Characteristics of MSX1 gene in Korean nonsyndromic cleft lip and palate individuals

Hae-Kyung Lee, DDS, MSD, PhD, a Seong-Sik Kim, DDS, MSD, PhD, b
Woo-Sung Son, DDS, MSD, PhD c

Objective: This study was performed to identify the characteristics of the MSX1 gene (locus chromosome 4p16) in Korean nonsyndromic cleft lip and palate (CL/P), which is assumed to be a major candidate gene acting as a causal factor in nonsyndromic CL/P and missing teeth. Methods: The 36 individuals (23 males and 13 females) who had visited the department of orthodontics at from 1998 to 2002 and who had nonsyndromic CL/P were included in the study. Using a PCR-based assay, the MSX1 gene was amplified, sequenced, and searched for inferred protein products (Reference: Homo sapiens MSX1, accession number AF426432 and NP_002439). The common single nucleotide polymorphisms were observed. Results: In exon 1, nucleotide "A" of the 253 basepair (bp) region was substituted for "G", and in the 255 bp region, nucleotide "G" was inserted. In exon 2, nucleotide "C" of the 11 bp region was substituted for "A", and "T" or "G" was inserted into the 351 bp region whereas "T" or "A" was inserted into the 352 bp region. In protein analysis, "Thr85Ala" missense mutation was found. The "Thr85Ala" missense mutation in this study is different from those of studies using subjects of other races. Conclusions: The results suggest that there is specific mutation of MSX1 in Korean and it plays an important role in Korean nonsyndromic CL/P. However, any distinct genetic polymorphisms between CL/P with missing teeth in the cleft region and CL/P without missing teeth could not be found. (Korean J Orthod 2008;38(2):133-43)

Key words: MSX1 gene, Korean nonsyndromic cleft lip and palate, Missing teeth

INTRODUCTION

Cleft lip with or without cleft palate (CL/P) is the most common congenital malformation in the head and neck region. The prevalence of CL/P varies depending on racial and ethnic backgrounds, geographic origin, and socioeconomic status. Its occurrence ranges from high prevalence in Native Americans (2.61 per 1,000) and Asians (2.02 per 1,000) to intermediate in Caucasians (1.45 per 1,000) and low in Blacks (0.61 per 1,000). According to Kim et al, the incidence of CL/P in Korean subjects was 1.81 per 1,000 live births.

There are more than 300 described syndromes that have a CL/P among their associated characteristics. Most CL/P are of the nonsyndromic type that does not accompany syndromes, and this nonsyndromic type is known to affect as much as 70% of the whole CL/P
phenotypes. Nonsyndromic oral clefts can be defined as complex traits, since they do not exhibit classic Mendelian recessive or dominant inheritance attributable to any single locus, but show strong familial aggregation and have a substantial genetic component. Nonsyndromic CL/P might be due to mutations in a number of different genes. Therefore, this research into CL/P focuses on understanding the etiology of the nonsyndromic form.

Gene-gene and gene-environmental interactions have been believed to cause CL/P, and identification of specific influences has expanded with the development of molecular and epidemiological studies. Several loci and genes have been suggested as candidates. The current list of candidate genes for nonsyndromic CL/P includes transforming growth factor beta 3 gene (TGFB3), orofacial cleft 1 gene (OFC1), transforming growth factor alpha gene (TGFα), methylenetetrahydrofolate reductase gene (MTHFR), muscle segment homeobox 1 gene (MSX1), retinoic acid receptor alpha gene (RARA), orofacial cleft 2 gene (OFC2) and orofacial cleft 3 gene (OFC3).

In recent years, MSX1 has been emerging as an especially strong candidate for nonsyndromic CL/P based on the complete secondary cleft palate, complete failure of incisor development, and the foreshortened maxilla phenotype shown in the mouse knockout model. There is considerable evidence suggesting that MSX1 plays an important role in both craniofacial and dental development, and many studies support the belief that MSX1 mutations result in both orofacial clefting and specific patterns of inherited tooth agenesis in humans. The patterns of MSX1 mutation vary according to race and geographical region and it is possible for Korean CL/P individuals to have distinct characteristics of the MSX1 gene. But there have been little studies conducted with exclusively Korean subjects.

