Association of Nucleotide Variants of GRHL3, IRF6, NAT2, SDC2, BCL3, and PVRL1 Genes with Nonsyndromic Cleft Lip With/Without Cleft Palate in Multigenerational Families: A Retrospective Study

Abstract

Background: Several genes are associated with the etiology of cleft lip and palate (CLP) in different populations. Many nucleotide variants on genes such as GRHL3, IRF6, NAT2, SDC2, BCL3, and PVRL1 were reported in different populations, but not studied in multigenerational cases in the Indian population. **Aim and Objective:** The aim of this study is to evaluate whether nucleotide variants rs41268753, rs861020, rs1041983, rs1042381, rs2965169, and rs10790332 are involved in the etiology of nonsyndromic CLP (NSCLP) in multigenerational Indian families. **Study Design:** Retrospective genetic study. **Materials and Methods:** Based on inclusion and exclusion criteria, 20 multigenerational families with nonsyndromic cleft lip with or without cleft palate (NSCL/P) were selected. Blood samples from both affected and unaffected participants were collected as a source of genomic DNA. Six nucleotide variants on these genes were genotyped to test for the association with NSCL/P. Genotyping was performed with the MassArray method. Genotype distribution was used to calculate the Hardy–Weinberg equilibrium using PLINK, a whole-genome association analysis toolset. The allelic association was compared among cases and controls using Chi-square test as implemented in PLINK. $P \leq 0.05$ indicates statistical differences between groups. **Results:** No significant associations were found between individual single-nucleotide polymorphisms and NSCL/P. The odds ratio was 1.531, 1.198, 0.8082, 1.418, 1, and 0.5929 for polymorphisms rs41268753, rs861020, rs1041983, rs1042381, rs2965169, and rs10790332, respectively. **Conclusion:** Our findings suggest that among the multigenerational families in our population, the high-risk nucleotide variants GRHL3 rs41268753, IRF6 rs861020, NAT2 rs1041983, SDC2 rs1042381, BCL3 rs2965169, and PVRL1 rs10790332 are not associated with increased risk of NSCL/P.

Keywords: *BCL3 and PVRL1, cleft lip and palate, gene, genotyping, GRHL3, IRF6, mass array method, NAT2, nonsyndromic cleft lip/palate, polymorphism, SDC2*

**Praveen Kumar Neela[1,2],
Srinivas Reddy Gosla[1,3],
Akhter Husain[4],
Vasavi Mohan[5],
Sravya Thumoju[5],
B. V. Rajeshwari[5,6]**

[1]GSR Institute of Craniomaxillofacial and Facial Plastic Surgery, Hyderabad, [2]Department of Orthodontics, Kamineni Institute of Dental Sciences, Nalgonda, [3]Department of Cranio Maxillofacial Surgery, AIIMS, Rishikesh, [4]Department of Orthodontics, Yenepoya Dental College, Yenepoya University, Mangalore, [5]Department of Genetics and Molecular Medicine, Vasavi Medical and Research Centre, Hyderabad, [6]Department of OBG, Surabhi Institute of Medical Sciences, Telangana, India

Address for correspondence:
Dr. Praveen Kumar Neela,
Department of Orthodontics,
Kamineni Institute of Dental Sciences, Narketpally,
Nalgonda - 508 254, Telangana, India.
E-mail: praveenneela@yahoo. com

Introduction

Cleft lip and palate (CLP) is among the most common congenital birth anomalies in humans. According to a study conducted by the WHO it was found that, at every 2 min an infant is born with cleft lip/palate in the world.[1] A study conducted by Reddy *et al*. showed that the incidence of clefts in India is around 1:800–1:1000, and three infants are born with some type of cleft every hour.[2] The care of patients with cleft palate remains a cause for concern, which will impose a substantial economic burden on society because of its increasing occurrence and costs of medical care.[3] CLP can be classified into syndromic and nonsyndromic, of which 70% are nonsyndromic and 30% are syndromic.[4] The etiology of CLP is multifactorial including genetic causes, malnutrition, endocrine disorders, infection, trauma, consanguinity, alcohol consumption, and some other environmental causes. About 20% of the CLP showed consanguinity of their parents, while the percentage of familial cases is 3.5% of all the cleft cases.[5] About 600 syndromes are characterized by some form of cleft phenotype.[6]

