HHS Public Access

Author manuscript

Am J Med Genet A. Author manuscript; available in PMC 2015 August 01.

Published in final edited form as:

Am J Med Genet A. 2015 May ; 167(5): 1054–1060. doi:10.1002/ajmg.a.36912.

Replication of 13q31.1 Association in Nonsyndromic Cleft Lip with Cleft Palate in Europeans

Zhonglin Jia[#1,2,3], **Elizabeth J. Leslie**[#4,*], **Margaret E. Cooper**[4], **Azeez Butali**[5], **Jennifer Standley**[2], **Jennifer Rigdon**[2], **Satoshi Suzuki1**[2,6], **Ayana Gongorjav**[7], **T. Enkhtur Shonkhuuz**[7], **Nagato Natsume**[6], **Bing Shi**[1,3], **Mary L. Marazita**[4], and **Jeffrey C. Murray**[2]

[1] State Key Laboratory of Oral Diseases, West China Hospital of Stomatology, Sichuan University, Chengdu, China

[2] Department of Pediatrics, University of Iowa, Iowa City, Iowa

[3] Department of Cleft Lip and Palate Surgery, West China Hospital of Stomatology, Sichuan University, Chengdu, China

[4] Center for Craniofacial and Dental Genetics, Department of Oral Biology, School of Dental Medicine, University of Pittsburgh, Pittsburgh, Pennsylvania

[5] Department of Oral Pathology, Radiology, and Medicine, College of Dentistry, University of Iowa, Iowa City, Iowa

[6] Division of Research and Treatment for Oral and Maxillofacial Congenital Anomalies, School of Dentistry, Aichi-Gakuin University, Japan

[7] National Center for Maternal and Child Health, Ulaanbaatar, Mongolia

[#] These authors contributed equally to this work.

Abstract

Genome wide association (GWA) studies have successfully identified at least a dozen loci associated with orofacial clefts. However, these signals may be unique to specific populations and require replication to validate and extend findings as a prelude to etiologic SNP discovery. We attempted to replicate the findings of a recent meta-analysis of orofacial cleft GWA studies using four different ancestral populations. We studied 946 pedigrees (3436 persons) of European (US white and Danish) and Asian (Japanese and Mongolian) origin. We genotyped six SNPs which represented the most significant P value associations identified in published studies: rs742071 (1p36), rs7590268 (2p21), rs7632427 (3p11.1), rs12543318 (8q21.3), rs8001641 (13q31.1) and

[*]Correspondence to: Elizabeth J. Leslie, PhD Center for Craniofacial and Dental Genetics University of Pittsburgh 100 Technology Drive, Suite 500 Pittsburgh, PA 15219 Phone: (412) 648-1488 ejl40@pitt.edu.

WEB RESOURCES

http://genome.ucsc.edu
http://biosun1.harvard.edu/_fbat/fbat/
http://www.1000genomes.org/
http://www.ncbi.nlm.nih.gov/SNP/
http://www.cbil.upenn.edu/cgi-bin/tess/tess

CONFLICT OF INTEREST DISCLOSURE

There are no conflicts to report.

rs7179658 (15q22.2). We directly sequenced three non-coding conserved regions 200kb downstream of *SPRY2* in 713 cases, 438 controls, and 485 trios from the US, Mongolia, and the Philippines. We found rs8001641 to be significantly associated with cleft lip with cleft palate (NSCLP) in Europeans (p-value=4×10^{-5}, $OR_{transmission}$=1.86 with 95% confidence interval: 1.38-2.52). We also found several novel sequence variants in the conserved regions in Asian and European samples, which may help to localize common variants contributing directly to the risk for NSCLP. This study confirms the prior association between rs8001641 and NSCLP in European populations.

Keywords

GWA study; Nonsyndromic orofacial clefting; replication; SPRY2

INTRODUCTION

Nonsyndromic orofacial clefts are common birth defects with a complex etiology due to both genetic and environmental risk factors [Dixon et al. 2011]. The prevalence of nonsyndromic orofacial clefts varies by ethnicity with a worldwide average prevalence of 1/700 live births [Rahimov et al. 2008]. Asians have the highest occurrence with 1/500 births, Africans the lowest with 1/2500 births, and Europeans an intermediate occurrence of 1/1000 births [Dixon et al. 2011]. Collectively, epidemiologic, embryologic, and genetic studies support dividing nonsyndromic orofacial clefts into two groups: cleft lip with or without palate (NSCL/P) and cleft palate only (NSCPO) [Dixon et al. 2011; Marazita 2012]. Within the NSCL/P category are individuals with cleft lip only (NSCLO) and cleft lip with cleft palate (NSCLP). Although these cleft types are historically grouped together, recent evidence suggests that NSCLO and NSCLP may have separate genetic etiologies [Ludwig et al. 2012; Rahimov et al. 2008].

