J Hum Genet (2006) 51:38–46
DOI 10.1007/s10038-005-0319-8

ORIGINAL ARTICLE

**Eisaburo Ichikawa · Akira Watanabe · Yoko Nakano
Sadanori Akita · Akiyoshi Hirano · Akira Kinoshita
Shinji Kondo · Tatsuya Kishino · Takeshi Uchiyama
Norio Niikawa · Koh-ichiro Yoshiura**

# *PAX9* and *TGFB3* are linked to susceptibility to nonsyndromic cleft lip with or without cleft palate in the Japanese: population-based and family-based candidate gene analyses

Received: 14 July 2005 / Accepted: 14 September 2005 / Published online: 25 October 2005


**Abstract** The prevalence of nonsyndromic cleft lip with or without cleft palate (CL/P) and cleft palate only (CPO) are believed to be higher in the Japanese than in Americans, Europeans or Africans. The purpose of this study was to investigate, in a Japanese population, relationships between CL/P or CPO and seven candidate genes (*TGFB3, DLX3, PAX9, CLPTM1, TBX10, PVRL1, TBX22*) that showed positive associations in other populations and are expressed in the oral/lip region in developing mice. We first searched for mutations in these genes among 112 CL/P and 16 CPO patients, and found a heterozygous missense mutation (640A > G, S214G) in exon 3 of *PAX9* in two sibs with CL/P and their phenotypically normal mother from a Japanese family. A population-based case-control analysis and a family-based transmission disequilibrium test (TDT), using single nucleotide polymorphisms (SNPs), and two-SNP haplotypes of the genes, between the 112 CL/P cases with their parents and 192 controls indicated a significant association at one SNP site, IVS1 + 5321, in *TGFB3* with a *P*-value of 0.0016. Population-based haplotyping revealed that the association was most significant for haplotype "A/A" consisting of IVS1 + 5321 and IVS1−1572; TDT also gave a *P*-value of 0.0252 in this haplotype.

**Keywords** Cleft lip with or without cleft palate · Association study · Mutation search · *PAX9* · *TGFB3*

E. Ichikawa · A. Watanabe · Y. Nakano · T. Uchiyama
Department of Oral and Maxillofacial Surgery,
Tokyo Dental College, Chiba, Japan

E. Ichikawa · A. Watanabe · A. Kinoshita · N. Niikawa
K. Yoshiura (✉)
Department of Human Genetics,
Nagasaki University Graduate School of Biomedical Sciences,
1-12-4 Sakamoto, Nagasaki 852-8523, Japan
E-mail: kyoshi@net.nagasaki-u.ac.jp
Tel.: +81-95-8497119
Fax: +81-95-8497121

S. Akita · A. Hirano
Division of Plastic and Reconstructive Surgery,
Nagasaki University Graduate School of Biomedical Sciences,
Nagasaki, Japan

S. Kondo · T. Kishino
Division of Functional Genomics, Center for Frontier Life Sciences, Nagasaki University, Nagasaki, Japan

A. Kinoshita · S. Kondo · T. Kishino · N. Niikawa · K. Yoshiura
Core Research for Evolutional Science and Technology (CREST),
Japan Science and Technology Agency (JST), Kawaguchi, Japan

## Introduction

Nonsyndromic oral clefts [cleft lip only (CLO, MIM 119530), cleft lip with cleft palate (CLP, MIM 119530), and cleft palate only (CPO, MIM 119540)] together represent one of the most frequently observed congenital anomalies. CLO and CLP are generally categorized into one entry, i.e., cleft lip with or without cleft palate (CL/P), because these two phenotypes are thought to have the same genetic etiology, while CPO may involve a set of genes different from those for CL/P. Oral clefts are classified into nonsyndromic and syndromic oral clefts (Schutte and Murray 1999). Approximately 70% of CL/P cases are nonsyndromic and 30% syndromic (Schutte and Murray 1999; Cobourne 2004). The prevalence of nonsyndromic oral clefts is 0.4–2.0/1,000 births (Schutte and Murray 1999; Natsume et al. 2000), with prevalence in Japanese seemingly higher than in other populations (Vanderas 1987; Tanabe et al. 2000). The occurrence of oral clefts is explained by the "multifactorial threshold model" involving both genetic and environmental factors (Cobourne 2004), but most of these factors remain

unknown. Some genes causing syndromic oral clefts as a single gene defect have been identified, e.g., *MSX1* (van den Boogaard et al. 2000), *IRF6* (Kondo et al. 2002), *PVRL1* (Suzuki et al. 2000), and *TBX22* (Braybrook et al. 2001). These are good candidate genes for some instances of nonsyndromic CL/P and CPO because it has been shown that reduced activity of the proteins they encode can affect oral development (Lidral et al. 1997; Jezewski et al. 2003; Zucchero et al. 2004). In addition, there are other potential candidate genes, although, with the exception of *TGFB3*, which showed a positive association with nonsyndromic oral clefts in some populations (Sato et al. 2001; Beaty et al. 2002; Kim et al. 2003), an association between these genes and oral cleft in humans has not yet been reported. Disruption of *CLPTM1* by a chromosomal translocation was reported in a patient with oral cleft (Yoshiura et al. 1998). Mutations in *PAX9* were reported in some syndromic oral cleft patients (Das et al. 2003). *DLX3* and *TBX10* are involved in the development of mouse oral cleft (Juriloff et al. 2001; Bush et al. 2004). Although population-based, genome-wide case-control analysis and the transmission disequilibrium test (TDT) are powerful methods with which to find genes responsible for susceptibility to nonsyndromic oral clefts, a genotype–phenotype association depends on the population history (Freedman et al. 2004).

The aim of our study was to investigate the contribution of seven candidate genes (*TGFB3*, *DLX3*, *PAX9*, *CLPTM1*, *TBX10*, *PVRL1*, and *TBX22*) to oral clefts in the Japanese. Here, we present the results of mutation searches and association studies. This is the first report on intensive candidate gene analysis of oral clefts in the Japanese.

## Materials and methods

### Subjects

The subjects studied included 112 nonsyndromic Japanese CL/P patients (45 females and 67 males) and 16 CPO patients (10 females and 6 males), their parents (256; a total of 128 trios), and 192 phenotypically normal adults for the association study. To verify that the base change in *PAX9* was not observed in healthy control samples, we searched for the mutation in 282 phenotypically normal controls. DNA from blood samples collected from patients and their parents at Tokyo Dental College Hospital and Nagasaki University Hospital, and samples from controls (volunteers) collected at Nagasaki City, was used for mutation searches and case-control studies. All sampling was performed with written informed consent. Diagnosis of CL/P or CPO was made through clinical inspections by well trained dentists and oral surgeons. The study protocol was approved by Institutional Review Board (IRB) for Ethical, Legal and Social Issues (ELSI) at each university/college.

