; Sun et al., 2015). The smaller number of NSCPO-specific risk loci identified to date could be attributed to its lower prevalence than NSCL/P, resulting in smaller sample sizes for NSCPO compared to NSCL/P in most genome-wide association studies (GWASs) (Moreno Uribe et al., 2017).

Co-occurrence of both NSCL/P and NSCPO within the same family suggests that some genetic variants may be common to both phenotypes (Rahimov et al., 2012). For example, variants in the *FOXE1* gene have been reported to be statistically significantly associated with both NSCL/P and NSCPO (Leslie et al., 2017). The hypothesis that a specific variant or gene can manifest pleiotropic effects has the potential to illuminate shared biological mechanisms contributing to the genesis of various phenotypes, thus playing an integral role in enhancing the accuracy of genetic counseling. However, analyzing NSCL/P and NSCPO data together is underpowered to identify pleiotropic loci if the effect of the variant on both phenotypes is in the opposite direction or if the sample size for one phenotype is considerably larger than that for the other phenotype (Moreno Uribe et al., 2017). To help overcome these limitations, we propose applying pleiotropic analysis under the composite null hypothesis (PLACO) (Ray & Chatterjee, 2020) to data obtained from an African GWAS of NSOFCs to identify gene variants associated with both NSCL/P and NSCPO. PLACO uses GWAS summary statistics data and has been applied successfully in determining pleiotropic loci for other traits, including NSOFCs (Ray & Chatterjee, 2020; Ray et al., 2021). Additionally, a substantial number of the pleiotropic loci for NSOFCs reported by Ray et al. were novel, not previously reported in GWASs analyzing NSCL/P and NSCPO independently or combined in any population. With the African population being the ancestral origin to modern humans and harboring the most extensive genetic variation (Conrad et al., 2006), we hypothesize that investigating pleiotropy in this population (with only one independent GWAS on NSOFCs to date) will lead to the discovery of novel loci/genes associated with both NSCL/P and NSCPO.

TABLE 1 Distribution of the study population by country, cleft status, and sex.

	NSCL/P cases		NSCPO cases		Controls		
Country	Males	Females	Males	Females	Males	Females	Total
Ghana	126	224	32	56	334	594	1366
Nigeria	79	141	22	40	210	373	865
Ethiopia	88	156	20	35	233	415	947
Total	293	521	74	131	777	1382	3178

Abbreviations: NSCL/P, nonsyndromic cleft lip with or without cleft palate; NSCPO, nonsyndromic cleft palate only.

2 | METHODS

2.1 | Study population

A detailed description of the African case-control GWAS study population has been published (Butali et al., 2019). To summarize, the GWAS data consisted of 1019 NSOFC cases (814 NSCL/P and 205 NSCPO) recruited from cleft clinic visits and during free surgical cleft repair missions in Ethiopia, Ghana, and Nigeria (Table 1). Eligible case children were patients diagnosed with NSCL/P or NSCPO and had both biological parents from Africa and still residing in Africa. Controls were children without a diagnosis of a birth defect from biological parents residing in Africa and attending immunization/welfare clinics in the same sites as the case children. Participants were recruited country-wide at several centers in Nigeria, at Kwame Nkrumah University in Ghana, and at Addis Ababa University in Ethiopia. Ethical approval was obtained from the local Institutional review board at the participating sites. Case and control child ascertainment was completed by pediatricians and cleft surgeons at the participating sites using a standardized phenotyping protocol to ensure NS phenotypes for case children and the absence of a birth defect for control children. Additionally, all eligible case children underwent echocardiography as part of presurgical planning to rule out congenital heart defects. Following ascertainment, the study was introduced to parents, who were provided with the details of the study and informed of their right to withdraw at any time.

2.2 | Data collection, DNA extraction, and genotyping

After obtaining informed consent from the case or control child's parent/guardian, demographic information (age, sex, and residential location), and limited exposure information (self-reported folic acid [yes/no], vitamins [yes/no], and other medications used during pregnancy) were obtained. However, exposure information was only available from the Lagos site (one of the six participating sites in Nigeria) and was not included further in data analyses. Saliva specimens were collected using the Oragene saliva kit, either through spitting or using a cheek swab. The specimens were deidentified and shipped to the Butali laboratory at the University of Iowa where DNA extraction was performed using the standard Oragene saliva DNA extraction protocol and quantified using Qubit (http://www.invitrogen.com/site/us/en/home/brands/Product-Brand/Qubit.html; Thermo Fisher Scientific). As part of preliminary quality control (QC), the reported sex of the participants was confirmed using Taqman XY genotyping. Subsequently, a 25 μL aliquot with a concentration ≥50 ng/μL was sent to the Center for Inherited Disease Research for genotyping using the Multi-Ethnic Genotyping Array (MEGAv2 15070954A2) (genome build 37) platform. The array consisted of over 2 million common and 60,000 rare variants selected from populations of African origin.

2.3 | Data cleaning and imputation

To ensure high-quality genotype data preimputation, the genotyped data were checked for relatedness, missing call rates, and chromosomal defects. A detailed description of the imputation and QC measures used were published elsewhere (Butali et al., 2019). Briefly, data were filtered for missing call rates ≥ 2%, Hardy–Weinberg equilibrium p value $< 10^{-3}$, and minor allele frequency (MAF) < 0.01. Continental ancestry confirmation was conducted by comparing GWAS data to HapMap specimens using a previously published approach (Laurie et al., 2010) and implemented in R packages: GWAS tools, SNPRelate, and GENESIS. Imputation was conducted using the IMPUTE2 program, and ~45 million single-nucleotide variants (SNVs) were inputted. Postimputation QC included restricting SNVs to those with an MAF ≥ 0.01 and info score ≥ 0.3. About 17 million imputed variants passed the postimputation QC and

were included in the downstream analyses, together with the 1,034,233 variants that passed the preimputation QC.

2.4 | Data analysis

Using GMMAT (Chen et al., 2016), SNV associations were analyzed using linear mixed models that adjusted for sex and study sites as detailed elsewhere (Butali et al., 2019). These analyses generated common SNV (MAF > 1%) association summary statistics for each phenotype (NSCL/P or NSCPO vs. controls). These summary statistics included the score test statistic and *p* value for each test, along with other information. The *p* values and signs (positive/negative) of the score test statistics were used to obtain *Z* scores for the SNVs, and data were harmonized to ensure that each phenotype association reported the same effect allele. Because the same control participants were included in the GWAS analysis of NSCL/P and NSCPO, *Z* scores were decorrelated to account for correlation and prevent *p* value inflation according to the method proposed by Ray and Chatterjee (2020). Subsequently, genetic overlap between NSCL/P and NSCPO was tested by implementing the PLACO method (Ray & Chatterjee, 2020) in R. This test is based on the null hypothesis that at most one of the phenotypes is associated with a genetic variant. The rejection of this null hypothesis indicates that the variant is associated with both phenotypes. A genome-wide significance threshold of 5E−08 was used, same as what we used for the published discovery GWAS (Butali et al., 2019). To examine systematic bias in the analysis, a Q–Q plot was constructed to evaluate the relationship between the observed and expected *p* values (Figure 1). Furthermore, spurious or mediated pleiotropy was examined by checking for the pleiotropic SNVs or SNVs in linkage disequilibrium ($r^2 \geq 0.5$) in the GWAS catalog database (Buniello et al., 2019) to determine if any had been previously reported to be associated with NSOFCs or known precursors (e.g., mandibular hypoplasia that causes secondary cleft palate).