Accordingly, the null hypothesis of this study was that the MSX1 gene in Korean nonsyndromic cleft lip and palate is identical to the reference gene from the human gene database bank (GENBANK). The purpose of this study was to identify the characteristics of the MSX1 gene as a causal factor to cleft lip and palate with missing teeth in Korean nonsyndromic cleft lip and palate patients.

**MATERIAL AND METHODS**

Subject ascertainment and chart review

Included in the study were 36 individuals who had nonsyndromic CL/P and who visited the Department of Orthodontics at Pusan National University Hospital (PNUH) from 1998 to 2002. They consisted of 23 males and 13 females. Their ages ranged from 5.1 years to 30.9 years (mean age 12.2 years). Additionally, for this study, the subjects were subdivided into subsets on the basis of their orthodontic records. Dental radiographs are made routinely on all patients seen at PNUH for diagnostic purposes, and a chart review of all of the subjects of the present study was conducted to verify the type of cleft present, missing teeth and other dental anomalies. Other dental anomalies (e.g. peg lateralis and supernumerary teeth) were also investigated, because of the potential correlation between these phenotypes and several distinct genetic polymorphisms. The findings from the chart review were confirmed by evaluation of available radiographs and clinical photographs. For every individual participating in this study, which was conducted with the permission of the ethical committee of PNUH, an informed written consent form approved by the institutional review board of the School of Dentistry at Pusan National University, was required.

DNA extraction and purification

DNA was extracted from whole blood of each subject. Thirty ml of whole blood was collected from each subject in order to yield large amounts of DNA (> 100 μg). Allelic variants of the MSX1 gene were analyzed using polymerase chain reaction (PCR)-based assays. Genomic DNA was obtained from lymphocytes in these samples using the Wizard Genomic DNA Purifying kit (Cat. #A1120, Promega, Madison, WI, USA).

The primer sequences for the PCR assay were
Table 1. MSX1 primers sequence, product size, and annealing temperature

<table>
<thead>
<tr>
<th>Region</th>
<th>Location</th>
<th>Primer sequence</th>
<th>Product size</th>
<th>Annealing Temperature</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exon 1F</td>
<td>136-158</td>
<td>5'-CTGACATGACTTCTTGTGCCACTC-3'</td>
<td>479</td>
<td>60°C for 30 seconds</td>
</tr>
<tr>
<td>Exon 1R</td>
<td>590-609</td>
<td>5'-TGGTTTCTGGCTACTCAGT-3'</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exon 2F</td>
<td>3249-3270</td>
<td>5'-GCTTCTCCTTAACCCCTTCT-3'</td>
<td>429</td>
<td>60°C for 30 seconds</td>
</tr>
<tr>
<td>Exon 2R</td>
<td>3718-3741</td>
<td>5'-GACCCCTCATGTCAGGTGCAT-3'</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

F, Forward; R, reverse; *Reference sequence, GENBANK accession number AF425672.

![Image](image1)

Fig 1. Two percent agarose gel electrophoresis of MSX1 gene PCR product. M, molecular weight marker; lane 1, 2, 3, 4, nonsyndromic cleft lip and palate subjects; bp, basepair.

obtained from the Primer 3 program (Whitehead Institute/MIT Center for Genome Research (WICGR), USA). The regions examined were exon 1 and exon 2 from 468 nucleotides 5' of the start codon to 1073 nucleotides past stop codon. 389 nucleotides section of the intron were excluded. Each primer sequence, PCR product size, and annealing temperature in this study were described in Table 1.