Genetic research on CLP uses various methods, including association analysis and linkage analysis, to determine the genetic determinants of oral and facial clefts. The results of candidate gene-based association studies, performed in diverse populations, have been mostly inconclusive or conflicting, with only a few candidate loci

Access this article online

Website:
www.contempclindent.org

DOI: 10.4103/ccd.ccd_329_20

Quick Response Code:

being implicated in cleft phenotypes. This inconsistency is due to genetic heterogeneity. These studies discovered multiple candidate genes linked to nonsyndromic CLP (NSCLP) such as IRF6, MSX1, CRISPLD2, ABC4, RARA, transforming growth factor (TGF)α, TGFβ, p63, MYH9, BCL3, MTHFR, TGFB2, SATB2, GRHL3, IRF6, NAT2, SDC2, BCL3 and PVRL1, P63, MSX2, FOXE1, BMP4, PAX7, PVRL1, TGFB3, RARA, RUNX2, BCL3, TGFB1, TBX1, and BCL3.[7-12] Some of the crucial genes responsible for cell adhesion, migration, proliferation, neural crest development, cell-matrix transcription factors during development include IRF6, NAT2, SDC2, BCL3, PVRL1, and GRHL3.

Many of the genomic studies revealed high-risk markers GRHL3 rs41268753, IRF6 rs861020, NAT2 rs1041983, SDC2 rs1042381, BCL3 rs2965169, and PVRL1 rs10790332 are associated with NSCLP in different populations.[13-19] The genes GRHL3, IRF6, NAT2, SDC2, BCL3, and PVRL1, are potentially functional, and the polymorphisms related to these genes were not studied in multigenerational families and also in our population. Therefore, the present study is aimed to evaluate the association of these high risk markers with NSCLP in multigenerational families.

Materials and Methods

The Institutional Review Board of GSR Institute of Craniofacial Surgery approved the research (GSR/IEC/2016/3).

Study population and sample

Multigenerational families with nonsyndromic cleft lip with or without cleft palate (NSCL/P) patients were selected. Inclusion criteria include cleft patients (affected) with any type of phenotypic feature, i.e., unilateral or bilateral cleft lip or cleft palate or both CLP. A proband is an individual with the condition being studied or reported on. Patients with a monogenic syndrome or chromosomal aberrations, associated malformations, and mental retardation were excluded from the study. Majority of the time taken for the study was for the identification of multigenerational families. The control sample comprised unaffected related individuals from these multiplex families. The study sample is selected from GSR Institute of Craniofacial Surgery who fulfilled the criteria. This center was chosen as it is a high volume cleft center where cleft patients from different states come for treatment. Based on the medical records of 13 years for positive family history for clefts, an invitation was sent to the families to participate in the study. People who agreed to participate voluntarily were recruited for the study. The age of the participants participated in the study ranged from 3 years to 67 years at the time of blood sample collection. Based on the power calculation for family-based association studies,[20] twenty multiplex families are selected randomly. The generations of families were two to four. Out of twenty families, one family had the history of cleft in all the four generations. These include one family with 5 affected, two families with 4 affected, five families with 3 affected, and 12 families with 2 affected with CLP. Four multigenerational families reported consanguinity. Informed consent was obtained from the patients', patient's parents in case of minor children and also the healthy controls from these multiplex families who participated in the study.

DNA isolation

About 3–5 ml of venous blood was taken in the ethylenediaminetetraacetic acid tubes. Genomic DNA was extracted from the blood lymphocytes using the salting-out method.[21] The DNA isolation was done at Vasavi Medical Research Centre, Hyderabad. An ultraviolet spectrometer was used to calculate the average 260/280 nm ratio to assess the purity and concentration of DNA. The ratio of absorbance readings at the two wavelengths should be between 1.8 and 2.0 (i.e., A260/A280 = 1.8–2.0). Later, the DNA was sent for single-nucleotide polymorphism (SNP) genotyping of the polymorphisms. The characteristics of the selected polymorphisms are represented in Table 1.

