OFFICIAL JOURNAL

ORIGINAL ARTICLE

Genetic Variations Associated with Non-Syndromic Cleft Lip and Palate in Malays with Whole Exome Sequencing: Case Report and Gene Review

Nurul Syazana Mohamad Shah[1], Sarina Sulong[2], Wan Azman Wan Sulaiman[1], Ahmad Sukari Halim[3]

[1]Reconstructive Science Unit, School of Medical Sciences, Health Campus, Universiti Sains Malaysia, 16150 Kubang Kerian, Kelantan, Malaysia
[2]Human Genome Centre, School of Medical Sciences, Health Campus, Universiti Sains Malaysia, 16150 Kubang Kerian, Kelantan, Malaysia
[3]Director Office, Hospital Universiti Sains Malaysia, Health Campus, 16150 Kubang Kerian, Kelantan, Malaysia

Abstract

Introduction: Exome sequencing technology which is part of Next Generation Sequencing (NGS) is known for detection of various disease mutations through commercially available platforms. Less reports in identifying genetic variation in non-syndromic cleft lip with or without cleft palate (NSCL/P) in Malaysia had embarked for discovery of susceptible genes to fill in the gaps with the healthcare delivery for a better treatment and management to the patients and family. **Methods**: Whole exome sequencing was carried out on two Malay NSCLP patients. Blood samples were withdrawn and intact DNA was extracted, fragmented, purified and hybridized using exome sequencing capture and sequenced with Agilent 2100 Bioanalyzer platform. Bioinformatic analyses were done and reviewed with GenBank and PubMed database. Variants were filtered based upon a high impact variant. **Results**: We have identified single nucleotide polymorphisms in 2 genes (*PDE4DIP* and *PDE11A*) and InDels frameshift mutations in 4 genes (*PDE4DIP*, *LTBP4*, *MMP12* and *MMP28*). Our preliminary study presents the successful application of whole exome sequencing to elucidate the genetic basis of NSCLP in Malays. **Conclusion**: Mutations that have been identified would shed more light on the susceptible genes to non-syndromic clefts and further investigation shall be carried out to confirm.

Keywords: Non-syndromic, cleft lip and palate, whole exome sequencing, next generation sequencing, gene

Introduction

Exome sequencing which is part of Next Generation Sequencing (NGS) technology has been widely used in detecting disease mutations. The exome represent a small scale, approximately 1-2% protein coding sequence of human genome, but a majority of known disease-causing variants could be found in the exome regions (Buermans and Den, 2014). Recently, the use of exome sequencing in identifying causing genes or gene mutation associated with cleft lip with or without cleft palate (CL/P) formation is increasing It was thought to be an efficient method to identify possible disease-causing mutations. CL/P is one of the most common human birth defects reported in Western countries and the second most common birth deformities that arises about 1 per 500-1000 live births through ethnic and geographic differences (Muhamad, 2012; Murray, 2002; Shaw et al., 1991). It was remarkable that Asian and Native North American descent population has highest birth prevalence with NSCL/P by 2 in 1000 births compared to the other

Received: 4 May 2020**; accepted revised manuscript:** 30 June 2020 **Published online:** 07 July 2020
***Corresponding author:** Dr. Nurul Syazana Mohamad Shah, Reconstructive Science Unit, School of Medical Sciences, Health Campus, Universiti Sains Malaysia, 16150 Kubang Kerian, Kelantan, Malaysia
Tel: +6097676894 **Email**: syazanashah@usm.my

populations (Cooper et al., 2006, Vieira, 2008). It is clear that CL/P incidence is varied due to multiple factors through genetic and environmental influences.