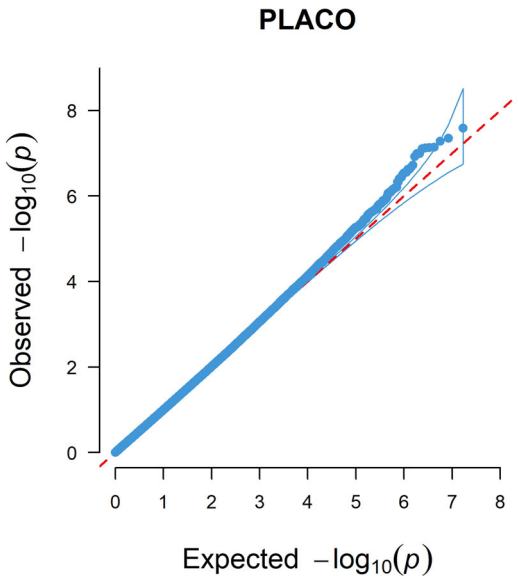

FIGURE 1 Q–Q plot for the genetic overlap analysis result. PLACO, pleiotropic analysis under the composite null.

3 | RESULTS

3.1 | SNVs showing statistical evidence of genetic overlap and direction of effects

Two genome-wide significant loci (rs181737795 $p = 2.58E-08$ and rs2221169 $p = 4.5E-08$) and one locus with marginal genome-wide significance (rs187523265

TABLE 2 Results from the PLACO analysis.

Chr	SNV ID	BP	Ref	Eff	NSCL/P_MAF	NSCPO_MAF	PLACO_pval
6	**rs181737795**	**90860856**	**A**	**G**	**0.01**	**0.01**	**2.58E−08**
10	**rs2221169**	**77558657**	**C**	**T**	**0.35**	**0.36**	**4.5E−08**
11	**rs187523265**	**118516932**	**C**	**T**	**0.02**	**0.02**	**5.22E−08**
1	rs6692379	99221575	A	G	0.54	0.55	7.29E−08
12	rs148804070	30826545	A	G	0.01	0.01	7.45E−08
7	rs115128289	32365116	C	G	0.04	0.04	7.51E−08
1	rs73110867	221915804	C	T	0.03	0.04	7.87E−08

Note: SNVs in bold texts showed statistically significant evidence of genetic overlap.

Abbreviations: BP, base pair position; Chr, chromosome; Eff, effect allele; MAF, minor allele frequency; NSCL/P, nonsyndromic cleft lip with or without cleft palate; NSCPO, nonsyndromic cleft palate only; PLACO, pleiotropic analysis under the composite null; pval, *p* value; Ref, reference allele; SNVs, single-nucleotide variants.

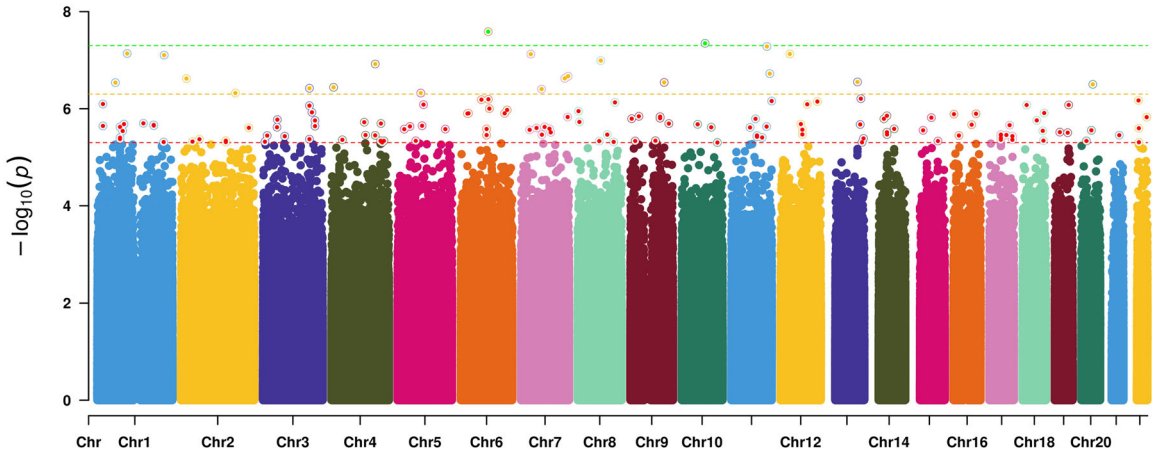

FIGURE 2 Manhattan plot showing the pleiotropic analysis results for NSCL/P and NSCPO. The chromosome numbers are colored differently and indicated on the *x*-axis. Dotted horizontal lines were placed at 5×10^{-6} (red), 5×10^{-7} (yellow), and 5×10^{-8} (green) to indicate suggestive, near, and genome-wide significance, respectively. NSCL/P, nonsyndromic cleft lip with or without cleft palate; NSCPO, nonsyndromic cleft palate only.

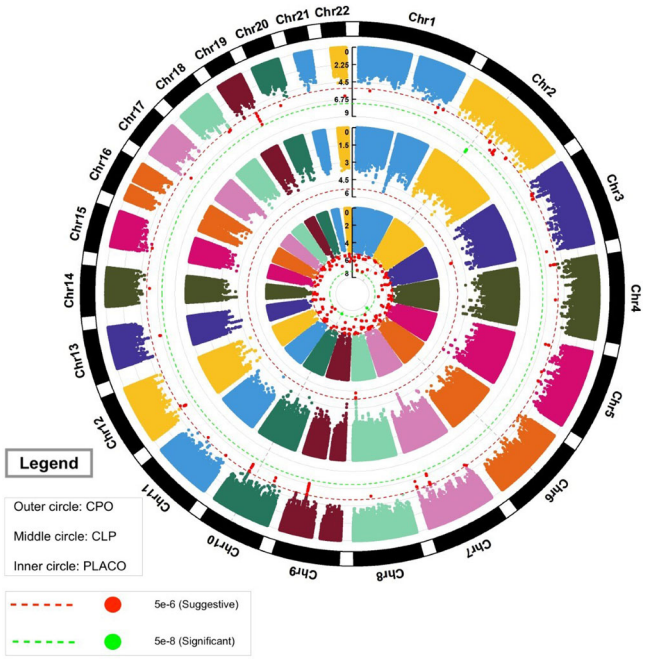

FIGURE 3 Circular Manhattan plot for the genome-wide results for NSCL/P, NSCPO, and overlap between the two phenotypes. The plots above show the log-transformed *p* values for the CPO GWAS results (outer circle), CLP GWAS results (middle circle), and the PLACO analysis results for the overlap between CLP and CPO (inner circle). The dotted red lines are used to indicate the different levels of significance (10^{-6} and 10^{-8}). Chromosome numbers are indicated along the outermost border. Variants below the genome-wide significant cut-off were colored green and those below the suggestive significant cut-off were colored red. CLP, nonsyndromic cleft lip with or without cleft palate; CPO, nonsyndromic cleft palate only; GWAS, genome-wide association study; NSCL/P, nonsyndromic cleft lip with or without cleft palate; NSCPO, nonsyndromic cleft palate only; PLACO, pleiotropic analysis under the composite null.