Genome-wide association (GWA) studies examine common genetic variants and are frequently used to identify genetic loci associated with disease. Currently, there have been four independent GWA studies and a meta-analysis to identify genetic risk factors for NSCL/P [Beaty et al. 2010; Birnbaum et al. 2009; Grant et al. 2009; Ludwig et al. 2012; Mangold et al. 2010]. The first NSCL/P GWA studies found very strong signals on 8q24 in populations of European decent [Birnbaum et al. 2009; Grant et al. 2009], while the third study, also in a European population, identified two additional loci at 17q22 and 10q25 [Mangold et al. 2010]. The fourth study was performed in case-parent trios from multiple populations of European and Asian ancestry and identified two new loci on 1p22 and 20q12 [Beaty et al., 2010]. The fifth genome-wide study was a meta-analysis based on the studies by Beaty et al. [2010] and Mangold et al. [2010]. In this study, the European case-control data was combined with the European-American trios, resulting in three new loci reaching genome-wide significance: rs7590268 (2p21), rs8001641 (13q31.1), and rs1873147 (15q22.2). Three additional loci were significant following addition of the Asian trios: rs742071 (1p36), rs7632427 (3p11.1), and rs12543318 (8q21.3). Ludwig and colleagues also performed analyses separating NSCLO and NSCLP, finding that rs8001641 (13q31.1) is uniquely associated with NSCLP [Ludwig et al. 2012].

A current challenge in human genetics is identifying specific etiologic variants following GWA studies, which identify regions or SNPs associated with the disease but may not have identified the causative variant due to linkage disequilibrium that results in tested SNPs serving as surrogates for other untested variants [Manolio 2010; Pearson and Manolio 2008]. Many of the SNPs associated with NSCL/P are located in gene deserts or far from the nearest gene, leading to the hypothesis that regulatory variants contribute to the risk of NSCL/P. Our study was designed to attempt replication of the six new loci from the Ludwig et al. meta-analysis (2p21, 13q31.1, 15q22.2, 1p36, 3p11.1, and 8q21.3) in families with NSCL/P from four different populations of Asian or European origin. We also included families with NSCPO to determine if these related disorders have similar genetic risk factors. We then performed additional Sanger sequencing of conserved regions near rs8001641 in a search for either the common etiologic variant(s) that might underlie the GWA signal or to find rare variants not detectable by GWAS with a large contributory effect to individual risk for clefting. The existence of such rare variants would also help focus the search for the common variant(s).

MATERIAL AND METHODS

Samples

Our samples are from individuals with nonsyndromic orofacial clefts without any other structural or cognitive deficits from five different populations: US (223 trios), Denmark (419 trios), Japan (97 trios), and Mongolia (207 trios). The European cohort consisted of US and Danish samples while the Asian cohort was comprised of samples from Mongolia and Japan. None of the samples were previously genotyped in GWA or replication studies. Some pedigrees (Denmark and Mongolia) also included genotyped affected and/or unaffected full or half-siblings. Cases included nonsyndromic cleft lip with cleft palate (NSCLP), cleft lip only (NSCLO), and cleft palate only (NSCPO). For analysis purposes, the pedigrees were classified by family history. Those pedigrees with a history of only NSCLO, were defined as NSCLO pedigrees. Similarly, those pedigrees with a history of only NSCLP were defined as NSCLP pedigrees. Pedigrees with a history of both NSCLO and NSCLP, while not numerous in this sample, were combined with the NSCLO and NSCLP pedigrees to form a subgroup of pedigrees defined as NSCL/P. Pedigrees with a history of only NSCPO were defined as NSCPO families. Finally, a few pedigrees with a history of both NSCPO and NSCL/P were included only in the overall pooled analysis along with pedigrees of unknown type of clefts (Table I). The University of Iowa Institutional Review Board (IRB) approved sample collection (University of Iowa approval numbers 199804080, 199804081, 201102785 and 201209749) and written informed consent was obtained for all participants.

Genotyping and Sequencing

We used Taqman SNP Genotyping Assays (Life Technologies, Carlsbad, CA) to genotype six SNPs (rs742071, rs7590268, rs7632427, rs12543318, rs8001641 and rs7179658). One of the significant SNPs from the Ludwig et al. meta-analysis paper, rs1873147, failed in the Taqman assay design process, so an alternative SNP, rs7179658, from the same linkage disequilibrium block (D'=1.0 and r^2=1.0 in Japanese, Chinese and European HapMap populations) was used as a surrogate.

