Nucleotide variants of the NAT2 and EGF61 genes in patients in Northern China with nonsyndromic cleft lip with or without cleft palate

Jiaqun Yan, PhD, Hongquan Song, PhD, Na Mi, PhD, Xiaohui Jiao, PhD, Yanru Hao, PhD

Abstract

Background: Non-syndromic cleft lip with or without cleft palate (NSCL/P) is a common orofacial congenital anomaly. The etiology of NSCL/P is multifactorial. It occurs in humans as a result of genetic and environmental factors. NSCL/P can be divided into 2 general categories: cleft lip with or without palate (CL/P) and cleft palate only (CPO). Cleft lip arises due to failure of the globular process to merge with the maxillary processes (MxP) during the seventh week of gestation, and clefts of the palate can arise from failure of the paired lateral palatine process to fuse with the nasal septum and then merge with the median palatine process during the 10th week of gestation. The factors that cause the failure of fusion are still unknown and may be related to genetic, nutrition, or endocrine disorders, infection, and trauma. The occurrence of CL/P greatly reduces the quality of life for patients and family members. C/CLP has become a prominent public health problem and social problem in China.

Methods: The frequencies of NAT2 (rs1799929) and EGF61 (rs4444903) gene variations were examined in a group of 285 NSCL/P patients and in 315 controls. Peripheral venous blood samples were collected for DNA extraction. Genotyping of the 2 SNPs was carried out using a mini sequencing (SNAPshot) method. Data were analyzed using the chi-square test.

Results: We found a significant association between the EGF61 (rs4444903) and NSCL/P (P = .01) genes. Conversely, NAT2 (rs1799929) was not significantly different between the cases and the control group. The genotype frequencies of rs4444903GA showed a significant difference compared with GG genotype as a reference (odds ratio = 0.59; 95% confidence interval: 0.42–0.84, P = .01).

Conclusion: Our study showed that the EGF61 rs4444903GA genotype had a decreased risk of NSCL/P. Our data provides further evidence regarding the role of EGF61 variations in the development of NSCL/P in families of the studied populations.

Keywords: EGF, NAT2, NSCL/P, polymorphisms, susceptibility gene

1. Introduction

Non-syndromic cleft lip with or without cleft palate (NSCL/P) is a common craniofacial congenital deformity. The etiology of NSCL/P is multifactorial. It occurs in humans as a result of genetic and environmental factors. NSCL/P can be divided into 2 general categories: cleft lip with or without palate (CL/P) and cleft palate only (CPO). Cleft lips arise due to failure of the globular process to merge with the maxillary processes (MxP) during the seventh week of gestation, and clefts of the palate can arise from failure of the paired lateral palatine process to fuse with the nasal septum and then merge with the median palatine process during the 10th week of gestation. The factors that cause the failure of fusion are still unknown and may be related to genetic, nutrition, or endocrine disorders, infection, and trauma. The occurrence of CL/P greatly reduces the quality of life for patients and family members. C/CLP has become a prominent public health problem and social problem in China.

The average prevalence of NSCL/P at birth is estimated to be 1/700, ranging from 1/500 to 1/1000, depending on the economic, geographical, and genetic background.

Compared with the other populations, Asian and Native North American populations have the highest birth prevalence of 2 of 1000 births. Individuals with NSCL/P may experience problems with feeding, digesting, speaking, hearing, and social engagement.

It was also reported that the highest incidence of clefts is among Chinese patients, with 1.42 per 1000 deliveries. Because of the high prevalence rate of CL/P births in China, numerous studies have been performed in Chinese patients.

To date, several association studies have discovered multiple candidate genes linked to NSCL/P formation, such as IRF6, TFP2A, CRISPLD2, ABC4, and JARID2.

Recently, Santos et al detected NAT2 rs1799929 by polymerase chain reaction (PCR) and evaluated the transmission disequilibrium tests (TDTs). TDT showed a positive association between NAT2 rs1799929 and NSCL/P in an Argentinean
population. The presence of NAT2 rs1799929 was significantly higher in cases with congenital NSCL/P.