Currently, there have been 52 genome-wide associations study reported regarding NSCL/P but very few putative functional variants were identified (Buniello et al., 2019; Mukhopadhyay et al., 2020). Several studies have been carried out on mice and human genome using exome sequencing approach. A previous study conducted on mice used targeted sequencing for the exome region of human genome to sequence mouse mutants associated with cleft (Fairfield et al., 2011). The vast technology of exome sequencing allows identification of human disease genes through the clinical phenotype and causative mutations in mice as both traits has > 90% similarities towards each other (Fairfield et al., 2011; Mouse Genome Informatics, 2015). Twenty-five candidate genes associated with NSCL/P was compared to variants found in 6500 exomes through 1000 Genomes Project and NHLBI Exome Sequencing Project (Leslie and Murray, 2013). Rare variant of muscle segment homeobox 1 (*MSX1*) and FGF family in NSCL/P has been detected, whereas the majority of affected individual with rare variants were inherited from unaffected parents or siblings. However, there was a case that certain rare variants such as GLI family zinc finger 2 (*GLI2*), Jagged 2 (*JAG2*) and Sprouty RTK signaling antagonist 2 (*SPRY2*) were not detected using exome sequencing, but it still serve as candidate genes through biological or statistical data (Jugessur and Murray, 2005). Thus, it is best to describe that NSCL/P discovery could be affected by multiple factors that caused variation in findings.

In this study, we had applied whole exome sequencing analysis on two Malay non-syndromic cleft lip and palate (NSCLP) patients, with one had bilateral CLP and another one unilateral CLP. In addition, one patient had a family history of non-syndromic cleft and another one had no background of cleft in the family. However, limitation existed on having limited sample size which was not including the parents too for testing. It was due to the financial constraint since the test was done overseas and the cost per sample was high at that time. Nevertheless, this is an attempt to assure if there is any differences of causative mutations and genes in different type of cleft and different family history in those two cases.

Methods

Experimental design

As mentioned before, due to its limitation on sample size, this study was focused on identifying genetic variations in affected patients only. Any similarities or differences of outcomes between the cases was further reviewed.

Since NSCL/P is known to be heterogenous throughout different population and geography, this study was aimed to set as a benchmark of genes causing this deformity in Malaysia population for a better management in future.

Subjects

Two 3-year-old boys, one boy with bilateral cleft lip and palate (BCLP) and another one boy had unilateral cleft lip and palate (UCLP). The BCLP boy (Patient 1) had no family history of cleft while the UCLP boy (Patient 2) had a father with BCLP. Both patients were born normal vaginal delivery at term with no other anomalies. Both patients claimed by parents to have no consanguinity in their family. This study was approved by the Research Ethics Committee (Human) of the Universiti Sains Malaysia, USMKK/PPP/JEPeM [258.3.(3)] dated February 2013 and written informed consent was obtained from the subjects.

Genomic DNA Samples

Intact genomic DNA and free from RNA contamination were selected. DNA concentration must be ≥ 37.5 ng/µl and their purity reading fell between 1.8 - 2.0. The genomic DNA quantity required for exome sequencing was ≥ 5 µg.

Agilent SureSelect System Exome Capture and Sequencing

Exome capture was carried out using Agilent SureSelect Human All Exon v4 (51 Mb) kit. First, the genomic DNA sample was fragmented with the size for the library fragments were ranged at 150-200 bp. Following adapter-ligated templates process, fragments with insert size about 176 bp were excised and amplified by ligation-mediated PCR (LM-PCR), purified and hybridized to the SureSelect Biotinylated RNA Library (BAITS) for enrichment.

 | ISSN (Online): 2716-649X

Hybridized fragments were bound to the streptavidin beads and captured LM-PCR products were subjected to Agilent 2100 Bioanalyzer to estimate the magnitude of enrichment. Each captured library was then loaded on Illumina HiSeq 2000 platform and high-throughput sequencing was performed. Raw image files were processed by Illumina base calling Software 1.7 for base calling with default parameters and the sequences of each individual were generated as 90/100 bp pair-end reads.

Bioinformatic Analyses

Quality Control

Raw reads that consists of sequence of adapter, high content of unknown bases and low quality reads are called "dirty reads". These dirty reads were removed before data analysis using 3 filtering steps. First, the adapter reads followed by low-quality reads were removed from the raw data. Adapter read is a read with the adapter bases meanwhile, base quality ≤ 5 represents low quality bases. If more than half of bases in a read were low-quality bases, it was categorized as low-quality reads. Finally, reads in which ≥ 10 % unknown bases were removed. After filtering, the remaining reads were called "clean reads" and used for downstream bioinformatics analysis.

Single Nucleotide Polymorphisms (SNPs) Analysis

SNPs were identified and AnnoDB was used to do annotation and classification. Two databases were used: Beijing Genomic Institute-Gene and Phenotype (BGI-GaP) database for annotation of diseases related gene/mutation and Population Variation Frequency Database (PVFD) for mutation frequency annotation.