### Candidate genes for nonsyndromic oral clefts

#### *TGFB3*

*TGFB3*, the transforming growth factor-beta 3 gene, located at 14q24, is especially expressed in medial edge epithelium cells of the palatal shelves for normal fusion of the palatal shelves, and is required for adhesion of the medial edge epithelium and elimination of the midline epithelial seam of the palatal shelves (Proetzel et al. 1995). In addition, *TGFB3* knock-out mice were previously reported to have a developmental defect of the secondary palate and delayed pulmonary development (Proetzel et al. 1995). Results of previous association studies on *TGFB3* have remained controversial, with reports of both significant association with nonsyndromic oral clefts in various populations (Sato et al. 2001; Beaty et al. 2002; Kim et al. 2003) and negative association (Lidral et al. 1997; Tanabe et al. 2000).

#### *DLX3*

*DLX* (distal-less homeobox) genes encode transcription factors containing the homeodomain that plays an important role in early patterning of embryonic structures such as craniofacial tissues. *Dlx3* belongs to the distal-less gene family in vertebrates, and its human homolog is *DLX3* at 17q21(Kraus and Lufkin 1999). The mouse *Dlx3* gene was identified from a candidate region for mouse nonsyndromic cleft lip by a linkage study (Juriloff et al. 2001). Although point mutations in human *DLX3* cause tricho-dento-osseous syndrome without CL/P (Price et al. 1998), one function of *DLX3* may be to interact with *MSX1*, which may be a gene causative for CL/P and CPO (Bryan and Morasso 2000; van den Boogaard et al. 2000). We chose *DLX3* as a candidate because of its chromosomal localization and relation to *MSX1*.

#### *PAX9*

*PAX9*, the paired box gene 9 at 14q12-q13, encodes a transcription factor containing the DNA-binding paired domain (Peters et al. 1998). Mouse *Pax9* is extensively expressed in the neural-crest-derived mesenchyme of the palatal shelves and tooth (Peters et al. 1998). *Pax9* knock-out mice presented with secondary cleft palate, tooth agenesis, and other abnormalities (Peters et al. 1998). *PAX9* mutations in human are reported to cause hypodontia involving molars that are frequently accompanied by CL/P (Das et al. 2003).

#### *CLPTM1*

*CLPTM1*, the cleft lip and palate associated transmembrane protein-1 gene at 19q13.2, was isolated from the breakpoint of a balanced chromosomal translocation [t(2;19)(q11.2;q13.3)] in a family where a CLP phenotype co-segregated with the translocation in a mother and her

children but not in the maternal grandmother (Yoshiura et al. 1998). *CLPTM1* may be involved with the immune system (Takeuchi et al. 1997), but its definitive function has not yet been identified. Eight rare and nine common variants of this gene were detected by a search for mutations in 74 unrelated patients with nonsyndromic CLP, but no significant association was obtained (Yoshiura et al. 1998). Nevertheless, as previous linkage analysis and TDT suggested that a CL/P locus maps to 19q13 close to *CLPTM1* (Wyszynski et al. 1997), *CLPTM1* is still a candidate gene for nonsyndromic oral clefts in some populations.

### *TBX10*

*TBX10*, the T-box transcription factor-10 gene at 11q13.1, is a member of the T-box gene family that encodes DNA-binding transcription factors (Law et al. 1998). Members of this family are known to play essential roles in mesoderm structures specific to early human developmental stages (Papaioannou and Silver 1998). Bush et al. (2004) found that Dancer (Dc) mice exhibiting CL/P carry a spontaneous mutation in *Tbx10*, which is located near the centromere of mouse chromosome 19. The localization of human *TBX10* syntenic to mouse *Tbx10* is 11q13, where a susceptibility locus to nonsyndromic CL/P was identified by genome-wide affected-sib pair analysis (Prescott et al. 2000). Moreover, *TBX10* strongly associates with *TBX22*, a known causative gene for CPO (Braybrook et al. 2001; Bush et al. 2004; Marcano et al. 2004).

### *PVRL1*

Autosomal recessive CL/P-ectodermal dysplasia syndrome (CLPED1, MIM 225000), previously called Margarita Island ectodermal dysplasia (MIM 225060), is characterized by CL/P, dental anomalies, hand anomalies, hidrotic ectodermal dysplasia, and occasionally mental retardation (Bustos et al. 1991). Suzuki et al. (2000) identified a nonsense mutation (G546A; W185X), a deletion (546delG), and a duplication (959dupG) of *PVRL1*, the poliovirus receptor like-1 gene at 11q23, in CLPED1 families from Margarita Island in north Venezuela, in patients from Israel, and patients from Brazil, respectively. *PVRL1* encodes a cell adhesion molecule, nectin-1, which is the principal receptor for alpha-herpes viruses (Suzuki et al. 2000). Sozen et al. (2001) found that the W185X mutation in *PVRL1* is one of risk factors for nonsyndromic CL/P in the Cumaná region of northern Venezuela and in Margarita Island.

### *TBX22*

*TBX22*, the T-box transcription factor-22 gene at Xq21.1 and a member of the T-box gene family, is important for both palatal and tongue development (Braybrook et al. 2001), and was shown to be expressed in the palatal shelves and tongue by in situ hybridization in both human and mouse (Braybrook et al. 2002). Functional loss of *TBX22* causes X-linked cleft palate with ankyloglossia (CPX, MIM 303400) (Braybrook et al. 2001). Marcano et al. (2004) found three *TBX22* mutations in CPX and CPO patients (Brazilians, North Americans, and Filipinos) from three geographically distinct populations, i.e., 105–106delGC and 581–582insCAG in one and two Brazilian CPX patients, respectively, and 548C > T (P183L) in an American CPO patient, indicating that *TBX22* contributes significantly to CPO patients. A genome-wide sib-pair analysis for nonsyndromic CL/P also identified susceptibility loci in the Xcen-q21 region in which *TBX22* is located (Prescott et al. 2000).

