J Hum Genet (2006) 51:671–676
DOI 10.1007/s10038-006-0006-4

ORIGINAL ARTICLE

MSX1 mutations contribute to nonsyndromic cleft lip in a Thai population

Siraprapa Tongkobpetch · Pichit Siriwan · Vorasuk Shotelersuk

Received: 17 January 2006 / Accepted: 26 April 2006 / Published online: 26 July 2006

Abstract Previous studies observed that *MSX1* mutations could contribute to nonsyndromic cleft lip with or without cleft palate (CL/P) in some populations. Of the proposed pathogenic mutations, the P147Q variant was predominant in Vietnamese and present in Filipino populations. We investigated whether *MSX1* mutations also contribute to nonsyndromic CL/P in the Thai population. Specifically, we performed mutation analysis covering all the coding regions of the *MSX1* gene for 100 Thai patients with nonsyndromic CL/P. A total of eight variant sites were identified. Six were in coding regions, including four nonsynonymous changes, 101C > G (A34G), 440C > A (P147Q), 799G > T (G267C), and 832C > T (P278S). The G267C and P278S variants were predicted to be ''probably damaging'' by PolyPhen, changed themselves as potential exonic splicing enhancers for serine/arginine-rich proteins, and were not present in 162 control individuals of Thai ethnic background. Unlike all of the previously reported potential missense mutations in *MSX1*, these two novel potential mutations were found in exon 2 on the C-terminal side of the homeodomain protein. Moreover, in contrast to previous reports, we found the P147Q variant in 8 out of 100 Thai controls and an association between the variant and CL/P in our population could not be detected, suggesting that it is not pathogenic. Our data support that *MSX1* mutations are found in 2% of cases of CL/P and should be considered for genetic counseling implications, but suggest that the P147Q variant is not pathogenic.

Keywords Nonsyndromic cleft lip · *MSX1* · Mutations · Association · Haplotype

S. Tongkobpetch · V. Shotelersuk
Division of Medical Genetics and Metabolism, Department of Pediatrics, Chulalongkorn University, Bangkok 10330, Thailand

P. Siriwan
Division of Plastic Surgery, Department of Surgery, Faculty of Medicine, Chulalongkorn University, Bangkok 10330, Thailand

V. Shotelersuk (✉)
Division of Medical Genetics and Metabolism, Department of Pediatrics, King Chulalongkorn Memorial Hospital, Sor Kor Building 11th floor, Bangkok 10330, Thailand
e-mail: vorasuk.s@chula.ac.th

Introduction

Nonsyndromic cleft lip with or without cleft palate (CL/P) is the most common craniofacial anomaly. It has a prevalence of approximately 1 in 600 among Thai newborns (Shotelersuk et al. 2003). Environmental and genetic factors have been implicated in CL/P and several different loci and genes have been associated with them (Jugessur and Murray 2005).

In 1994, *MSX1* first emerged as a candidate based on the CL/P and foreshortened maxilla phenotype in the knockout mouse (Satokata and Maas 1994). Several association studies of the gene with CL/P and cleft palate only (CPO) further supported the role of *MSX1* in nonsyndromic clefting in different populations (Lidral et al. 1998; Blanco et al. 2001; Jugessur et al. 2003). Later on, a study of a Dutch family with tooth agenesis and various combinations of CL/P and CPO showed a nonsense mutation in *MSX1*, suggesting that

disease-causing mutations in *MSX1* might be etiological in a portion on nonsyndromic CL/P cases (van den Boogaard et al. 2000). Two previous studies, reporting that 2% of cases of clefting had *MSX1* mutations, supported this hypothesis (Jezewski et al. 2003; Suzuki et al. 2004). Of the proposed pathogenic mutations, the P147Q variant was found in approximately 2% of Vietnamese (Suzuki et al. 2004) and 0.15% of Filipino cases (Vieira et al. 2005). In the current study, we used direct sequencing covering all the coding sequences of *MSX1* to determine whether *MSX1* mutations might be etiological in some cases of Thai patients with nonsyndromic CL/P.