Genes were extracted depend upon their impact variants comprised of high, moderate, modifier or low impact, Sorting Intolerant From Tolerant (SIFT) score and PolyPhen2 score. SIFT predicts whether an amino acid substitution could affect protein function based on scoring. SIFT score was defined as damaging (D) and tolerated (T), with score < 0.05 it was predicted as "D (damaging)". PolyPhen2 score predicted the "probably damaging" (D) if the score > 0.909 while others were "possibly damaging (P) or "benign" (B).

Insertions and Deletions (InDels) Analysis

Pair-end reads for gap alignment was used to detect the InDels (small Insertion/Deletion). GATK and SAMtools software were used to detect InDels. After InDels were identified, AnnoDB was used to do annotation and classification. Genes were extracted depends upon the function (Frameshift variant, 3′ UTR variant, 5′ UTR variant, disruptive inframe deletion, downstream gene variant, intergenic region, intron variant, non coding exon variant, splice acceptor variant, splice donor variant, splice region variant, stop lost, upstream gene variant) and impact (high, moderate, modifier, low) condition. The selected genes were then reviewed using GenBank and PubMed to compare and relate the function of each gene with any craniofacial deformities associated with clefting.

Results

Identification of Novel Mutations in NSCLP Patients using Exome Sequencing

Quality Control

Quality controls removed the adapter reads, low-quality reads and unknown bases in reads. The final remaining reads is called "clean reads". This step is important to obtain high quality data prior to bioinformatics analysis. Table 1 showed the final quality control data for the samples. The raw and clean data indicate the number of reads before and after filtering removal steps. Most of the removed reads were due to low-quality reads and the remaining clean data was ≈89%. Five Gb clean data was generated and aligned to the human reference sequence (GRCh37/hg19; 2,861,327,131 non-N bases) using Burrows-Wheeler Aligner (BWA). This provided an efficient program that aligns short nucleotide sequences relatively against a long reference producing accurate and fast results with low error rates.

 | ISSN (Online): 2716-649X

Table 1. Data statistics showed raw and clean data, before and after the quality control process.

Cases	Patient 1		Patient 2	
	Raw data	Clean data	Raw data	Clean data
Number of reads	76724732	68900176	75293848	67231840
Data size	6905225880	6201015840	6776446320	6050865600
Discard reads related to N (Unknown)	1637890		1593586	
Discard reads related to low quality	6122258		6401658	
Discard reads related to adapter	64408		66764	
Clean data	89.80 %		89.29 %	

SNPs Calling

By using human reference genome (GRCh37, hg19) as a comparison, a total of 78,804 and 79,341 single nucleotide variants were identified in Patient 1 with BCLP and Patient 2 with UCLP respectively. A total of 1,094 and 1,081 novel SNPs were detected in both patients respectively. It was found that Patient 1 and Patient 2 had 1,743 and 1,869 SNPs mapped to intergenic regions, 20,205 and 20,387 SNPs in coding exonic region, 49,072 and 49,182 SNPs in coding intronic region, 1,338 and 1,422 SNPs in 5'UTR, and 3,011 and 2,984 SNPs in 3'UTR respectively. Among the SNPs in coding regions, ≈ 10,000 synonymous substitution and ≈ 9,000 nonsynonymous changes were identified.

A total of 65536 genes were identified under various parameters. Based on 141 genes detected under high impact variants, five genes were found exclusive for Patient 1 (BCLP) with SIFT score. Meanwhile, under the same parameters, 9 out of 142 genes were exclusively detected in Patient 2 (UCLP) and eight genes were similarly identified in both patients regardless type of cleft and family history. Based on PolyPhen score, only one gene, *C2orf70* was detected in Patient 1 which had achieved > 0.909 (0.954) score stated as having "probably damaging" condition. Several genes were selected for review and none of the genes were found associated with craniofacial deformities previously, but taking into consideration that novel gene as a risk factor for clefting was possibly elucidated. Three gene of interests were reviewed: *PDE4DIP* at chromosome 1, *PDE11A* at chromosome 2 and *LTBP4* at chromosome 19 (Table 2).

 | ISSN (Online): 2716-649X

Table 2. List of genes identified from SNP analysis with high impact condition.

