

## RESEARCH ARTICLE

## Analysis of Single Nucleotide Polymorphisms on Locus 13q33.1-34 in Multigenerational Families of Cleft Lip Palate using MassArray

Praveen Kumar Neela<sup>1,2\*</sup>, Srinivas Reddy Gosla<sup>1,3</sup>, Akhter Husain<sup>4</sup>, Vasavi Mohan<sup>5</sup>,  
Sravya Thumuju<sup>5</sup>, Rajeshwari<sup>5</sup>

<sup>1</sup>GSR Institute of Craniomaxillofacial and Facial Plastic Surgery, Vinay Nagar Main Rd, Vinay Nagar Colony, Hyderabad, Telangana 500059, India

<sup>2</sup>Department of Orthodontics, Kamineni Institute of Dental Sciences, Sreepuram, Dist. Nalgonda, Telangana 508254, India

<sup>3</sup>Department of Cranio Maxillofacial Surgery, All India Institute of Medical Science (AIIMS), Virbhadra Road Shivaji Nagar, Rishikesh, Uttarakhand 249203, India

<sup>4</sup>Yenepoya Dental College, Yenepoya University, University Road, Deralakatte, Mangalore, Karnataka 575018, India

<sup>5</sup>Department of Genetics and Molecular Medicine, Vasavi Medical and Research Centre, Khairatabad, Hyderabad, Telangana 500007, India

\*Corresponding author. E-mail: praveenneela@yahoo.com

Received date: Apr 17, 2020; Revised date: Aug 22, 2020; Accepted date: Aug 31, 2020

### Abstract

**BACKGROUND:** Cleft lip palate is a common congenital anomaly with multifactorial etiology. Many high-risk markers at different loci were reported to be involved in its etiology. Advanced genetic research led to the discovery of evidence of a new linkage on 13q33.1-34 region at marker rs1830756 in two multigenerational Indian families. However, no further study was reported to confirm or validate this linkage in other families. Hence, the present study was designed.

**METHODS:** Twenty multigenerational families affected by non-syndromic cleft lip palate were selected for the study. Polymorphisms, rs1830756, rs1323672, rs1935135 of *FAM155A* gene; rs1961495, rs953386, rs1411040 of *COL4A1* gene; and rs726449, rs984300 of *MYO16* gene were selected. Genomic DNA was isolated and sent for genetic analysis by single nucleotide polymorphism (SNP) genotyping using the MassArray method. Statistical analysis

of the genomic data was done by PLINK. Bonferroni correction was applied and haplotype analysis was done using Haploview software.

**RESULTS:** Polymorphisms followed the Hardy Weinberg Equilibrium. In the allelic association, all the polymorphisms analysed showed no statistical significance. Hence, there was no significant difference in the allelic frequencies between non-syndromic cleft lip palate patients and healthy controls. The odds ratio was not more than 1.6 for all the SNPs. Haplotype analysis showed that haplotypes were not significantly higher in non-syndromic cleft patients than in control subjects.

**CONCLUSION:** There is no association between SNPs analysed in the locus 13q33.1-34 with cleft lip palate.

**KEYWORDS:** cleft lip palate, chromosome, polymorphism

*Indones Biomed J. 2021; 13(1): 27-33*

### Introduction

Clefts in the oro-facial region is an important congenital anomaly affecting humanity. An infant is born with a cleft lip or palate somewhere on the planet every 2 minutes according to a WHO study.(1) Global surveys have shown that the frequency of cleft lip and palate varies significantly from one country to another. Cleft lip palate (CLP) is lowest

in Africans (1:2500), while the North American Indians and East Asians have the highest prevalence (1:500). According to a study the incidence of clefts in India is around 1:800 to 1:1000 and 3 infants are born with some type of cleft every hour.(2) Cleft lip palate can be syndromic or non-syndromic. Approximately 70 % of the cleft lip and palate cases are non-syndromic and occur as isolated cases, whereas the remaining 30% of clefts are syndromic and are associated with some other anomalies. Our knowledge of

the etiology and pathogenesis of non-syndromic variants yet remains relatively poor. Etiology of CLP is multifactorial, which includes genetic causes, malnutrition, endocrine disorders, infection, trauma, consanguinity. Roughly 20% of the CLP showed consanguinity of their parents, while the percentage of familial cases is 3.5% of all the cleft cases.(3) About 600 syndromes are characterized by some form of cleft phenotype.(4)