Meanwhile, Falagan-Lotsch et al[14] have provided evidence that the EGF61 A>G polymorphism was not related to susceptibility to nonsyndromic oral clefts in a Brazilian population, but supported a different genetic background between CL+P and CP.

The main objective of our study was to investigate the role of NAT2 and EGF61 gene polymorphisms in the pathogenesis of NSCL/P in a northern Chinese population.

2. Materials and methods

2.1. Samples

In all, 600 individuals, including 285 NSCL/P patients and 315 normal controls, were included in our study. All the study participants were recruited between 2010 and 2013 from the Smile Train Project: the Affiliated Stomatology Hospital of Harbin Medical University, Harbin Children’s Hospital, and the Second Affiliated Hospital of Harbin Medical University. In all, 315 healthy individuals without a family history of genetic disease were recruited as controls, which were matched to cases for age, sex, and ethnic background. In our study, patients with NSCL/P were divided into 2 subtypes: 189 CL/P and 96 CPO.

The study was approved by the Ethics Committee of the First Affiliated Hospital of Harbin Medical University, and informed consent was obtained from each participant. The study followed the Declaration of Helsinki’s recommendations for medical protocol and ethics.

2.2. Genotype analysis

We selected 2 single-nucleotide polymorphisms (SNPs), rs1799929 in NAT2 and rs4444903 in EGF61, according to the research of Santos et al[11] and Falagan-Lotsch et al[14].

DNA in peripheral blood samples from participants were prepared with DNA kits (Applied Axygen Biosciences, Union City, CA) and stored at −80°C for further analysis. The purity and concentration were measured at 260 and 280nm absorbance by spectrophotometric methods. The target DNA sequences were amplified by using multiplex PCRs. The primer pairs for each SNP were designed by Primer3.0 software (Premier Biosoft International, Palo Alto, CA). The PCR system (15 μL) included 10× buffer, 1.5 μL MgCl2, 5 μL of Platinum Taq DNA polymerase (Invitrogen, Carlsbad, CA), 0.3 μL dNTP, 1 μL genomic DNA, and optimal concentrations of primers. The genotyping of 2 SNPs was performed with the ABI PRISM 3730 DNA Sequencer (Applied Biosystems, Foster City, CA), by the mini-sequencing (SNaPshot) method. For rs1799929CT, the 212-bp fragments of rs1799929 were as follows: 5'-GACGGCGAGAATCATGCTAGCTGAC-3' (forward) and 5'-TTTCTCCTTTGGCGAGGAGATGAG-3' (reverse); the probe was AGAAGAGAGGAATCTGGTAC. For rs4444903AG, the 211-bp fragments of rs4444903 were as follows: 5'-GTCATCAAGGAAAGGAGGTG-3' (forward) and 5'-ACTTCAGTGCGACGCGAGGAGT-3' (reverse); the probe was TTTCAGCCCAATCCTCAAAGGTGTG. Approximately 20% of the blood samples were repeated randomly for verification, and the results were consistent with genotype identification.

2.3. Statistical analysis

Hardy–Weinberg equilibrium (HWE) was assessed for all analyzed markers in cases and controls. The chi-square test was used to compare differences in the genotype and allele frequency between patients and control subjects. Allelic odds ratios (ORs) and 95% confidence intervals (CIs) were calculated by the standard chi-square test. P < .05 (2-sided) was regarded as significant, and the OR < 1 was regarded as a decreased risk of disease. Statistical analyses were performed by SAS v9.13 software (SAS Institute Inc). The linkage disequilibrium analysis was performed using Hapview analysis software and SNPStats online software.

3. Results

The characteristics and minor allele frequencies of the 2 SNPs are shown in Table 1. The HWE of the SNPs was calculated among controls. None of the SNPs among these groups deviated from the HWE. The polymorphic allele frequencies of the case group were compared with the controls. The minor allele frequency (MAF) of SNPs (ie, rs1799929 T allele and rs4444903 A allele) among the controls was 0.052 and 0.337. The differences were not statistically significant (P = .05 for rs1799929 and rs4444903).