Three conserved regions in the 13q31.1 region (around rs8001641) were Sanger sequenced (Supplementary Figure SI) and the potential effects of each variant on transcription factor binding were characterized using TESS (Supplemental Table SI). For clarity, we have named the sequenced elements based on their respective distances in kilobases from the *SPRY2* transcription start site. We sequenced the +222 region (hg19: chr13:80692633-80693142) in a total of 713 cases (142 US cases, 173 Mongolian cases and 398 Filipino cases) and 438 controls (185 Mongolian controls and 253 Filipino controls). The +221.2 region (hg19:chr13:80693238-80693818) was sequenced in 485 Filipino trios. The +213.5 region (hg19:chr13:80699863-80701605) was sequenced in 108 US cases with NSCLP. Chromatograms were transferred to a Unix workstation, base-called with PHRED (v.0.961028), assembled with PHRAP (v. 0.960731), scanned by POLYPHRED (v. 0.970312), and viewed with the CONSED program (v. 4.0). Transcription Element Search System (TESS), (http://www.cbil.upenn.edu/cgi-bin/tess/tess) was used to determine if the variants identified by sequencing created or destroyed transcription factor binding sites.

Statistical analysis

All populations underwent Hardy–Weinberg Equilibrium (HWE) analysis and minor allele frequency (MAF) determination. HWE, MAF, and parent-of-origin effects, were performed using PLINK [Purcell et al. 2007]. All SNPs were in HWE ($p>0.08$). Allelic and genotypic TDT was performed using FBAT (v1.73) [Horvath et al. 2004]. We used a Bonferonni correction for 36 tests to determine a threshold for formal significance of $p=0.0014$ (36 independent tests: 6 SNPs x 3 cleft groups x 2 racial groups). This p-value threshold does not consider the combined analyses in the calculations as they are not truly independent tests. A more conservative approach would be a correction for 90 total tests ($p=0.00056$). Rare variants (MAF<1%) from sequencing in trios were combined for TDT-based “burden” test using FBAT (v2.0.4) [De et al. 2013]. Rare variants from case sequencing were compared with variants from the 1000 Genomes Project and analyzed by Fisher's exact test using R (version 3.0.2). Gene by gene interactions were determined with R Package *trio* (v1.4.23) [Schwender et al. 2012].

RESULTS

Association analyses

We replicated the association with one of the six SNPs and found nominal significance with several others (Table II). Specifically, allelic TDT results showed the A allele of rs8001641 was associated with NSCLP in Europeans ($p=4.02 \times 10^{-5}$, $OR_{transmission}=1.86$, 95%CI: 1.38-2.52). This SNP was also associated with NSCLP ($p=8.36 \times 10^{-5}$) and NSCL/P ($p=0.00042$) in the pooled sample. Genotypic TDT analysis further confirmed that rs8001641 was associated with European NSCLP (global p-value=0.00004) and the individual genotypic TDT analysis showed that the homozygous A/A genotype was the risk-associated genotype ($p=0.002$). There were no significant findings for parent-of-origin effects nor for gene by gene interactions (data not shown). SNPs reaching nominal significance included rs742071 ($p=0.0037$ in the combined analysis of NSCL/P), rs7632427 ($p=0.015$ in the combined NSCL/P), and rs12543318 ($p=0.009$ in all clefts in the Asian group).

Sequencing conserved regions at 13q31.1

We did not identify any novel rare variants (MAF < 1%) in the US cases. However, we found three novel rare variants (chr13:80692698A>C, chr13:80692752G>T and chr13:80692951-80692955delAAATT) in Filipino cases and one novel variant (chr13:80693009T>C) in a Mongolian case. None of these variants was seen in 253 unrelated Philippines controls and 185 unrelated Mongolian controls

Although the rare variants were not found to be statistically enriched in Filipino cases versus controls (p=0.65), the case-control results motivated additional sequencing of the +222 region and a nearby conserved element (+221.2). We identified five additional novel variants that were all inherited from unaffected parents and did not segregate with clefting when we sequenced additional family members (Supplementary Table SI). In total, there were nine rare variants with MAF <1% and one variant (chr13:80693720G>A) was more common with a MAF of 2.60%. We note that the three variants found in the Filipino case-control cohort were also found in these independent Filipino trios. We used the rare variant extension of the family-based association test to determine if rare variants in the +222 and +221.2 were cumulatively overtransmitted, however the results were not significant under the weighted and un-weighted models (p>0.18).

Our screen of the +222 and +221.2 regions did not identify any novel rare or common variants in European samples. We also examined a region +213.5kb from the *SPRY2* transcription start site containing rs11842594, which is in high linkage disequilibrium with rs8001641 in European HapMap samples (D'=0.98 and r^2=0.94) and our European samples (D'=0.92 and r^2=0.84). We identified six novel rare variants (Supplementary Table SI). We used 379 European samples sequenced by the 1000 Genomes Project as controls; there were 8 rare variants in these samples within the +213.5 region. There was no significant enrichment of rare variants in NSCLP cases in the +213.5 region (p=0.09).