MassARRAY analysis

Agena Bio MassARRAY (Agena Bioscience, Inc., San Diego, CA, USA) platform using iPLEX Gold technology was utilized for the SNP genotyping at Genes2Me, a subsidiary of Imperial Life Sciences, Delhi. This system is a nonfluorescent, highly accurate detection platform utilizing matrix-assisted laser desorption/ionization-time of flight mass spectrometry. The assay was designed using proprietary Agena software (Assay Design Suite 2.0). The assay design was used to design primers. Table 2 shows the assay and the primer sequences. Follow the correct workflow according to the MassArray protocol, and finally run the sample through the analyzer. Agena's SpectroTyper 4.0 software (San Diego, CA, USA) was used, which automatically generates reports that identify the SNP alleles (homozygous or heterozygous). The data obtained from the analyzer software are sent for the statistical analysis.

Statistical analysis

The SNP allele data of the affected and controls obtained from the MassArray system was subjected to the statistical analysis. PLINK software (Version 1.09) (USA) was used for this study.[22] It is a free, open-source whole-genome association toolset, designed to perform a range of basic to large-scale analyses in a computationally efficient manner. Genotype distribution was used to calculate the Hardy–Weinberg equilibrium (HWE) using the same PLINK. Statistical comparisons between the affected and unaffected were carried out using PLINK software. Odds ratio (OR) and 95% confidence intervals were provided. Allelic Association was analyzed using the Chi-square test.

Table 1: Characteristics of nucleotide variants selected for the study

Gene	Polymorphism	Type of alteration	Alleles	Ancestral allele	Global MAF
GRHL3	rs41268753	Missense variant	C/T	C	0.01
IRF6	rs861020	Intron variant	A/G	G	0.18
NAT2	rs1041983	Synonymous variant	C/T	C	0.4
SDC2	rs1042381	Missense variant	T/A	T	0.48
BCL3	rs2965169	Missense variant	A/C	C	0.47
PVRL1	rs10790332	Intron variant	C/T	C	0.33

Source of information: https://www.ncbi.nlh.nih.gov/snp/, http://ensembl.org/Homo_sapiens. MAF: Minor allele frequency; A: Adenine; C: Cytosine; T: Thymine; G: Guanine

Table 2: Primer sequences

Assay	Nucleotide variant/SNP	Forward primer	Reverse primer
Assay 1	rs1041983	1-CAGACCACAATGTTAGGAGG	2-CCATGCCAGTGCTGTATTTG
Assay 2	rs1042381	2-AGGATGTAGAGAGTCCAGAG	1-TTCCACTTTTGGAGCAGCAC
Assay 3	rs10790332	2-TCCTACTCAGCTGACAGTTC	1-AGGCTAGACACATATGCCAG
Assay 4	rs861020	1-AAAACAGGCAGAGATGGAAC	2-TATTGCAGCCTGTGGTTGTG
Assay 5	rs41268753	1-CAATGAGACGACCTACCTTC	2-CTGGAGAAGTGCACATTGGG
Assay 6	rs2965169	1-ACCTAGGATTTTCCGAGCAC	2-AGCGCTCCACTGATTGTGTC

#rs: Reference ID of the SNP. SNP: Single-nucleotide polymorphism

For nominal association, the statistical significance level is set to $\alpha = 0.05$. $P \leq 0.05$ indicates statistical differences between groups.

Results

Six SNPs present on six important genes were genotyped in twenty multigenerational families. All the polymorphisms followed HWE. In the allele association analysis [Table 3], we observed that none of the variants showed any association with NSCLP ($P > 0.05$). The OR also was <2. An OR is a statistic that quantifies the strength of the association between two events such as exposure and outcome. OR >1 means greater odds of association with the exposure and outcome. OR = 1 means there is no association between exposure and outcome and if OR <1 means there is a lower odds of association between the exposure and outcome.