$p = 5.22E{-}08$) were identified (Table 2 and Figures 2 and 3). When analyzed separately, the effect allele at rs181737795 showed a null effect on both NSCL/P (odds ratio [OR]: 1.0, 95% confidence interval [CI]: 0.5–1.8) and NSCPO (OR: 1.0, 95% CI: 0.3–3.8). For rs222116, the effect allele was associated with an increased risk of NSCPO (OR = 1.5, 95% CI: 1.2–1.8) and a null effect on NSCL/P (OR = 1.0, 95% CI: 0.9–1.1). The region plots showing the other SNPs within the same region as the lead SNPs and their associated *p* values can be found in Supporting Information S1: Figures A–C. On evaluation of mediated or spurious pleiotropy, none of these three SNVs showed any evidence of horizontal/spurious pleiotropy (Supporting Information S1: Table 2). Four additional SNVs showed near genome-wide significant associations ($5 \times 10^{-8} < p < 9 \times 10^{-8}$). These SNVs included rs6692379 ($p = 7.3E{-}08$), rs148804070 ($p = 7.5E{-}08$), rs115128289 ($p = 7.5E{-}08$), and rs73110867 ($p = 7.9E{-}08$). The full results of all the SNPs showing significant/suggestive associations ($p \leq E{-}07$) can be found in Supporting Information S1: Table 1.

3.2 | Identification and prioritization of candidate genes

The SNVs, rs181737795, rs2221169, and rs187523265 are noncoding; thus, we constructed a topologically associated domain (TAD) around each SNV to identify potential genes interacting within the same TAD (<1 MB) of the SNV. This approach is based on the premise that noncoding regulatory regions of the human genome influence the expression of genes within the same TAD (Dixon et al., 2012). Subsequently, SysFACE (system tool

	Expression					Enriched Expression				
Frontonasal	E10.5	E11.0	E11.5	E12.0	E12.5	E10.5	E11.0	E11.5	E12.0	E12.5
Bach2	339.39	567.12	626.92	586.36	670.10	-2.00	-1.28	-1.13	-1.20	-1.02
Casp8ap2	1269.41	1449.72	1501.76	1164.42	1230.45	-1.20	-1.06	-1.03	-1.30	-1.24
Map3k7	893.82	1476.96	1201.60	1159.61	1296.45	1.04	1.61	1.36	1.33	1.51
Mdn1	429.96	527.17	370.75	404.33	358.76	1.04	1.27	-1.14	-1.03	-1.15
Maxilla	E10.5	E11.0	E11.5	E12.0	E12.5	E10.5	E11.0	E11.5	E12.0	E12.5
Bach2	300.34	383.07	511.08	435.61	563.65	-2.31	-1.78	-1.36	-1.61	-1.21
Casp8ap2	1214.12	1480.14	1547.92	1556.21	1426.76	-1.25	-1.04	1.002	-1.01	-1.07
Map3k7	859.75	1200.66	1381.16	1153.4	1435.82	-1.02	1.37	1.61	1.33	1.68
Mdn1	408.96	512.67	425.43	424.01	426.75	-1.002	1.23	1.02	1.02	1.04
Mandible	E10.0	E11.0	E11.5	E12.0	E12.5	E10.0	E11.0	E11.5	E12.0	E12.5
Bach2	528.05	665.82	729.41	657.69	664.01	-1.29	-1.03	1.06	-1.05	-1.02
Casp8ap2	1493.97	1722.3	1548.93	1461.45	1422.36	-1.01	1.13	-1.01	-1.04	-1.07
Map3k7	1858.63	1192.75	1083.55	1301.33	1359.92	2.17	1.38	1.24	1.51	1.59
Mdn1	1003.17	535.94	383.36	455.07	396.81	2.43	1.29	-1.08	1.09	-1.04
Palate	E13.5a	E13.5b	E14.5a	E14.5b	E14.5c	E13.5a	E13.5b	E14.5a	E14.5b	E14.5c
Bach2	772.75	1228.89	359.62	469.01	613.36	1.13	1.81	-1.88	-1.44	-1.11
Casp8ap2	2620.82	2801.98	1415.57	1784.63	1734.91	1.74	1.86	-1.06	1.18	1.15
Map3k7	511.3	764.09	1031.14	650.79	975.43	-1.66	-1.11	1.21	-1.31	1.14
Mdn1	424.77	639.81	566.33	566.26	427.22	1.03	1.56	1.38	1.36	1.04

FIGURE 4 SysFACE-based expression analysis of genes in 1 MB region flanking rs181737795 (GRCh37/hg19). Expression of four mouse ortholog genes in frontonasal, palate, maxillary, and mandible tissue in embryonic (E) and postnatal (P) stages. Heat map shows the extent of expression and values represent the average fluorescence intensity for individual genes. The heat map on the right side shows the tissue-enriched expression of the genes in comparison to the whole embryo body. MB, megabase; SysFACE, system tool for craniofacial expression-based gene discovery.

for craniofacial expression-based gene discovery) (https://bioinformatics.udel.edu/research/sysface/) was used to interrogate transcriptomics data from the frontonasal, maxilla, palatal, and mandibular tissues from the developing mouse face. Mouse transcriptomics data available through the National Center for Biotechnology Information gene Expression Omnibus (GEO) (https://www.ncbi.nlm.nih.gov/geo/) and FaceBase (https://www.facebase.org/) databases were meta-analyzed for this analysis, as previously published (Cox et al., 2018; Liu et al., 2017). Additionally, tissue enrichment in gene expression was examined by comparing the gene expression level in the tissue of interest to that of the whole embryo body. Tissue-enriched gene expression may be a good indicator of potential function, especially for constitutively expressed genes (Anand & Lachke, 2017). Following gene-level expression and enrichment analyses (Figures 4–6), *MAP3k7*, *MDN1*, *CASP8AP2*, and *BACH2* were identified as potential candidate genes for rs181737795 (Figure 4), *BCL9L*, *H2AFX*, *HINFP*, *KMT2A*, *RPS25*, *UPK2*, *VPS11*, *HYOU1*, *PHLDB1*, *ARCN1*, *ATP5L*, *CCDC84*, *DDX6*, and *MPZ12* for rs187523265 (Figure 5), and *VADC2* for rs2221169 (Figure 6). Furthermore, *MDN1*, *MAP3k7*, *KMT2A*, and *ARCN1* were prioritized based on the presence of clefting phenotypes in mice or humans with mutations affecting these genes.

4 | DISCUSSION

We identified two genome-wide significant loci, rs181737795 (chr6q15) and rs2221169 (chr10q22), and one locus, rs187523265 (chr11q23), with marginal genome-wide significance of genetic overlap between NSCL/P and NSCPO. On evaluation for mediated or spurious pleiotropy, none of the genome-wide significant SNVs or SNVs in linkage disequilibrium have been reported to be associated with known cleft predisposing factors or directly with cleft phenotypes. Using the expression and enrichment information from mouse face development, we identified 19 potential candidate genes (*MAP3k7*, *MDN1*, *CASP8AP2*, *BACH2*, *BCL9L*, *H2AFX*, *HINFP*, *KMT2A*, *RPS25*, *UPK2*, *VPS11*, *HYOU1*, *PHLDB1*, *ARCN1*, *ATP5L*, *CCDC84*, *DDX6*, *MPZ12*, and *VADC2*) within the TAD of the significantly associated SNVs. Subsequently, we prioritized *MDN1*, *MAP3k7*, *KMT2A*, and *ARCN1* as the top candidates at rs181737795 and rs187523265 loci based on the presence of cleft phenotypes in mouse models or individuals with reported variants in these genes.