DISCUSSION

In this study, we successfully replicated the association between rs8001641 and NSCLP from the genome-wide meta-analysis of NSCL/P [Ludwig et al. 2012]. The association between rs8001641 and NSCLP in the European population was confirmed by allelic and genotypic TDT analyses. Simultaneously, we found no indication of association in the Japanese and Mongolian populations. The p-value for our pooled European and Asian trios was higher (less significant) than in the European trios alone, indicating that the significance of rs8001641 in the pooled samples is driven by the Europeans. The counts for informative families contributing to the TDT statistic reflect the robustness of this result (Table II).

rs8001641 is located in a highly conserved region 222kb downstream of the *SPRY2* gene, a strong candidate gene for orofacial clefting based on studies of animal models [Goodnough et al. 2007; Matsumura et al. 2011; Smith et al. 2012; Welsh et al. 2007]. Numerous studies have shown that highly conserved non-coding elements act as developmental enhancers *in vivo* [Pennacchio et al. 2006; Visel et al. 2007; Visel et al. 2009]. Non-coding conserved elements around rs8001641 therefore represent putative regulatory elements for *SPRY2*, so we sequenced three conserved regions to determine if they contain functional variants

involved in human craniofacial development. In the present study, we sequenced US, Mongolian and Filipino populations to search for rare variants that would not be detected by GWA but might have large individual impact and could target regions in which common, contributory variants might lie. Although in aggregate we found no evidence for an accumulation of rare variants in these conserved regions, there were several variants predicted to create or destroy binding sites of transcription factors related to craniofacial development. Specifically, three variants were predicted to create or destroy POU1F1a binding sites. Notably, POU1F1a interacts with PITX2, a gene that acts in a pathway to regulate cell proliferation in murine palatal mesenchyme [Iwata et al. 2012] [Amendt et al. 1998]. Another variant was predicted to destroy several binding sites including LEF-1. LEF1 is regulated by Tgfβ3 which appears to play a major role in transformation of the medial edge epithelial seam into mesenchyme during mouse palatal development [Nawshad et al. 2004]. LEF1 also binds in response to stimulation through the WNT signaling pathway which also plays a crucial role during craniofacial developmental processes.

Although the absence of these rare variants from public databases suggests that some of these variants could be etiologic, the Mongolian and Filipino populations are not represented in the 1000 Genomes Project. Furthermore, we were only able to sequence a relatively small number of controls. Previous work suggests that these numbers may be insufficient to distinguish private variants from very rare polymorphisms [Leslie and Murray 2013]. Future sequencing efforts utilizing additional samples or experiments testing the effects of these variants on transcription factor binding or gene expression (perhaps *SPRY2* or adjacent genes) will be necessary to determine their true impact. It is also possible that these conserved regions are not involved in craniofacial development. Nearby elements with characteristic profiles of active enhancer elements from the ENCODE project should also be considered.

We were unable to replicate reported associations for rs7590268 (2p21; *THADA*) and rs7179658 (15q22.2; *TPM1*) in individual populations or in pooled analyses. For rs742071 (1p36; *PAX7*), rs7632427 (3p11.1; *EPHA3*), and rs12543318 (8q21.3), we observed nominal significance that could be considered evidence of replication. That we didn't observe stronger association for these variants might be due to inadequate power of our sample sizes or failure of some SNPs to associate with clefting in some ancestral groups as we have previously noted between Asian and European populations [Beaty et al. 2010]. We did not find any association between these SNPs and NSCPO, suggesting that NSCL/P and NSCPO have distinct genetic etiologies, which is consistent with the results of recent genetic studies [Beaty et al. 2011]. This study further confirms that rs8001641 is associated specifically with NSCLP in European populations, which is consistent with the results from GWA study meta-analysis paper [Ludwig et al. 2012].

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGEMENTS

We are particularly thankful to the patients and their families who participated in this study. We greatly appreciate the support in Iowa of Tamara Busch and Adela Mansilla. We are also thankful to the China Scholarship Council for supporting the Joint-Ph.D program. This project was supported by the NIDCR (NIH) grants R37-DE08559, U01-DE20057, K99/R00-DE022378, and Grant-in-Aid for Scientific Research (Category A, No. 24249092) JSPS KAKENHI grant-in-Aid for Young Scientists B (Category B, No. 26861757) from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

REFERENCES

Amendt BA, Sutherland LB, Semina EV, Russo AF. The molecular basis of Rieger syndrome. Analysis of Pitx2 homeodomain protein activities. J Biol Chem. 1998; 273(32):20066–20072. [PubMed: 9685346]