## Mutation search and genotyping

We performed mutation searches in all 128 patients with CL/P or CPO by sequencing all exons and part of the introns of the seven candidate genes. PCR amplification was performed as follows: denaturation at 94°C for 1 min, followed by 40 cycles of 94°C for 30 s, 60–65°C for 30 s, 72°C for 30 s, and a final extension at 72°C for 7 min, using TaKaRa Ex TaqHS (TaKaRa, Shiga, Japan) according to the manufacturer's protocol. Dimethyl sulfoxide (5–10%, v/v) was added for the amplification of GC-rich regions. PCR products were sequenced using a BigDye Terminator Cycle Sequencing Ready Reaction Kit v3.1 (Applied Biosystems, Foster City, CA), and run on an ABI 3100 automated sequencer (Applied Biosystems). Sequencing primers were designed inside the amplification primers. Sequencing electropherograms were aligned by ATGC software version 3.0 (Genetyx, Tokyo, Japan), and single nucleotide polymorphisms (SNPs) or mutations were analyzed visually. Patient's parents and normal control individuals were genotyped using TaqMan assay on ABI 7900HT (Applied Biosystems) or direct sequencing on an ABI3100 sequencer (Applied Biosystems).

## Statistical calculation for case-control study and TDT

Case-control analysis was performed for individual SNPs detected during the mutation search, or for haplotypes constructed from SNPs identified within a linkage disequilibrium block. All data from the analysis were calculated using SNPAlyze version 4.0 statistical software package (Dynacom, Mobara, Japan). Because many low-frequency haplotypes tended to show significant *P*-values, we adopted relatively common haplotypes (frequency $>0.05$) for further statistical analysis. In case-control analysis, we used the standard Bonferroni correction to adjust for multiple testing. We divided a type I error significant at level 0.05 by the number of the independent tests to give a Bonferroni-corrected *P*-value. TDT using individual SNPs or haplotypes was

performed with FBAT version 1.5.5 software (http://www.biostat.harvard.edu/~clange/default.htm).

## Results

Our mutation search revealed a possibly causative missense mutation, 640A > G, in exon 3 of *PAX9* in two children from one family with nonsyndromic CL/P and their phenotypically normal mother. This mutation results in an amino acid change from serine to glycine at position 214 (S214G), and was not found among a total of 474 control samples (Fig. 1).

A total of 66 SNPs (23 SNPs in *TGFB3*, 4 in *DLX3*, 4 in *PAX9*, 12 in *CLPTM1*, 10 in *TBX10*, 9 in *PVRL1*, and 4 SNPs in *TBX22*), including 23 novel polymorphisms other than those registered in the NCBI (http://www.ncbi.nlm.nih.gov/SNP) and JSNP (http://snp.ims.u-tokyo.ac.jp/) databases were found during the mutation search (Table 1). Among these genes, those whose frequencies corresponded to a Hardy–Weinberg distribution in the normal control, and those showing linkage disequilibrium (LD) block structures, were used for case-control analysis in CL/P cases (Fig. 2). Data for CPO cases is not presented here as the number of such samples was too small.

The case-control analysis revealed that 4 *TGFB3* SNPs showed a significant association ($P < 0.01$) with CL/P (Table 1). The lowest *P*-value was obtained at IVS1 + 5321 ($P = 0.0016$) and the second lowest at IVS1 + 2118 ($P = 0.0024$). Association at IVS1 + 5321 was not significant after the standard Bonferroni correction ($0.0016 > 0.0008 = 0.05/66$), but it was still significant when we divided by the number of SNPs in *TGFB3* ($0.0016 < 0.0022 = 0.05/23$). Under such correction, IVS1 + 2118 is unlikely to hold statistical significance. We used these two SNPs for subsequent TDT, because their minor allele frequencies of > 0.05 were sufficient to obtain statistically reliable analysis. Consequently, a significant association ($P = 0.0412$) was obtained at IVS1 + 5321, albeit insignificant at IVS1 + 2118 ($P = 0.7630$) (Table 2). The odds ratio for each SNP genotype was calculated, assuming that SNPs in *TGFB3* have dominant or recessive effects. Table 3 shows the odds ratio from *TGFB3*-SNPs with the minor allele frequency > 0.05 among CL/P patients and with $P < 0.05$ in case-control analysis. Eight SNPs showed significant associations between "major-homozygotes" and "major-heterozygotes plus minor-homozygotes". Moreover, IVS1 + 5321 showed the highest odds ratio of 2.34 [95% confidence interval (CI) = 1.40–3.93]. These results suggest that *TGFB3* plays a role with a recessive effect in the development of CL/P in the Japanese. In the other six candidate genes analyzed, no association was observed.

We next performed haplotype-based association analysis by selecting SNPs with minor allele frequency > 0.2 and D′-value > 0.5 within an LD block at the *TGFB3* locus. Table 4 shows the results for two-SNP haplotypes including IVS1 + 5321 with a frequency of > 0.05 in CL/P patients and with $P < 0.01$. A haplotype "A/A" for the IVS1 + 5321/VS1−1572 loci gave the lowest *P*-value ($P = 0.00055$), being lower than the Bonferroni-corrected *P*-value ($0.00104 = 0.05/48$). Haplotype "A/A" constructed from IVS1 + 5321/IVS1−1572 was more frequent in CL/P patients, although IVS1−1572 alone did not show a significant *P*-value in either the population-based analysis or TDT (Tables 1, 2). TDT using haplotypes "A/A" for IVS1 + 5321 and IVS1−1572 showed a significant association with CL/P ($P = 0.0252$) (Table 5). The *P*-values obtained from haplotype-based association studies were lower than analyses using individual SNPs.

