Damaging Mutations in AFDN Contribute to Risk of Nonsyndromic Cleft Lip With or Without Cleft Palate

Awotoye, Waheed; Mossey, Peter A.; Hetmanski, Jacqueline B.; Gowans, Lord J. J.; Eshete, Mekonen A.; Adeyemo, Wasiu L.

Published in:
Cleft Palate-Craniofacial Journal

DOI:
10.1177/10556656221135926

Publication date:
2022

Document Version
Peer reviewed version

Damaging mutations in *AFDN* contribute to risk of nonsyndromic cleft lip with or without cleft palate

Waheed Awotoye[1,2], Peter A. Mossey[3], Jacqueline B. Hetmanski[4], Lord J.J Gowans[5], Mekonen A. Eshete[6], Wasiu L. Adeyemo[7], Azeez Alade[2,8], Erliang Zeng[9], Olawale Adamson[7], Olutayo James[7], Azeez Fashina[7], Modupe O Ogunlewe[7],Thirona Naicker[10], Chinyere Adeleke[2], Tamara Busch[2], Mary Li[2], Aline Petrin[1], Abimbola Oladayo[2], Sami Kayali[2], Joy Olotu[11], Veronica Sule[12], Mohaned Hassan[2], John Pape[2], Emmanuel T. Aladenika[1,2], Peter Donkor[13], Fareed K.N. Arthur[5], Solomon Obiri-Yeboah[14], Daniel K. Sabbah[15], Pius Agbenorku[13], Debashree Ray[4], Gyikua Plange-Rhule[16], Alexander Acheampong Oti[14], Daniah Albokhari[17], Nara Sobreira[17], Martine Dunnwald[18], Terri H.Beaty[4], Margaret Taub[4], Mary L. Marazita[19], , Adebowale A. Adeyemo[20], Jeffrey C. Murray[21], Azeez Butali[1,2]

[1]Iowa Institute for Oral Health Research, University of Iowa, Iowa City, IA, USA
[2]Department of Oral Pathology, Radiology and Medicine, College of Dentistry, University of Iowa, Iowa City, IA, USA
[3]Department of Orthodontics, University of Dundee, Dundee, UK
[4]Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA
[5]Department of Biochemistry and Biotechnology, Kwame Nkrumah University of Science and Technology, Kumasi, Ghana
[6]Addis Ababa University, School Medicine, Surgical Department, Addis Ababa, Ethiopia
[7]Department of Oral and Maxillofacial Surgery, University of Lagos, Lagos Nigeria
[8]Department of Epidemiology, College of Public Health, University of Iowa, Iowa City, IA, USA.
[9]Division of Biostatistics and Computational Biology, College of Dentistry, University of Iowa, Iowa City, IA, USA
[10]Department of Pediatrics, University of KwaZulu-Natal, South Africa
[11]Department of Anatomy, University of Port Harcourt
[12]Department of Operative Dentistry, College of Dentistry, University of Iowa, Iowa City, IA, USA
[13]Department of Surgery, School of Medicine and Dentistry, Kwame Nkrumah University of Science and Technology, Kumasi, Ghana
[14]Department of Maxillofacial Sciences, School of Medicine and Dentistry, Kwame Nkrumah University of Science and Technology, Kumasi, Ghana
[15]Department of Child Oral Health and Orthodontics, School of Medicine and Dentistry, Kwame Nkrumah University of Science and Technology, Kumasi, Ghana
[16]Department of Child Health, School of Medicine and Dentistry, Kwame Nkrumah University of Science and Technology, Kumasi, Ghana
[17]McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University
[18]Department of Anatomy and Cell Biology, University of Iowa`
[19]Center for Craniofacial and Dental Genetics, Department of Oral and Craniofacial Sciences, School of Dental Medicine, and Department of Human Genetics, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA, USA
[20]National Human Genomic Research Institute, Bethesda, MD, USA
[21]Department of Pediatrics, University of Iowa, Iowa City, IA, USA

Corresponding author: Azeez Butali, Department of Oral Pathology, Radiology and Medicine/Iowa Institute for Oral Health Research, College of Dentistry, University of Iowa, Butali Laboratory, ML2198, 500 Newton Road, Iowa City, IA 52242, USA. Email: azeez-butali@uiowa.edu

Awotoye, W. et al. 'Damaging Mutations in AFDN Contribute to Risk of Nonsyndromic Cleft Lip With or Without Cleft Palate', Cleft Palate-Craniofacial Journal (2022). DOI: https://doi.org/10.1177/10556656221135926

ABSTRACT

Objectives: Novel or rare damaging mutations have been implicated in the developmental pathogenesis of non-syndromic cleft lip with or without cleft palate (nsCL±P). Thus, we investigated the human genome for high impact mutations that could explain the risk of nsCL±P in our cohorts.