Beaty TH, Murray JC, Marazita ML, Munger RG, Ruczinski I, Hetmanski JB, Liang KY, Wu T, Murray T, Fallin MD, Redett RA, Raymond G, Schwender H, Jin SC, Cooper ME, Dunnwald M, Mansilla MA, Leslie E, Bullard S, Lidral AC, Moreno LM, Menezes R, Vieira AR, Petrin A, Wilcox AJ, Lie RT, Jabs EW, Wu-Chou YH, Chen PK, Wang H, Ye X, Huang S, Yeow V, Chong SS, Jee SH, Shi B, Christensen K, Melbye M, Doheny KF, Pugh EW, Ling H, Castilla EE, Czeizel AE, Ma L, Field LL, Brody L, Pangilinan F, Mills JL, Molloy AM, Kirke PN, Scott JM, Arcos-Burgos M, Scott AF. A genome-wide association study of cleft lip with and without cleft palate identifies risk variants near MAFB and ABCA4. Nat Genet. 2010; 42(6):525–529. [PubMed: 20436469]

Beaty TH, Ruczinski I, Murray JC, Marazita ML, Munger RG, Hetmanski JB, Murray T, Redett RJ, Fallin MD, Liang KY, Wu T, Patel PJ, Jin SC, Zhang TX, Schwender H, Wu-Chou YH, Chen PK, Chong SS, Cheah F, Yeow V, Ye X, Wang H, Huang S, Jabs EW, Shi B, Wilcox AJ, Lie RT, Jee SH, Christensen K, Doheny KF, Pugh EW, Ling H, Scott AF. Evidence for gene-environment interaction in a genome wide study of nonsyndromic cleft palate. Genetic Epidemiology. 2011; 35(6):469–478. [PubMed: 21618603]

Birnbaum S, Ludwig KU, Reutter H, Herms S, Steffens M, Rubini M, Baluardo C, Ferrian M, Almeida de Assis N, Alblas MA, Barth S, Freudenberg J, Lauster C, Schmidt G, Scheer M, Braumann B, Berge SJ, Reich RH, Schiefke F, Hemprich A, Potzsch S, Steegers-Theunissen RP, Potzsch B, Moebus S, Horsthemke B, Kramer F-J, Wienker TF, Mossey PA, Propping P, Cichon S, Hoffmann P, Knapp M, Nothen MM, Mangold E. Key susceptibility locus for nonsyndromic cleft lip with or without cleft palate on chromosome 8q24. Nat Genet. 2009; 41(4):473–477. [PubMed: 19270707]

De G, Yip WK, Ionita-Laza I, Laird N. Rare variant analysis for family-based design. PLoS One. 2013; 8(1):e48495. [PubMed: 23341868]

Dixon MJ, Marazita ML, Beaty TH, Murray JC. Cleft lip and palate: understanding genetic and environmental influences. Nature Reviews Genetics. 2011; 12(3):167–178.

Goodnough LH, Brugmann SA, Hu D, Helms JA. Stage-dependent craniofacial defects resulting from Sprouty2 overexpression. Dev Dyn. 2007; 236(7):1918–1928. [PubMed: 17576140]

Grant SF, Wang K, Zhang H, Glaberson W, Annaiah K, Kim CE, Bradfield JP, Glessner JT, Thomas KA, Garris M, Frackelton EC, Otieno FG, Chiavacci RM, Nah HD, Kirschner RE, Hakonarson H. A genome-wide association study identifies a locus for nonsyndromic cleft lip with or without cleft palate on 8q24. J Pediatr. 2009; 155(6):909–913. [PubMed: 19656524]

Horvath S, Xu X, Lake SL, Silverman EK, Weiss ST, Laird NM. Family-based tests for associating haplotypes with general phenotype data: application to asthma genetics. Genet Epidemiol. 2004; 26(1):61–69. [PubMed: 14691957]

Iwata J, Tung L, Urata M, Hacia JG, Pelikan R, Suzuki A, Ramenzoni L, Chaudhry O, Parada C, Sanchez-Lara PA, Chai Y. Fibroblast growth factor 9 (FGF9)-pituitary homeobox 2 (PITX2) pathway mediates transforming growth factor beta (TGFbeta) signaling to regulate cell proliferation in palatal mesenchyme during mouse palatogenesis. J Biol Chem. 2012; 287(4): 2353–2363. [PubMed: 22123828]

Leslie EJ, Murray JC. Evaluating rare coding variants as contributing causes to non syndromic cleft lip and palate. Clin Genet. 2013; 84(5):496–500. [PubMed: 22978696]