MAP3K7 encodes a serine/threonine kinase protein, and mutations in this gene have been implicated in fronto-metaphyseal dysplasia with CPO as one of the clinical presentations (Wade et al., 2016). Additionally, mice with mutations in *MAP3K7* present with cleft palate

	Expression					Enriched Expression				
Frontonasal	E10.5	E11.0	E11.5	E12.0	E12.5	E10.5	E11.0	E11.5	E12.0	E12.5
Arcn1	1517.08	1813.97	2090.73	1846.44	2271.39	-2.04	-1.75	-1.51	-1.69	-1.36
Atp5l	2326.84	3149.55	2653.14	3174.58	3430.19	-1.90	-1.39	-1.67	-1.37	-1.27
Bcl9l	289.97	276.76	372.21	298.08	373.56	1.04	-1.03	1.33	1.04	1.33
C2cd2l	263.12	248.22	256.39	252.73	234.48	-1.06	-1.12	-1.09	-1.10	-1.19
Ccdc84	124.44	143.15	123.94	149.11	131.77	-1.21	-1.06	-1.24	-1.01	-1.14
Cd3g	21.68	19.46	18.51	19.39	17.75	-1.06	-1.18	-1.25	-1.19	-1.30
Cxcr5	42.70	39.15	35.57	37.66	33.05	1.05	-1.04	-1.15	-1.08	-1.22
Ddx6	740.75	1005.32	1301.37	1289.00	1358.70	-3.75	-2.71	-2.12	-2.12	-1.99
Dpagt1	489.42	599.31	509.19	547.15	510.51	-1.22	1.01	-1.19	-1.09	-1.16
H2afx	3635.27	4868.01	5752.81	6099.34	5478.70	-1.26	1.08	1.27	1.35	1.21
Hinfp	279.64	257.34	236.93	274.94	278.98	1.20	1.11	1.01	1.18	1.21
Hmbs	1557.20	1741.84	1325.33	1507.03	1136.26	-2.27	-2.01	-2.69	-2.33	-3.11
Hyou1	1072.59	1575.30	1365.42	1607.27	1143.92	-1.54	-1.07	-1.22	-1.03	-1.47
Ift46	909.93	1005.76	1039.83	1163.95	1183.88	-1.52	-1.38	-1.33	-1.19	-1.16
Kmt2a	323.36	376.77	447.49	451.73	554.23	-1.35	-1.18	-1.01	1.01	1.26
Mpzl2	230.49	256.27	280.97	283.25	343.64	-1.31	-1.17	-1.10	-1.07	1.15
Mpzl3	91.65	89.53	78.59	77.99	73.41	1.10	1.07	-1.07	-1.08	-1.14
Phldb1	307.81	443.39	393.41	456.92	408.68	-1.57	-1.08	-1.23	-1.05	-1.17
Prpf31	1536.46	1490.76	1262.53	1388.38	1086.84	-1.04	-1.11	-1.27	-1.16	-1.48
Rps25	754.56	769.65	596.19	670.74	598.76	1.26	1.29	-1.00	1.12	-1.00
Scn4b	53.55	46.21	44.12	43.60	47.41	1.04	-1.11	-1.17	-1.18	-1.08
Slc37a4	234.54	203.98	188.80	194.37	171.42	-1.34	-1.55	-1.70	-1.62	-1.86
Tmem25	59.20	50.69	48.67	47.88	47.47	1.04	-1.12	-1.16	-1.18	-1.19
Tmprss4	88.14	77.88	80.12	82.25	88.50	1.32	1.16	1.20	1.22	1.31
Trappc4	1324.86	1471.72	1119.64	1486.90	1286.07	-1.26	-1.12	-1.49	-1.12	-1.30
Ttc36	74.53	66.66	61.68	52.21	47.24	-1.24	-1.33	-1.43	-1.70	-1.87
Ube4a	324.29	337.88	320.51	321.48	284.43	-1.13	-1.09	-1.14	-1.14	-1.29
Upk2	425.83	402.39	469.58	366.35	437.82	1.47	1.41	1.63	1.29	1.55
Vps11	351.55	416.71	367.46	460.07	460.57	-1.29	-1.07	-1.23	1.03	1.02

	Expression					Enriched Expression				
Maxilla	E10.5	E11.0	E11.5	E12.0	E12.5	E10.5	E11.0	E11.5	E12.0	E12.5
Arcn1	1438.32	1696.25	1998.61	2066.47	2565.45	-2.18	-1.86	-1.57	-1.54	-1.2
Atp5l	1894.15	3037.8	3166.82	3491.89	3746.58	-2.31	-1.45	-1.38	-1.25	-1.17
Bcl9l	326.37	278.8	400.84	349.66	473.64	1.14	-1.03	1.42	1.22	1.7
C2cd2l	266.4	243.42	232.75	250.53	221.61	-1.05	-1.15	-1.21	-1.12	-1.26
Ccdc84	126.62	144.22	130.35	133.51	141.9	-1.19	-1.05	-1.16	-1.12	-1.06
Cd3g	22.37	20.85	20.04	23.29	19.4	-1.031	-1.11	-1.15	1	-1.19
Cxcr5	41.27	39.75	33.77	36.82	31.57	1.02	-1.02	-1.2	-1.1	-1.23
Ddx6	737.25	1045.14	1306.47	1294.3	1447.28	-3.74	-2.64	-2.09	-2.18	-1.86
Dpagt1	477.15	651.72	552.18	506.26	562.97	-1.25	1.1	-1.08	-1.18	-1.06
H2afx	3945.12	4890.77	6433.37	5423.17	6482.27	-1.15	1.08	1.42	1.19	1.45
Hinfp	231.98	272.73	277.37	246.99	287.3	-1.01	1.18	1.19	1.06	1.24
Hmbs	1244.87	1650.68	1549.56	1527.61	1236.65	-2.85	-2.13	-2.27	-2.3	-2.83
Hyou1	1186.1	1658.4	1275.66	1512.98	1269.59	-1.39	-1.01	-1.32	-1.11	-1.31
Ift46	752.48	978.89	1137.08	1047.24	1226.95	-1.84	-1.41	-1.22	-1.32	-1.12
Kmt2a	388.43	384.09	477.11	443.71	659.96	-1.13	-1.17	1.08	-1.02	1.47
Mpzl2	167.31	199.1	214.01	226.91	306.35	-1.85	-1.52	-1.42	-1.33	1.03
Mpzl3	100.39	82.67	71.72	78.17	69.94	1.19	-1.01	-1.17	-1.08	-1.2
Phldb1	321.48	407.06	455.75	422.62	485.47	-1.48	-1.17	-1.05	-1.12	1.03
Prpf31	1414.98	1675.35	1407.11	1403.29	1122.11	-1.13	1.05	-1.14	-1.15	-1.42
Rps25	674.6	818.15	637.4	670.88	714.63	1.13	1.37	1.06	1.13	1.2
Scn4b	61.98	53.35	49.82	50.63	41.97	1.2	1.04	-1.03	-1.02	-1.23
Slc37a4	214.74	198.72	159.5	144.52	183.72	-1.47	-1.58	-2.03	-2.2	-1.73
Tmem25	48.28	51.3	47.7	44.98	49.5	-1.17	-1.1	-1.19	-1.26	-1.14
Tmprss4	82.3	81.67	70.57	74.92	86.38	1.24	1.21	1.06	1.12	1.2
Trappc4	1158.08	1491.74	1413.89	1381.27	1413.49	-1.42	-1.1	-1.18	-1.21	-1.16
Ttc36	71.96	73.19	59.89	62.63	52.28	-1.23	-1.24	-1.49	-1.43	-1.69
Ube4a	255.75	298.93	323.2	308.61	317.97	-1.44	-1.23	-1.13	-1.19	-1.15
Upk2	447.76	373.33	363.45	327.16	382.4	1.57	1.31	1.28	1.16	1.35
Vps11	338.94	440.26	418.83	390.55	444.59	-1.32	-1.02	-1.07	-1.15	-1.07

FIGURE 5 SysFACE-based expression analysis of genes in 1 MB region flanking rs187523265 (GRCh37/hg19). Expression of 24 mouse ortholog genes in frontonasal, palate, maxillary, and mandible tissue in embryonic (E) and postnatal (P) stages. Heat map shows the extent of expression and values represent the average fluorescence intensity for individual genes. The heat map on the right side showed the tissue-enriched expression of the genes in comparison to the whole embryo body. MB, megabase; SysFACE, system tool for craniofacial expression-based gene discovery.