Discussion

The etiology of the CLP is complex and multifactorial. They involve the influence of genetic, environmental, causes. The importance of genetic studies on the etiology of CLP is ever increasing due to the advent of different technologies. With advancements in the field of molecular biology, our envelope of research has grown. The identification of genetic polymorphisms in our population would be invaluable in understanding the developmental mechanisms involved in causing the disease. Data from animal models (e.x., zebrafish), in which clefts arise either spontaneously or as a result of mutagenesis experiment, combined with an analysis of how expression patterns correlate with gene function and examining the effects of gene-environment interactions in nonsyndromic clefts. Importantly, they also contribute to our knowledge of normal craniofacial development and the molecular pathogenesis of CL/P. Several recent studies have also provided strong evidence that syndromic forms having Mendelian patterns of inheritance may provide insights into the genetic etiology of nonsyndromic types of clefting.

Phenomenal advances in gene identification studies on CLP identified numerous new potential candidate genes associated with this condition. Associations between various nucleotide or polymorphic markers and risk of clefts have been identified in different populations. Genetic studies were conducted on diverse populations both on syndromic and nonsyndromic cases. Only IRF6 variants showed consistency in the etiology of CLP in different populations. Studies were conducted on case-parent trios, isolated clefts, and familial cases. To date, this is one of the few CLP genetic studies conducted on multigenerational families in our population. When we take the percentage of familial cases, it comes to a meager 3.5% of the total cleft cases. GSR Inst of Craniofacial surgery is a high-volume cleft center. The participants were selected from this center, as patients from different states are offered treatment with the generous help of various national and international agencies. The nonsyndromic and familial cases were identified after a thorough medical history and examination of the patients. All syndromic cases excluded during the data collection stage.

The genes selected (GRHL3, IRF6, NAT2, SDC2, BCL3, and PVRL1) are having an essential role in protein-coding, neural crest development, embryogenesis, cell adhesion, and differentiation. Six polymorphisms on these genes reported as significant markers in the etiology of CLP in different populations were selected for the study. The aim of this study was to investigate whether high-risk variants of GRHL3 rs41268753, IRF6 rs861020, NAT2 rs1041983,

Table 3: Association between polymorphisms and nonsyndromic cleft lip and palate

CHR	SNP	BP	A1	F_A	F_U	A2	χ^2	*P*	OR (CI -95%)
1	rs41268753	20	T	0.02	0.01316	C	0.1207	0.7283	1.531
1	rs861020	30	A	0.37	0.3289	G	0.3189	0.5722	1.198
8	rs1041983	1	C	0.46	0.5132	T	0.4887	0.4845	0.8082
8	rs1042381	2	A	0.16	0.1184	T	0.6126	0.4338	1.418
11	rs10790332	3	T	0.29	0.4079	C	2.671	0.1022	0.5929
19	rs2965169	18	C	0.5	0.5	A	0	1	1

Chi-square test. CHR: Chromosome number; SNP: Single-nucleotide polymorphism; A1: Major allele (wild allele); F_A: Minor allele frequency affected; F_ U: Minor allele frequency unaffected; A2: Minor allele (mutant); OR: Odds ratio; CI: Confidence interval. The *P*<0.05 is significant

SDC2 rs1042381, BCL3 rs2965169, and PVRL1 rs10790332 genes are associated with increased risk of NSCL/P in multigenerational families.

GRHL3, grainyhead-like transcription factor 3 is a critical transcription factor associated with neural tube defects. Missense variant rs41268753 of GRHL3 reported being associated with cleft palate in the Europeans.[13] However, in a machine learning model of genetic risk of infants the OR was 2.043, and the risk allele was seen in 0.5%–1.5% in the Chinese population.[23]

The findings of two GWAS studies and animal studies on zebrafish reported that rs41268753 as being pathogenic rather than merely being in linkage disequilibrium with a pathogenic variant.[19,24] These findings support rs41268753 as being pathogenic rather than merely being in linkage disequilibrium. In a Chinese study, the authors reported that rs41268753 did not report the increased risk of NSCL/P. However, they advised examining its role in other ethnic backgrounds.[25] However, in our study too, we could not see an increased risk for NSCL/P.

IRF6 is a vital gene which has confirmed to be associated with CLP than any other gene in many ethnicities. One of the nucleotide variant rs861020 reported being significantly associated with NSCL/P in Chinese patients (*P* = 0.001).[26] A GWAS study reported that in European ancestry, rs861020 of IRF6 is significantly associated with NSCLP.[27] The results of the study revealed that the high-risk nucleotide variant rs861020 of IRF6 is not associated with our multigenerational families.