Materials and Methods: We conducted next-generation sequencing (NGS) analysis of 130 nsCL±P case-parent African trios to identify pathogenic variants that contribute to the risk of clefting. We replicated this analysis using whole-exome sequence data from a Brazilian nsCL±P cohort. Computational analyses were then used to predict the mechanism by which these variants could result in increased risks for nsCL±P.

Results: We discovered damaging mutations within the *AFDN* gene, a cell adhesion molecule (CAMs) that was previously shown to contribute to cleft palate in mice. These mutations include p.Met1164Ile, p.Thr453Asn, p.Pro1638Ala, p.Arg669Gln, p.Ala1717Val and p.Arg1596His. We also discovered a novel splicing p.Leu1588Leu mutation in this protein. Computational analysis suggests that these amino acid changes affect the interactions with other cleft associated genes including nectins (PVRL1, PVRL2, PVRL3, PVRL4) CDH1, CTNNA1 and CTNND1.

Conclusion: This is the first report of the contribution of *AFDN* to the risk for nsCL±P in humans. *AFDN* encodes AFADIN, an important CAM that forms calcium-independent complexes with nectins 1 and 4 (encoded by the genes *PVRL1* and *PVRL4*). This discovery shows the power of NGS analysis of multi-ethnic cleft samples in combination with a computational approach in the understanding of the pathogenesis of nsCL±P.

Keywords: Next-generation sequencing, craniofacial genetics, pathogenic variants,

Thermodynamics, damaging mutations, cleft, cell adhesion molecule, AFDN

INTRODUCTION

Nonsyndromic orofacial clefts (nsOFCs) are the most common craniofacial defect and occur due to failure of fusion of embryonic facial prominences[1,2]. Globally, this birth defect has an incidence of 1 in 700 livebirths[3]. The management of this nsOFCs requires a multidisciplinary approach to correct the defect surgically, the medical complications such as speech defects, feeding problems, malocclusion, as well as psychosocial and economic impacts on the affected family[4]. Studies have also shown that this defect has a negative impact on oral health related quality of life of the affected family[5]. The economic cost due to the multidisciplinary management, is estimated to be about $200,000USD per affected individual per year, and the absenteeism of the affected parents from work during the management period compounds the burden this defect has on public health[5].

Genetic factors that contribute to the risk of nsOFCs have been studied across most world populations. Common variants analysis through genome-wide association studies have identified a number of risk variants associated with nsCL±P [6-11]. The role of rare coding mutations in the etiology of this birth defect have also been reported[12-14]. However, ~75% of the genetic heritability of this nsOFCs remain unknown. Whole-genome sequencing (WGS) is a powerful sequencing strategy that provides the opportunity to identify novel/ rare variants present across the entire genome. Analyses of such sequencing data have been applied to large nsCL±P cohorts resulting in the identification of some novel loci and candidate genes that are associated with increased risk for nsOFCs[15]. WGS also allows us to screen for variants in genes that are involved in several molecular processes that contribute to the development of human anatomical structures required for the formation of the face. Albeit there are still challenges with the interpretation of some of these variants. In-silico tools can serve as screening to identify damaging or deleterious variants thus providing evidence on how these variants contribute to the pathogenesis if nsOFCs.

Many cellular processes have been involved in the morphogenesis of the lip and the palate, cell proliferation, migration and apoptosis, and the cell-cell adhesion [16-20]. Thus, pathogenic variants

in genes that play roles in these processes may contribute to the etiology of nsCL±P. To identify new candidate genes, we explored next-generation sequencing using our African nsCL±P case-parent trios and replicated findings in Brazilian nsCL±P cohort.

METHODS

Study Participants and Samples

Ethical approvals were obtained from the local institutional review boards (IRBs) (ADM/DCST/HREC/VOL.XV/321, ERC/2011/12/01, CHRPE/RC/018/130, and IRB ID #: 201101720).

The African Craniofacial Anomalies Network (AfriCRAN) was established to recruit and investigate factors contributing to the risk of craniofacial anomalies within the African population[9,21]. Using the protocol designed by AfriCRAN, case-parent trios were recruited for this study. In some cases, recruitment also included affected child and mother (dyads) in situations where the father was not available, as well as other family members. Each family unit (case-parent trio, dyad) consisted of a proband with nsCL±P and the unaffected parents (Father and mother for trio, mother for dyad) (Figure 1A). These individuals were recruited from 2 sub-Saharan African countries (Ghana and Nigeria) during the preparation for cleft repair surgery.