	Expression					Enriched Expression				
Mandible	E10.0	E11.0	E11.5	E12.0	E12.5	E10.0	E11.0	E11.5	E12.0	E12.5
Arcn1	2285.75	2066.91	1952.77	2032.38	2454.16	-1.34	-1.51	-1.65	-1.53	-1.27
Atp5l	3441.38	3155.82	2621.06	3543.06	3537.76	-1.27	-1.4	-1.73	-1.23	-1.24
Bcl9l	288.55	319.94	390.13	305.01	383.75	1.03	1.12	1.39	1.08	1.37
C2cd2l	243.27	267.29	252.13	253.4	210.3	-1.15	-1.05	-1.11	-1.1	-1.33
Ccdc84	210.43	144.67	143.35	146.42	132.39	1.4	-1.05	-1.06	-1.03	-1.14
Cd3g	18.42	19.2	21.42	23.03	18.43	-1.25	-1.2	-1.08	-1.03	-1.26
Cxcr5	41.88	40.87	36.22	36.77	36.12	1.04	1.01	-1.11	-1.1	-1.13
Ddx6	1031.22	1283.03	1137.32	1323.7	1347.35	-2.62	-2.15	-2.42	-2.06	-2.02
Dpagt1	907.03	580.87	459.81	514.8	503.37	1.53	-1.03	-1.31	-1.16	-1.18
H2afx	5587.72	5997.04	5223.47	5777.84	5173.73	1.25	1.34	1.15	1.29	1.16
Hinfp	411.71	256.67	243.54	248.89	275.65	1.78	1.1	1.05	1.06	1.17
Hmbs	2455.64	1534.45	1204.19	1561.77	1211.77	-1.43	-2.3	-2.94	-2.26	-2.9
Hyou1	2017.57	1702.82	1179.41	1561.98	1161.16	1.22	1.02	-1.42	-1.07	-1.44
Ift46	1176.75	947.74	917.94	1092.18	1145.62	-1.17	-1.46	-1.52	-1.27	-1.2
Kmt2a	403.13	372.13	460.53	454.61	598.23	-1.09	-1.2	1.04	1.02	1.36
Mpzl2	304.02	196.63	150.23	186.41	249.83	1.01	-1.55	-2.08	-1.67	-1.2
Mpzl3	100.85	82.36	84.1	78.71	69.55	1.19	-1.02	1.01	-1.07	-1.21
Phldb1	541.01	524.55	368.02	413.66	403.42	1.15	1.11	-1.28	-1.17	-1.18
Prpf31	2057.97	1501.87	1201.02	1350.05	1095.75	1.29	-1.06	-1.34	-1.18	-1.46
Rps25	1238.51	717.2	606.71	687.7	665.24	2.08	1.2	1.02	1.16	1.11
Scn4b	90.77	49.91	52.21	42.78	42.82	1.77	-1.03	1.01	-1.2	-1.2
Slc37a4	336.78	186.88	184.85	171.9	172.16	1.07	-1.69	-1.73	-1.83	-1.84
Tmem25	60.63	49.97	51.11	51.81	47.38	1.07	-1.13	-1.11	-1.09	-1.2
Tmprss4	73.22	74.6	76.98	77.3	70.18	1.1	1.12	1.15	1.16	1.05
Trappc4	1925.73	1080.16	908.1	1164.53	1207.47	1.17	-1.55	-1.84	-1.43	-1.38
Ttc36	62.84	64.76	66.82	65.16	50.11	-1.42	-1.37	-1.34	-1.38	-1.77
Ube4a	410.97	310.63	280.59	325.77	278.72	1.12	-1.2	-1.31	-1.12	-1.34
Upk2	272.31	355.08	449.52	356.06	383.1	-1.04	1.26	1.58	1.25	1.35
Vps11	547.68	407.56	336.18	420.55	459.2	1.23	-1.12	-1.33	-1.09	1.03

	Expression					Enriched Expression				
Palate	E13.5a	E13.5b	E14.5a	E14.5b	E14.5c	E13.5a	E13.5b	E14.5a	E14.5b	E14.5c
Arcn1	3205.19	3342.15	3255.5	3801.08	4042.79	1.045	1.09	1.06	1.24	1.32
Atp5l	7272.5	4768.01	4424.1	5399	6003.43	1.67	1.1	1.02	1.24	1.38
Bcl9l	415.77	468.61	427.22	343.55	294.32	1.49	1.68	1.48	1.23	1.05
C2cd2l	223.39	230.65	237.94	264.19	269.77	-1.26	-1.21	-1.17	-1.05	-1.03
Ccdc84	197.37	196.08	224.62	287.03	192.82	1.31	1.31	1.5	1.91	1.28
Cd3g	17.61	17.58	17.26	16.47	17.42	-1.31	-1.31	-1.35	-1.4	-1.32
Cxcr5	29.3	28.44	33.55	32.22	38.7	-1.37	-1.42	-1.2	-1.25	-1.05
Ddx6	3085.8	3691.88	1612.46	1949.75	2067.48	1.15	1.37	-1.67	-1.38	-1.3
Dpagt1	532.63	520.52	521.4	628.94	515	-1.11	-1.14	-1.15	1.06	-1.15
H2afx	3735.28	3126.43	4941.8	4868.99	4839.91	-1.19	-1.43	1.1	1.09	1.09
Hinfp	301.57	214.38	224.05	259.43	243.71	1.31	-1.08	-1.04	1.12	1.05
Hmbs	1391.45	1166.62	1095.71	1095.38	1127.09	-2.52	-3.01	-3.2	-3.2	-3.11
Hyou1	942.47	1056.3	1443.34	1636.6	947.07	-1.75	-1.56	-1.15	-1.02	-1.75
Ift46	1191.74	880.89	1460.76	1235.43	1199.5	-1.15	-1.57	1.06	-1.12	-1.15
Kmt2a	346.66	468.38	797.51	845.88	440.94	-1.27	1.07	1.82	1.9	1.01
Mpzl2	763.44	671	364.91	578.25	553.76	2.56	2.24	1.22	1.88	1.85
Mpzl3	77.8	71.98	69.25	78.3	88.63	-1.08	-1.17	-1.21	-1.07	1.06
Phldb1	808.32	592.83	617.43	615.7	703.53	1.72	1.26	1.32	1.31	1.49
Prpf31	1047.6	839.71	917.19	819.85	1014.91	-1.52	-1.9	-1.74	-1.94	-1.57
Rps25	837.06	982.97	721.27	964.77	624.12	1.41	1.64	1.15	1.59	1.05
Scn4b	47.76	49.88	49.68	42.81	49.16	-1.07	-1.03	-1.04	-1.2	-1.05
Slc37a4	320.6	223.08	327.56	279.07	363.12	1.02	-1.41	1.03	-1.13	1.15
Tmem25	48.47	49.23	55.94	53.72	60.55	-1.16	-1.15	-1.01	-1.05	1.07
Tmprss4	71.88	63.58	110.34	99.08	98.33	1.08	-1.05	1.64	1.48	1.45
Trappc4	1740.46	1647.87	1049.51	1246.85	1575.78	1.06	1.01	-1.57	-1.32	-1.04
Ttc36	46.47	44.41	44.25	37.21	41.69	-1.9	-1.99	-2	-2.37	-2.12
Ube4a	242.24	207.5	307.11	225.7	211.68	-1.51	-1.76	-1.19	-1.62	-1.72
Upk2	292.2	294.46	252.67	272.2	289.58	1.04	1.04	-1.12	-1.04	1.02
Vps11	302.6	253.12	407.42	438.69	444.68	-1.48	-1.77	-1.09	-1.02	-1

FIGURE 5 (Continued).