Gene	Chromosome	Function	SIFT	PolyPhen
PDE4DIP	1	Stopgain, Het	1(T)	-
**C2orf70*	2	Startloss, Het	0(D)	0.954(D)
**PDE11A*	2	Stopgain, Het	0.01(D)	-
HIBCH	2	Startloss, Hom	0.01(D)	0.04,0.018(B,B)
#*NFU1*	2	Startloss, Het	0.01(D)	0.019,0.008(B,B)
ATP6V1B1	2	Startloss, Het	0(D)	0.405,0.599(B,P)
#*C3orf79*	3	Startloss, Het	0(D)	0.011(B)
#*TMPRSS11A*	4	Startloss, Het	0.05(D)	0.001,0.001(B,B)
**C5orf20*	5	Stopgain, Hom	0(D)	-
**IFNE*	9	Stopgain, Het	0.03(D)	-
**ERCC6L2*	9	Startloss, Het	0(D)	0.0(B)
OR51I1	11	Stopgain, Het	0(D)	-
OR4C16	11	Stopgain, Het	0.02(D)	-
OR8K1	11	Startloss, Hom	0(D)	0.006(B)
OR5AR1	11	Stopgain, Hom	0.02(D)	-
#*OR4D10*	11	Stopgain, Het	0.03(D)	-
#*CCDC77*	12	Startloss, Het	0.01(D)	0.025(B)
VDR	12	Startloss, Hom	0(D)	0.275,0.848(B,P)
#*COPZ1*	12	Startloss, Het	0(D)	0.0(B)
#*ING1*	13	splice donor variant, Het	0.01(D)	0.054(B)
#*SLFN12L*	17	Startloss, Hom	0.03(D)	-
#*KRT37*	17	Stopgain, Het	0(D)	-

SIFT score indicates D; damaging, T; tolerated
PolyPhen score indicates B; benign, P; possibly damaging, D; probably damaging
Hom; Homozygous, Het; Heterozygous
*Genes found only in Patient 1
#Genes found only in Patient 2

InDels Calling

Insertions and Deletions (InDels) is insertion or deletion of bases in the DNA. Total InDels of 9,236 and 484 novel InDels was found in Patient 1 compared with human genome reference (hg19). Meanwhile, 9,466 total InDels and 426 novel InDels were detected in Patient 2. It was observed that both patients had 313 and 294 InDels mapped to intergenic regions, 673 and 676 InDels in coding exonic region, 7,375 and 7,124 InDels in coding intronic region, 143 and 174 InDels in 5'UTR, and 479 and 456 InDels in 3'UTR respectively. Of the total InDels, 170 and 190 frameshift mutations, and 247 and 241 non-frameshift mutations were identified in both patients respectively.

Based on high impact variants and frameshift mutations, fourteen genes were exclusively detected in Patient 1 and 16 genes were found exclusive in Patient 2. Several genes were selected and reviewed using GenBank and PubMed to determine possible association with orofacial cleft. In extension, the genes were also reviewed based on the previous finding by Sarah et al. (2015) . Several genes extracted from the lists were as follow: *PDE4DIP*, *LTBP4*, MMP12 and MMP28 (Table 3).

Table 3. List of possible genes to be associated with orofacial cleft. Genes with true positive mutations were highlighted in bold as reviewed based on Online Gene Prioritization Tools (Sarah et al., 2015).

Chromosome	Genes
	Patient 1 and Patient 2
1	**PLA2G2F,* ***PDE4DIP****, ACTN2, TTLL10, HRNR*
2	*FAM228B, CD207, ARMC9, #TNFAIP6*
3	*SLC38A3, SEMA3B, ZNF717, *GPR156, *COL6A5, #WWTR1*
4	*ZNF595, MAML3, DCHS2*
5	*SCAMP1, *ZNF474, SMAD5, SRA1, TIGD6, CYFIP2, NOP16*
6	*PRIM2, AK9, PBOV1, #ECT2L, #MYCT1*
7	*LFNG, ZNF3, ZAN, KCP, SSPO, ATG9B*
8	*PRKDC, *ADCK5, #PEBP4,*
9	*FAM166B, OR1B1, #PDCD1LG2, #GLT6D1*
10	*PTCHD3, *RTKN2, ZNF518A, ATRNL1*
11	*SPON1, CREB3L1, MS4A14, B3GNT6, CWC15, MMP12, DIXDC1, EI24, #NCR3LG1*
12	*WNK1, EMG1, CELA1, OR6C76, *DEPDC4, #DUSP16, *ADAMTS20*
13	*#OXGR1*
14	*OR4L1, #LTB4R2*
15	*#SYNM*
16	*NPRL3, CNOT1, CTU2, #ZNF720*
17	*PIK3R6, RAI1, SARM1, MMP28, KRT24, MAP3K14, *EFCAB13, PPP1R9B, #ABCA10*
18	**TGIF1*
19	*UHRF1, OR7G3, *ZNF599, KMT2B,* ***LTBP4****, *ZNF480, *GP6, *ZNF211, ZNF274, #DMKN, #ZNF880*
21	*DNAJC28, COL18A1, PCNT*
22	*GAS2L1, MAPK8IP2*

