Published in final edited form as:
Am J Med Genet A. 2013 May ; 161(5): 965–972. doi:10.1002/ajmg.a.35749.

Replication of Genome Wide Association Identified Candidate Genes Confirm the Role of Common and Rare Variants in *PAX7* and *VAX1* in the Etiology of Non-syndromic CL(P)

Azeez Butali[1,*], **Satoshi Suzuki**[1,2,6,*], **Margaret E. Cooper**[3], **Adela M. Mansilla**[1], **Karen Cuenco**[3], **Elizabeth J Leslie**[1], **Yasushi Suzuki**[2], **Teruyuki Niimi**[2], **Masahiko Yamamoto**[6], **Gongorjav Ayanga**[9], **Tudevdorj Erkhembaatar**[9], **Hiroo Furukawa**[2,6], **Kumiko Fujiwawa**[2], **Hideto Imura**[2], **Aline L. Petrin**[1], **Nagato Natsume**[2], **Terri H. Beaty**[4], **Mary L. Marazita**[3,5], and **Jeffery C. Murray**[1,7,8]

[1]Department of Pediatrics, University of Iowa

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

[3]Johns Hopkins University, School of Public Health, Baltimore, Maryland, USA

[4]Center for Craniofacial and Dental Genetics, Department of Oral Biology, School of Dental Medicine, University of Pittsburgh, Pittsburgh, PA 15260

[5]Department of Human Genetics, School of Public Health, University of Pittsburgh, Pittsburgh, PA 15219

[6]Faculty of Psychological and Physical Science, Aichi-Gakuin University, Japan

[7]University of Iowa, Departments of Pediatric Dentistry, Epidemiology and Biology, Iowa City, IA 52242

[8]College of Nursing, University of Iowa, Iowa City, IA 52242

[9]Maternal and Children's Health Research Center Hospital, Ulaanbaatar, Mongolia

Abstract

Following recent genome wide association studies (GWAS), significant genetic associations have been identified for several genes with non-syndromic cleft lip with or without cleft palate (CL(P). To replicate two of these GWAS signals, we investigated the role of common and rare variants in the *PAX7* and *VAX1* genes. TaqMan genotyping was carried out for SNPs in *VAX1* and *PAX7* and Transmission Disequilibrium Test (TDT) was performed to test for linkage and association in each population. Direct sequencing in and around the *PAX7* and *VAX1* genes in 1,326 individuals

Corresponding Author: Jeffrey C. Murray, MD, Professor, Departments of Pediatrics, Epidemiology and Biological Sciences, University of Iowa Carver College of Medicine, 500 Newton Road, 2182 ML, Iowa City, IA 52242-1181, 319-335-6897 phone; 319-335-6970 fax, jeff-murray@uiowa.edu.

*Both authors contributed equally to the overall project.

SUPPLEMENTAL DATA DESCRIPTION

Supplemental data include two tables.

WEB RESOURCES

http://www.uiowa.edu/~genetics/.
http://biosun1.harvard.edu/~fbat/fbat/)
http://pngu.mgh.harvard.edu/purcell/plink/).
http://genetics.bwh.harvard.edu/pph/
(http://www.1000genomes.org/)
(http://snp.gs.washington.edu/EVS/).

of European and Asian ancestry was done. TDT analysis showed strong associations with markers in *VAX1* (rs7078160, p=2.7E-06 and rs475202, p=0.0002) in a combined sample of Mongolian and Japanese CL (P) case-parent triads. Analyses using parent-of-origin effects showed significant excess transmission of the minor allele from both parents with the effect in the mothers (p=6.5E-05, OR (transmission) =1.91) more striking than in the fathers (p=0.004, OR (transmission) =1.67) for *VAX1* marker rs7078160 in the combined Mongolian and Japanese samples when all cleft types were combined. The rs6659735 trinucleotide marker in *PAX7* was significantly associated with all the US cleft groups combined (p=0.007 in all clefts and p=0.02 in CL(P)). Eight rare missense mutations found in *PAX7* and two rare missense mutations in *VAX1*. Our study replicated previous GWAS findings for markers in *VAX1* in the Asian population, and identified rare variants in *PAX7* and *VAX1* that may contribute to the etiology of CL(P). Determining the role of rare variants clearly warrants further investigation.