	Expression					Enriched Expression				
Frontonasal	E10.5	E11.0	E11.5	E12.0	E12.5	E10.5	E11.0	E11.5	E12.0	E12.5
Lrmda	157.67	134.35	127.42	117.43	90.29	1.33	1.13	1.05	-1.03	-1.31
Vdac2	5340.85	5824.36	5419.94	5344.07	5708.11	-1.28	-1.17	-1.27	-1.28	-1.19
Maxilla	E10.5	E11.0	E11.5	E12.0	E12.5	E10.5	E11.0	E11.5	E12.0	E12.5
Lrmda	154.25	154.15	106.45	132.65	91.93	1.3	1.29	-1.12	1.1	-1.29
Vdac2	4846.35	6025.7	5889.59	5129.93	5880.12	-1.4	-1.14	-1.16	-1.33	-1.16
Mandible	E10.0	E11.0	E11.5	E12.0	E12.5	E10.0	E11.0	E11.5	E12.0	E12.5
Lrmda	101.63	144.2	135.85	125.2	98.43	-1.16	1.22	1.15	1.05	-1.21
Vdac2	8956.08	5975.02	5050.41	5599.95	6036.75	1.32	-1.14	-1.36	-1.22	-1.13
Palate	E13.5a	E13.5b	E14.5a	E14.5b	E14.5c	E13.5a	E13.5b	E14.5a	E14.5b	E14.5c
Lrmda	96.58	97.076	117.39	92.8	111.36	-1.22	-1.21	-1.01	-1.28	-1.06
Vdac2	5363.27	5459.71	6438.18	6655.88	6746.42	-1.27	-1.24	-1.05	-1.02	-1.01

FIGURE 6 SysFACE-based expression analysis of genes in a 1 MB region flanking rs2221169 (GRCh37/hg19). Expression of one mouse ortholog gene in frontonasal, palate, maxillary, and mandible tissue in embryonic (E) and postnatal (P) stages. Heat map shows the extent of expression and values represent the average fluorescence intensity for individual genes. The heat map on the right side showed the tissue-enriched expression of the genes in comparison to the whole embryo body. MB, megabase; SysFACE, system tool for craniofacial expression-based gene discovery.

(Yumoto et al., 2013). *MDN1* functions in the maturation and export of the pre-60S ribosome subunits from the nucleus (Raman et al., 2016), with protein-altering de novo mutations in this gene previously reported in individuals with NSCL/P (Ishorst et al., 2022). Germline mutations in *KMT2A* have been reported in individuals with Wiedemann–Steiner syndrome—a condition characterized by craniofacial dysmorphism, including thin upper vermilion and wide nasal bridge, among others (Jones et al., 2012). *ARCN1* functions in the intracellular trafficking of proteins between the endoplasmic reticulum and the Golgi apparatus. Mutations in the gene have been reported to cause facial dysmorphisms, including CPO in humans (Izumi et al., 2016). *VDCA2* is a mitochondria membrane protein that, together with other proteins, controls apoptosis and autophagy (Zhou et al., 2018); apoptosis is one of the processes needed for facial development.

Considering that PLACO analysis relies on the GWAS summary statistics, the limitations of our primary GWAS also apply to our current findings. First, NSOFCs have multifactorial etiology that includes several suspected environmental exposures. The available environmental exposure data were obtained from only one clinic, self-reported, and mostly qualitative (yes/no). Considering the timing of NSOFC formation (early in the first trimester), and the wide age range at presentation, the obtained exposure information collected at the surgical clinic would not be reliable. Hence, we were unable to control for the impact of environmental exposures in our analysis and could not evaluate the possibility of an environment-dependent pleiotropic effect or pleiotropy by environmental interactions in our PLACO analysis. Secondly, our GWAS SNV association analyses assume an additive genetic effect for all the SNVs tested; however, nonadditive genetic effects (e.g., dominant or recessive effects) also contribute to complex trait etiology, including NSOFCs. The preference for additive genetic effect stems from the realization that, for variants with MAF of 1%–5%, those homozygous for the minor allele will be small and may not be informative for genetic effect estimation (Cantor et al., 2010). Additionally, from a statistical point of view, an additive genetic model is more efficient because it requires one degree of freedom compared to the two degrees of freedom required for a general model, which allows for the consideration of the different genetic effect models (Cantor et al., 2010). Additionally, our GWAS SNV analyses focused on common ($MAF > 5\%$) and low-frequency variants ($1 \leq MAF \geq 5\%$) rather than rare variants ($MAF < 1\%$), because of the employed genotyping approach (array-based augmented with imputation). With this approach, rare variants imputation, particularly in populations with limited reference samples in available databases, has been shown to perform poorly (Asimit & Zeggini, 2012).

Our findings further support the utility of PLACO for identifying pleiotropic SNVs whose effect on either phenotype might have been nonsignificant due to the limited sample size of each phenotype. Although we tried to rule out mediated and spurious pleiotropy using the GWAS catalog database, the paucity of genetic studies in the African population makes this challenging. Hence, we used all populations for our search to capture all available data. Moreover, the African population harbors the most extensive genetic variations, and findings from a

previous study suggest the presence of population-specific risk loci for pleiotropic signals associated with NSOFCs (Debashree Ray et al., 2021). Furthermore, we focused on two cleft phenotypes (NSCL/P and NSCPO) in the current study and could not consider NSCLO separately. Recent evidence suggests subtle differences in the genetic etiology of NSCLO and NSCLP (Carlson et al., 2017). However, the African GWAS data grouped NSCLO with NSCLP as NSCL/P to maximize power considering the small sample of NSCLO cases.

In conclusion, our study combines the power of PLACO and the understudied nature of the African population to identify novel pleiotropic variants influencing the risk of both NSCL/P and NSCPO. We leveraged mouse transcriptomics data from relevant craniofacial structures during mouse face development to identify potential cleft candidate genes in and around identified SNVs. These genes were either associated with syndromic/non-syndromic forms of OFCs or involved in biological processes crucial to face formation, providing additional evidence for their involvement in NSOFCs etiology.

AUTHOR CONTRIBUTIONS

Azeez Alade and Azeez Butali contributed to the conception, design, data acquisition, analysis, and interpretation, drafted, and critically revised the manuscript. Tabitha Peter, Waheed Awotoye, Deepti Anand, Oladayo Abimbola, Emmanuel Aladenika, Mojisola Olujitan, Oscar Rysavy, Phuong Fawng Nguyen, Erliang Zeng, Peter A. Mossey, Lord J. J. Gowans, Mekonen A. Eshete, Wasiu L. Adeyemo, Thirona Naicker, Tamara Busch, Eric Van Otterloo, Michael O'Rorke, Salil A. Lachke, Paul A. Rommitti, Adebowale Adeyemo, and Jeffrey C. Murray contributed to the conception, data acquisition, analysis, and interpretation, critically revised the manuscript. All authors gave final approval and agreed to be accountable for all aspects of the work.