*Genes found only in Patient 1

#Genes found only in Patient 2

Discussion

SNPs for PDE4DIP and PDE11A

As far as we are aware, this was the first study applying whole exome sequencing to elucidate the genetics of NSCL/P in Malaysia particularly among the Malays. The use of exome sequencing was beneficial as it is feasible to sequence large amounts of DNA and focus only on exons that make up for proteins. Differ from whole genome sequencing that sequence the entire genomes which is far challenging, whole exome sequencing merely focus on protein coding sequences. It allows variations in protein-coding region of any genes instead of only in selected genes. Previously, we had successfully identified novel genes in large extended families with NSCL/P using microarray platform (Shah et al., 2016). However, various genes were detected using this platform as each family showed different contributing genes to cleft pathogenesis. However, in this preliminary study, we found mostly similar genes in both patients despite different cleft type and family background of cleft they had. There are 6 different type of clefts based on classification but none can consider perfect to classify the cleft type precisely. Pathogenesis would be different as they had different type of deformity either unilateral or bilateral cleft, cleft lip alone or with cleft palate, or cleft palate alone. On contrary, a previous study conducted had identified rare damaging variants in different family with NSCL/P (Basha et al., 2018). Plus, another one study carried out in determining candidate genes on non-syndromic cleft palate only (NSCPO) had found several novel rare risk variants in the family trait (Hoebel et al., 2017). Therefore, all these findings depicted that cleft type or family history background could not be concluded as the cause of gene variation.

SNP and InDels analysis had identified phosphodiesterase 4D Interacting protein gene (*PDE4DIP*) which is located at chromosome 1 in

both patients. Similar finding had been found in NSCLP Honduran families that detected true positive mutation associated with clefting in *PDE4DIP* when using Toppgene and Endeavor as reference (Sarah et al., 2015). PDE4DIP protein also known as myomegalin acts as an anchor phosphodiesterase that is localized in the Golgi/centrosome region of the cell (Li et al., 2015). Several studies have reported the function of *PDE4DIP* in various diseases such as lung squamous cancer and asthma susceptibility (Li et al., 2015; DeWan et al., 2012). Although evidence for *PDE4DIP*-associated cleft is not convincing, the occurrence as nonsynonymous SNV and InDels frameshift mutation have given an understanding of disease risk and a causing gene of *PDE4DIP* to NSCLP formation.

Nonsynonymous SNV had identified phosphodiesterase 11A (*PDE11A*) in Patient 1, which is mapped at 2q31-35 region. It has a role as a second messenger in a wide variety of signal transduction pathways and involves in normal human tissues of various organs (D'Andrea et al., 2005). High expression of *PDE11A* increased susceptibility to prostate cancer as it has been detected in patients with prostate cancer (Faucz et al., 2011). Recently, none has reported the function of *PDE11A* in association with human development or craniofacial formation.