Subjects and methods

The subjects of this study were 88 sporadic cases of nonsyndromic CL/P and 12 additional cases with a positive family history. Details of their characteristics and recruitment have been previously reported (Leoyklang et al. 2006). The study was approved by the institutional review board of the Faculty of Medicine of Chulalongkorn University, and written informed consent was obtained from each person included in the study. The control samples were Thai blood donors with no oral clefts who denied history of oral clefts in other family members.

Genomic DNA was isolated from peripheral blood, according to established protocols. Primers in noncoding regions were used to specifically amplify fragments encompassing coding regions in both exons (primers 1F and 1R for exon 1 and primers 2F and 2R for exon 2; Table 1) of the *MSX1* gene. Polymerase chain reactions (PCR) were carried out in a 20-μl volume containing 50 ng genomic DNA, 1× PCR buffer, 1.9–2.0 mM $MgCl_2$, 0.2 mM dNTPs, 0.2 μM of each primer, and 0.5 U Taq polymerase, using the following parameters: 40 s at 94°C, 40 s at the annealing temperature (Table 1), and 40 s at 72°C for 35 cycles. PCR products were treated with ExoSAP-IT (USP, Cleveland, OH, USA) according to the manufacturer's recommendations, and sent for direct sequencing to Macrogen, Seoul, Korea. Primers used for sequencing were the same as those for PCR reactions, except the primer 1RS (Table 1), which was used for sequencing exon 1 in the 3′–5′ direction. Analyses were performed using Sequencher 4.2. When the results indicated a possible new variant, the sample was resequenced. The position of variants corresponds to the coding sequence for the nucleotide position within the Genbank entry AF426432.

The P147Q variant was verified in the patients by polymerase chain reaction-restriction fragment length polymorphism (PCR-RFLP) using a mutagenesis primer P147Q-F and the primer 1RS (Table 1) and restriction endonuclease, *Dde*I. In addition, to determine whether the variant was associated with cleft lip in our Thai population, we performed an association study by genotyping 50 more patients with cleft lip with or without CL/P (bringing the total number of participants to 150) and 100 Thai controls.

Standard Chi-squared and *p* values were calculated by a program available at http://www.unc.edu/~preacher/chisq/chisq.htm. Odds ratio and 95% confidence intervals (95%CI) were calculated from the Epi Info 2000 program downloaded from http://www.cdc.gov/epiinfo/.

We determined whether the P147Q variant in the Thai population was on the same haplotype as the Vietnamese by typing three single nucleotide

Table 1 Oligonucleotides and polymerase chain reaction (*PCR*) conditions for *MSX1* mutation analysis. *SNP* single nucleotide polymorphism

Name	Primer sequences for PCR 5′–3′	Product size (bp)	Annealing temperature (°C)
1F	CCAGTGCTGCGGCAGAAGG	848	62
1R	ATTCATCCGCTGGGGTGAA		
2F	GGCTGATCATGCTCCAATGC	556	58
2R	CACCAGGGCTGGAGGAAT		
1RS	TGGAACCTTCTTCCTGGGTG	–	–
P147Q-F	CCGAGAGGACCCCGTGGATGCAGAGCCCCCGCTTCTCTC	(with primer 1RS) 231	58
G267C-R	CAGGAAACAGCTATGACCCTGGAAGGGGCCAGAGGCTC	(with primer 2F) 448	60
SNP1-F	TAGGGCTTCTCAGGGAATCA	230	55
SNP1-R	TTGCGTGGTTTCCCGTATAC		
SNP4/5-F	AAGTCCAAAGGATCGTTGTG	960	57
SNP4/5-R	GGGAAGATGTGAAATCACCT		

polymorphisms (SNPs), snp1, snp4, and snp5 (SNPs were designated in accordance with a previous study, Suzuki et al. 2004) using primers SNP1-F and SNP1-R (Table 1) and *BstB*I for snp1; primers SNP4/5-F and SNP4/5-R (Table 1) and *BseR*I for snp4; and primers SNP4/5-F and SNP4/5-R and *Mbo*II for snp5, in 50 Thai controls. Haplotype frequencies were estimated by the EH program, which tested and estimated linkage disequilibrium between different markers, downloaded from http://www.linkage.rockefeller.edu/ott/eh.htm.