Keywords

GWAS; *PAX7*; *VAX1*; CL(P)

INTRODUCTION

Clefts of the lip and cleft palate are the most common craniofacial birth defects with a worldwide birth prevalence of approximately 1/700 [Mossey and Little, 2002]. These complex traits are associated with an increased infant mortality and significant morbidity through adult life and require multidisciplinary treatment for their management [Jugessur *et al.*, 2009]. The development of the craniofacial structures is a complex process that involves the coordinated growth of multiple independently derived primordia. Genetic and environmental factors or their interactions may influence the growth of these primordia and lead to abnormal development of facial structures which then results in clefting of the lip, the primary or secondary palate, or a combination of these sites [Jugessur *et al.*, 2009].

Although the etiology of orofacial clefts is complex, several genetic and environmental risk factors have been identified [Dixon *et al.*, 2011]. Linkage and association studies have provided significant statistical evidence for candidate genes such as interferon regulatory factor six (*IRF6*) [Rahimov *et al.*, 2008] and forkhead box E1 (*FOXE1*) [Marazita *et al.*, 2004; Moreno *et al.*, 2009]. Other candidate genes suggested to play a role include bone morphogenetic protein four (*BMP4*) [Suzuki *et al.*, 2009], tumour protein 63 (*TP63*) [Scapoli *et al.*, 2008], the jagged2 gene (*JAG2*) [Vieira *et al.*, 2005], fibroblast growth factor and their receptors (*FGFs, FGFRs*)[Riley *et al.*, 2007], poliovirus receptor-like 1 (*PVRL1*) [Turhani *et al.*,2005; Avila *et al.*, 2006], cysteine-rich secretory protein LCCL domain containing 2 (*CRISPLD2*) [Chiquet *et al.*, 2007] and muscle segment homeobox one (*MSX1*) [Jezewski *et al.*, 2003].

Genome wide association studies (GWAS) have also identified new candidate genes and regions that are associated with non-syndromic cleft lip with or without cleft palate (CL(P)). The first GWAS for CL(P) identified a strong association of non-syndromic CL(P) to markers in a gene desert on chromosome 8q24.21[Birnbuam *et al.*, 2009] in a German case-control sample, which was replicated in another study of European Americans [Grant *et al.*, 2009]. In an extension of the German study, Mangold *et al.* [2010] reported two new loci associated with non syndromic CL(P) at chromosome 17q.22 and chromosome 10q25.3. A GWAS carried out on CL(P) case-parent trios from an international consortium [Beaty *et al.*, 2010]showed evidence of linkage and association at a genome-wide level of significance for chromosome 8q24 (consistent with previous studies) and for the *IRF6* gene. In the same case-parent trio study, two other loci with evidence for association at genome-wide

significance were identified in or near two novel candidate genes: *ABCA4* on chromosome 1p22.1 and *MAFB* on chromosome 20q.12. They also reported three potential candidate genes where markers attained or approached genome wide significance: *PAX7* on chromosome 1p.36, *VAX1* on chromosome 10q25.3 and *NTN1* on chromosome 17p.13.

A replication study carried out using Mesoamerican populations revealed a significant association between CL(P) and genetic risk factors in *IRF6*, chr.8q24 and 10q25 loci using single SNP association analysis [Rojas-Martinez *et al.*,2010]. Similarly, another replication study in Estonia reported a significant association with a single nucleotide polymorphism (SNP) on chr.10q25, confirming its association with CL(P) in this Baltic population [Nikopensius *et al.*, 2010].