The genomic research uses both 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 being implicated in cleft phenotypes. These studies discovered multiple candidate genes linked to non-syndromic cleft lip palate (NSCLP) such as interferon regulatory factor (IRF)-6, Msh homeobox (MSX)-1, cysteine-rich secretory protein LCCL domain-containing 2 (CRISPLD2), ATP-Binding Cassette member 4 (ABC4), retinoic acid receptor alpha (RARA), transforming growth factor (TGF)- $\alpha$ , TGF- $\beta$ , transformation-related protein 63 (P63), myosin heavy chain 9 (MYH9), B cell lymphoma 3 protein (BCL3), methylene tetra hydro-folate reductase (MTHFR), TGF- $\beta$ 2, special AT-rich sequence-binding protein 2 (SATB2), MSX-2, forkhead box protein E1 (FOXE1), bone morphogenetic protein 4 (BMP4), paired box 7 (PAX7), polio virus receptor-related 1 (PVRL1), TGF- $\beta$ 3, Runt-related transcription factor 2 (RUNX2), TGF- $\beta$ 1, T-box transcription factor (TBX1).(5-10)

Many genome wide association studies (GWAS) were done, which revealed significant evidence of linkage at different loci on different chromosomes. Some of them include 7p21.3, 14q32.32, 9p23, 3q26.33, 10q25.1, 18q21.1, 6p12.3, 4q28.1. In a GWAS study conducted in 2006 on two Indian multigenerational families, the authors reported evidence of linkage at 13q33.1-34, and most significant evidence at marker rs1830756.(11) Hence, analysing high risk SNPs at this locus 13q33.1-34 in many families is recommended. Though many GWAS were done previously, they were done on other non-Indian populations. We intended to select a susceptibility region that is identified in our population. The locus 13q33.1-34 is one among them, because this locus revealed significant evidence of linkage over 15 years ago. However, no further study was reported in this susceptible region for its involvement in familial cases of NSCLP. Therefore, the rationale for the present study was to know the association of high-risk markers at locus 13q33.1-34 in other Indian multiplex families of NSCLP. The study hypothesized that SNPs selected from

this high-risk region for NSCLP are involved in the etiology of NSCLP in multiplex families of the population.

## Methods

### Ethical Approval and Sample Selection

The Institutional Review Board of GSR Institute of Craniofacial Surgery approved the research with Approval number: GSR-IEC-1/2016. All study procedures were performed in accordance with the principles of the Declaration of Helsinki. Medical records of GSR Institute of Craniofacial Surgery, India, a high-volume cleft center was searched. Patients with monogenic syndrome or chromosomal aberrations, associated malformations, and mental retardation were excluded from the study. Based on the power calculation for family-based association studies, twenty multiplex families were selected. All these cleft families were selected from this high volume cleft centre where patients from different parts of the country come for treatment. These 20 affected families include one family with five affected, two families with four affected, five families with three affected and 12 families with two affected cleft patients. Thus, a total of 88 subjects were included with 50 affected patients and 38 unaffected subjects from these 20 multiplex families were taken for this study. All the selected families were from same ethnic background, *i.e.*, Dravidian. Informed consent was obtained from the patients', patient's parents in case of minor children and also the healthy subjects from these multiplex families who participated in the study.

### DNA Isolation

Three milliliters of venous blood was taken in the EDTA tubes. Genomic DNA was extracted from the blood lymphocytes using the salting-out.(12) A UV 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.*,  $A_{260}/A_{280} = 1.8 - 2.0$ ). Later, the DNA was sent for SNP genotyping of the polymorphisms. The characteristics of the selected polymorphisms are given in Table 1.