ACKNOWLEDGMENTS

The authors are grateful to all members of the Butali Laboratory for their helpful comments and suggestions at laboratory meetings. Additionally, we thank all the families in Ethiopia, Nigeria, and Ghana who voluntarily participated in this study. This study was supported by funds from the IADR/Smile Train grant for cleft research (2022) to A. Alade, the National Institutes of Health/National Institute of Dental and Craniofacial Research grants DE022378 and DE28300 to A. Butali, and DE024776 to S. A. Lachke.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

Data available through dbGAP Accession Number: phs001090.v1. p1.

WEB RESOURCES

Qubit http://www.invitrogen.com/site/us/en/home/brands/Product-Brand/Qubit.html
SysFACE http://bioinformatics.udel.edu/research/sysface/
GEO http://www.ncbi.nlm.nih.gov/geo/
FaceBase http://www.facebase.org/aasa

ORCID

Azeez Alade http://orcid.org/0000-0001-9176-0221

REFERENCES

Anand, D., & Lachke, S. A. (2017). Systems biology of lens development: A paradigm for disease gene discovery in the eye. *Experimental Eye Research*, *156*, 22–33. https://doi.org/10.1016/j.exer.2016.03.010

Asimit, J. L., & Zeggini, E. (2012). Imputation of rare variants in next-generation association studies. *Human Heredity*, *74*(3–4), 196–204. https://doi.org/10.1159/000345602

Buniello, A., MacArthur, J. A. L., Cerezo, M., Harris, L. W., Hayhurst, J., Malangone, C., McMahon, A., Morales, J., Mountjoy, E., Sollis, E., Suveges, D., Vrousgou, O., Whetzel, P. L., Amode, R., Guillen, J. A., Riat, H. S., Trevanion, S. J., Hall, P., Junkins, H., ... Parkinson, H. (2019). The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Research*, *47*(D1), D1005–D1012. https://doi.org/10.1093/nar/gky1120

Butali, A., Mossey, P. A., Adeyemo, W. L., Eshete, M. A., Gowans, L. J. J., Busch, T. D., Jain, D., Yu, W., Huan, L., Laurie, C. A., Laurie, C. C., Nelson, S., Li, M., Sanchez-Lara, P. A., Magee, III, W. P., Magee, K. S., Auslander, A., Brindopke, F., Kay, D. M., ... Adeyemo, A. A. (2019). Genomic analyses in African populations identify novel risk loci for cleft palate. *Human Molecular Genetics*, *28*(6), 1038–1051. https://doi.org/10.1093/hmg/ddy402

Calzolari, E., Pierini, A., Astolfi, G., Bianchi, F., Neville, A. J., & Rivieri, F. (2007). Associated anomalies in multi-malformed infants with cleft lip and palate: An epidemiologic study of nearly 6 million births in 23 EUROCAT registries. *American Journal of Medical Genetics, Part A*, *143A*(6), 528–537. https://doi.org/10.1002/ajmg.a.31447

Cantor, R. M., Lange, K., & Sinsheimer, J. S. (2010). Prioritizing GWAS results: A review of statistical methods and recommendations for their application. *The American Journal of Human Genetics*, *86*(1), 6–22. https://doi.org/10.1016/j.ajhg.2009.11.017

Carinci, F., Scapoli, L., Palmieri, A., Zollino, I., & Pezzetti, F. (2007). Human genetic factors in nonsyndromic cleft lip and palate: An update. *International Journal of Pediatric Otorhinolaryngology*, *71*(10), 1509–1519. https://doi.org/10.1016/j.ijporl.2007.06.007

Carlson, J. C., Taub, M. A., Feingold, E., Beaty, T. H., Murray, J. C., Marazita, M. L., & Leslie, E. J. (2017). Identifying genetic

sources of phenotypic heterogeneity in orofacial clefts by targeted sequencing. *Birth Defects Research*, *109*(13), 1030–1038.

Chen, H., Wang, C., Conomos, M. P., Stilp, A. M., Li, Z., Sofer, T., Szpiro, A. A., Chen, W., Brehm, J. M., Celedón, J. C., Redline, S., Papanicolaou, G. J., Thornton, T. A., Laurie, C. C., Rice, K., & Lin, X. (2016). Control for population structure and relatedness for binary traits in genetic association studies via logistic mixed models. *The American Journal of Human Genetics*, *98*(4), 653–666. https://doi.org/10.1016/j.ajhg.2016.02.012

Conrad, D. F., Jakobsson, M., Coop, G., Wen, X., Wall, J. D., Rosenberg, N. A., & Pritchard, J. K. (2006). A worldwide survey of haplotype variation and linkage disequilibrium in the human genome. *Nature Genetics*, *38*(11), 1251–1260. https://doi.org/10.1038/ng1911

Cox, L. L., Cox, T. C., Moreno Uribe, L. M., Zhu, Y., Richter, C. T., Nidey, N., Standley, J. M., Deng, M., Blue, E., Chong, J. X., Yang, Y., Carstens, R. P., Anand, D., Lachke, S. A., Smith, J. D., Dorschner, M. O., Bedell, B., Kirk, E., Hing, A. V., … Roscioli, T. (2018). Mutations in the epithelial cadherin-p120-catenin complex cause mendelian non-syndromic cleft lip with or without cleft palate. *The American Journal of Human Genetics*, *102*(6), 1143–1157.

Dixon, J. R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J. S., & Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature*, *485*(7398), 376–380. https://doi.org/10.1038/nature11082

Ishorst, N., Henschel, L., Thieme, F., Drichel, D., Sivalingam, S., Mehrem, S. L., Fechtner, A. C., Fazaal, J., Welzenbach, J., Heimbach, A., Maj, C., Borisov, O., Hausen, J., Raff, R., Hoischen, A., Dixon, M., Rada-Iglesias, A., Bartusel, M., Rojas-Martinez, A., … Mangold, E. (2023). Identification of de novo variants in nonsyndromic cleft lip with/without cleft palate patients with low polygenic risk scores. *Molecular Genetics & Genomic Medicine*, *11*, e2109. https://doi.org/10.1002/mgg3.2109

Izumi, K., Brett, M., Nishi, E., Drunat, S., Tan, E.-S., Fujiki, K., Lebon, S., Cham, B., Masuda, K., Arakawa, M., Jacquinet, A., Yamazumi, Y., Chen, S.-T., Verloes, A., Okada, Y., Katou, Y., Nakamura, T., Akiyama, T., Gressens, P., … Shirahige, K. (2016). ARCN1 mutations cause a recognizable craniofacial syndrome due to COPI-mediated transport defects. *The American Journal of Human Genetics*, *99*(2), 451–459. https://doi.org/10.1016/j.ajhg.2016.06.011

Jones, W. D., Dafou, D., McEntagart, M., Woollard, W. J., Elmslie, F. V., Holder-Espinasse, M., Irving, M., Saggar, A. K., Smithson, S., Trembath, R. C., Deshpande, C., & Simpson, M. A. (2012). De novo mutations in MLL cause Wiedemann-Steiner syndrome. *The American Journal of Human Genetics*, *91*(2), 358–364.