Insertion/Deletions (InDels) Frameshift Mutation in LTBP4

Latent TGF-β Binding Protein (*LTBP*) consists of four members; *LTBP1-4* whereas *LTBP1* and *LTBP3* bind all three isoforms of *TGF-β*, meanwhile *LTBP4* binds only *TGF-β1* (Saharinen and Keski-Oja, 2000). *LTBP* acts as a molecular modulator in directing latent *TGF-β* and propeptide to fibrillin microfibrils in extracellular matrix (ECM) (Urban et al., 2009; Rifkin, 2005). A previous *in vitro* study has indicated *LTBP4* function is an association between *LTBP* and *TGF-β* activity. They found that cells derived from *LTBP4* mutant mice have decreased level of active extracellular *TGF-β* (Koli et al., 2004). *TGF-β* is an important growth factor during normal embryonic development. Altered *TGF-β* is known causing syndromic and nonsyndromic cleft palate (Iwata et al., 2012). Therefore, it was clear that LTBP4 mutation modulates the normal TGF-β activity and contribute to pathogenesis of cleft.

InDels analysis detected frameshift mutation of *LTBP4* at chromosome 19 in both NSCLP patients. Similarly, Sarah et al. (2015) had identified insertion/deletions in *LTBP4* in 5 NSCLP individuals of the Honduran families that underwent whole exome sequencing (Sarah et al., 2015). Surprisingly, another 21 unaffected family members sequenced for the *LTBP4* indels showed positive mutations too. It was concluded that these mutations were unlikely to cause disease but it may be a contributing gene if NSCLP has a polygenic inheritance (Sarah et al., 2015). In addition, a separate study on patients with craniofacial deformities and from different ethnic groups; Palestinian, Mexican and Hispanic have been found positive for *LTBP4* mutations (Urban et al., 2009). We could not conclude *LTBP4* mutation in our population since this is a small scale preliminary analysis, limited to the number of affected patients and we did not perform analysis on the other affected or unaffected family members. But, since NSCLP is known to be heterogenous due to variation differences, it would shed light for future direction in confirming *LTBP4* mutation among the NSCLP.

InDels Frameshift Mutation in MMP12 and MMP28

Frameshift mutations of *MMP12* at chromosome 11 and *MMP28* at chromosome 17 were detected from both NSCLP patients. Exome analysis in NSCLP among the Honduran families have detected a gene, *MMP13* of the same family (Sarah et al., 2015). However, they concluded that these mutated gene was unlikely a disease causing since they are known single nucleotide polymorphisms (Sarah et al., 2015). Matrix metalloproteinase (MMP) is a proteolytic enzyme that play a role for degradation of ECM components that allow cell migration and differentiation, cell-cell interactions and tissue resorption (Letra et al., 2007; Nagase and Woessner, 1999). The ECM breakdown is also essential for embryonic development, morphogenesis and reproduction (Nagase and Woessner, 1999). Evidence of linkage and association was identified among the NSCLP for *MMP25* mapped at chromosome 16p13 (Blanton et al., 2004). Nevertheless, *MMP3* and *MMP13* on chromosomal region of 11q22.3, *MMP25* at loci 16p13.3 region, *MMP2* at loci 16q13-q21 and *MMP9* at loci 20q11.2-q13.1 were confirmed for evidence of linkage and association in NSCLP as well as in gene expression analysis from human and animal embryonic tissue studies (Lace et al.,

 | ISSN (Online): 2716-649X

2011). None has reported the association of *MMP12* to CL/P so far. But, the crucial role of *MMP12* in mediating the synthesis and breakdown of ECM components as well as its expression in developing bones during human fetal development shall be concerned as it may be related to orofacial development (Yang et al., 2007; Kerkelä et al., 2001).

MMP28 is located at the 17q11.2 region and its functional role in embryogenesis and tumor development was significant. High expression of *MMP28* has been detected in cancerous diseases such as colon carcinoma, ovarian carcinoma, pancreatic adenocarcinoma, lung carcinoma and prostate carcinoma (Marchenko and Strongin, 2001). *MMP28* overexpression has also been identified in chondrosarcoma cells that caused cell morphology change, increased adhesion and decreased migration in ECM (Rodgers et al., 2009). Despite no reports of the functional role of *MMP12* and *MMP28* in NSCLP formation, significant finding of *MMP*s in relation to human early development was elucidated. Further analysis to larger groups and family members of NSCLP shall be carried to shed light on the mutation associated to cleft formation.