### SNP Genotyping and Data Analysis

Agena Bio MassARRAY (Agena Bioscience, Inc., San Diego, CA, USA) platform using iPLEX Gold technology was utilised for the SNP genotyping. This system is a non-fluorescent, highly accurate detection platform utilising

**Table 1. List of SNPs analysed in the locus 13q33.1-34.(22-24)**

SNP	Candidate Gene	Normal sequence	Ancestral Allele
rs1830756	<i>FAM155A</i>	A/C-FWD	C
rs1323672	<i>FAM155A</i>	C/T-FWD	C
rs1935135	<i>FAM155A</i>	C/T-FWD	T
rs726449	<i>MYO16</i>	A/G-REV	C
rs984300	<i>MYO16</i>	C/T-REV	G
rs1961495	<i>COL4A1</i>	C/T-FWD	C
rs953386	<i>COL4A1</i>	A/G-REV	T
rs1411040	<i>COL4A1</i>	A/G-REV	T

rs: reference sequence; FWD: forward; REV: reverse.

Matrix-Assisted Laser Desorption/Ionisation-Time Of Flight (MALDI-TOF) mass spectrometry. Using proprietary Agena software, with Assay Design Suite 2.0, (Agena Bioscience, Inc., San Diego, CA, USA) the assay was designed for primers. Table 2 shows the primers and reverse sequence for all the SNPs along with the Assay numbers. Following the correct workflow, according to the MassArray protocol, the samples are run through the analyser. 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 analyser software is sent for statistical analysis.

### Statistical Analysis

The SNP allele data of the affected patients and controls obtained from the MassArray system was subjected to statistical analysis. PLINK software Version 1.09 (Harvard University, Cambridge, MA, USA) was used for this study. (13) 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 and 95% confidence intervals were provided. Allelic Association was

analysed using the chi-square test. For nominal association, the statistical significance level is set to  $\alpha=0.05$ . Bonferroni correction was applied, and  $p \leq 0.05$  indicates statistical differences between groups. Haplotype analysis was done using Haploview software (version 4.2; Broad Institute, Cambridge, MA, USA)

## Results

Eight SNPs present in the locus 13q33.1-34 were genotyped in 20 multiplex families. All the polymorphisms follow Hardy–Weinberg equilibrium. In the allele association analysis (Table 3), rs1961495 of *COL4A1* showed a  $p$ -value of 0.02493. When Bonferroni correction was applied, the SNP rs1961495 also showed no significance. However, for other remaining polymorphisms, the  $p > 0.05$ . Hence, there was no significant difference in the allelic frequencies between NSCLP patients and healthy controls. The odds ratio (OR) also was less than 2 in all the polymorphisms. The linkage disequilibrium (LD) plot consisting of eight single nucleotide polymorphisms (SNPs) generated using the genotype data from the sample was shown in the Figure 1. The names of the examined SNPs were enclosed in boxes and were indicated at the top of the figure. Block 1 corresponds to SNPs rs1323672 and rs1411040 and