Laurie, C. C., Doheny, K. F., Mirel, D. B., Pugh, E. W., Bierut, L. J., Bhangale, T., Boehm, F., Caporaso, N. E., Cornelis, M. C., Edenberg, H. J., Gabriel, S. B., Harris, E. L., Hu, F. B., Jacobs, K. B., Kraft, P., Landi, M. T., Lumley, T., Manolio, T. A., McHugh, C., … Weir, B. S. (2010). Quality control and quality assurance in genotypic data for genome-wide association studies. *Genetic Epidemiology*, *34*(6), 591–602. https://doi.org/10.1002/gepi.20516

Leslie, E. J., Carlson, J. C., Shaffer, J. R., Butali, A., Buxó, C. J., Castilla, E. E., Christensen, K., Deleyiannis, F. W. B., Leigh Field, L., Hecht, J. T., Moreno, L., Orioli, I. M., Padilla, C., Vieira, A. R., Wehby, G. L., Feingold, E., Weinberg, S. M., Murray, J. C., Beaty, T. H., & Marazita, M. L. (2017). Genome-wide meta-analyses of nonsyndromic orofacial clefts identify novel associations between FOXE1 and all orofacial clefts, and TP63 and cleft lip with or without cleft palate. *Human Genetics*, *136*(3), 275–286. https://doi.org/10.1007/s00439-016-1754-7

Liu, H., Busch, T., Eliason, S., Anand, D., Bullard, S., Gowans, L. J. J., Nidey, N., Petrin, A., Augustine-Akpan, E. A., Saadi, I., Dunnwald, M., Lachke, S. A., Zhu, Y., Adeyemo, A., Amendt, B., Roscioli, T., Cornell, R., Murray, J., & Butali, A. (2017). Exome sequencing provides additional evidence for the involvement of ARHGAP29 in Mendelian orofacial clefting and extends the phenotypic spectrum to isolated cleft palate. *Birth Defects Research*, *109*(1), 27–37. https://doi.org/10.1002/bdra.23596

Marazita, M. L., Field, L. L., Cooper, M. E., Tobias, R., Maher, B. S., Peanchitlertkajorn, S., & Liu, Y. E. (2002). Nonsyndromic cleft lip with or without cleft palate in China: Assessment of candidate regions. *The Cleft Palate-Craniofacial Journal*, *39*(2), 149–156. https://doi.org/10.1597/1545-1569_2002_039_0149_nclwow_2.0.co_2

Moreno Uribe, L. M., Fomina, T., Munger, R. G., Romitti, P. A., Jenkins, M. M., Gjessing, H. K., Gjerdevik, M., Christensen, K., Wilcox, A. J., Murray, J. C., Lie, R. T., & Wehby, G. L. (2017). A population-based study of effects of genetic loci on orofacial clefts. *Journal of Dental Research*, *96*(11), 1322–1329. https://doi.org/10.1177/0022034517716914

Rahimov, F., Jugessur, A., & Murray, J. C. (2012). Genetics of nonsyndromic orofacial clefts. *The Cleft Palate-Craniofacial Journal*, *49*(1), 73–91. https://doi.org/10.1597/10-178

Raman, N., Weir, E., & Müller, S. (2016). The AAA ATPase MDN1 acts as a SUMO-targeted regulator in mammalian pre-ribosome remodeling. *Molecular Cell*, *64*(3), 607–615.

Ray, D., & Chatterjee, N. (2020). A powerful method for pleiotropic analysis under composite null hypothesis identifies novel shared loci between type 2 diabetes and prostate cancer. *PLoS Genetics*, *16*(12), e1009218. https://doi.org/10.1371/journal.pgen.1009218

Ray, D., Venkataraghavan, S., Zhang, W., Leslie, E. J., Hetmanski, J. B., Weinberg, S. M., Murray, J. C., Marazita, M. L., Ruczinski, I., Taub, M. A., & Beaty, T. H. (2021). Pleiotropy method reveals genetic overlap between orofacial clefts at multiple novel loci from GWAS of multi-ethnic trios. *PLoS Genetics*, *17*(7), e1009584. https://doi.org/10.1371/journal.pgen.1009584

Slavec, L., Karas Kuželički, N., Locatelli, I., & Geršak, K. (2022). Genetic markers for non-syndromic orofacial clefts in populations of European ancestry: A meta-analysis. *Scientific Reports*, *12*(1), 1214. https://doi.org/10.1038/s41598-021-02159-5

Sperber G. H. (2002). Palatogenesis: Closure of the secondary palate. In D. F. Wyszynski (Ed.), *Cleft lip and palate: From origin to treatment* (pp. 14–24). Oxford University Press.

Sun, Y., Huang, Y., Yin, A., Pan, Y., Wang, Y., Wang, C., Du, Y., Wang, M., Lan, F., Hu, Z., Wang, G., Jiang, M., Ma, J., Zhang, X., Ma, H., Ma, J., Zhang, W., Huang, Q., Zhou, Z., ... Yang, Y. (2015). Genome-wide association study identifies a new susceptibility locus for cleft lip with or without a cleft palate. *Nature Communications*, *6*(1), 6414. https://doi.org/10.1038/ncomms7414

Wade, E. M., Daniel, P. B., Jenkins, Z. A., McInerney-Leo, A., Leo, P., Morgan, T., Addor, M. C., Adès, L. C., Bertola, D., Bohring, A., Carter, E., Cho, T. J., Duba, H. C., Fletcher, E., Kim, C. A., Krakow, D., Morava, E., Neuhann, T., Superti-Furga, A., ... Robertson, S. P. (2016). Mutations in MAP3K7 that alter the activity of the TAK1 signaling complex cause frontometaphyseal dysplasia. *The American Journal of Human Genetics*, *99*(2), 392–406. https://doi.org/10.1016/j.ajhg.2016.05.024

Yumoto, K., Thomas, P. S., Lane, J., Matsuzaki, K., Inagaki, M., Ninomiya-Tsuji, J., Scott, G. J., Ray, M. K., Ishii, M., Maxson, R., Mishina, Y., & Kaartinen, V. (2013). TGF-β-activated kinase 1 (Tak1) mediates agonist-induced Smad activation and linker region phosphorylation in embryonic craniofacial neural crest-derived cells. *Journal of Biological Chemistry*, *288*(19), 13467–13480.

Zhou, K., Yao, Y.-L., He, Z.-C., Chen, C., Zhang, X.-N., Yang, K.-D., Liu, Y.-Q., Liu, Q., Fu, W.-J., Chen, Y.-P., Niu, Q., Ma, Q.-H., Zhou, R., Yao, X.-H., Zhang, X., Cui, Y.-H., Bian, X.-W., Shi, Y., & Ping, Y.-F. (2018). VDAC2 interacts with PFKP to regulate glucose metabolism and phenotypic reprogramming of glioma stem cells. *Cell Death & Disease*, *9*(10), 988. https://doi.org/10.1038/s41419-018-1015-x

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Alade, A., Peter, T., Busch, T., Awotoye, W., Anand, D., Abimbola, O., Aladenika, E., Olujitan, M., Rysavy, O., Nguyen, P. F., Naicker, T., Mossey, P. A., Gowans, L. J. J., Eshete, M. A., Adeyemo, W. L., Zeng, E., Van Otterloo, E., O'Rorke, M., Adeyemo, A., ... Butali, A. (2024). Shared genetic risk between major orofacial cleft phenotypes in an African population. *Genetic Epidemiology*, 1–12. https://doi.org/10.1002/gepi.22564