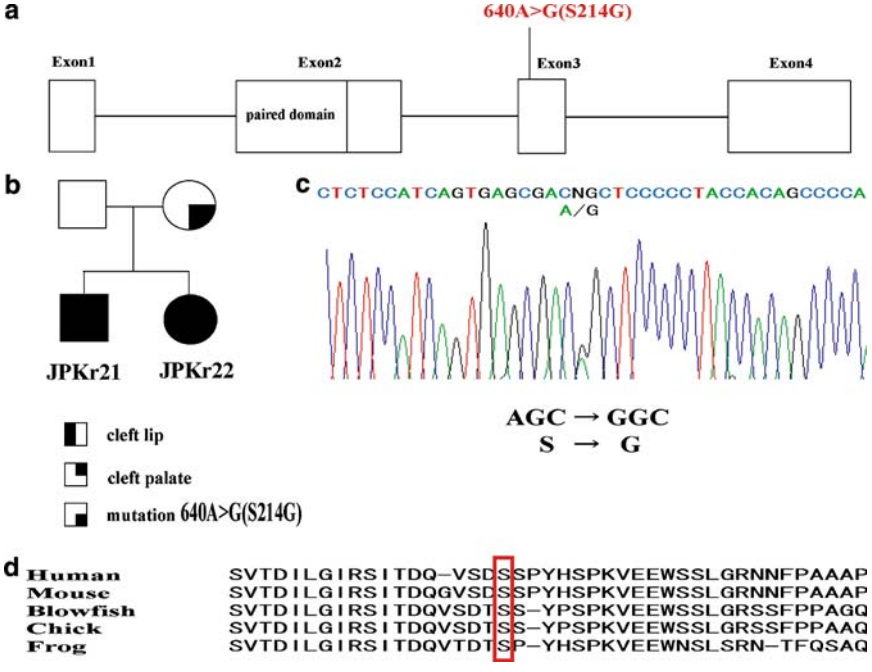


**Fig. 1a–d** A missense mutation found in *PAX9* in a Japanese family. **a** Schematic representation of the structure of *PAX9* showing the position of the mutation 640A > G (S214G) in exon 3. **b** Two children (JPKr21 and JPKr22) with CL/P and their mother without CL/P had the 640A > G (S214G) mutation. **c** Sequencing profile showing the A to G substitution in exon 3 of *PAX9*, resulting in a serine to glycine substitution (S214G). **d** Consensus amino acid sequence at the mutation site (*boxed in red*) of *PAX9* in human, mouse, blowfish, chick, and frog

**Table 1** Single nucleotide polymorphisms (SNPs) in seven candidate genes and their allele frequencies between CL/P patients and normal control individuals

| Gene and SNP position | SNP ID in NCBI | Nucleotide change in major/minor alleles | Minor allele frequency | | *P*-value[a] |
|---|---|---|---|---|---|
| | | | CL/P | Control | |
| *TGFB3* | | | | | |
| 88T > C | rs11466415 | T/C | 0.053 | 0.057 | 0.8620 |
| IVS1 + 2118 | rs2268626 | A/G | 0.317 | 0.359 | 0.0024 |
| IVS1 + 2183 | | C/T | 0.000 | 0.034 | 0.0047 |
| IVS1 + 3137 | rs3917161 | T/C | 0.050 | 0.052 | 0.8787 |
| IVS1 + 3306 | rs2284792 | T/C | 0.394 | 0.466 | 0.0233 |
| IVS1 + 3336 | rs3917162 | C/T | 0.041 | 0.049 | 0.6380 |
| IVS1 + 5321 | rs3917168 | A/T | 0.377 | 0.462 | 0.0016 |
| IVS1 + 5417 | rs3917169 | G/A | 0.377 | 0.462 | 0.0016 |
| IVS1−1778 | | T/C | 0.014 | 0.014 | 0.9490 |
| IVS1−1779 | | C/T | 0.014 | 0.014 | 0.9490 |
| IVS1−1572 | rs2268625 | A/G | 0.344 | 0.408 | 0.0239 |
| IVS1−1283 | rs2268624 | G/C | 0.344 | 0.408 | 0.0239 |
| IVS1−1146 | | A/T | 0.010 | 0.008 | 0.8504 |
| IVS1−1092 | rs3917172 | G/A | 0.043 | 0.051 | 0.6762 |
| IVS1−952 | rs2268623 | G/C | 0.344 | 0.408 | 0.0239 |
| IVS3 + 1041 | rs2268622 | A/G | 0.347 | 0.412 | 0.0271 |
| IVS4 + 256 | rs3917192 | G/A | 0.350 | 0.422 | 0.0922 |
| IVS4 + 397 | rs3917194 | A/T | 0.055 | 0.054 | 0.9488 |
| IVS5 + 104 | rs3917201 | A/G | 0.477 | 0.508 | 0.0942 |
| IVS5 + 289 | rs2359994 | T/G | 0.371 | 0.398 | 0.3219 |
| IVS5 + 527 | rs3917205 | C/T | 0.110 | 0.120 | 0.9378 |
| IVS5 + 621 | rs2284791 | C/G | 0.384 | 0.354 | 0.3273 |
| IVS5−360 | rs3917210 | T/A | 0.339 | 0.384 | 0.5065 |
| *DLX3*[b] | | | | | |
| –64T > C | rs2278163 | T/C | 0.300 | 0.286 | 0.9174 |
| 138C > T | | C/T | 0.222 | 0.247 | 0.3529 |
| IVS1 + 195 | | G/A | 0.232 | 0.224 | 0.1110 |
| IVS1 + 347 | rs11079883 | G/A | 0.329 | 0.410 | 0.5062 |
| *PAX9* | | | | | |
| IVS1−54 | rs12883049 | G/A | 0.276 | 0.303 | 0.2890 |
| IVS1−41 | rs12882923 | A/G | 0.280 | 0.239 | 0.4698 |
| 717C > T | rs12881240 | C/T | 0.281 | 0.254 | 0.7720 |
| 718G > C | rs4904210 | G/C | 0.437 | 0.445 | 0.1417 |
| *CLPTM1* | | | | | |
| 264T > C | rs204481 | T/C | 0.101 | 0.146 | 0.1812 |
| IVS4 + 5 | | G/A | 0.032 | 0.032 | 0.9999 |
| IVS4 + 46 | | A/C | 0.063 | 0.080 | 0.6244 |
| IVS4 + 97 | rs204479 | T/C | 0.104 | 0.139 | 0.2305 |
| IVS5−42 | rs204472 | A/G | 0.101 | 0.146 | 0.1812 |
| IVS6 + 6 | rs204471 | T/C | 0.101 | 0.146 | 0.1812 |
| 927A > G | rs3786565 | A/G | 0.464 | 0.446 | 0.5546 |
| 993T > C | rs204468 | T/C | 0.432 | 0.403 | 0.7587 |
| IVS9−18 | rs875255 | C/G | 0.417 | 0.383 | 0.6845 |
| IVS9−11 | | C/T | 0.009 | 0.003 | 0.5639 |
| IVS13 + 10 | rs2075618 | G/C | 0.459 | 0.434 | 0.5155 |
| IVS13 + 24 | rs2075619 | A/G | 0.473 | 0.444 | 0.6513 |
| *TBX10* | | | | | |
| IVS1−86 | | G/A | 0.005 | 0.000 | 0.4057 |
| IVS2 + 12 | | G/A | 0.009 | 0.000 | 0.1834 |
| 302A > C | rs3758938 | A/C | 0.085 | 0.113 | 0.1365 |
| 549C > T | | C/T | 0.027 | 0.016 | 0.6572 |
| IVS4 + 25 | rs12787511 | C/G | 0.085 | 0.113 | 0.1365 |
| IVS6−87 | rs3765088 | G/A | 0.064 | 0.086 | 0.4629 |
| IVS6−31 | rs11227869 | C/A | 0.378 | 0.406 | 0.5591 |
| IVS7 + 37 | | G/T | 0.027 | 0.013 | 0.4814 |
| IVS8 + 195 | rs4394845 | T/C | 0.062 | 0.086 | 0.4414 |
| IVS8 + 214 | rs1531514 | T/A | 0.384 | 0.406 | 0.5455 |