To perform the PCR assay, the Accupower Hotstart PCR Premix Ki (Cat. #K-2016, Bioeer, Daejeon, Korea) was used. A reaction mix was prepared with 2 μl of DNA, 1 μl of 10 μM sense primer, 1 μl of 10 μM antisense primer, and distilled water. A total volume of 20 μl mixture was added to the Accupower PCR Premix kit tube; then PCR was initiated. The steps of 1 cycle are as follows: predenaturation at 94°C for 5 minutes; denaturation at 95°C for 30 seconds; annealing at 60°C for 30 seconds; elongation at 72°C for 60 seconds; and postelongation at 72°C for 10 minutes. The genes were amplified for 35 cycles (MWG-Biotech AG, Ebersberg, Germany). All of the PCR products were separated by using 2% agarose gel electrophoresis, stained with ethidium bromide 0.2 g/ml, and visualized under ultraviolet light. All of the PCR screening methods used in this study had been validated (Fig 1).
Table 2. Characteristics of MSX1 gene as a reference

<table>
<thead>
<tr>
<th>Genomic regions, transcricting, and products</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exon 1 (72.4270), Exon 2 (3247.3689)</td>
</tr>
<tr>
<td><strong>Gene</strong></td>
</tr>
<tr>
<td>MSX1</td>
</tr>
<tr>
<td><strong>Chromosome</strong></td>
</tr>
<tr>
<td>4p16</td>
</tr>
<tr>
<td><strong>Organism</strong></td>
</tr>
<tr>
<td>Homo sapiens</td>
</tr>
<tr>
<td><strong>Molecular Type</strong></td>
</tr>
<tr>
<td>genomic DNA</td>
</tr>
<tr>
<td><strong>Protein ID</strong></td>
</tr>
<tr>
<td>AAL17870.1</td>
</tr>
<tr>
<td><strong>Product</strong></td>
</tr>
<tr>
<td>muscle segment homeobox 1 protein</td>
</tr>
<tr>
<td><strong>DB_xref</strong></td>
</tr>
<tr>
<td>GI:16326739</td>
</tr>
</tbody>
</table>

Sequencing analysis

Sequencing was performed in both directions on the DNA samples of the 36 subjects. Templates included PCR products purified using a QIAEX II Gel Extraction kit (Cat. No. 20021, QIAGEN, Valencia, CA, USA). The templates were sequenced by the ABI 3700 sequence (Applied Biosystems Inc., Foster City, CA, USA) machine.

Vector NTI 5.0 (Invitrogen, Carlsbad, CA, USA) sequencing analysis software were used for base calling, assembling, scanning, and reviewing. Then, the sequences of MSX1 gene, exon 1 and exon 2, were compared to the human sequences available via GENBANK accession number AF426432 (Table 2). To minimize sequencing errors, these procedures were repeated 6 times.

Protein sequence comparisons

MSX1 sequences of all patients were first identified through a BLAST search using Homo sapiens MSX1, accession NP_002439, as the reference sequence. All known and complete MSX1 sequences were indicated from the vertebrate lineage. These files in FASTA format were then manipulated in Jalview (Jalview: http://www.ebi.ac.uk/jalview/) and submitted for remote alignment at the EBI using a ClustalW algorithm (version 1.82). The exon 1 and exon 2 sequences of the MSX1 gene were aligned using remote ClustalW alignment from EBI.

RESULTS

Subject distributions

The chart review that had been completed on the 36 subjects was consulted to confirm the type of clefting present and the type and location of missing teeth, peg lateralis, and supernumerary tooth (Table 3). These findings from the chart review were confirmed by evaluation of available radiographs and clinical photographs.

Regarding the type of clefting, 27 (75%) of the subjects had unilateral cleft lip and palate (UCLP). Subjects affected on the left side outnumbered those affected on the right side. Twenty-three were affected on the left and 4 on the right; 6 had bilateral cleft lip and palate (BCLP); 3 had cleft lip and alveolus (CLA, 1 affected on the left side, 1 affected on the right side, and 1 affected bilaterally.).

Regarding the prevalence of missing teeth, 21 (58.4%) of the subjects had missing teeth either inside or outside the cleft region, 12 (33.3%) had missing teeth in the region of the cleft (that is, the lateral incisors on the cleft side), 4 had missing teeth inside and outside the cleft region simultaneously, and 5 had missing teeth outside the cleft region. Most of the missing teeth outside the cleft region were premolars. The remainder, 15 (41.6%) of the patients exhibited no evidence of missing teeth.