Both of the novel nonsynonymous coding variants, 799G > T and 832C > T, were verified by PCR-RFLP, using the primer 2F and a mutagenesis primer G267C-R (Table 1) and *Dde*I for the 799G > T, and primers 2F and 2R and *Mwo*I for the 832C > T. One hundred and sixty-two Thai control individuals were also examined for the variants by restriction enzyme analysis.

For protein sequence comparisons, MSX1 orthologs were first identified through a BLAST search of the nonredundant database using Homo sapiens MSX1, accession NP_002439, as the reference sequence. All known and complete MSX1 sequences were included from the vertebrate lineage. These files in FASTA format were then analyzed by ClustalX 1.81 program. The human MSX1 was aligned with cow (accession NP_777223), Norway rat (accession NP_112321), house mouse (accession NP_034965), red jungle fowl (accession XP_444660), African clawed frog (accession AAH81101), and zebrafish (accession NP_571348). The program classified amino acids by the variation in polarity, assessing both amino acid class conservation and evolutionary conservation at any given site.

PolyPhen (http://www.bork.embl-heidelberg.de/PolyPhen/) was used to predict the effect of the nonsynonymous mutations. ESEfinder software (http://www.rulai.cshl.edu/tools/ESE/) was used to predict potential exonic splicing enhancers (ESEs; Shotelersuk et al. 2004).

Results

The sequencing effort concentrated on the coding regions of the *MSX1* gene. In 100 DNA samples from subjects with nonsyndromic CL/P, 8 variant sites were identified. Seven were single nucleotide changes, comprising 3 transitions (2 in coding regions) and 4 transversions (all 4 in a coding region). The other variant was a single nucleotide deletion in intron 1 (Table 2).

The coding regions of *MSX1* contained 6 different variants, 2 synonymous and 4 nonsynonymous. Two nonsynonymous variants, 101C > G (A34G) and 440C > A (P147Q), were in exon 1 and have been previously reported. The A34G variant, previously reported as a nonpathogenic polymorphism (Suzuki et al. 2004), was found in 8 of our patients, 4 were heterozygous and the other 4 were either homozygous or hemizygous. Individuals who were homozygous/hemizygous for A34G have never been reported previously. The P147Q variant, previously proposed to be a pathogenic missense mutation in Vietnamese and Filipino cases (Suzuki et al. 2004; Vieira et al. 2005), was found in 3 of our 100 patients, all in the heterozygous state. Because of its relatively high prevalence in our patient group, we determined its frequencies in 50 more patients with nonsyndromic CL/P and 100 normal Thai controls. The observed frequencies of the 440C and 440A alleles, and the CC and CA genotypes in affected subjects and controls are shown in Table 3. The observed distribution of genotypes among controls was compared with that expected according to the Hardy–Weinberg equilibrium: no difference was found ($\chi^2 = 0.007$, $df = 1$, $P = 0.93$). Genotype frequencies of the patients also followed the Hardy–Weinberg equilibrium ($\chi^2 = 0.002$, $df = 1$, $P = 0.96$). The distributions of alleles and genotypes among patients were compared with those among controls: no differences of either allelic ($P = 0.285$) or genotypic ($P = 0.277$) distributions between patients and controls were found.