Ludwig KU, Mangold E, Herms S, Nowak S, Reutter H, Paul A, Becker J, Herberz R, AlChawa T, Nasser E, Bohmer AC, Mattheisen M, Alblas MA, Barth S, Kluck N, Lauster C, Braumann B, Reich RH, Hemprich A, Potzsch S, Blaumeiser B, Daratsianos N, Kreusch T, Murray JC, Marazita ML, Ruczinski I, Scott AF, Beaty TH, Kramer F-J, Wienker TF, Steegers-Theunissen RP, Rubini M, Mossey PA, Hoffmann P, Lange C, Cichon S, Propping P, Knapp M, Nothen MM. Genome-wide meta-analyses of nonsyndromic cleft lip with or without cleft palate identify six new risk loci. Nat Genet. 2012; 44(9):968–971. [PubMed: 22863734]

Mangold E, Ludwig KU, Birnbaum S, Baluardo C, Ferrian M, Herms S, Reutter H, de Assis NA, Chawa TA, Mattheisen M, Steffens M, Barth S, Kluck N, Paul A, Becker J, Lauster C, Schmidt G, Braumann B, Scheer M, Reich RH, Hemprich A, Potzsch S, Blaumeiser B, Moebus S, Krawczak M, Schreiber S, Meitinger T, Wichmann H-E, Steegers-Theunissen RP, Kramer F-J, Cichon S, Propping P, Wienker TF, Knapp M, Rubini M, Mossey PA, Hoffmann P, Nothen MM. Genome-wide association study identifies two susceptibility loci for nonsyndromic cleft lip with or without cleft palate. Nat Genet. 2010; 42(1):24–26. [PubMed: 20023658]

Manolio TA. Genomewide Association Studies and Assessment of the Risk of Disease. New England Journal of Medicine. 2010; 363(2):166–176. [PubMed: 20647212]

Marazita ML. The evolution of human genetic studies of cleft lip and cleft palate. Annu Rev Genomics Hum Genet. 2012; 13:263–283. [PubMed: 22703175]

Matsumura K, Taketomi T, Yoshizaki K, Arai S, Sanui T, Yoshiga D, Yoshimura A, Nakamura S. Sprouty2 controls proliferation of palate mesenchymal cells via fibroblast growth factor signaling. Biochem Biophys Res Commun. 2011; 404(4):1076–1082. [PubMed: 21195053]

Nawshad A, LaGamba D, Hay ED. Transforming growth factor beta (TGFbeta) signalling in palatal growth, apoptosis and epithelial mesenchymal transformation (EMT). Arch Oral Biol. 2004; 49(9): 675–689. [PubMed: 15275855]

Pearson TA, Manolio TA. How to interpret a genome-wide association study. JAMA. 2008; 299(11): 1335–1344. [PubMed: 18349094]

Pennacchio LA, Ahituv N, Moses AM, Prabhakar S, Nobrega MA, Shoukry M, Minovitsky S, Dubchak I, Holt A, Lewis KD, Plajzer-Frick I, Akiyama J, De Val S, Afzal V, Black BL, Couronne O, Eisen MB, Visel A, Rubin EM. In vivo enhancer analysis of human conserved non-coding sequences. Nature. 2006; 444(7118):499–502. [PubMed: 17086198]

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007; 81(3):559–575. [PubMed: 17701901]

Rahimov F, Marazita ML, Visel A, Cooper ME, Hitchler MJ, Rubini M, Domann FE, Govil M, Christensen K, Bille C, Melbye M, Jugessur A, Lie RT, Wilcox AJ, Fitzpatrick DR, Green ED, Mossey PA, Little J, Steegers-Theunissen RP, Pennacchio LA, Schutte BC, Murray JC. Disruption of an AP-2alpha binding site in an IRF6 enhancer is associated with cleft lip. Nat Genet. 2008; 40(11):1341–1347. [PubMed: 18836445]

Schwender H, Taub MA, Beaty TH, Marazita ML, Ruczinski I. Rapid testing of SNPs and gene-environment interactions in case-parent trio data based on exact analytic parameter estimation. Biometrics. 2012; 68(3):766–773. [PubMed: 22150644]

Smith TM, Lozanoff S, Iyyanar PP, Nazarali AJ. Molecular signaling along the anterior- posterior axis of early palate development. Front Physiol. 2012; 3:488. [PubMed: 23316168]

Visel A, Minovitsky S, Dubchak I, Pennacchio LA. VISTA Enhancer Browser--a database of tissue-specific human enhancers. Nucleic acids research. 2007; 35(Database issue):D88–92. [PubMed: 17130149]

Visel A, Rubin EM, Pennacchio LA. Genomic views of distant-acting enhancers. Nature. 2009; 461(7261):199–205. [PubMed: 19741700]

Welsh IC, Hagge-Greenberg A, O'Brien TP. A dosage-dependent role for Spry2 in growth and patterning during palate development. Mech Dev. 2007; 124(9-10):746–761. [PubMed: 17693063]