To compare the genotype frequencies in the NSCL/P and control groups, the ORs and 95% CI calculations for SNPs of NAT2 and EGF61 are shown in Table 2. The frequencies of the CC, CT, and TT genotypes of rs1799929 in the controls were 89.84%, 9.84%, and 0.32%, respectively. The genotype frequencies of CC, CT, and TT in the NSCL/P group were 94.04%, 5.96%, and 0%, respectively. Similarly, the frequencies of the rs4444903 genotypes GG, GA, and AA in the controls were 44.44%, 43.81%, and 11.75%, respectively. The genotype frequencies of GG, GA, and AA in NSCL/P group were 55.44%, 32.28%, and 12.28%, respectively (P = .012). For rs1799929, there was no significant difference in the genotype distribution between cases and controls. The result was statistically significant in the genotype frequencies of the rs4444903GA = 0.59, 95% CI 0.42–0.84; OR 4444903-GA+AA = 0.64, 95% CI 0.47–0.89, P = .012), and the genotype was associated with decreased risks of NSCL/P.

Likewise, for rs4444903 in the CL/P subgroup, the result was statistically significant between the patients and the controls (OR

### Table 1

| SNP ID   | Chromosome position | Base change | MAF for control | HWE P | MAF | P     | MAF | P     | MAF | P     |
|----------|---------------------|-------------|-----------------|--------|-----|-------|-----|-------|-----|-------|
| rs1799929 | 18400484            | C>T         | 0.052           | .878   | 0.03 | .05   | 0.034 | .185  | 0.021 | .065 |
| rs4444903 | 100912054           | G>A         | 0.337           | .737   | 0.284 | .05   | 0.265 | .017  | 0.323 | .727 |

CL/P = cleft lip with or without cleft palate, CPO = cleft palate only, HWE = Hardy–Weinberg equilibrium, MAF = minor allele frequency.

P<.05 is bolded.

*P value by chi-square test.
which is located on chromosome 8 (8p22), together with the rs4444903 — between NAT2 and EGF61 markers.

environmental factors, and their association with a wide range of inactivation of many compounds. NAT2 variants have been

Results of association tests with NAT2 and EGF61 genes single-nucleotide polymorphisms (SNPs) in NSCL/P groups.

| SNP     | Genotype | NSCL/P     | Case/control | OR  | 95% CI | P    | CL/P     | Case/control | OR  | 95% CI | P    | CPO     | Case/control | OR  | 95% CI | P    |
|---------|----------|------------|--------------|-----|--------|------|----------|--------------|-----|--------|------|---------|--------------|-----|--------|------|
| rs1799929 | CC       | 268/283    | 1             | Ref | .135   |      | 176/283  | 1             | Ref | .381   |      | 92/283  | 1             | Ref | .185   |      |
|         | CT       | 17/31      | 0.58          | 0.31–1.07 | 0.67  | 0.34–1.32 |      | 13/31      | 0.65  | 0.33–1.28 |      | 4/31     | 0.4          | 0.14–1.15 |      |
|         | TT       | 0/1        | 0/1           |      | 0/1    |      | 0/1      |      | 0/1    |      |         |      |         |      |         |      |
| rs4444903 | CT + TT  | 17/32      | 0.56          | 0.30–1.03 | 13/32 | 0.65  | 0.33–1.28 |      | 4/32  | 0.38  | 0.13–1.12 |      |
|          | GG       | 158/140    | 1             | Ref | .012   |      | 110/140  | 1             | Ref | .008   |      | 48/140  | 1             | Ref | .33    |      |
|          | GA       | 92/138     | 0.59          | 0.42–0.84 | 58/138 | 0.53  | 0.36–0.79 |      | 34/138 | 0.72  | 0.44–1.18 |      |
|          | AA       | 35/37      | 0.84          | 0.50–1.40 | 21/37 | 0.72  | 0.40–1.30 |      | 14/37  | 1.10  | 0.55–2.22 |      |
|          | GA + AA  | 127/175    | 0.64          | 0.47–0.89 | 79/175 | 0.57  | 0.40–0.83 |      | 48/175 | 0.8   | 0.51–1.26 |      |