Prior to recruitment of the individuals into the study, we ascertain their ancestry by ensuring that all parents and grandparents were of African descent. Among the team involved in the recruitment are clinicians who conduct deep phenotyping to ensure there are no other structural birth defects. This screen includes parents and relatives where available, to identify undiagnosed syndromic clefts which may have been missed because of variation in the expressivity of the associated phenotypes such as lip pits.

During surgical preparation of these cases, clinical investigations are carried out to rule out potential perioperative complications. We leveraged the results of these investigations to rule out other congenital anomalies such as heart defects. Informed consent was obtained from the parents individually, jointly for the child. All case-parent trios were natives of the participating countries.

All the individuals recruited in this study had their saliva samples collected using the Oragene saliva tool kits. Each participant was assigned a unique identifier number (UNID) and their epidemiological as well as clinical data were obtained at the participating institutions. The UNID and other information were remotely uploaded in the REDCap database while their saliva samples were shipped to a laboratory in the USA for further analyses.

DNA Extraction, Next-generation Sequencing, Variant Calling and Quality Control

Details of these processes have been previously published [22]. Briefly, following the shipment of the samples to the US laboratory, DNA was extracted using the Oragene DNA extraction protocol. Quantifications of the extracted DNA were done using Qubit (http://www.invitrogen.com/site/us/en/home/brands/Product-Brand/Qubit.html; Thermo Fisher Scientific, Grand Island, NY). Stocks and working aliquots of each DNA sample was made for downstream analyses. TaqMan XY genotyping was done to ascertain the reported sex as a quality control step.

Whole-genome sequencing was done as part of the Gabriella Miller Kids First genomic sequencing cohorts. The mean coverage depth was 30x and the genomic sequence of each sample was aligned to the Human genome assembly 38(GRCh38/hg38). The binary alignment map (BAM) and sequence alignment map (SAM) files obtained were used to call variants using the GenomeAnalysisToolKit (GATK) pipelines developed at the Broad Institute (https://software.broadinstitute.org/gatk/best-practices/workflow). The variants which include single nucleotide polymorphisms (SNPs) and Insertions/Deletions (Indels) were called using the HaplotypeCaller in GVCF mode for single-sample variant calling and the GenotypeGVCFs for

multiple-sample joint variant calling; respectively. The called variants were stored in a variant call format (VCF) file which was used for downstream analysis.

The quality control (QC) after the WGS include elimination of data (loci and samples) with genomic missingness > 10% and Hardy-Weinberg equilibrium (HWE) < 1E-06. We also excluded from our analysis samples with Mendelian errors outside of 3 standard deviations and with calls outside 4 standard deviations from mean heterozygote/homozygote ratio. Additionally, we checked for relatedness between the proband and parents, and did a sex confirmation based on genomic sequence. All samples that passed QC were included in the downstream analysis.

In our investigation of the coding regions of the Brazilian nsCL±P cohorts, whole-exome sequencing (WES) of the probands and their parents, where available was done. The WES sequence data were initially aligned to the Human genome assembly 19 (GRCh37/hg19) and the BAM and SAM files generated were used to call the variants and downstream analyses were done using the VCF files. For the variants annotated, we used the UCSC lift-over tool to convert the genome coordinates from the GRCh37/hg19 assembly to GRCh38/hg38 assembly. This ensured consistency in variant annotation (using the same genome assembly) across all analyzed samples. It is important to note that the analyses of the African and Brazilian cohorts were done separately.

Screening for Pathogenic Novel/Rare Variants

To identify novel candidate genes that play roles in lip and palate development and whose pathogenic variants contribute to the risk of cleft; we screened the entire genome of the African nsCL±P cohort for high impact protein-altering variants. Firstly, we filtered out the low-quality variants by setting quality metrics with genotype quality (GQ) threshold of 20 and a read depth (RD) threshold of 10. Variants with GQ ≥ 20 and RD ≥ 10 were considered as high quality with an exceptionally low false positive probability[23,24]. We then filtered these high-quality variants for rarity: variants with minor allele frequency (MAF) < 1% and those variants that are within the coding regions whose effect include loss of function mutations and missense mutations (Figure 1B).

Among these rare and novel protein-altering variants, we filtered for those variants in genes that play roles in craniofacial development by reviewing the literatures, and publicly available databases such as DECIPHER (https://www.deciphergenomics.org/) and Mouse Genome Informatics (http://www.informatics.jax.org/) databases. For database mining, we prioritized those genes with novel / rare variants in at least two affected individuals. Following this, we ascertained the parent from which the variants were inherited from through segregation analysis and used insilico tools to predict the effect of the pathogenicity of the variants.

To evaluate the replicability of our findings in other populations, we investigated the coding regions of the Brazilian nsCL±P cohorts for variants within those genes identified in the African cohort.