Table 2 Variant sites of *MSX1* found in 100 Thai patients with nonsyndromic cleft palate (CL/P)

Nucleotide position	Exon/intron	Nucleotide change	Expected amino acid change	Frequencies of heterozygotes	Frequencies of homozygotes/hemizygotes
90	Exon 1	C > A	A30A	0	2
101	Exon 1	C > G	A34G	4	4
330	Exon 1	C > T	G110G	15	15
440	Exon 1	C > A	P147Q	3	0
452–14	Intron 1	del T	–	8	0
799	Exon 2	G > T	G267C	1	0
832	Exon 2	C > T	P278S	1	0
894 + 6	3′UTR	C > T	–	0	1

Table 3 Genotypic and allelic distributions and comparisons of the *MSX1* 440C > A in patients with CL/P and controls

	Patients (n = 150)	Controls (n = 100)	χ^2 (P value, df = 1)	Odds ratio (95% CI)
Alleles				
C	0.977	0.96		
A	0.023	0.04	1.145 (0.285)	
Genotypes				
CC	0.953 (143)	0.92 (92)		
CA	0.047 (7)	0.08 (8)	1.182 (0.277)	0.56 (0.18–1.78)

In addition, no association was found with the CA genotype compared with the CC (odds ratio 0.56, 95% CI 0.18–1.78; Table 3). In contrast to a previous study in a Vietnamese population (Suzuki et al. 2004), the P147Q variant was not associated with cleft lip in the Thai population. Next, haplotype analysis was performed in 50 unrelated control Thai. The results are shown in Table 4. Of the 8 440A alleles found in Thai controls, 7 were on the haplotype #5 (included in the haplotype #4 of a previous study, Suzuki et al. 2004), the same haplotype as in the Vietnamese. Due to the unavailability of samples from the relatives of our CL/P patients, we did not determine the haplotype of their P147Q variants.

The two nonsynonymous variants in exon 2, 799G > T (G267C) and 832C > T (P278S), have not been previously reported (Fig. 1). One variant was found in each of 2 patients; both were sporadic cases, with normal development, no anomalies besides the oral clefts, no consanguinity and no mutations in the coding region of *p63* (Leoyklang et al. 2006). Clinical and molecular features of these two patients with potential mutations are shown in Table 5.

Discussion

The sequencing analysis of the *MSX*1 gene found four nonsynonymous variants. Two of them, the 799G > T (G267C) and 832C > T (P278S) variants, were predicted to be ''probably damaging'' by PolyPhen, changed (either created or eliminated) themselves as potential ESEs for serine-arginine (SR) proteins, and not present in 162 control individuals of Thai ethnic background. Both were found in exon 2 and have not been previously reported. Exon 2 of the *MSX1* was mostly conserved with significantly fewer sequence variations compared with exon 1 (Jezewski et al. 2003). These two potential mutations were found on the C-terminal side of the homeodomain protein while all of the previously reported potential missense mutations [E78V (Jezewski et al. 2003), G91D (Jezewski et al. 2003), G98E (Suzuki et al. 2004), V114G (Jezewski et al. 2003), G116E (Jezewski et al. 2003), and R151S (Jezewski et al. 2003)] were on the N-terminal side of the homeodomain protein. Moreover, for the P278S mutation, the proline 278 is conserved in MSX1 to the cow, rat, mouse, chicken, and frog protein sequence; and there is a substantial change in amino acid class, from nonpolar proline to polar serine.

Previous studies suggested that the P147Q variant is etiologically based on the strong conservation of the amino acid and the surrounding amino acids, the segregation analysis, and its absence in over 1,600 control individuals of various ethnic backgrounds (Suzuki et al. 2004; Vieira et al. 2005). Nevertheless, the fact that in some cleft families with the P147Q variant it was found in unaffected members while some affected did not carry it (Vieira et al. 2005) makes its role arguable. A previous observation suggested that it was a founder mutation in the Vietnamese population. In this study, the P147Q variant has been found in 8 out of 100 Thai controls, suggesting that it is not etiologic. In addition, an association between the variant and cleft lip in our population could not be detected (Table 3). Next, we determined whether the P147Q variant in our population was on the same haplotype as those in Vietnamese by genotyping three SNPs, which would be able to identify the four most common haplotypes in the Vietnamese (Suzuki et al. 2004). We did not genotype the snp2; therefore, haplotype #1 in