D: confidence interval, CL/P = cleft lip with or without palate, CPO = cleft palate only, NSCL/P = nonsyndromic cleft lip with or without cleft palate, OR = odds ratio. P < .05 is bold.

P-value by chi-square test.

4444903-GA=0.53, 95% CI 0.36–0.79; OR 4444903-GA + AA=0.57, 95% CI 0.40–0.83, P=.008). The data did not show any significant differences in the 2 SNPs in the CPO group.

The analysis of pair-wise linkage disequilibrium is shown in Table 3. There was no strong linkage disequilibrium between rs1799929 and rs4444903. Therefore, haplotype analysis was not performed in our study.

4. Discussion

In essence, the occurrence of NSCL/P was a result of the failure of the maxillary processes or palatal shelves between the fourth and twelfth weeks of embryogenesis.[13]

Patients experienced speech, hearing, and dental problems, and also difficulty consuming food. Clefts could be repaired with surgery, and multiple craniofacial and dental surgeries were often performed, along with treatment for speech and hearing.[15,16] Patients may experience the lifelong psychosocial effects of malformations, despite these interventions. Furthermore, individuals born with clefts have an increased incidence of mental health problems throughout their lives.[17]

In our study, we investigated the impact of polymorphisms of the NAT2 (rs1799929) and EGF 61 (rs4444903) genes and the susceptibility to NSCL/P in a northern Chinese Han population.

The human arylamine N-acetyltransferases, NAT1 and NAT2, are important xenobiotic-metabolizing enzymes involved in the detoxification and metabolic activation of numerous drugs and chemicals.[18] The NAT2 isoform is encoded by the NAT2 gene, which is located on chromosome 8 (8p22), together with the NAT1 gene. Erickson et al.[19] found that NAT2 significantly influences teratogen-induced orofacial clefting. NAT2 is an important detoxification enzyme, which is mainly expressed in the liver and intestine, and is involved in the activation or inactivation of many compounds. NAT2 variants have been widely studied to assess their distribution among different ethnic groups, which can also be related to the exposure to different environmental factors, and their association with a wide range of diseases, including those of autoimmune origin.[20] Song et al.[21] found that the NAT2 rs1041983 polymorphism is associated with nonsyndromic orofacial clefts for the first time. A study by Santos et al.[22] reported that the presence of the NAT2 rs1799929CT allele is significantly higher in cases with congenital NSCL/P. The NAT2 gene was related to a high risk of NSCL/P (P = .012) in an Argentinean population, which is different from the value of 0.135 found in our controls. In the present research, the SNPs rs1799929 showed no association with the development of NSCL/P in a northern Chinese population.

Epidermal growth factor (EGF) is encoded by a single gene on chromosome 4q25–q27. EGF is 1 of the ligands that bind to ErbB receptors. Mature EGF is widely present in the extracellular fluid of human tissue. The function of EGF is mainly a mitogenic agent of epidermal cells, mesenchymal cells, and tumor cells, and is 1 of the most stable proteins. It plays a role in promoting growth. EGF activates the EGF pathway through its EGF receptor on the cell membrane, which can promote cell proliferation and differentiation, and inhibit apoptosis. As a potent mitogenic factor for the formation of a variety of cultured cells in the ectoderm and mesoderm, EGF is expressed in several developmental organs. Moreover, EGF is closely related to the development of teeth[23] and mandibular morphology,[24] and also the normal palate. It has been proven that EGF facilitates the fusion of the maxillofacial region and the growth of the primary palate.[25]