Table I

Pedigree counts for TDT by population and cleft family history

Population	Cleft Type				
	CLO	CLP	CPO	CL/P[a]	All Clefts[b]
European	194	259	158	470	642
Iowa	58	86	54	157	223
Denmark	136	173	104	313	419
Asian	65	165	43	230	304
Japan	23	54	20	77	97
Mongolia	42	111	23	153	207
TOTAL	259	424	201	700	946

[a] CL/P includes CLO and CLP pedigrees well as pedigrees with a family history of both CLO and CLP

[b] All clefts includes CL/P, CPO, and those with history of both CPO and CL/P or an unknown type of clefting

Table II

Association Analyses

SNP[a]	Population		Freq.[b] (n)[c]	P (OR; 95% CI)
rs742071 (G/**T**) 1p36	European	NSCLO	0.45 (97)	0.06 (1.39; 0.98-1.99)
		NSCLP	0.46 (129)	0.31 (1.17; 0.86-1.58)
		NSCL/P	0.46 (229)	0.03 (1.29; 1.03-1.62)
		NSCPO	0.44 (82)	0.62 (1.12; 0.75-1.64)
		All clefts	0.45 (317)	0.05 (1.21; 0.99-1.47)
	Asian	NSCLO	0.06 (16)	0.32 (1.67; 0.61-4.59)
		NSCLP	0.09 (48)	0.057 (1.7; 0.98-2.95)
		NSCL/P	0.08 (79)	0.02 (1.66; 1.07-2.57)
		NSCPO	0.06 (8)	0.48 (0.6; 0.14-2.51)
		All clefts	0.08 (64)	0.03 (1.69; 1.04-2.75)
	Combined Asian and European	NSCLO	0.34 (113)	0.04 (1.42; 1.02-1.99)
		NSCLP	0.29 (177)	0.069 (1.28; 0.98-1.67)
		NSCL/P	0.32 (293)	*0.0037 (1.35; 1.10-1.66)*
		NSCPO	0.34 (90)	0.77 (1.06; 0.72-1.55)
		All clefts	0.31 (396)	*0.0069 (1.28; 1.07-1.53)*
rs7590268 (T/**G**) 2p21	European	NSCLO	0.27 (84)	0.84 (0.96; 0.66-1.41)
		NSCLP	0.26 (126)	0.62 (1.08; 0.79-1.49)
		NSCL/P	0.27 (216)	0.90 (0.99; 0.77-1.25)
		NSCPO	0.22 (59)	0.54 (0.86; 0.53-1.39)
		All clefts	0.25 (280)	0.78 (0.97; 0.78-1.20)
	Asian	NSCLO	0.07 (14)	0.22 (1.83; 0.68-4.96)
		NSCLP	0.08 (41)	0.13 (0.63; 0.34-1.16)
		NSCL/P	0.08 (77)	0.91 (0.98; 0.64-1.49)
		NSCPO	0.07 (12)	0.56 (1.4; 0.44-4.41)
		All clefts	0.08 (55)	0.52 (0.85; 0.51-1.40)
	Combined Asian and European	NSCLO	0.22 (98)	0.79 (1.05; 0.74-1.50)
		NSCLP	0.18 (167)	0.77 (0.96; 0.72-1.27)
		NSCL/P	0.19 (271)	0.69 (0.96; 0.77-1.19)
		NSCPO	0.18 (71)	0.73 (0.93; 0.59-1.45)
		All clefts	0.19 (357)	0.77 (0.97; 0.80-1.18)
rs7632427 (**T**/C) 3p11.1	European	NSCLO	0.63 (106)	0.04 (1.44; 1.01-2.05)
		NSCLP	0.61 (122)	0.13 (1.29; 0.93-1.79)
		NSCL/P	0.61 (236)	0.02 (1.34; 1.06-1.69)
		NSCPO	0.62 (63)	0.23 (0.74; 0.46-1.20)
		All clefts	0.62 (307)	0.19 (1.15; 0.93-1.41)
	Asian	NSCLO	0.81 (23)	0.72 (0.87; 0.43-1.79)