Table 4 Top 92% of haplotypes from 50 unrelated control Thai subjects

Haplotype # (designated in Suzuki et al. 2004)	Snp1[a] (–8796A > G)	P147Q (440C > A)	Snp4[a] (452–667T > G)	Snp5[a] (452–402G > T)	Frequency (#/total chromosomes)
1 (1 and 5)	A	C	T	T	27/100
2 (3)	G	C	T	T	27/100
3 (2)	A	C	G	G	22/100
4 (4)	A	C	T	G	9/100
5 (4)	A	A	T	G	7/100

[a]Numbers of SNPs are those used in Suzuki et al. (2004)

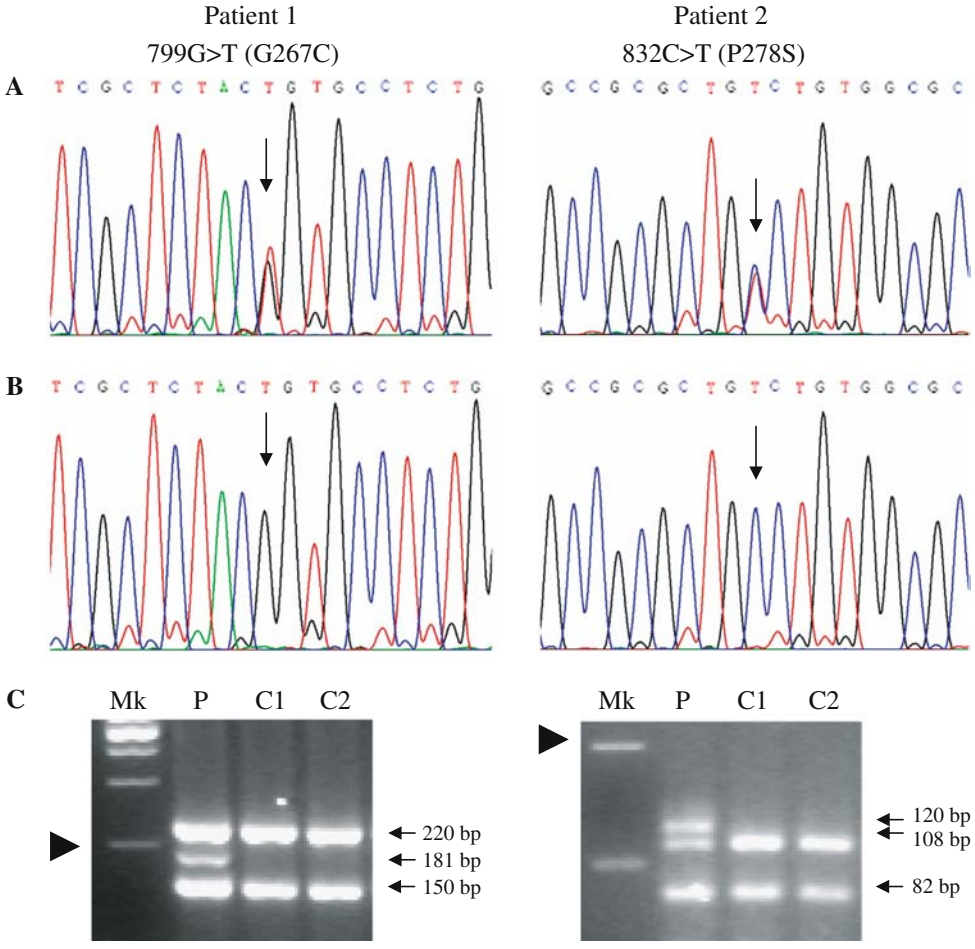

Fig. 1 Mutation analysis. The *left and right panels* relate to patients 1 and 2 respectively. **a** Electropherograms of patients, showing 799G > T and 832C > T (*arrows*) in patients 1 and 2 respectively. **b** Electropherograms of controls showing normal genotypes at nucleotide 799 as GG and 832 CC (*arrows*). **c** Restriction enzyme digestion analysis. C1 and C2 denote controls, Mk 100-bp marker, and *P* patient. The *arrowhead* indicates the 200-bp marker. In the *left panel*, *Dde*I digested the 448-bp product of controls into 220-, 150-bp, and other smaller products (not shown). The 799G > T mutation in patient 1 creates another *Dde*I restriction site. Therefore, the 220-bp PCR product of the mutant allele of the patient is further digested into 181- and 39-bp (not shown) products, present with the 220-bp product of the normal allele, indicating that patient 1 is heterozygous for the 799G > T mutation. In the *right panel*, *Mwo*I digested the 556-bp product of a control into 108-, 82-bp, and other smaller products (not shown). The 832C > T mutation in patient 2 eliminates a *Mwo*I site, leaving the 120-bp product, present with the digested 108- and 12-bp (not shown) of the normal alleles, indicating that patient 2 is heterozygous for the 832C > T mutation

Table 5 Clinical and molecular features of patients with potential mutations

	Patient 1	Patient 2
Age (years)	45	1.3
Sex	Female	Male
Province	Trang	Trang
Cleft type	Left complete cleft lip and palate	Left complete cleft lip
Nucleotide change (heterozygous)	799G > T	832C > T
Exon	2	2
Expected amino acid change	G267C	P278S
Animals with the same amino acid at the codon	Cow, chicken, frog[a]	Cow, rat , mouse, chicken, frog[a]
Position in MSX1	42 AA 3′ to homeobox	53 AA 3′ to homeobox
PolyPhen prediction	Probably damaging	Probably damaging