**Table 1** (Contd.)

| Gene and SNP position | SNP ID in NCBI | Nucleotide change in major/minor alleles | Minor allele frequency | | *P*-value[a] |
|---|---|---|---|---|---|
| | | | CL/P | Control | |
| *PVRL1* | | | | | |
| IVS2 + 53 | | G/C | 0.009 | 0.008 | 0.9918 |
| IVS2−131 | rs2292293 | A/C | 0.468 | 0.470 | 0.2490 |
| 626G > A | | G/A | 0.014 | 0.003 | 0.2943 |
| IVS3−123 | rs10892429 | C/T | 0.191 | 0.176 | 0.8352 |
| 826G > A | | C/A | 0.009 | 0.013 | 0.8929 |
| IVS4−107 | | C/G | 0.005 | 0.000 | 0.4215 |
| IVS5−19 | | A/G | 0.005 | 0.000 | 0.4248 |
| 1326C > T | | C/T | 0.005 | 0.000 | 0.4248 |
| 1337G > A | | G/A | 0.005 | 0.000 | 0.4248 |
| *TBX22*[c] | | | | | |
| IVS4 + 121 | rs195294 | T/C | 0.114 | 0.069 | 0.4069 |
| IVS5−61 | | T/C | 0.011 | 0.008 | 0.9487 |
| IVS7 + 56-57del | | C/- | 0.489 | 0.467 | 0.9274 |
| 1169C > A | | C/A | 0.023 | 0.008 | 0.4547 |

[a]Significant Bonferroni-corrected *P*-value is 0.00075, after dividing 0.05 by the number of all SNPs (0.05/66)
[b]Results of locus after exclusion of exon 2 of *DLX3* due to its high GC component rate and because the Hardy–Weinberg equilibrium was skewed in controls
[c]Statistical analysis of *TBX22* shown applies only to females

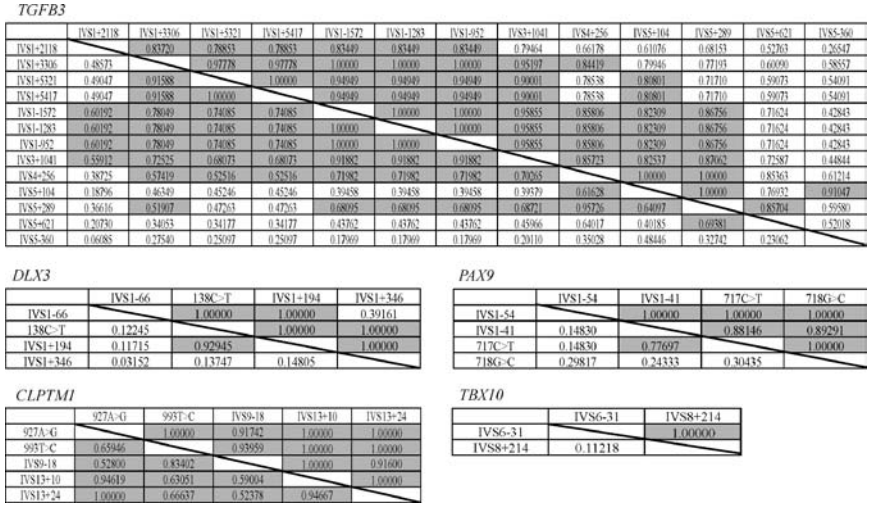


**Fig. 2** Linkage disequilibrium between single nucleotide polymorphisms (SNPs) of five of the seven candidate genes in Japanese CL/P patients. Data for two other genes or for SNPs with minor allele frequency (< 0.2 in CL/P patients) were excluded. Absolute values of D′ and $r^2$ calculated using the statistical software package SNPAlyze version 4.0 (Dynacom, Mobara, Japan) are shown above and below the diagonal, respectively. Values of $|D'| > 0.8$ and of $r^2 > 0.5$ are *highlighted in gray*

## Discussion

In the present study, we performed comprehensive genetic analysis, including mutation searches and association studies, of seven candidate genes linked to CL/P and CPO in a Japanese population. The mutation search identified a missense mutation, 640A > G (S214G), in exon 3 of *PAX9* in two children with CL/P and in their mother without CL/P. The mutation site in this family is located at an amino acid residue that is conserved across species, but is outside the obvious functional domain of *PAX9*. Since it was not found among 474 Japanese controls, the 640A > G mutation is likely causative for CL/P in these children, albeit inconclusively. Since many *PAX9* mutations have been reported to be associated with molar missing (Stockton et al. 2000; Das et al. 2002) or with molar-hypodontia with CL/P (Das et al. 2003), *PAX9* may play a role in molar genesis in humans. Unfortu-

**Table 2** Results of transmission disequilibrium test (TDT) using three SNPs in *TGFB3*

| SNP position | SNP ID in NCBI | Major/minor alleles | *P*-value[a] |
|---|---|---|---|
| IVS1 + 2118 | rs2268626 | A/G | 0.7630 |
| IVS1 + 5321 | rs2300607 | A/T | 0.0412 |
| IVS1−1572 | rs2268625 | A/G | 0.5601 |

[a]Calculated using FBAT version 1.5.5

**Table 3** Odds ratios with 95% confidence intervals (CI) of SNP genotypes in *TGFB3*. Only those SNPs with frequencies >0.05 in CL/P patients and $P < 0.05$ in case-control analysis are listed. Two (IVS1 + 5321 and IVS1 + 5417) and three SNPs (IVS1−1572 and IVS1−1283 and IVS1−952) in *TGFB3* are in complete linkage disequilibrium (LD; $r^2 = 1$)