Effects of pathogenic variants on protein structures and functions

Using the computational methods to ascertain the effect of the pathogenic variants on the protein functionality hence predicting the mechanism of cleft developmental pathogenesis, we investigated the effect of the amino acid changes on the protein structures. Here, we used a bioinformatic tool called Have (y)Our Protein Explained (HOPE). This is a web-based tool which

helps to interpret the structural and molecular effects of point mutations that results in amino acid changes[25].

For each of the pathogenic amino acid changes discovered, we mutated the protein sequence to generate the predicted pathogenic protein variant and then determined the effect on the physicochemical properties of the protein. This analysis also predicted the effect of these structural change on the protein function[25].

RESULTS

Study Samples and Quality Control

Among the families that have been recruited through the AfriCRAN, we selected 150 case-parent trios that were sent for whole-genome sequencing (WGS) at the Broad Institute. We defined the case-parent trios here as an affected child (with nsCL±P) and unaffected parents (Figure 1A). These WGS samples were part of the Gabriella-Miller Kids First (GMKF) Pediatric Research Consortium (https://kidsfirstdrc.org/) which was established and funded to address the knowledge gaps in the understanding role of the genetics in the etiology of structural birth defects and pediatric cancers.

Following quality control (QC) process, 20 trios were excluded after checking for completeness of the sequenced genomes, Mendelian errors and relatedness. The remaining 130 case-parent trios had over 25 million variants (both common and rare) which were analyzed following our data filtration steps (Figure 1B).

Variants in *AFDN* discovered in nsCL±P cohorts

Among the protein-altering variants identified in our cohorts, we screened for those in genes that contribute to craniofacial development. In addition to the *de novo* variants that were identified in our previous study, we found protein-altering variants in a few other genes that contribute to craniofacial development (Table 1). Notably, we found protein-altering variants in *AFDN,* a protein-coding gene which functions in cell-adhesion. This gene is highly constrained with metrics

that show its intolerance to loss of function mutations - the Loss-of-function Observed / Expected Upper bound Fraction (LOEUF) is 0.22 for this gene and the missense metrics measured by the "Z" score is 2.21. These metrics suggest that it is a less mutable gene. The top significant craniofacial anomalies (with strong association with orofacial clefts) associated with this gene include micrognathia, microcephaly, high palate, long philtrum, and cleft palate.

The variants we identified in *AFDN* were missense and splice site variants (Table 1). These variants have a combined annotation dependent depletion (CADD) score ranging from 7.95 to 31 (Table 1). The high CADD scores of some of the variants indicates they are among the topmost deleterious mutations in the human genome (Table 1).

In the African cohorts, we investigated the segregation patterns of the variants in the families with the burden. We identified that 2 of the variants (c.35A>G; p.Lys12Arg and c.1358C>A; p.Thr453Asn) were inherited from the mother in a dominant heterozygous model (Figure 3A and 3B) while c.3492G>A; p.Met1164Ile and c.3559A>G; p.Ser1187Gly were inherited in a compound heterozygous model (Figure 3).

Other protein-altering variants identified in genes that play roles in craniofacial development include *ADAM23* and *ITGA6*. These genes including *AFDN,* were not among the cleft candidates published on the CleftGeneDB (https://bioinfo.uth.edu/CleftGeneDB/) as at our last search in May 2022. The CleftGeneDB is an updated database that records genes that cause cleft in humans and mice[26].

Damaging mutations in *AFDN* are predicted to affect the protein interaction with other cleft-associated proteins

Our detailed analysis of the effect of the amino acid changes on the protein structures showed that some of these residues are highly conserved and the genetic mutations leading to amino acid change may distort the Afadin structure (Figure 2). The p.Lys12Arg mutation can result in disruption of the α-helix secondary structure of Afadin and the size of the mutant residue could potentially cause a bump in the Afadin. The p.Met1164Ile mutation could result in the loss of protein interactions due to the change in size of the residues. Finally, the p.Thr453Asn mutation occur within the Forkhead-Associated (FHA) domain. This change in amino acid residue within this domain affect the hydrophobic interactions within the core and surface of the protein. The effect of this change may abolish the domain function.

The p.Ser1187Gly variant could lead to the loss of interaction of Afadin with other proteins. This is due to the smaller size and the flexibility of the mutant residue. The damaging p.Pro1681Ala variant affect the special conformation of the protein due to the loss of the proline rigidity. The change in size in the mutant protein might lead to loss of interactions with other protein. The damaging p.Arg710Glu variant occurs within a domain and the arginine residue is highly conserved (100%). Due to the mutation within a domain and highly conserved location, change in the charge and the smaller size of the mutant residue; significant interactions of Afadin with other molecules will be lost. The p.Ala1760Val variant result in introduction of a larger residue and a bump at amino acid position 1760 in the Afadin protein. The deleterious p.Arg1596His variant occur at a highly conserved location and result in change in the charge at this location. The change in charge combined with the change in size might contribute to the loss of interactions between Afadin and other molecules.