Frequency in 324 control chromosome	0	0

[a]Codon 267 of rat, mouse, and zebrafish is serine, which is a polar, uncharged amino acid; and codon 278 of zebrafish is asparagine, which is polar, uncharged

our study included both haplotypes #1 and #5 of the previous study (Suzuki et al. 2004). We showed that the P147Q variant in our control Thai subjects (Table 4) was on the same haplotype as those in Vietnamese (Suzuki et al. 2004). This observation strongly suggests that this P147Q variant in both the Thai and Vietnamese populations is inherited by descent from a founder genetic change. PolyPhen predicted the P147Q variant to be benign. This evidence suggests that the P147Q is neither pathogenic by itself, nor associated with cleft lip in the Thai population.

The A34G variant was a change within amino acid class, previously reported in cases and controls, and previously proposed to be benign (Jezewski et al. 2003; Suzuki et al. 2004). The roles of the two synonymous variants and the two variants in noncoding regions need further investigation.

Cleft palate is a multifactorial disorder caused by a combination of genes and environmental interactions. These factors may contribute differently to CL/P in different populations. In the Thai population, we have shown that there were associations between cleft lip and the maternal *MTHFR* 677CT/1298AC genotype, with an odds ratio of 4.43 and a 95%CI of 1.33–15.10 (Shotelersuk et al. 2003), and the *IRF6* 820G > A SNP of the proband, with an odds ratio of 1.67 and a 95% CI of 1.13–2.47 (Srichomthong et al. 2005). We also demonstrated that the *p63* mutation was responsible for approximately 1% of nonsyndromic cleft lips in the Thai population (Leoyklang et al. 2006). In this study, we have shown that *MSX1* mutations cause 2% of Thai cases of nonsyndromic cleft lip. Other populations may have different percentages of contributions from these genes.

This report demonstrates that the *MSX1* mutations appear to contribute about 2% of cases of nonsyndromic CL/P, consistent with two previous reports (Jezewski et al. 2003; Suzuki et al. 2004), but suggests that the P147Q variant is not pathogenic. Further studies of the full phenotypic spectrum and penetrance of *MSX1* mutations may improve genetic counseling in families with mutations.