In 2002, Shahbazi et al.[25] identified a G-to-A polymorphism at position 61 in the 5’ untranslated region of EGF that is significantly correlated with the production of EGF. Individuals with the GG or AG genotype exhibited higher EGF production in vitro than individuals with the AA genotype, and homozygosity of the G allele was shown to be associated with the risk of developing malignant melanoma and several other malignancies, including gastric cancer[26] and glioma.[27]

Furthermore, an association between rs4444903 and NSCL/P was not identified by Falagan-Lotsch et al.[14]. and rs4444903 was not shown to be associated with NSCL/P in a Brazilian population. In the present research, we found that the SNP rs4444903 was significantly associated with NSCL/P (P = .012), and the EGF 61 (rs4444903) gene was related to a decreased risk of NSCL/P. When we further stratified the NSCL/P groups into 2 subgroups (CL/P and CPO), the results showed that the EGF 61 rs4444903 was significantly associated with CL/P (P = .008), which was inconsistent with the results of Falagan-Lotsch et al.[13].

Our study revealed that these 2 SNPs have different roles in the NSCL/P, and the 2 SNPs failed to show any association with the CPO subgroup. This is consistent with the results of recent genetic studies,[28] which indicated that CL/P and CPO have significantly different associations with a variety of genetic factors.
different genetic causes. Compared with CL/P, CPO shows less variation in newborns. Nearly half of the live births in CPO have another congenital anomaly or some identifiable malformation syndrome. EGF61 may be highly correlated with NSCL/P, particularly in CL/P, in this northern Chinese population, which has not been previously reported. This result may be because the object of the study was a different population with environmental exposure and multiple genetic backgrounds.

There were several limitations in our research. The major one was that the sample size was relatively small. Secondly, our research population was ethnically relatively homogeneous; therefore, further investigations are required.

5. Conclusions
In summary, despite the limitations caused by our small sample size, our data suggested that the EGF61 G allele is strongly associated with functional effects on the NSCL/P in a northern Chinese population. The EGF61 gene was demonstrated to play a role in the development of NSCL/P. Although NAT2 rs1799929 had no association with NSCL/P in a northern Chinese population, other susceptibility genes loci need to be discovered. However, allele frequencies vary among different populations and ethnic backgrounds. Therefore, further studies are required to investigate the functional significance with different ethnic groups.

Acknowledgment
The authors thank all participants who donated blood samples for this research.