To identify the specific interactions between Afadin and other cleft associated molecules, we constructed the Afadin protein network using the String database (https://string-db.org/). The

string database is one of the databases of known and predicted protein-protein interactions. These interactions are determined from high-throughput wet lab experiments which are extracted from published works, established interactions from organisms which are transferred to others based on homology and computational prediction. The computational prediction used co-expression knowledge, text mining of queried proteins and data from other databases. The result from this shows an organism specific network of protein interactions and statistical values of the strength of the prediction and the false discovery rate (FDR) which is a test of the significance of the association of processes, pathways and diseases as determined[27].

Our analysis to determine what other proteins interact with Afadin resulted in the identification of interactions with several known cleft candidate such as the Nectins (PVRLs), CDH1, CTNNA1 and CTNND1 (Figure 4A). Further analysis showed that the interaction between afadin and nectins1 and 4 (Figure 4B) is critical for the organization of the adherens junction (GO:0034332) with an FDR<0.05 (FDR = 0.00045). Additionally, our disease-gene association analysis showed that this interaction is significantly associated with a syndromic cleft: cleft lip-palate-ectodermal dysplasia syndrome (DOID:0060773; OMIM:225060) with an FDR of 0.00041[28-30]. In summary, our analyses suggest that the disruption of Afadin interactions by these identified variants contribute to the risk of cleft.

Discussion

Our results from the *de novo* analysis reported novel genetic variants that contribute to the risk of nsCL±P in the African population[22]. To identify other genes, we analyzed for novel or rare inherited protein-altering variants in genes reported to contribute to craniofacial development. These analyses are premised on the fact that some risk genes have low penetrance where parents carry the variant but not the nsCL±P phenotype and, proband has both variant and phenotype.

Based on our analysis, we identified a novel cleft candidate gene which encodes one of the cell adhesion molecules (CAMs). Afadin is a cell adhesion molecule occurring as a complex with the nectins at intercellular junctions. Many cell adhesion molecules have been well-reported in the etiology of nsCL±P in humans[31]. One of these CAMs constitute the E-cadherin-catenin complex found at the adherens and tight junctions (AJ and TJ) albeit predominant at the tight junction (TJ). This E-cadherin-catenin complex are well reported in nsCL±P pathogenesis. Mutations in *CDH1* (encodes the E-cadherin) and *CTNNB1, CTNND1, CTNNA1* and *CTNNA2 (*encodes the catenins), have been shown to contribute to the risk of clefts[9,32-35]. Afadin forms a complex with the nectins (*PVRL1* and *PVRL4*) which is predominant at the adherens junction (AJ).

The nectins belong to the family of transmembrane cell adhesion proteins which interacts with the afadin to maintain the integrity of the adherens and tight junctions [36,37]. In addition to our findings from the protein interaction network analysis, reports have also established that afadin binds exclusively to the cytoplasmic tail of the nectins 1 and 4 and anchor the filamentous actin[34]. This interaction links the F-actin to the adherens junction (AJ) and it is crucial for normal development of the palate [38]. Studies have reported associations between these nectins and low penetrant cleft palate[39,40]. Knockout of the *Afdn(Afdn$^{fl/fl}$)* in the epithelial cells in developing mice resulted in a highly penetrant cleft palate among other defects[41]. Murine studies have thus concluded that the nectin-afadin complex is critical in the development of the lip and palate[37]. However, despite mutations in the nectins being reported in the risk of nsCL±P, human genetic studies haven not

reported pathogenic mutations in Afadin. Specific afadin (MLLT4) interaction with Nectin1 (PVRL1) and Nectin4 (PVRL4) is significantly associated with Cleft lip-palate-ectodermal dysplasia (CLPED) syndrome (FDR = 0.0075) [40]. Individuals from the present study with Afadin mutations have non-sydromic clefts are clinically do not present with features suggestive of CLPED.

The result from our analysis give evidence supporting pathway analysis in discovery of novel candidate genes. Investigation of pathogenic mutation in genes that are involved in the pathways and biological processes that contribute to lip and palate development may help discover some of the unknown genetic risk liability of clefts.

To the best of our knowledge, the pathogenic variants we identified in *AFDN* is the first report implicating the role of the molecule in nsCL±P etiology in humans. These variants are rare, or novel based on their minor allele frequency of less than 1% in all population. The fact that these variants are either rare or novel in addition to the *in-silico* tools predictions provide evidence suggesting that these variants contribute to the pathogenesis of nsCL±P in the African population.