| SNP position | Genotype | Number of individuals (%) | | Odds ratio (95% CI) |
|---|---|---|---|---|
| | | CL/P | Control | |
| IVS1 + 2118 | AA | 57 (51) | 70 (37) | 1.81 (1.13–2.90)[a] |
| | AG | 39 (35) | 106 (55) | 0.54 (0.26–1.14)[b] |
| | GG | 16 (14) | 16 (8) | |
| IVS1 + 3306 | TT | 42 (39) | 47 (24) | 1.96 (1.18–3.26)[a] |
| | TC | 47 (43) | 111 (58) | 1.00 (0.54–1.87)[b] |
| | CC | 19 (18) | 34 (17) | |
| IVS1 + 5321 | AA | 44 (43) | 45 (24) | 2.34 (1.40–3.93)[a] |
| | AT | 39 (38) | 108 (59) | 1.05 (0.54–2.03)[b] |
| | TT | 19 (19) | 31 (17) | |
| IVS1 + 5417 | GG | 44 (43) | 45 (24) | 2.34 (1.40–3.93)[a] |
| | GA | 39 (38) | 108 (59) | 1.05 (0.54–2.03)[b] |
| | AA | 19 (19) | 31 (17) | |
| IVS1−1572 | AA | 50 (46) | 60 (32) | 1.83 (1.13–2.98)[a] |
| | AG | 43 (39) | 105 (55) | 0.88 (0.45–1.73)[b] |
| | GG | 16 (15) | 25 (13) | |
| IVS1−1283 | GG | 50 (46) | 60 (32) | 1.83 (1.13–2.98)[a] |
| | GC | 43 (39) | 105 (55) | 0.88 (0.45–1.73)[b] |
| | CC | 16 (15) | 25 (13) | |
| IVS1−952 | GG | 50 (46) | 60 (32) | 1.83 (1.13–2.98)[a] |
| | GC | 43 (39) | 105 (55) | 0.88 (0.45–1.73)[b] |
| | CC | 16 (15) | 25 (13) | |
| IVS3 + 1041 | AA | 50 (45) | 58 (31) | 1.90 (1.17–3.09)[a] |
| | AG | 45 (41) | 105 (56) | 0.91 (0.46–1.79)[b] |
| | GG | 16 (14) | 25 (13) | |

[a]Major/Major (M/M) genotype versus Major/minor (M/m) plus minor/minor (m/m) genotypes
[b]M/M plus M/m genotypes versus m/m genotype

nately, re-evaluation of dental anomalies in the mother in this case was not possible as her clinical and examination data were unavailable. It is not surprising that a single gene exhibits variable expression affecting dental anomalies and CL/P, e.g., *MSX1* mutations lead to various phenotypes of hypodontia, CL/P and/or CPO (van den Boogaard et al. 2000). Variable expression levels depend on penetrance, pleiotropic effects of the gene, or expression of modifier gene(s) (Carinci et al. 2003). A new conception recently proposed that phenotypic effects in most single-gene defects may result from the combined actions of oligo-locus alleles (Badano and Katsanis 2002).

Among the seven candidate genes analyzed for a possible association with CL/P, several SNPs in *TGFB3* showed significant results, especially at the IVS1 + 5321 ($P = 0.0016$) and IVS1 + 2118 ($P = 0.0024$) sites (Table 1). Allele A at IVS1 + 5321 and allele A at IVS1 + 2118 were both seen more frequently in CL/P patients than allele G and allele T, respectively. Moreover, IVS1 + 5321 showed the highest odds ratio of 2.34 (95% CI = 1.40–3.93) under an assumption of a recessive effect (Table 3). SNPs surrounding IVS1 + 5321 also showed high odds ratio under the same assumption. From these results, it is suggested that the major SNP allele in *TGFB3* has a recessive effect for a risk of CL/P in the Japanese. The association at IVS1 + 5321 was also confirmed by both TDT (Table 2) and the case-control analysis using two-SNP haplotypes that include IVS1 + 5321 (Table 4). All these results

**Table 4** Case-control analysis using haplotypes including IVS1 + 5321 in *TGFB3*. Only haplotypes with frequencies >0.05 in CL/P patients and $P < 0.01$ are listed. Three SNPs (IVS1−1572, IVS1−1283 and IVS1−952) in *TGFB3* are in complete LD ($r^2 = 1$). Significant Bonferroni-corrected *P*-value is 0.00104, after dividing 0.05 by the number of haplotypes constructed from SNPs in *TGFB3* (0.05/48)

| Haplotype-constructing SNPs | Haplotype | Frequency of haplotype in | | *P*-value |
|---|---|---|---|---|
| | | CL/P | Control | |
| IVS1 + 5321/IVS1−1572 | A/A | 0.6183 | 0.4669 | 0.00055 |
| | T/A | 0.0549 | 0.1265 | 0.00677 |
| IVS1 + 5321/IVS1−1283 | A/G | 0.6183 | 0.4669 | 0.00055 |
| | T/G | 0.0549 | 0.1265 | 0.00677 |
| IVS1 + 5321/IVS1−952 | A/G | 0.6183 | 0.4669 | 0.00055 |
| | T/G | 0.0549 | 0.1265 | 0.00677 |
| IVS1 + 5321/IVS3 + 1041 | A/A | 0.6079 | 0.4636 | 0.00102 |
| IVS1 + 5321/IVS4 + 256 | A/G | 0.5712 | 0.4291 | 0.00159 |

**Table 5** TDT using four haplotypes consisting of IVS1 + 5321 and IVS1−1572 in *TGFB3*

| Haplotype[a] | *P*-value[b] |
|---|---|
| A/A | 0.0252 |
| T/G | 0.2811 |
| T/A | 0.0898 |
| A/G | 0.3992 |

[a]IVS1 + 5321/IVS1−1572
[b]Calculated using FBAT version 1.5.5

indicate that *TGFB3* is one of the genes linked to susceptibility to CL/P in the Japanese. Because *P*-values obtained by haplotype analysis were much lower than those obtained by analysis with individual SNPs, the most important SNP affecting CL/P may be located somewhere in an LD block defined by the IVS1 + 5321/IVS1−1572 haplotype in *TGFB3*.

There was a discrepancy between the case-control analysis and TDT in the results of IVS1 + 2118: the former indicated a significant association ($P = 0.0024$), while the latter gave an insignificant result ($P = 0.7630$) (Tables 1, 2). This may reflect the different statistical power between case-control analysis and TDT. With this assumption, a power calculation was performed using 112 trios (or cases) and 192 controls using Genetic Power Calculator (http://statgen.iop.kcl.ac.uk/gpc/; Purcell et al. 2003). When we applied a relative risk of 1.8 with a recessive effect, the 80% power was 0.62 for case-control analysis and 0.52 for TDT. The discrepancy may be due to the small number of samples examined, and thus we may have to collect more subjects for association analysis and TDT. In other words, we need 180 samples for the case-control test and 220 trios to achieve 80% power at type I error of 0.05. Population-based association studies are more sensitive than family-based studies such as TDT; however, if a positive association is ever obtained from a family-based study, it will provide strong evidence (Mitchell et al. 2002).