Regarding peg lateralis, 22 (61%) of the subjects suffered from this condition. Eighteen of those subjects had peg lateralis in the cleft side, whereas 3 had peg lateralis outside the cleft region and missing lateral incisors inside the cleft region. Only 1 had peg lateralis in and outside the cleft region simultaneously.
Table 3. Frequency of cleft type and occurrence of missing teeth and peg lateralis related to the cleft region

<table>
<thead>
<tr>
<th>Type of Cleft</th>
<th>Frequency</th>
<th>Missing teeth</th>
<th>No missing teeth</th>
<th>Peg lateralis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>in the cleft</td>
<td>outside the cleft</td>
<td></td>
</tr>
<tr>
<td>UCLP</td>
<td>27</td>
<td>13</td>
<td>4</td>
<td>10</td>
</tr>
<tr>
<td>BCLP</td>
<td>6</td>
<td>2</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>CLA</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Total</td>
<td>36</td>
<td>16</td>
<td>5</td>
<td>21</td>
</tr>
</tbody>
</table>

*UCLP, unilateral cleft lip and palate; BCLP, bilateral cleft lip and palate; CLA, cleft lip and alveolus.

Reference →
Pattern 1 →
Pattern 2 →
Pattern 3 →

Fig 2. Three different genetic features were exhibited, and were divided into patterns I, II and III. The sequences of pattern I were similar to the reference sequences, and most of the subjects belonged to pattern I. However, the sequences of pattern II and pattern III were very different from the reference sequences, and they also differed from each other.

<table>
<thead>
<tr>
<th>Group</th>
<th>Patient Number (#)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alignment</td>
<td></td>
</tr>
<tr>
<td>reference</td>
<td></td>
</tr>
<tr>
<td>#19, 25</td>
<td></td>
</tr>
</tbody>
</table>

Pattern 1
#1, 2, 5, 6, 7, 8, 9, 12, 13, 14, 16, 17, 18, 21,
22, 23, 26, 28, 29, 30, 31, 32, 33, 36

Pattern 2
#3, 20, 24, 34, 35

Pattern 3
#4, 10, 11, 15

Fig 3. Neighbour joining tree of MSX1 gene from ClustalW multiple sequence alignment program. Most of the subjects except 9 subjects belonged to pattern I. *#* means patient number in raw data.

Genomic structures of exon 1 and exon 2 in MSX1 gene

Single nucleotide polymorphisms (SNPs)
The sequencing analysis software (version 5.0) was used for base calling, assembling, and scanning. Then the sequences of the MSX1 gene, exon 1 and exon 2, were compared with the human sequences of the MSX1 gene available via GENBANK accession number AF426432 (Table 2).
Fig 4. The common single nucleotide polymorphisms (SNPs) of pattern I. In exon 1, nucleotide "A" of the 253 basepair region was substituted for "G", and in the 255 basepair region, nucleotide "G" was inserted.

Fig 5. The common SNPs of exon 2. In exon 2, nucleotide "C" of the 11 basepair region was substituted for "A".

Three different genetic features were exhibited, and were divided into patterns I, II and III (Fig 2). The sequences of pattern I were similar to the reference sequences, and most of the subjects belonged to pattern I (Fig 3). However, the sequences of pattern II (patient number 3, 20, 24, 34, 35) and pattern III (patient number 4, 10, 11, 15) were very different from the reference sequences, and they also differed from each other. To identify the phenotypes of the three patterns, the raw data was examined. The common phenotypes related to dental anomalies were not found in pattern I. But most of the subjects of patterns II and III had common phenotypes with peg lateralis and no missing tooth inside the cleft region.

The common single nucleotide polymorphisms (SNPs) of all of the subjects of pattern I were observed. In exon 1, nucleotide "A" of the 253 basepair region was substituted for "G", and in the 255 basepair region, nucleotide "G" was inserted (Fig 4). In exon 2, nucleotide "C" of the 11 basepair region was substituted for "A" (Fig 5), and "T" or "G" was inserted into the 351 basepair region whereas "T" or "A" was inserted into the 352 basepair region (Fig 6).