Among the network of AFADIN interactions from our computational analysis, the predicted disruption of the AFADIN and NECTIN due to amino acid changes is the most probable mechanism of cleft developmental pathogenesis. The *AFDN* variants discovered in this study affect the protein structure resulting in a disruption of its molecular interaction profile. In this report, we have detailed the result of our findings from the analysis of the entire genome for inherited variants that contribute to the risk of low penetrance cleft. However, an experimental analysis in model organism will be required to elucidate the pathways by which these variants result in nsCL±P.

References

1. Mossey PA, Little J, Munger RG, Dixon MJ, Shaw WC. Cleft lip and palate. *Lancet (London, England).* 2009;374(9703):1773-1785.
2. Smarius B, Loozen C, Manten W, Bekker M, Pistorius L, Breugem C. Accurate diagnosis of prenatal cleft lip/palate by understanding the embryology. *World journal of methodology.* 2017;7(3):93-100.
3. Rahimov F, Jugessur A, Murray JC. Genetics of nonsyndromic orofacial clefts. *Cleft Palate Craniofac J.* 2012;49(1):73-91.
4. Hunt O, Burden D, Hepper P, Johnston C. The psychosocial effects of cleft lip and palate: a systematic review. *European journal of orthodontics.* 2005;27(3):274-285.
5. Wehby GL, Cassell CH. The impact of orofacial clefts on quality of life and healthcare use and costs. *Oral diseases.* 2010;16(1):3-10.
6. van Rooij IA, Ludwig KU, Welzenbach J, et al. Non-Syndromic Cleft Lip with or without Cleft Palate: Genome-Wide Association Study in Europeans Identifies a Suggestive Risk Locus at 16p12.1 and Supports SH3PXD2A as a Clefting Susceptibility Gene. *Genes (Basel).* 2019;10(12).
7. Yu Y, Zuo X, He M, et al. Genome-wide analyses of non-syndromic cleft lip with palate identify 14 novel loci and genetic heterogeneity. *Nat Commun.* 2017;8:14364.
8. Leslie EJ, Carlson JC, Shaffer JR, et al. 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. *Hum Genet.* 2017;136(3):275-286.
9. Butali A, Mossey PA, Adeyemo WL, et al. Genomic analyses in African populations identify novel risk loci for cleft palate. *Hum Mol Genet.* 2019;28(6):1038-1051.
10. Howe LJ, Lee MK, Sharp GC, et al. Investigating the shared genetics of non-syndromic cleft lip/palate and facial morphology. *PLoS Genet.* 2018;14(8):e1007501.
11. Rojas-Martinez A, Reutter H, Chacon-Camacho O, et al. Genetic risk factors for nonsyndromic cleft lip with or without cleft palate in a Mesoamerican population: Evidence for IRF6 and variants at 8q24 and 10q25. *Birth Defects Res A Clin Mol Teratol.* 2010;88(7):535-537.
12. Al Mahdi HB, Edris S, Bahieldin A, et al. Identification of Causative Variants Contributing to Nonsyndromic Orofacial Clefts Using Whole-Exome Sequencing in a Saudi Family. *Genet Test Mol Biomarkers.* 2020;24(11):723-731.
13. Aylward A, Cai Y, Lee A, Blue E, Rabinowitz D, Haddad J, Jr. Using Whole Exome Sequencing to Identify Candidate Genes With Rare Variants In Nonsyndromic Cleft Lip and Palate. *Genet Epidemiol.* 2016;40(5):432-441.
14. Liu H, Busch T, Eliason S, et al. 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 Res.* 2017;109(1):27-37.
15. Bishop MR, Diaz Perez KK, Sun M, et al. Genome-wide Enrichment of De Novo Coding Mutations in Orofacial Cleft Trios. *The American Journal of Human Genetics.* 2020;107(1):124-136.
16. Murray JC, Schutte BC. Cleft palate: players, pathways, and pursuits. *J Clin Invest.* 2004;113(12):1676-1678.
17. Riley BM, Mansilla MA, Ma J, et al. Impaired FGF signaling contributes to cleft lip and palate. *Proceedings of the National Academy of Sciences.* 2007;104(11):4512-4517.
18. Suzuki K, Hu D, Bustos T, et al. Mutations of PVRL1, encoding a cell-cell adhesion molecule/herpesvirus receptor, in cleft lip/palate-ectodermal dysplasia. *Nat Genet.* 2000;25(4):427-430.
19. Yu W, Serrano M, Miguel SS, Ruest LB, Svoboda KK. Cleft lip and palate genetics and application in early embryological development. *Indian J Plast Surg.* 2009;42 Suppl(Suppl):S35-50.
20. Angelo Antiguas, Brian J Paul, Martine Dunnwald. To Stick or Not to Stick: Adhesions in Orofacial Clefts. Review. Biology (Basel). 2022;11(2):153.