**Table 2. Primers and reverse sequence for all the SNPs.**

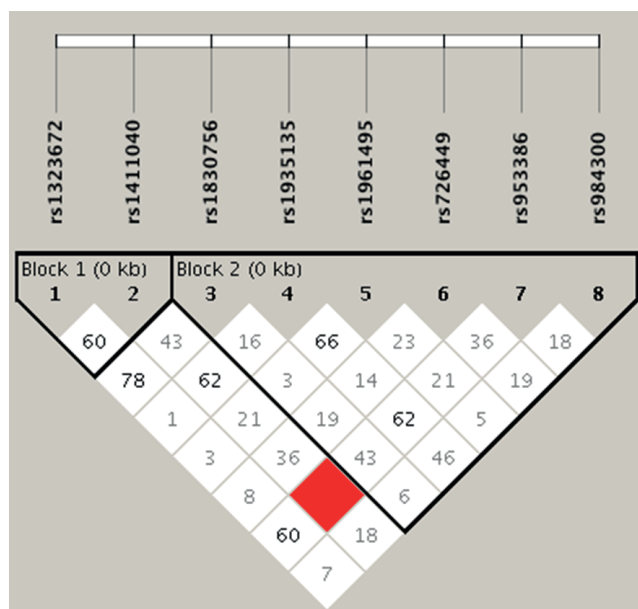
Assay	SNP	Primer Sequence	Reverse Primer Sequence
Assay 1	rs1830756	1-TGAAGCCTCCATCTAAATAC	2-GAATTGCATTGTTCTTTAGG
Assay 2	rs1323672	1-TCCCAGAGAGAACTGAAAG	2-GCTATTGCTGACTATAAAAAC
Assay 3	rs1935135	2-TAGAGCCAATCCAGATGAGG	1-GGAGCCAAGAAGAAGGTATG
Assay 4	rs726449	1-GTGCTTTCAGAACATGGTTG	2-AGTCCATGCTTAAAAACAGGC
Assay 5	rs984300	1-ACAGGCTGTCAAAGACTGAG	2-TCATTGCATCCGGTCAAAGC
Assay 6	rs1961495	2-AGGAAAACCAGATGGCATTG	1-TTCATCCTCCCTGTCTCTTC
Assay 7	rs953386	2-TCAGCAACCAAGAATCTGCC	1-GATGGCATGGTAAAAGCTCC
Assay 8	rs1411040	1-ACGTGGCAGGGACTAGACTT	2-GGCCATTGCATTTTCAGCAC

**Table 3. Allelic association for SNPs on chromosome 13.**

CHR	SNP	MAF	A1	F_A	F_U	A2	CHISQ	P	BONF	OR
13	rs1323672	0.18	C	0.22	0.18	T	0.33	0.55	1	1.24
13	rs1411040	<b>0.26</b>	T	0.28	0.26	C	0.06	0.80	1	1.08
13	rs1830756	<b>0.14</b>	C	0.15	0.21	A	1.09	0.29	1	0.66
13	rs1935135	<b>0.30</b>	T	0.29	0.21	C	1.43	0.23	1	1.53
13	rs1961495	<b>0.19</b>	C	0.12	0.25	T	5.02	0.02	0.19	0.40
13	rs726449	<b>0.12</b>	C	0.25	0.22	T	0.16	0.68	1	1.15
13	rs953386	<b>0.27</b>	T	0.28	0.26	C	0.06	0.80	1	1.08
13	rs984300	<b>0.49</b>	G	0.43	0.39	A	0.22	0.63	1	1.15

CHR: chromosome number; SNP: single nucleotide polymorphism; MAF: minor allele frequency (South Asian); A1: major allele (wild allele); F\_A: minor allele frequency affected; F\_U: minor allele frequency unaffected; A2: minor allele (mutant); CHISQ: chi-square, P: *p*-value (<0.05 is significant); BONF: Bonferroni correction; OR: odds ratio.

Block 2 corresponds to the remaining of the eight SNPs. The numbers in the diamond indicate percentage of LD between a given pair of SNPs. Linkage disequilibrium was displayed as pairwise *D'* values (*D'*=standardized coefficient of LD). White diamonds indicated absence of LD  $r^2=0$  or *D'*=0, where  $r^2$  was the pairwise correlation coefficient. Red diamond without a number corresponds to *D'* values of 1.0. The comparison of haplotype frequencies between cleft patients and controls was shown Table 4. This haplotype-based case-control analysis had also not revealed any significant finding wherein the haplotypes were not significantly higher in non-syndromic cleft patients than in control subjects.



**Figure 1. The LD plot consisting of eight single nucleotide polymorphisms (SNPs) generated using the genotype data from the sample.**

## Discussion

The importance of genetic studies on the etiology of cleft lip and palate is ever increasing. With advancements in the field of molecular biology, our envelope of research has grown. Identification of genetic polymorphisms in our population would be invaluable in understanding the developmental mechanisms involved in causing the disease. Data from animal models, 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 have proven themselves as powerful tools for identifying candidate genes for complex traits, like non-syndromic clefts. Importantly, they also contribute to our knowledge of normal craniofacial development and the molecular pathogenesis of CL/P, taking into account that facial development in mice mirrors human craniofacial development. Several recent studies have also provided strong evidence that syndromic forms having Mendelian patterns of inheritance may provide insights into the genetic etiology of non-syndromic types of clefting.