Acknowledgements We wish to thank the medical staff of the Thai Red Cross and the Provincial Hospitals of Nakornratchaseema, Nan, Uthaithanee, Maehongsorn, Trang, Srakaew, Kalasin, Nongkhai, and Mahasarakam for the excellent care of their patients. This study was supported by the Research Unit Grant from Chulalongkorn University, the National Center for Genetic Engineering and Biotechnology, and the Thailand Research Fund.

References

Blanco R, Chakraborty R, Barton SA, Carreno H, Paredes M, Jara L, Palomino H, Schull WJ (2001) Evidence of a sex-dependent association between the MSX1 locus and nonsyndromic cleft lip with or without cleft palate in the Chilean population. Hum Biol 73:81–89

Jezewski PA, Vieira AR, Nishimura C, Ludwig B, Johnson M, O'Brien SE, Daack-Hirsch S, Schultz RE, Weber A, Nepomucena B, Romitti PA, Christensen K, Orioli IM, Castilla EE, Machida J, Natsume N, Murray JC (2003) Complete sequencing shows a role for MSX1 in non-syndromic cleft lip and palate. J Med Genet 40:399–407

Jugessur A, Murray JC (2005) Orofacial clefting: recent insights into a complex trait. Curr Opin Genet Dev 15:270–278

Jugessur A, Lie RT, Wilcox AJ, Murray JC, Taylor JA, Saugstad OD, Vindenes HA, Abyholm F (2003) Variants of developmental genes (TGFA, TGFB3, and MSX1) and their associations with orofacial clefts: a case-parent triad analysis. Genet Epidemiol 24:230–239

Leoyklang P, Siriwan P, Shotelersuk V (2006) A mutation of the p63 gene in non-syndromic cleft lip. J Med Genet 43:e28

Lidral AC, Romitti PA, Basart AM, Doetschman T, Leysens NJ, Daack-Hirsch S, Semina EV, Johnson LR, Machida J, Burds A, Parnell TJ, Rubenstein JL, Murray JC (1998) Association of MSX1 and TGFB3 with nonsyndromic clefting in humans. Am J Hum Genet 63:557–568

Satokata I, Maas R (1994) Msx1 deficient mice exhibit cleft palate and abnormalities of craniofacial and tooth development. Nat Genet 6:348–356

Shotelersuk V, Ittiwut C, Siriwan P, Angspatt A (2003) Maternal 677CT/1298AC genotype of the MTHFR gene as a risk factor for cleft lip. J Med Genet 40:e64

Shotelersuk V, Desudchit T, Tongkobpetch S (2004) ASA E382K disrupts a potential exonic splicing enhancer and causes exon skipping, but missense mutations in ASA are not associated with ESEs. Int J Mol Med 14:683–689

Srichomthong C, Siriwan P, Shotelersuk V (2005) Significant association between IRF6 820G > A and non-syndromic cleft lip with or without cleft palate in the Thai population. J Med Genet 42:e46

Suzuki Y, Jezewski PA, Machida J, Watanabe Y, Shi M, Cooper ME, Viet le T, Nguyen TD, Hai H, Natsume N, Shimozato K, Marazita ML, Murray JC (2004) In a Vietnamese population, MSX1 variants contribute to cleft lip and palate. Genet Med 6:117–125

Van den Boogaard MJ, Dorland M, Beemer FA, van Amstel HK (2000) MSX1 mutation is associated with orofacial clefting and tooth agenesis in humans. Nat Genet 24:342–343

Vieira AR, Avila JR, Daack-Hirsch S, Dragan E, Felix TM, Rahimov F, Harrington J, Schultz RR, Watanabe Y, Johnson M, Fang J, O'Brien SE, Orioli IM, Castilla EE, Fitzpatrick DR, Jiang R, Marazita ML, Murray JC (2005) Medical sequencing of candidate genes for nonsyndromic cleft lip and palate. PLoS Genet 1:e64