Protein structure prediction
MSX1 protein sequences of all patients were first identified through a BLAST search using Homo sapiens MSX1, accession NP_002439, as the reference sequences. One missense mutation (Thr85Ala) was found in all of the subjects, which was caused by nucleotide "A" being substituted to "G" in the 253 basepair region and the insertion of nucleotide "G" into the 255 basepair region shown in Figure 4. "ACC (Threonine)" was changed to "GCG (Alanine)".
Fig 6. The common SNPs in exon 2 of pattern I in 351,352 basepair region. "T" or "G" was inserted into the 351 basepair region whereas "T" or "A" was inserted into the 352 basepair region.

Fig 7. Protein sequence comparison in exon 1 of the MSX1 gene. In 85 region, Threonine was substituted for Alanine in all subjects. Sequence changes of amino acid by this protein mutation in the 85 region were continued into 279 region and then reverted to the normal structure from the 280 region.

Sequence changes of amino acid by this protein mutation in the 85 region were continued into 279 region and then reverted to the normal structures from the 280 region (Fig 7).

The protein sequences of patient number 32 were considerably different from those of other subjects in pattern 1 (Fig 8). In the MSX1 protein sequences of patient number 32, the "Lys20Gln" missense mutation was found, and was caused by the insertion of nucleotide "T" into the 57 basepair region. "AAG (Lysine)" was changed to "CAA (Glutamine)". Compared with the phenotype of this subject, patient number 32 had BCLP and had 5 missing teeth in the cleft region (two maxillary lateral incisors bilaterally) and in the outside cleft region.

DISCUSSION

Nonsyndromic cleft lip and palate are common congenital anomalies with genetically complex traits. Segregation analyses and epidemiological studies have shown that 25% - 35% of CL/P patients have a family history of clefting and that simple Mendelian inheritance models are insufficient to explain the mode of inheritance in families with segregating clefting. In addition, CL/P are heterogeneous traits, with an estimated 2 - 20 genes interacting multiplicatively to cause clefts, including a possible major gene that might account for 10% - 50% of the incidences of these birth defects.16
Numerous studies have reported dental anomalies of permanent teeth in association with various forms of cleft lip, cleft palate, or both. These anomalies consist of variations in tooth number and position and reduced tooth dimensions. The prevalence of the congenital absence of permanent teeth inside the cleft region in CL/P individuals is clinically higher than that in the general population without clefts. In the general population, the prevalence of congenital missing teeth has been reported to range from 3% to 10% of the subjects examined.26,27 Studies of subjects with clefts, by contrast, have found the prevalence of missing teeth to range from 18% to 30%, and missing teeth was found to occur approximately three times as frequently on the cleft side as on the noncleft side.28,31 In the present study, 16 of the 36 subjects (44.5%) had missing permanent lateral incisors in the cleft region and 5 of the 36 subjects (13.9%) had missing teeth outside the cleft region. The prevalence (44.5%) of missing teeth in CL/P in this study was higher than that in past studies, ranging from 18% to 30%.28-33 With this in mind, these two phenotypes, clefting and missing teeth inside the cleft region, might be related in their genetic etiology. Therefore, in the present study, we tried to discover the genetic relations between clefting and missing teeth in the cleft region. However, we could not find any distinct genetic polymorphisms between CL/P with missing teeth in the cleft region and CL/P without missing teeth, suggesting that the missing teeth inside the cleft region were due to other causes. Whereas the etiology of missing teeth outside the cleft region is genetic in origin, the etiology of missing teeth inside the cleft region might not be genetic in origin, but environmental, and related to cleft formation.
In human, by genetic linkage analyses in a family with autosomal dominant agenesis of second premolars and third molars, Vastardis et al. identified a locus on 4p16.1 where the MSX1 gene resides. Sequence analyses demonstrated an arg31-to-pro missense mutation (Arg31Pro) in the homeodomain of MSX1 in all affected family members. Jumlongras et al. used candidate-gene linkage analysis in a 3-generation family to identify the gene responsible for Witkop syndrome, also known as tooth-and-nail syndrome. Direct sequencing and restriction enzyme analysis revealed that a heterozygous ser202-to-ter mutation (S202X) in the homeodomain of MSX1 segregated with the phenotype. The resemblance between the tooth-and-nail phenotype in the human family strongly supported the conclusion that the S202X nonsense mutation in MSX1 causes Witkop syndrome and that MSX1 is critical for both tooth and nail development. Van den Boogaard et al. identified a nonsense mutation (Ser104stop) of the MSX1 gene in a family with autosomal dominant tooth agenesis and combinations of cleft palate only and cleft lip and palate. Recently, a large-scale sequence analysis of the MSX1 gene was performed on 917 persons of various ethnicities who had nonsyndromic cleft lip/palate, and identified potentially etiologic mutations (Glu78Val, Gly116Glu) in 16 individuals. The authors estimated that MSX1 mutations contribute to 2% of all nonsyndromic cleft lip and palate cases.