Conclusion

Exome sequencing analysis through NGS platform revealed variations in protein-coding regions for orofacial cleft formation. Several nonsynonymous variants and InDels frameshift mutations were identified in NSCLP patients including *PDE4DIP*, *PDE11A*, *LTBP4*, *MMP12* and *MMP28*. Although *LTBP4* was reported to unlikely associated with clefting and *PDE4DIP* was positive to orofacial clefts in other population, we believed that heterogeneity exists throughout different population, thus we cannot rule out the possibility as a causal gene in Malay population. While there is no clear consensus of which method is the best, whole exome sequencing has its advantages and disadvantages. The limitations in this study were exome sequencing did not sequence DNA variations outside the exons and sample size was small. Further study of these genes in larger sampling including family members within local population shall be carried out in future to obtain significant findings that may have an impact on healthcare delivery which will aid disease diagnosis in future. It is also would help in genetic understanding for precautionary measures in affected family members, psychological management and psychosocial effect to the patients and parents.

Acknowledgements

We thank the voluntary patients for their full co-operation given. This work was supported by Research University (RU) Grant: 1001/PPSP/812083 and student was funded by a scholarship SLAB/SLAI KPM/USM.

Conflict of Interests

The authors declare there are no competing interests.

References

Basha M, Demeer B, Revencu N, Helaers R, Theys S, Saba SB, et al. Whole exome sequencing identifies mutations in 10% of patients with familial non-syndromic cleft lip and/or palate in genes mutated in well-known syndromes. *Journal of medical genetics* 2018; jmedgenet-2017-105110.

Blanton SH, Bertin T, Serna ME, Stal S, Mulliken JB, Hecht JT. Association of chromosomal regions 3p21. 2, 10p13, and 16p13. 3 with nonsyndromic cleft lip and palate. *American Journal of Medical Genetics Part A* 2004; 125(1): 23-7.

Buermans H, Den Dunnen J. Next generation sequencing technology: advances and applications. Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease 2014; 1842(10): 1932-41.

Buniello, A., MacArthur, J.A.L., Cerezo, M., Harris, L.W., Hayhurst, J., Malangone, C., McMahon, A., Morales, J., Mountjoy, E., Sollis, E. and Suveges, D., 2019. The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic acids research*, *47*(D1), pp.D1005-D1012.

Cooper ME, Ratay JS, Marazita ML. Asian oral-facial cleft birth prevalence. *The Cleft palate-craniofacial journal* 2006; 43(5): 580-9.

D'Andrea MR, Qiu Y, Haynes-Johnson D, Bhattacharjee S, Kraft P, Lundeen S. Expression of PDE11A in normal and malignant human tissues. *Journal of Histochemistry & Cytochemistry* 2005; 53(7): 895-903.

DeWan AT, Egan KB, Hellenbrand K, Sorrentino K, Pizzoferrato N, Walsh KM, et al. Whole-exome sequencing of a pedigree segregating asthma. *BMC medical genetics* 2012; 13(1): 95.

Fairfield H, Gilbert GJ, Barter M, Corrigan RR, Curtain M, Ding Y, et al. Mutation discovery in mice by whole exome sequencing. *Genome biology* 2011; 12(9): R86.

Faucz FR, Horvath A, Rothenbuhler A, Almeida MQ, Libé R, Raffin-Sanson M-L, et al. Phosphodiesterase 11A (PDE11A) genetic variants may increase susceptibility to prostatic cancer. *The Journal of Clinical Endocrinology & Metabolism* 2011; 96(1): E135-E40.

Hoebel A, Drichel D, van de Vorst M, Böhmer A, Sivalingam S, Ishorst N, et al. Candidate genes for nonsyndromic cleft palate detected by exome sequencing. *Journal of dental research* 2017; 96(11): 1314-21.

Iwata JI, Hacia JG, Suzuki A, Sanchez-Lara PA, Urata M, Chai Y. Modulation of noncanonical TGF-β signaling prevents cleft palate in Tgfbr2 mutant mice. *The Journal of clinical investigation* 2012; *122*(3), pp.873-885.

Jugessur A, Murray JC. Orofacial clefting: recent insights into a complex trait. *Current opinion in genetics & development* 2005; 15(3): 270-8.

Kerkelä E, Böhling T, Herva R, Uria J, Saarialho-Kere U. Human macrophage metalloelastase (MMP-12) expression is induced in chondrocytes during fetal development and malignant transformation. *Bone* 2001; 29(5): 487-93.