21. Butali A, Mossey P, Tiffin N, et al. Multidisciplinary approach to genomics research in Africa: the AfriCRAN model. *Pan Afr Med J.* 2015;21:229.
22. Awotoye W, Mossey P, Hetmanski J, et al. Whole-genome Sequencing Reveals De-novo Mutations Associated with Nonsyndromic Cleft Lip/Palate. In: Research Square; 2021.
22. Pedersen BS, Brown JM, Dashnow H, et al. Effective variant filtering and expected candidate variant yield in studies of rare human disease. *NPJ Genom Med.* 2021;6(1):60.
23. Gonzalez-Pena V, Natarajan S, Xia Y, et al. Accurate genomic variant detection in single cells with primary template-directed amplification. *Proceedings of the National Academy of Sciences.* 2021;118(24):e2024176118.
24. Venselaar H, Te Beek TA, Kuipers RK, Hekkelman ML, Vriend G. Protein structure analysis of mutations causing inheritable diseases. An e-Science approach with life scientist friendly interfaces. *BMC Bioinformatics.* 2010;11:548.
25. Yan F, Dai Y, Iwata J, Zhao Z, Jia P. An integrative, genomic, transcriptomic and network-assisted study to identify genes associated with human cleft lip with or without cleft palate. *BMC Med Genomics.* 2020;13(Suppl 5):39.
26. Szklarczyk D, Gable AL, Nastou KC, et al. The STRING database in 2021: customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Res.* 2021;49(D1):D605-d612.
27. Ji Y, Garland MA, Sun B, et al. Cellular and developmental basis of orofacial clefts. *Birth Defects Res.* 2020;112(19):1558-1587.
28. Pascale Gaudet , Michael S Livstone, Suzanna E Lewis, Paul D Thomas. Phylogenetic-based propagation of functional annotations within the Gene Ontology consortium. Brief Bioinform. 2011 Sep;12(5):449-62.
29. Miyahara M, Nakanishi H, Takahashi K, Satoh-Horikawa K, Tachibana K, Takai Y. Interaction of nectin with afadin is necessary for its clustering at cell-cell contact sites but not for its cis dimerization or trans interaction. J Biol Chem. 2000 Jan 7;275(1):613-8. doi: 10.1074/jbc.275.1.613. PMID: 10617658.
30. Takai, Y., & Nakanishi, H. (2003). Nectin and afadin: novel organizers of intercellular junctions. *Journal of cell science*, *116*(1), 17-27.
31. Cox LL, Cox TC, Moreno Uribe LM, et al. Mutations in the Epithelial Cadherin-p120-Catenin Complex Cause Mendelian Non-Syndromic Cleft Lip with or without Cleft Palate. *Am J Hum Genet.* 2018;102(6):1143-1157.
32. Kievit A, Tessadori F, Douben H, et al. Variants in members of the cadherin-catenin complex, CDH1 and CTNND1, cause blepharocheilodontic syndrome. *Eur J Hum Genet.* 2018;26(2):210-219.
33. Du S, Yang Y, Yi P, et al. A Novel CDH1 Mutation Causing Reduced E-Cadherin Dimerization Is Associated with Nonsyndromic Cleft Lip With or Without Cleft Palate. *Genet Test Mol Biomarkers.* 2019;23(11):759-765.
34. Vijayan V, Ummer R, Weber R, Silva R, Letra A. Association of WNT Pathway Genes With Nonsyndromic Cleft Lip With or Without Cleft Palate. *Cleft Palate Craniofac J.* 2018;55(3):335-341.
35. Mandai K, Nakanishi H, Satoh A, et al. Afadin: A novel actin filament-binding protein with one PDZ domain localized at cadherin-based cell-to-cell adherens junction. *J Cell Biol.* 1997;139(2):517-528.
36. Miyahara M, Nakanishi H, Takahashi K, Satoh-Horikawa K, Tachibana K, Takai Y. Interaction of nectin with afadin is necessary for its clustering at cell-cell contact sites but not for its cis dimerization or trans interaction. *J Biol Chem.* 2000;275(1):613-618.