In a study on the Chinese population, NAT2 rs1041983 polymorphism reported to be associated with nonsyndromic orofacial clefts for the first time.[15] In a study, the authors confirmed the accuracy of the tag SNP approach for rs1041983 in Brazilians of European & Middle Eastern descent, but not in Brazilians of African ancestry, sub-Saharan Mozambicans, and Guarani Amerindians.[28] However, in the present population, the association with NSCLP is insignificant. In a study, SDC2 rs1042381 reported to be significantly associated with CLP in Hispanic families.[16] In a Korean cleft study, SNP rs2965169 of BCL3 appeared to be associated with increased risk with the excess maternal transmission.[17]

Many studies reported with the possibility and conclusion that PVRL1 is also a candidate gene for the etiology of CLP. Some studies on Thai people, Caucasians, and Venezuelans reported with some high-risk markers. In a study, rs10790332 was reported to be significantly associated with the CLP.[19] However, in the present study, none of the high-risk nucleotide variants was associated with the etiology of CLP in the multigenerational families in our population.

The variation or inconsistent results for the different populations or ethnicities in the etiology of CLP could be due to multifactorial etiology, epigenetic causes, gene-gene interactions.

Strengths and limitations

Only twenty multi-generational families were recruited to this study, but these families are rare and a small proportion of CLP cases.

Future work

Future research should focus on the functional study of nucleotide variants and epigenetics.

Conclusion

Our study suggests that no significant associations found between GRHL3 rs41268753, IRF6 rs861020, NAT2 rs1041983, SDC2 rs1042381, BCL3 rs2965169, and PVRL1 rs10790332 variants with NSCL/P in multigenerational families. It shows that the marker identified as a risk, in one particular ethnicity/population or one family may or may not be a significant marker again in the same or different ethnicity or the same or different family. The variation or inconsistent results could be due to multifactorial etiology, epigenetics, gene-gene interactions. However, it should be our endeavor to continue to research on cleft lip palate etiology. Additional studies in other populations, along with increased samples, are required to understand the basis of the etiology of CLP.

Conflicts of interest

There are no conflicts of interest.