Koli K, Wempe F, Sterner-Kock A, Kantola A, Komor M, Hofmann WK, von Melchner H, Keski-Oja J. Disruption of LTBP-4 function reduces TGF-β activation and enhances BMP-4 signaling in the lung 2004. *Journal of Cell Biology* 2004; *167*(1), pp.123-133.

Lace B, Kempa I, Piekuse L, Grinfelde I, Klovins J, Pliss L, et al. Association studies of candidate genes and cleft lip and palate taking into consideration geographical origin. *European journal of oral sciences* 2011; 119(6): 413-7.

Leslie EJ, Murray JC. Evaluating rare coding variants as contributing causes to non-syndromic cleft lip and palate. *Clinical genetics* 2013; 84(5): 496-500.

Letra A, Silva RA, Menezes R, Astolfi CM, Shinohara A, de Souza AP, et al. MMP gene polymorphisms as contributors for cleft lip/palate: association with MMP3 but not MMP1. *archives of oral biology* 2007; 52(10): 954-60.

Li S, Wang L, Ma Z, Ma Y, Zhao J, Peng B, et al. Sequencing study on familial lung squamous cancer. *Oncology letters* 2015; 10(4): 2634-8.

Marchenko GN, Strongin AY. MMP-28, a new human matrix metalloproteinase with an unusual cysteine-switch sequence is widely expressed in tumors. *Gene* 2001; 265(1): 87-93.

Mouse Genome Informatics. http://www.informatics.jax.org/mgihome/homepages/stats/all_stats.shtml. Accessed June 14, 2015.

Muhamad A-H. Genetic of Non-syndromic Cleft Lip and Palate. 2012.

Mukhopadhyay, N., Bishop, M., Mortillo, M., Chopra, P., Hetmanski, J.B., Taub, M.A., Moreno, L.M., Valencia-Ramirez, L.C., Restrepo, C., Wehby, G.L. and Hecht, J.T., 2020. Whole genome sequencing of orofacial cleft trios from the Gabriella Miller Kids First Pediatric Research Consortium identifies a new locus on chromosome 21. *Human genetics*, *139*(2), pp.215-226.

Murray J. Gene/environment causes of cleft lip and/or palate. *Clinical genetics* 2002; 61(4): 248-56.

Nagase H, Woessner JF. Matrix metalloproteinases. *Journal of Biological Chemistry* 1999; 274(31): 21491-4.

Rifkin DB. Latent transforming growth factor-β (TGF-β) binding proteins: orchestrators of TGF-β availability. *Journal of Biological Chemistry* 2005; 280(9): 7409-12.

Rodgers UR, Kevorkian L, Surridge AK, Waters JG, Swingler TE, Culley K, et al. Expression and function of matrix metalloproteinase (MMP)-28. *Matrix Biology* 2009; 28(5): 263-72.

Saharinen J, Keski-Oja J. Specific sequence motif of 8-Cys repeats of TGF-β binding proteins, LTBPs, creates a hydrophobic interaction surface for

 | ISSN (Online): 2716-649X

binding of small latent TGF-β. *Molecular biology of the cell* 2000; 11(8): 2691-704.

Sarah KE, Christen LJ, José Arturo PN, Dario G-C, Carlos F, Wendy CK, et al. *39th Annual Eastern-Atlantic Student Research Forum*. 2015.

Shah NSM, Salahshourifar I, Sulong S, Sulaiman WAW, 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(1): 39.

Shaw GM, Croen LA, Curry CJ. Isolated oral cleft malformations: associations with maternal and infant characteristics in a California population. *Teratology* 1991; 43(3): 225-8.

Urban Z, Hucthagowder V, Schürmann N, Todorovic V, Zilberberg L, Choi J, et al. Mutations in LTBP4 cause a syndrome of impaired pulmonary, gastrointestinal, genitourinary, musculoskeletal, and dermal development. *The American Journal of Human Genetics* 2009; 85(5): 593-605.

Vieira A. Unraveling human cleft lip and palate research. *Journal of dental research* 2008; 87(2): 119-25.

Yang X, Dong Y, Zhao J, Sun H, Deng Y, Fan J, et al. Increased expression of human macrophage metalloelastase (MMP-12) is associated with the invasion of endometrial adenocarcinoma. *Pathology-Research and Practice* 2007; 203(7): 499-505.

 | ISSN (Online): 2716-649X