There have been many association studies on *TGFB3* (Lidral et al. 1997; Tanabe et al. 2000; Sato et al. 2001; Beaty et al. 2002; Kim et al. 2003). Among them, four reports showed a positive association (Sato et al. 2001; Beaty et al. 2002; Kim et al. 2003), while others obtained negative results. Beaty et al. (2002) revealed a significant association between the *D14S61* locus that is 100-kb from the 3′ end of *TGFB3* and nonsyndromic CPO in Caucasian, African American, and other populations by TDT. In the Japanese population, Sato et al. (2001) revealed a positive association between nonsyndromic CL/P and a CA repeat polymorphic marker 60-kb from the 5′ end of *TGFB3* (Lidral et al. 1997). Kim et al. (2003) reported that the allele G at the IVS5 + 104 site in *TGFB3* increased the risk (odds ratio = 15.92) of nonsyndromic CL/P in the Korean population, although their finding conflicts with the result of our case-control analysis in the Japanese (Table 1). To our knowledge, our study is the first to report a positive association between CL/P and *TGFB3*-SNPs reproduced by both population- and family-based analyses. Therefore, the results of our study are expected to be reliable and reproducible among Japanese CL/P patients.

Regarding the five genes examined in addition to *PAX9* and *TGFB3* (*DLX3*, *CLPTM1*, *TBX10*, *PVRL1*, and *TBX22*), mutations have been reported in some syndromic oral clefting patients, such as in familial CLPED1 patients from Margarita Island, Israel and from Brazil (Suzuki et al. 2000) as well as in patients with X-linked cleft palate with ankyloglossia (CPX) and CPO (Braybrook et al. 2001; Marcano et al. 2004). However, in the present study, no causative mutation or positive association was observed between oral clefts and individual SNPs in these genes. Because *TBX22* maps to the X-chromosome, we performed statistical calculations for males and females individually, but no significant association was observed in either case. It is likely that these genes are not major factors playing a role in oral clefts in the Japanese.

In conclusion, we identified a novel mutation in *PAX9* that may contribute to the development of nonsyndromic CL/P. We also demonstrated, by both population- and family-based analyses, a positive association between *TGFB3* and nonsyndromic CL/P in the Japanese. Our results will assist in understanding the development and prevention of CL/P and CPO.

**Acknowledgements** The participation of patients, their families and control individuals in the study is greatly appreciated. We also thank staff members of Tokyo Dental College Hospital and Nagasaki University Hospital who assisted in collecting blood samples and those of the Department of Human Genetics at Nagasaki University for excellent technical assistance. K. Yoshiura was supported in part by a Grant-in-Aid for Scientific Research on Priority Areas (Medical Genome Science, No. 16012101); N. Niikawa by a Grant-in-Aid for Scientific Research (Category S, No. 13854024) and CREST from the Japan Science and Technology Agency Corporation (JST).

## References

Badano JL, Katsanis N (2002) Beyond Mendel: an evolving view of human genetic disease transmission. Nat Rev Genet 3:779–789

Beaty TH, Hetmanski JB, Zeiger JS, Fan YT, Liang KY, VanderKolk CA, McIntosh I (2002) Testing candidate genes for non-syndromic oral clefts using a case-parent trio design. Genet Epidemiol 22:1–11

Braybrook C, Doudney K, Marcano AC, Arnason A, Bjornsson A, Patton MA, Goodfellow PJ, Moore GE, Stanier P (2001) The T-box transcription factor gene *TBX22* is mutated in X-linked cleft palate and ankyloglossia. Nat Genet 29:179–183

Braybrook C, Lisgo S, Doudney K, Henderson D, Marcano AC, Strachan T, Patton MA, Villard L, Moore GE, Stanier P, Lindsay S (2002) Craniofacial expression of human and murine *TBX22* correlates with the cleft palate and ankyloglossia phenotype observed in CPX patients. Hum Mol Genet 11:2793–2804

Bryan JT, Morasso MI (2000) The *Dlx3* protein harbors basic residues required for nuclear localization, transcriptional activity and binding to *Msx1*. J Cell Sci 113:4013–4023

Bush JO, Lan Y, Jiang R (2004) The cleft lip and palate defects in Dancer mutant mice result from gain of function of the *Tbx10* gene. Proc Natl Acad Sci USA 101:7022–7027

Bustos T, Simosa V, Pinto-Cisternas J, Abramovits W, Jolay L, Rodriguez L, Fernandez L, Ramela M (1991) Autosomal recessive ectodermal dysplasia: I. An undescribed dysplasia/malformation syndrome. Am J Med Genet 41:398–404

Carinci F, Pezzetti F, Scapoli L, Martinelli M, Avantaggiato A, Carinci P, Padula E, Baciliero U, Gombos F, Laino G, Rullo R, Cenzi R, Carls F, Tognon M (2003) Recent developments in orofacial cleft genetics. J Craniofac Surg 14:130–143

Cobourne MT (2004) The complex genetics of cleft lip and palate. Eur J Orthod 26:7–16

Das P, Stockton DW, Bauer C, Shaffer LG, D'Souza RN, Wright T, Patel PI (2002) Haploinsufficiency of *PAX9* is associated with autosomal dominant hypodontia. Hum Genet 110:371–376

Das P, Hai M, Elcock C, Leal SM, Brown DT, Brook AH, Patel PI (2003) Novel missense mutations and a 288-bp exonic insertion in *PAX9* in families with autosomal dominant hypodontia. Am J Med Genet A 118:35–42

Freedman ML, Reich D, Penney KL, McDonald GJ, Mignault AA, Patterson N, Gabriel SB, Topol EJ, Smoller JW, Pato CN, Pato MT, Petryshen TL, Kolonel LN, Lander ES, Sklar P, Henderson B, Hirschhorn JN, Altshuler D (2004) Assessing the impact of population stratification on genetic association studies. Nat Genet 36:388–393

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

Juriloff DM, Harris MJ, Brown CJ (2001) Unravelling the complex genetics of cleft lip in the mouse model. Mamm Genome 12:426–435

Kim MH, Kim HJ, Choi JY, Nahm DS (2003) Transforming growth factor-beta3 gene SfaN1 polymorphism in Korean nonsyndromic cleft lip and palate patients. J Biochem Mol Biol 36:533–537

Kondo S, Schutte BC, Richardson RJ, Bjork BC, Knight AS, Watanabe Y, Howard E, de Lima RL, Daack-Hirsch S, Sander A, McDonald-McGinn DM, Zackai EH, Lammer EJ, Aylsworth AS, Ardinger HH, Lidral AC, Pober BR, Moreno L, Arcos-Burgos M, Valencia C, Houdayer C, Bahuau M, Moretti-Ferreira D, Richieri-Costa A, Dixon MJ, Murray JC (2002) Mutations in *IRF6* cause Van der Woude and popliteal pterygium syndromes. Nat Genet 32:285–289