Advances in the genomic arena led to whole-genome studies, linkage studies, targeted studies. These led to the identification of many high-risk markers, candidate genes in the etiology of CLP. Associations between polymorphic markers in RUNX2, BMP4, TGF- $\beta$ 3 PAX7, NTN1, IRF-6, PTHFR, GHR, *etc.* and risk of clefts were identified in different populations. Various genetic studies conducted on diverse populations both on syndromic and non-syndromic cases revealed significant linkages on multiple chromosomes. Studies were conducted on case-parent trios, isolated clefts and familial cases. However, studies on

**Table 4. Comparison of haplotype frequencies between cleft patients and controls.**

Haplotype	Freq.	Cleft Case, Control Ratios	Chi Square	p-value
<b>Haplotype Associations</b>				
<b>Block 1</b>				
CT	0.545	0.522, 0.574	0.463	0.49
CC	0.251	0.258, 0.242	0.057	0.81
TT	0.183	0.198, 0.163	0.348	0.55
TC	0.022	0.022, 0.021	0.002	0.96
<b>Block 2</b>				
CTCCCA	0.172	0.177, 0.166	0.041	0.84
CTCCCG	0.133	0.140, 0.124	0.097	0.75
CTCCTG	0.078	0.096, 0.054	1.019	0.31
CCCTCA	0.062	0.083, 0.034	1.797	0.18
CTCCTA	0.055	0.061, 0.047	0.168	0.68
CCCCCA	0.054	0.065, 0.040	0.538	0.46
CTCCA	0.054	0.025, 0.091	3.633	0.05
CCCCTA	0.049	0.053, 0.044	0.064	0.8
ACCCCG	0.048	0.041, 0.058	0.274	0.60
ATCTCA	0.036	0.042, 0.027	0.289	0.59
ATTCA	0.027	0.010, 0.050	2.563	0.10
CCCCCG	0.026	0.021, 0.032	0.207	0.64
CTCTCG	0.023	0.030, 0.014	0.472	0.49
CTTTCA	0.022	0.016, 0.029	0.353	0.55
CTCTTG	0.020	0.025, 0.013	0.348	0.55
ATCCCG	0.019	0.013, 0.027	0.456	0.49
CTTCTG	0.019	0.012, 0.028	0.609	0.43
ATCCTA	0.018	0.011, 0.026	0.581	0.44
CCTTCG	0.016	0.027, 0.001	1.814	0.17
CTCTCA	0.013	0.006, 0.023	0.974	0.32
CTTCCG	0.012	0.009, 0.016	0.178	0.67
ATCCCA	0.011	0.015, 0.007	0.27	0.60
ATTCTG	0.009	0.007, 0.010	0.044	0.83
CTCTTA	0.007	0.001, 0.015	1.099	0.29
ATTTG	0.007	0.010, 0.004	0.200	0.65
CTTTG	0.006	0.000, 0.013	1.332	0.24
CTTCTA	0.005	0.003, 0.006	0.083	0.77

multigenerational families or multiplex families are few. In India, research on familial non-syndromic cases and the trio of case-parents is minimal. When we take the percentage of familial cases, it comes to a meagre 3.5 % of the total cleft cases.(3) This study sample was taken from a large cleft centre in India as people come from different states are treated. The non-syndromic and familial cases were identified after a thorough medical history and examination of the patients. All syndromic cases were excluded during the data collection stage.

In an article published on the Indian Collaboration on Craniofacial Anomalies (INDIANCRAN) initiative, one of the recommendations was to study genetic polymorphisms and their disease susceptibility and association in various

genetically identifiable groups in India.(4) Following the recommendations of INDIACRAN, we intended to select a susceptibility region. The locus 13q33.1-34 is one among them because this locus revealed significant evidence of linkage over 15 years ago. However, no further study was reported in this susceptible region for its involvement in familial cases of NSCLP. The SNPs evaluated are from candidate genes *FAM155A*, *MYO16* and *COL4A1*. *COL4A1* gene provides instructions for making one component of type IV collagen, *FAM155A* in protein-coding, and *MYO16* in protein synthesis. The genetic studies in the short term will help in the evaluation and improve non-invasive methods of screening for the disorder in at-risk family members. In the medium term helps in developing and establishing



a knowledge-based approach to the management and treatment of individuals experiencing with CL/P.