References

1. Mossey P. Global strategies to reduce the healthcare burden of craniofacial anomalies. Br Dent J 2003;195:613.
2. Reddy SG, Reddy RR, Bronkhorst EM, Prasad R, Ettema AM, Sailer HF, *et al*. Original article incidence of cleft lip and palate in the state of Andhra Pradesh, South India. Indian J Plast Surg 2010;43:184-9.
3. Wehby GL, Cassell CH. The impact of orofacial clefts on quality of life and healthcare use and costs. Oral Dis 2010;16:3-10.
4. Dixon MJ, Marazita ML, Beaty TH, Murray JC. Cleft lip and palate: Understanding genetic and environmental influences. Nat Rev Genet 2011;12:167-78.
5. Neela PK, Reddy SG, Husain A, Mohan V. Association of cleft lip and/or palate in people born to consanguineous parents: A 13-year retrospective study from a very high-volume cleft center. J Cleft Lip Palate Craniofac Anomal 2019;6:33-7.
6. Mossey P, Little J. Addressing the challenges of cleft lip and palate research in India. Indian J Plast Surg 2009;42 Suppl: S9-S18.
7. Leslie EJ, Taub MA, Liu H, Steinberg KM, Koboldt DC, Zhang Q, *et al*. Identification of functional variants for cleft lip with or without cleft palate in or near PAX7, FGFR2, and NOG by targeted sequencing of GWAS loci. Am J Hum Genet 2015;96:397-411.
8. Mohamad Shah NS, Salahshourifar I, Sulong S, Wan Sulaiman WA, Halim AS. Discovery of candidate genes for nonsyndromic cleft lip palate through genome-wide linkage analysis of large extended families in the Malay population. BMC Genet 2016;17:39.
9. Mehrotra D. Genomic expression in non syndromic cleft lip and palate patients: A review. J Oral Biol Craniofac Res 2015;5:86-91.
10. Vieira AR. Unraveling human cleft lip and palate research. J Dent Res 2008;87:119-25.
11. Sull JW, Liang KY, Hetmanski JB, Fallin MD, Ingersoll RG, Park J, *et al*. Differential parental transmission of markers in RUNX2 among cleft case-parent trios from four populations. Genet Epidemiol 2008;32:505-12.
12. Funato N, Nakamura M. Identification of shared and unique gene families associated with oral clefts. Int J Oral Sci 2017;9:104-9.
13. Mangold E, Böhmer AC, Ishorst N, Hoebel AK, Gültepe P, Schuenke H, *et al*. Sequencing the GRHL3 coding region reveals rare truncating mutations and a common susceptibility variant for nonsyndromic cleft palate. Am J Hum Genet 2016;98:755-62.
14. Ludwig KU, Mangold E, Herms S, Nowak S, Reutter H, Paul A, *et al*. Genome-wide meta-analyses of nonsyndromic cleft lip with or without cleft palate identify six new risk loci. Nat Genet 2012;44:968-71.
15. Song T, Wu D, Wang Y, Li H, Yin N, Zhao Z. Association of NAT1 and NAT2 genes with nonsyndromic cleft lip and palate. Mol Med Rep 2013;8:211-6.
16. Chiquet BT, Hashmi SS, Henry R, Burt A, Mulliken JB, Stal S, *et al*. Genomic screening identifies novel linkages and provides further evidence for a role of MYH9 in nonsyndromic cleft lip and palate. Eur J Hum Genet 2009;17:195-204.
17. Park BY, Sull JW, Park JY, Jee SH, Beaty TH. Differential parental transmission of markers in BCL3 among Korean cleft case-parent trios. J Prev Med Public Health 2009;42:1-4.
18. Widdershoven JC, Bowser M, Sheridan MB, McDonald-McGinn DM, Zackai EH, Solot CB, *et al*. A candidate gene approach to identify modifiers of the palatal phenotype in 22q11.2 deletion syndrome patients. Int J Pediatr Otorhinolaryngol 2013;77:123-7.
19. Leslie EJ, Liu H, Carlson JC, Shaffer JR, Feingold E, Wehby G, *et al*. A genome-wide association study of nonsyndromic cleft palate identifies an etiologic missense variant in GRHL3. Am J Hum Genet 2016;98:744-54.
20. Peng B, Li B, Han Y, Amos CI. Power analysis for case-control association studies of samples with known family histories. Hum Genet 2010;127:699-704.
21. Miller SA, Dykes DD, Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Res 1988;16:1215.
22. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, *et al*. PLINK: A tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 2007;81:559-75.
23. Zhang SJ, Meng P, Zhang J, Jia P, Lin J, Wang X, *et al*. Machine learning models for genetic risk assessment of infants with non-syndromic orofacial cleft. Genomics Proteomics Bioinformatics 2018;16:354-64.
24. Duncan KM, Mukherjee K, Cornell RA, Liao EC. Zebrafish models of orofacial clefts. Dev Dyn 2017;246:897-914.
25. He M, Bian Z. Lack of association between missense variants in GRHL3 (rs2486668 and rs545809) and susceptibility to non-syndromic orofacial clefts in a Han Chinese population. PLoS One 2016;11:e0159940.
26. Song T, Wu D, Wang Y, Li H, Yin N, Zhao Z. SNPs and interaction analyses of IRF6, MSX1 and PAX9 genes in patients with non-syndromic cleft lip with or without palate. Molecular Med Rep 2013;8:1228-34.
27. Beaty TH, Murray JC, Marazita ML, Munger RG, Ruczinski I, Hetmanski JB, *et al*. A genome-wide association study of cleft lip with and without cleft palate identifies risk variants near MAFB and ABCA4. Nat Genet 2010;42:525-9.
28. Suarez-Kurtz G, Sortica VA, Vargens DD, Bruxel EM, Petzl-Erler ML, Tsuneto LT, *et al*. Impact of population diversity on the prediction of 7-SNP NAT2 phenotypes using the tagSNP rs1495741 or paired SNPs. Pharmacogenet Genomics 2012;22:305-9.