References
[1] Asghar E, Roya H, Hamid R, et al. Association of transforming growth factor alpha polymorphisms with nonsyndromic cleft lip and palate in Iranian population. Avicenna J Med Biotech 2015;7:168–72.
[2] Mohamad Shah NS, Salahshourifar I, Sulong S, et al. Discovery of candidate genes for nonsyndromic cleft lip palate through genome-wide linkage analysis of large extended families in the Malay population. BMC Genet 2016;17:39.
[3] Mossey PA, Little J, Munger RG, et al. Cleft lip and palate. Lancet 2009;374:1773–85.
[4] Cooper ME, Ratay JS, Marazita ML. Asian oral-facial cleft birth prevalence. Cleft Palate Craniofac J 2006;43:580–9.
[5] Vieira AR, McHenry TG, Daack-Hirsch S, et al. A genome wide linkage scan for cleft lip and palate and dental anomalies. Am J Med Genet A 2008;146:1406–13.
[6] Carter TC, Molloy AM, Pangilinan F, et al. Testing reported associations of genetic risk factors for oral clefts in a large Irish study population. Birth Defects Res A Clin Mol Teratol 2010;88:84–93.
[7] Simioni M, Acauso TK, Montillo IL, et al. Investigation of genetic factors underlying typical orofacial clefts: mutational screening and copy number variation. J Hum Genet 2015;60:17–25.
[8] Dai L, Zhu J, Mao M, et al. Time trends in oral clefts in Chinese newborn: data from the Chinese National Birth Defects Monitoring Network. Birth Defects Res A Clin Mol Teratol 2010;88:41–7.
[9] Li H, Sheridan R, Williams T. Analysis of TPAP2A mutations in branchio-ocular-facial syndrome indicates functional complexity within the AP-2alpha DNA binding domain. Hum Mol Genet 2013;22:3195–206.
[10] Shi J, Jiao X, Song T, et al. CRISPLD2 polymorphisms are associated with non-syndromic cleft lip with or without cleft palate in a northern Chinese population. Eur J Oral Sci 2010;118:430–3.
[11] Fontoura C, Silva RM, Granjeiro JM, et al. Further evidence of association of the ABCA4 gene with cleft lip/palate. Eur J Oral Sci 2012;120:553–7.
[12] Hao Y, Mi N, Jiao X, et al. Association of JARID2 polymorphisms with non-syndromic orofacial clefts in northern Chinese Han population. J Oral Pathol Med 2015;44:386–91.
[13] Santos MR, Ramallo V, Muzio M, et al. Association between NAT2 polymorphisms with non-syndromic cleft lip with or without cleft palate in Argentina. Rev Med Chile 2013;143:444–50.
[14] Falagan-Losch P, Lopes TX, Kühler EC, et al. The functional EGF+61 Polymorphism and nonsyndromic oral clefts susceptibility in a Brazilian population. J Appl Oral Sci 2015;23:390–6.
[15] Christensen K, Juel K, Herskind AM, et al. Long term follow up study of survival associated with cleft lip and palate at birth. BMJ 2004;328:1405.
[16] Wehby GL, Cassell CH. The impact of orofacial clefts on quality of life and healthcare use and costs. Oral Dis 2010;16:3–10.
[17] Bille C, Knudsen LB, Christensen K. Changing lifestyles and oral clefts occurrence in Denmark. Cleft Palate Craniofac J 2005;42:255–9.
[18] Sim F, Fakis G, Laurieri N, et al. Arylamine N-acetyltransferases: from drug metabolism and pharmacogenomics to identification of novel targets for pharmacological intervention. Adv Pharmacol 2012;65:169–205.
[19] Erickson RP, Cao W, Acuña DK, et al. Confirmation of the role of N-acetyltransferase 2 in teratogen-induced cleft palate using transgenics and knockouts. Mol Reprod Dev 2008;75:1071–6.
[20] Santore EC, Pinto AC, Klumb EM, et al. Polymorphisms in NAT2 (N-acetyltransferase 2) gene in patients with systemic lupus erythematosus. Rev Bras Reumatol 2016;56:521–9.
[21] Song T, Wu D, Wang Y, et al. Association of NAT1 and NAT2 genes with nonsyndromic cleft lip and palate. Mol Med Rep 2013;8:211–6.
[22] Krommiller JE. Spatial distribution of epidermal growth-factor transcripts and effects of exogenous epidermal growth factor on the pattern of the mouse dental lamina. Arch Oral Biol 1995;40:137–43.
[23] Shum I, Sakakura Y, Bringas PJr, et al. EGF abrogation-induced fusiform dysmorphismogenesis of Meckel’s cartilage during embryonic mouse mandibular morphogenesis in vitro. Development 1995;118:903–17.
[24] Iamaroon A, Wallon UM, Overall CM, et al. Expression of 72-kD acetyltransferase 2) gene in patients with systemic lupus erythematosus. Rev Bras Reumatol 2011;51:359–64.
[25] Hamay Y, Matsumura S, Matsuaki K, et al. A single nucleotide polymorphism in the 3’ untranslated region of the EGF gene is associated with occurrence and malignant progression of gastric cancer. Pathobiology 2005;72:133–8.
[26] Costa BM, Ferreira P, Costa S, et al. Association between functional EGF +61 polymorphism and glioma risk. Clin Cancer Res 2007;13:2621–6.
[27] Beatty TH, Ruczinski I, Murray JC, et al. Evidence for gene-environment interaction in a genome wide study of nonsyndromic cleft palate. Genet Epidemiol 2011;35:469–78.