In the present study, 20 multiplex families were selected. The distribution of variants in all the 20 families included 24 subjects are with unilateral cleft, 2 subjects with bilateral cleft lip and 24 subjects with both cleft lip and palate. There were no subjects with cleft palate only, without involving cleft lip in the selected families. The affected and unaffected subjects were 50 and 38, respectively. Eight SNPs in the locus, were analysed for their association to cleft lip and palate in multiplex families. MassArray was selected for genotyping. The advantages include high throughput, high multiplex capability, flexibility and adaptability, and the high level of accuracy. The marker rs1961495 of gene *COL4A1* showed a *p*-value of 0.02493. Though it is statistically significant, on applying Bonferroni correction, this marker also became statistically not significant. The *p*-value of the remaining 7 polymorphisms, rs1830756, rs1323672, rs1935135 of *FAM155A* gene; rs953386, rs1411040 of *COL4A1* gene; and rs726449, rs984300 of *MYO16* gene was more than 0.05. The findings of the present study are different when compared to findings of a GWAS study done previously (11), wherein the most significant evidence was found for marker rs1830756 of candidate gene *FAM155A*. This shows that the marker identified as significant risk, in two large multigenerational families may not be a significant marker in the same population for different families. This could be due to the complex and multifactorial etiology of cleft lip palate, epigenetic causes or gene-gene interactions.

In a previous study on Fine-Mapping of 18q21.1 locus, *MYO5B* SNP rs183559995 was revealed to be a high-risk marker with an odds ratio of over 18.(14) *MYO5B* belongs to the same group of like *MYO16*. However, in the present study, *MYO16* has not shown any significance. In one of the candidate gene replication studies on *BMP4*, located at 14q22-q23 in humans, in one of the earlier studies conducted previously in the Chinese population, polymorphisms on this gene were reported to be clinically significant in the development and progression of NSCLP.(15) However, in a study of this locus on NSCLP, there was a significantly increased risk for the Chinese population. However, it had a protective effect on the Brazilian population.(16) According to literature reports, in the population of southeastern Iran, the *BMP4* rs17563 variant has a protective effect on the occurrence of NSCLP.(17) In a separate case-control study on NSCLP, one study reported an increased risk of NSCLP in the Indian population.(18)

In another candidate gene study, it was reported that in a case-parent trio study on four populations,

*i.e.*, Taiwan, Singapore, Korea, and Maryland, RUNX2 rs2819861 polymorphism influenced the risk of NSCLP.(9) PAX genes are transcription factors that play crucial roles during embryonic development in the form of regulation of differentiation programs and organogenesis.(19,20) In a replication study in Poland on its population, the authors confirmed that PAX7 is a strong candidate gene for NSCLP/P as they revealed the high-risk nature of rs2743218 of PAX7.(21)

This shows that the marker identified as a risk, in one particular population or one large multigenerational family may not be a significant marker in the same population or different population or other families of the same population. The variation or inconsistent results could be due to multifactorial etiology, epigenetic causes, gene-gene interactions. The implications of the study suggest the multifactorial nature of the disease. The limitation of the study: All the multiplex families selected for the study should have undergone GWAS. It does not put light on the role of epigenetics, gene-gene interaction.

## Conclusion

Our research showed that polymorphisms rs1830756, rs1323672, rs1935135 of candidate *FAM155A* gene; rs1961495, rs953386, rs1411040 of *COL4A1* gene and rs726449, rs984300 of *MYO16* gene in the locus 13q33.1-34 are not associated with an increased risk of NSCLP in the multiplex families.

## Authors Contribution

PKN, SRG and AH were involved in conceiving and planning the research. PKN, SRG, AH and RBV performed the data acquisition/collection. PKN, VM and ST calculated the experimental data and performed the analysis. VM and ST interpreted the results. PKN, SRG, AH, VM and RBV drafted the manuscript PKN. RBV designed the Figure and/or Table Design. All authors took parts in giving critical revision of the manuscript.