In the present study, direct sequencing of the MSX1 gene was performed, which revealed a disease-causing mutation. Three different genetic features were exhibited, and were divided into patterns I, II, and III. The sequences of pattern I were similar to the reference sequences, and most of the subjects belonged to pattern I. The sequences of patterns II and III were very different from the reference sequences, and they also differed from each other (Figs 2 and 3). The common phenotypes related to dental anomalies were not found in pattern I. Some subjects in pattern I had missing teeth inside or outside the cleft region and the others had no missing teeth. And some subjects in pattern I had peg lateralis and the others had no peg lateralis. All of the subjects in patterns II and III except only one subject had no missing teeth but peg lateralis inside the cleft region. The different genetic features of patterns II and III might be affected by the genetic characteristics of peg lateralis. Although the high prevalence of peg lateralis of nonsyndromic cleft lip and palate was shown, there have been no genetic studies of the etiology of microdontia like a peg lateralis; that is, the etiology has not yet been established. Identifying the genetic etiology of clefting and peg lateralis will make for interesting studies hereafter.

A common single nucleotide polymorphisms of all of the subjects of pattern 1 were observed. In exon 1, nucleotide "A" of the 253 basepair region substituted to "G", and in the 255 basepair region, nucleotide "G" inserted (Fig 4). This substitution affected protein sequences and made one missense mutation (Thr85Ala). The "Thr85Ala" missense mutation was found in 27 of the total 36 subjects who had nonsyndromic cleft lip and palate. As in previous studies, the MSX1 mutation was found in Korean nonsyndromic cleft lip and palate, and the characteristics of the MSX1 mutation in Koreans were found to be different from those of other races.

Particularly, the sequence of patient number 32 in this study was considerably different from those of the other subjects of pattern I (Fig 8). In the MSX1 protein sequences of patient number 32, the "Lys20Gln" missense mutation was found. The phenotypes of this subject were BCLP and 5 missing teeth inside and outside the cleft region (two maxillary lateral incisors bilaterally and 3 premolars in the maxillary and mandibular arches). Only that subject of all the subjects, had multiple missing teeth in the cleft region and outside the cleft region simultaneously. This specific sequence feature might be related to multiple missing teeth inside and outside the cleft region. Further studies on the genotypes of multiple missing teeth inside and outside the cleft region simultaneously are necessary.

For better identification of the polymorphism, family-based studies can be useful in cases of population stratification. The case-parents trios design is a concern because of the spurious results, due to population stratification within a sample, primarily because the observed case is always compared with
ethnically matched 'pseudocontrols' (parents). However, only the patient sample and the GENE database bank as the reference were used in this study. To compare with the wild type gene of GENBANK in this study is a basic experiment to elucidate the role of MSX1 in orofacial clefting. Further studies using large sample sizes and normal Korean subjects unrelated to CL/P and missing tooth as controls, are necessary to identify more clearly the role and characteristics of the MSX1 gene in Koreans.

CONCLUSION

This study was performed to identify the characteristics of the MSX1 gene (locus chromosome 4p16) in Korean nonsyndromic cleft lip and palate, which is assumed to be a major candidate gene acting as a causative factor in nonsyndromic cleft lip and palate and missing teeth. Using a PCR-based assay, the MSX1 gene was amplified, sequenced, and searched for inferred protein products. The "Thr85Ala" missense mutation was found in this study and it was different from those studies using subjects of other races. However, any distinct genetic polymorphisms between CL/P with missing teeth and CL/P without missing teeth could not be found.

REFERENCES

14. Kim SS, Son WS. Sequencing analysis of the OFC1 gene on the
Characteristics of MSX1 gene in cleft lip and palate individuals