Kraus P, Lufkin T (1999) Mammalian *Dlx* homeobox gene control of craniofacial and inner ear morphogenesis. J Cell Biochem 32–33[Suppl]:133–140

Law DJ, Garvey N, Agulnik SI, Perlroth V, Hahn OM, Rhinehart RE, Gebuhr TC, Silver LM (1998) *TBX10*, a member of the *Tbx1*-subfamily of conserved developmental genes, is located at human chromosome 11q13 and proximal mouse chromosome 19. Mamm Genome 9:397–399

Lidral AC, Murray JC, Buetow KH, Basart AM, Schearer H, Shiang R, Naval A, Layda E, Magee K, Magee W (1997) Studies of the candidate genes *TGFB2*, *MSX1*, *TGFA*, and *TGFB3* in the etiology of cleft lip and palate in the Philippines. Cleft Palate Craniofac J 34:1–6

Marcano AC, Doudney K, Braybrook C, Squires R, Patton MA, Lees MM, Richieri-Costa A, Lidral AC, Murray JC, Moore GE, Stanier P (2004) *TBX22* mutations are a frequent cause of cleft palate. J Med Genet 41:68–74

Mitchell LE, Beaty TH, Lidral AC, Munger RG, Murray JC, Saal HM, Wyszynski DF (2002) International consortium for oral clefts genetics. Guidelines for the design and analysis of studies on nonsyndromic cleft lip and cleft palate in humans: summary report from a workshop of the international consortium for oral clefts genetics. Cleft Palate Craniofac J 39:93–100

Natsume N, Kawai T, Kohama G, Teshima T, Kochi S, Ohashi Y, Enomoto S, Ishii M, Nakano Y, Matsuya T, Kogo M, Yoshimura Y, Ohishi M, Nakamura N, Katsuki T, Goto M, Shimizu M, Yanagisawa S, Mimura T, Sunakawa H (2000) Incidence of cleft lip or palate in 303738 Japanese babies born between 1994 and 1995. Br J Oral Maxillofac Surg 38:605–607

Papaioannou VE, Silver LM (1998) The T-box gene family. Bioessays 20:9–19

Peters H, Neubuser A, Kratochwil K, Balling R (1998) *Pax9*-deficient mice lack pharyngeal pouch derivatives and teeth and exhibit craniofacial and limb abnormalities. Genes Dev 12:2735–2747

Prescott NJ, Lees MM, Winter RM, Malcolm S (2000) Identification of susceptibility loci for nonsyndromic cleft lip with or without cleft palate in a two stage genome scan of affected sib-pairs. Hum Genet 106:345–350

Price JA, Bowden DW, Wright JT, Pettenati MJ, Hart TC (1998) Identification of a mutation in *DLX3* associated with tricho-dento-osseous (TDO) syndrome. Hum Mol Genet 7:563–569

Proetzel G, Pawlowski SA, Wiles MV, Yin M, Boivin GP, Howles PN, Ding J, Ferguson MW, Doetschman T (1995) Transforming growth factor-beta 3 is required for secondary palate fusion. Nat Genet 11:409–414

Purcell S, Cherny SS, Sham PC (2003) Genetic power calculatior: design of linkage and association genetic mapping studies of complex traits. Bioinformatics 19:149–150

Sato F, Natsume N, Machido J, Suzuki S, Kawai T (2001) Association between transforming growth factor beta 3 and cleft lip and/or palate in the Japanese population. Plast Reconstr Surg 107:1909–1910

Schutte BC, Murray JC (1999) The many faces and factors of orofacial clefts. Hum Mol Genet 8:1853–1859

Sozen MA, Suzuki K, Tolarova MM, Bustos T, Fernandez Iglesias JE, Spritz RA (2001) Mutation of *PVRL1* is associated with sporadic, non-syndromic cleft lip/palate in northern Venezuela. Nat Genet 29:141–142

Stockton DW, Das P, Goldenberg M, D'Souza RN, Patel PI (2000) Mutation of *PAX9* is associated with oligodontia. Nat Genet 24:18–19

Suzuki K, Hu D, Bustos T, Zlotogora J, Richieri-Costa A, Helms JA, Spritz RA (2000) Mutations of *PVRL1*, encoding a cell–cell adhesion molecule/herpes virus receptor, in cleft lip/palate-ectodermal dysplasia. Nat Genet 25:427–430

Takeuchi T, Kuro-o M, Miyazawa H, Ohtsuki Y, Yamamoto H (1997) Transgenic expression of a novel thymic epithelial cell antigen stimulates abberant development of thymocytes. J Immunol 159:726–733

Tanabe A, Taketani S, Endo-Ichikawa Y, Tokunaga R, Ogawa Y, Hiramoto M (2000) Analysis of the candidate genes responsible for non-syndromic cleft lip and palate in Japanese people. Clin Sci 99:105–111

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

Vanderas AP (1987) Incidence of cleft lip, cleft palate, and cleft lip and palate among races: a review. Cleft Palate J 24:216–225

Yoshiura K, Machida J, Daack-Hirsch S, Patil SR, Ashworth LK, Hecht JT, Murray JC (1998) Characterization of a novel gene disrupted by a balanced chromosomal translocation t(2;19)(q11.2;q13.3) in a family with cleft lip and palate. Genomics 54:231–240

Wyszynski DF, Maestri N, McIntosh I, Smith EA, Lewanda AF, Garcia-Delgado C, Vinageras-Guarneros E, Wulfsberg E, Beaty TH (1997) Evidence for an association between markers on chromosome 19q and non-syndromic cleft lip with or without cleft palate in two groups of multiplex families. Hum Genet 99:22–26

Zucchero TM, Cooper ME, Maher BS, Daack-Hirsch S, Nepomuceno B, Ribeiro L, Caprau D, Christensen K, Suzuki Y, Machida J, Natsume N, Yoshiura K, Vieira AR, Orioli IM, Castilla EE, Moreno L, Arcos-Burgos M, Lidral AC, Field LL, Liu YE, Ray A, Goldstein TH, Schultz RE, Shi M, Johnson MK, Kondo S, Schutte BC, Marazita ML, Murray JC (2004) Interferon regulatory factor 6 (*IRF6*) gene variants and the risk of isolated cleft lip or palate. N Engl J Med 351:769–780