## References

1. Mossey P. Global strategies to reduce the healthcare burden of craniofacial anomalies. *Br Dent J.* 2003; 195(10): 613. doi: 10.1038/sj.bdj.4810738.

2. Reddy SG, Reddy RR, Bronkhorst EM, Prasad R, Ettema AM, Sailer HF, *et al.* Incidence of cleft lip and palate in the state of Andhra Pradesh, South India. *Indian J Plast Surg.* 2010; 43: 184-9.
3. 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.
4. Mossey P, Little J. Addressing the challenges of cleft lip and palate research in India. *Indian J Plast Surg.* 2009; 42(Suppl): S9-S18.
5. 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.
6. 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 Genetics.* 2016; 17: 39. doi: 10.1186/s12863-016-0345-x.
7. Mehrotra D. Genomic expression in non-syndromic cleft lip and palate patients: A review. *J Oral Biol Craniofac Res.* 2015; 5: 86-91.
8. Vieira AR. Unraveling human cleft lip and palate research. *J Dent Res.* 2008; 87: 119-25.
9. 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.
10. Funato N, Nakamura M. Identification of shared and unique gene families associated with oral clefts. *Int J Oral Sci.* 2017; 9: 104-9.
11. Radhakrishna U, Ratnamala U, Gaines M, Beiraghi S, Hutchings D, Golla J, *et al.* Genomewide scan for nonsyndromic cleft lip and palate in multigenerational Indian families reveals significant evidence of linkage at 13q33.1-34. *Am J Hum Genet.* 2006; 79: 580-5.
12. Miller SA, Dykes DD, Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res.* 1988; 16: 1215. doi: 10.1093/nar/16.3.1215.
13. 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.
14. Mitra AK, Stessman HAF, Schaefer RJ, Wang W, Myers CL, Van Ness, *et al.* Fine-mapping of 18q21.1 locus identifies single nucleotide polymorphisms associated with non-syndromic cleft lip with or without cleft palate. *Front Genet.* 2016; 7: 88. doi: 10.3389/fgene.2016.00088.
15. Lin JY, Chen YJ, Huang YL, Tang GP, Zhang L, Deng B, *et al.* Association of bone morphogenetic protein 4 gene polymorphisms with non-syndromic cleft lip with or without cleft palate in Chinese children. *DNA Cell Biol.* 2008; 27: 601-5.
16. Hao J, Gao R, Wu W, Hua L, Chen Y, Li F, *et al.* Association between BMP4 gene polymorphisms and cleft lip with or without cleft palate in a population from South China. *Arch Oral Biol.* 2018; 93: 95-9.
17. Rafighdoost H, Hashemi M, Danesh H, Bizhani F, Bahari G, Taheri M. Association of single nucleotide polymorphisms in AXIN2, BMP4, and IRF6 with non-syndromic cleft lip with or without cleft palate in a sample of the southeast Iranian population. *J Appl Oral Sci.* 2017; 25: 650-6.
18. Savitha S, Sharma S M, Veena S, Rekha R. Single nucleotide polymorphism of bone morphogenetic protein 4 gene: A risk factor of non-syndromic cleft lip with or without palate. *Indian J Plast Surg.* 2015; 48: 159-64.
19. Lang D, Powell SK, Plummer RS, Young KP, Ruggeri, BA. PAX genes: Roles in development, pathophysiology, and cancer. *Biochem Pharmacol.* 2007; 73: 1-14. doi: 10.1016/j.bcp.2006.06.024.
20. Wang Q, Fang WH, Krupinski J, Kumar S, Slevin M, Kumar P. Pax genes in embryogenesis and oncogenesis. *J Cell Mol Med.* 2008; 12: 2281-94.
21. Gaczkowska A, Biedziak B, Budner M, Zadurska M, Lasota A, Hozyasz KK. PAX7 nucleotide variants and the risk of non-syndromic orofacial clefts in the Polish population. *Oral Diseases.* 2019; 25: 1608-18.