37. Takahashi K, Nakanishi H, Miyahara M, et al. Nectin/PRR: an immunoglobulin-like cell adhesion molecule recruited to cadherin-based adherens junctions through interaction with Afadin, a PDZ domain-containing protein. *J Cell Biol.* 1999;145(3):539-549.
38. Suzuki K, Hu D, Bustos T, et al. Mutations of PVRL1, encoding a cell-cell adhesion molecule/herpesvirus receptor, in cleft lip/palate-ectodermal dysplasia. *Nature Genetics.* 2000;25(4):427-430.
39. Brancati F, Fortugno P, Bottillo I, et al. Mutations in PVRL4, encoding cell adhesion molecule nectin-4, cause ectodermal dysplasia-syndactyly syndrome. *Am J Hum Genet.* 2010;87(2):265-273.
40. Lough KJ, Spitzer DC, Bergman AJ, Wu JJ, Byrd KM, Williams SE. Disruption of the nectin-afadin complex recapitulates features of the human cleft lip/palate syndrome CLPED1. *Development.* 2020;147(21).

Table 1: List of identified genetic variants and the prediction scores of their pathogenicity. These genes have evidence supporting their roles in craniofacial development.

Chromosome Location (Hg 38)	**MAF**	**Gene**	**HGVS consequence**	**Effect**	**CADD Score**
2:172465556[a]	6.57E-6	*ITGA6*	p.Pro67Leu	missense	31
2:206443783[a]	0	*ADAM23*	p.Met1ext-28	Premature start codon	20.6
6:167827167[a]	0	*AFDN*	p.Lys12Arg	Missense	19.82
6:167898245[a]	3.94E-5	*AFDN*	p.Thr412Asn	missense	20.6
6:167914259[b]	1.30E-3	*AFDN*	p.Arg669Gln	missense	25.8
6:167946861[a]	3.94E-5	*AFDN*	p.Met1123Ile	missense	22.9
6:167947879[a]	7.45E-3	*AFDN*	p.Ser1146Gly	missense	7.95
6:167952162[b]	8.54E-5	*AFDN*	p.Arg1555His	missense	31
6:167962433[b]	2.70E-4	*AFDN*	c.4690-6T>C	Synonymous (Splice variant)	15.75
6:167965850[b]	6.57E-6	*AFDN*	p.Pro1638Ala	missense	23.3
6:167969156[b]	6.57E-6	*AFDN*	p.Ala1717Val	missense	12.58

*Pathogenicity prediction: Non-pathogenic Pathogenic

[a]Variants discovered in African (discovery) cohort; [b]Variants discovered in Brazilian (replication) cohort

MAF: Minor allele frequency (source: https://gnomad.broadinstitute.org/)

Figure Legends

Figure 1A: Pedigree of a Case-parent trio. The male child has nsCL/P while the father and mother are unaffected.

Figure 1B: Data filtration pipeline used to identify the high confidence protein-altering variants that contribute to the risk of nsCL±P.

Figure 2: Protein structural and functional analysis indicating the effect of the amino acid changes. **A**. p.Lys12Arg variant disrupts the Afadin α-helix structure. The mutant residue is bigger and might lead to bumps. **B**. p.Met1123Ile variant result in a smaller size Afadin. Loss of interactions with other molecules is highly likely due to this change. **C**. p.Thr412Asn variant affects the size and hydrophobic interactions of Afadin. The mutation occurs in the Forkhead-associated (FHA) domain which is critical for the formation of complexes. This mutation disrupts the interaction and formation of complexes with other molecules thus abolish the function of the domain. **D**. p.Ser1146Gly variant cause a reduction in the size of the Afadin. The glycine mutant affects the rigidity of Afadin and might lead to loss of molecular interactions. **E**. p.Pro1638Ala variant has a distorted rigidity and may force the backbone in a specific conformation. The disruption of the structure might lead to loss of interactions. **F**. p.Arg669Gln variant has a disrupted charge and the mutation occurred in the Dilute domain of Afadin. These and the change in size might result in loss of interactions. The function of the Dilute domain might also be abolished due to this mutation. **G**. p.Ala1717Val variant have a bigger size protein product. This change in size of the residue might lead to bumps. **H**. p.Arg1555His mutation occurred at a highly conserved location. There was also a change in the protein charge and size. These changes can cause a loss of interactions.

Figure 3: Segregation analysis showing: **A.** p.Lys12Arg and **B.** p.Thr412Asn were inherited by the proband from the unaffected mother. **C:** Segregation analysis showing the compound heterozygous variants p.Met1123Ile and p.Ser1146Gly were inherited by the proband from the father and mother, respectively.

Figure 4: Protein Interaction network showing: **A.** Afadin (MLLT4) interaction with cleft candidates: Nectins (PVRL1-4), CDH1, CTNNA1, CTNND1. **B.** Specific Afadin(MLLT4) interaction with Nectin1 (PVRL1) and Nectin4 (PVRL4) is significantly associated with Cleft lip-palate-ectodermal dysplasia (CLPED) syndrome (FDR = 0.0075).