SNP[a]	Population		Freq.[b] (n)[c]	P (OR; 95% CI)
		NSCLP	0.82 (74)	0.29 (1.26; 0.83-1.91)
		NSCL/P	0.82 (129)	0.17 (1.24; 0.91-1.70)
		NSCPO	0.83 (19)	0.68 (1.18; 053-2.64)
		All clefts	0.82 (97)	0.46 (1.14; 0.79-1.64)
	Combined Asian and European	NSCLO	0.68 (129)	0.09 (1.31; 0.95-1.79)
		NSCLP	0.69 (196)	0.07 (1.27; 0.98-1.65)
		NSCL/P	0.69 (333)	0.015 (1.28; 1.05-1.55)
		NSCPO	0.67 (82)	0.40 (0.84; 0.56-1.26)
		All clefts	0.69 (436)	0.066 (1.17; 0.99-1.39)
rs12543318 (A/**C**) 8q21.3	European	NSCLO	0.36 (103)	0.44 (1.14; 0.82-1.60)
		NSCLP	0.37 (140)	0.49 (0.90; 0.67-1.21)
		NSCL/P	0.37 (251)	0.82 (1.03; 0.82-1.28)
		NSCPO	0.32 (79)	0.68 (0.92; 0.62-1.37)
		All clefts	0.36 (336)	0.92 (0.99; 0.82-1.19)
	Asian	NSCLO	0.56 (52)	0.29 (1.28; 0.81-2.03)
		NSCLP	0.53 (123)	0.015 (1.48; 1.08-2.02)
		NSCL/P	0.54 (233)	0.03 (1.28; 1.03-1.61)
		NSCPO	0.55 (34)	0.65 (0.87; 0.48-1.58)
		All clefts	0.54 (175)	0.009 (1.41; 1.09-1.83)
	Combined Asian and European	NSCLO	0.42 (155)	0.21 (1.19; 0.91-1.56)
		NSCLP	0.44 (263)	0.23 (1.14; 0.92-1.41)
		NSCL/P	0.43 (426)	0.06 (1.17; 0.99-1.39)
		NSCPO	0.38 (113)	0.55 (0.90; 0.65-1.26)
		All clefts	0.43 (569)	0.17 (1.11; 0.96-1.28)
rs8001641 (**A**/G) 13q31.1	European	NSCLO	0.48 (120)	0.69 (1.07; 0.78-1.47)
		NSCLP	0.54 (155)	**4.02E-05 (1.86; 1.38-2.52)**
		NSCL/P	0.51 (281)	*0.0019 (1.39; 1.13-1.73)*
		NSCPO	0.49 (90)	0.58 (1.11; 0.77-1.59)
		All clefts	0.51 (381)	0.006 (1.29; 1.07-1.54)
	Asian	NSCLO	0.15 (25)	0.09 (1.9; 0.88-4.09)
		NSCLP	0.17 (73)	0.33 (1.23; 0.81-1.88)
		NSCL/P	0.17 (131)	0.01 (1.49; 1.08-2.06)
		NSCPO	0.16 (14)	0.44 (1.5; 0.53-4.21)
		All clefts	0.17 (98)	0.095 (1.37; 0.95-1.98)
	Combined Asian and European	NSCLO	0.39 (145)	0.29 (1.17; 0.87-1.56)
		NSCLP	0.39 (228)	**8.36E-05 (1.63; 1.27-2.08)**
		NSCL/P	0.39 (379)	**0.00042 (1.39; 1.16-1.67)**
		NSCPO	0.41 (104)	0.43 (1.15; 0.81-1.61)
		All clefts	0.39 (512)	**3.44E-04 (1.33; 1.14-1.56)**

SNP[a]	Population		Freq.[b] (n)[c]	P (OR; 95% CI)
rs7179658 (**C**/T) 15q22.2	European	NSCLO	0.69 (100)	0.29 (1.21; 0.85-1.72)
		NSCLP	0.72 (127)	0.34 (0.86; 0.63-1.17)
		NSCL/P	0.71 (233)	0.95 (1.00; 0.8-1.27)
		NSCPO	0.73 (75)	0.25 (0.79; 0.52-1.187)
		All clefts	0.71 (317)	0.51 (0.94; 0.77-1.14)
	Asian	NSCLO	0.17 (27)	0.37 (1.30; 0.68-2.83)
		NSCLP	0.22 (89)	0.33 (0.82; 0.56-1.21)
		NSCL/P	0.23 (164)	0.77 (0.96; 0.72-1.27)
		NSCPO	0.28 (25)	0.47 (1.31; 0.64-2.69)
		All clefts	0.21 (116)	0.67 (0.93; 0.66-1.30)
	Combined Asian and European	NSCLO	0.55 (127)	0.17 (1.24; 0.91-1.70)
		NSCLP	0.51 (216)	0.17 (0.84; 0.66-1.08)
		NSCL/P	0.53 (349)	0.85 (0.98; 0.81-1.19)
		NSCPO	0.62 (100)	0.53 (0.89; 0.63-1.27)
		All clefts	0.54 (481)	0.48 (0.94; 0.80-1.11)

P-values, odds ratios (ORs), and 95% confidence intervals (CIs) were calculated in PLINK and are presented based on the Ludwig et al. associated allele. Allele frequencies and counts of informative families were calculated in FBAT. P-values are in bold if significance was reached (p>0.0014).

[a] The bolded allele for each SNP is the NSCL/P associated allele from Ludwig et al.

[b] Allele frequency for the Ludwig et al. associated allele

[c] Number of informative families.