The present study tested for the effects of common and rare variants in *PAX7* and *VAX1* genes as a follow up to the GWAS signals reported by Beaty *et al.* [2010] and recently confirmed in a meta-analysis by Ludwig *et al* [2012]. The associated SNPs in chromosome 1p36 previously reported in GWAS results [Beaty *et al.*, 2010] are approximately 2 kb and 20 kb upstream of the *PAX7* gene itself and over 100 kb centromeric to the *KLHDC7A* gene. The *KLHCD7A* gene does not have any known biological function in the craniofacial region. Therefore we choose to examine the *PAX7* gene because of its proximity to the signal and the biological role it plays in neural crest development [Mansouri *et al.*, 1996]. The most strongly associated SNP, rs7078160 on chr10q25 reported in the GWAS by Beaty *et al.* [2010] ($p=1.07x10^{-7}$) and Mangold *et al.* [2010] ($p=1.1x10^{-7}$) lies 50 kb 3′ of *VAX1* and about 50 kb 5′ of *KIAA1598*. The *KIAA1598* gene product is predicted to be involved in signals for neuronal polarization. It is possible the causative SNPs are in either of these genes or in other nearby genes. However, we choose *VAX1* as a strong candidate because the SNP rs7078160 is in strong linkage disequilibrium (LD) with SNPs in and around *VAX1* and also based on the biological role it plays in the craniofacial region. Replication of the GWAS signals using independent samples complemented by direct sequencing of these two genes to identify novel sequence variations in and around each gene in cases and controls was carried out. *PAX7* is a member of the paired box containing gene family which includes nine different genes. These *PAX* genes are regulatory transcriptional proteins encoding a family of highly conserved DNA-binding transcription factors [Underhill, 2000]. The *VAX1* gene is a homeodomain transcription factor which encodes the ventral anterior homeobox 1 protein [Hallonet *et al.*, 1998].

MATERIALS AND METHODS

Study design

The patients were examined to identify the presence of other congenital anomalies and other major structural anomalies and excluded if these were present [Murray *et al.*, 1997]. Parents were interviewed to obtain information on family history of orofacial clefts. The University of Iowa Institutional Review Board (IRB) gave approval for sample collection (approval number 199804080) in conjunction with local approval in the Philippines (approval number 199804081) after approval by the IRB's of all local sites. The Ethics Committee of the School of Dentistry, Aichi-Gakuin University, approved sample collection in Mongolia and Japan (approval number 11) in conjunction with local IRB approval.

A total of 5,421 individuals (including triads from Japan, Iowa, Mongolia and family members from Philippines multiplex families) were genotyped and all analyses were performed over sub-groups by cleft type and by family history(Table I). Often families had both CLO and CLP in their family history and they formed a separate subgroup (CLO and CLP). An overall cleft lip with or without cleft palate (CL ± CP) subgroup was formed (CL(P)) from the three distinct groups of families mentioned above. The CL(P) group was

distinct from those families with a history of cleft palate only (CPO). The CPO group included families with a history of CPO as well as those with a history of cleft palate as well as CL ± CP. Some analyses were conducted on the overall all cleft group formed from the CL(P) and the CPO group. Analyses were also carried out for the overall Asian group by pooling the data from the Philippines, Japanese, and Mongolian studies. Finally, an overall test was conducted by pooling Asian and Iowan samples. In the Beaty *et al*. [2010] GWAS some genes were shown to be population specific while others played a role in multiple populations suggesting that ancestral background is a factor that should be considered in cleft etiology. For example, in Beaty *et al*. [2010] study, a GWAS significant SNP in the *MAFB* gene was observed only in the Asian population, while the GWAS significant SNP in the chromosome 8q.24 locus was observed only in the European population.

We sequenced a total of 1,326 individuals with clefts in all four populations (Tables II and III). Missense mutations were compared to missense mutations present in the 1000 genome database [http://www.1000genomes.org/] and exome variant sequence database [http://snp.gs.washington.edu/EVS/].

Genotyping analysis

We included 206 non-syndromic case-parent triads from Mongolia, 98 case-parent triads from Japan, 157 case-parent triads from Iowa, and 190 extended pedigrees from the Philippines multiplex families in this study. We performed TaqMan genotyping (Applied biosystems) using four SNPs in potential candidate genes from previous GWAS (*VAX1* rs7078160, rs4752028 and *PAX7* rs4920520, rs766325) reported to show association (Beaty *et al*., 2010; Mangold *et al*., 2010). TaqMan genotyping of a variant rs14874160>T in the 5′ UTR found through sequencing of the *VAX1* gene was also performed on 644 Philippine cases and 112 Philippine controls.

Sequencing Analyses

Direct sequencing was used to search for sequence variations in coding regions and conserved non-coding regions within 1Kb of the *PAX7* (NM_001135254.1 8 exons) and the *VAX1* genes (NM_001112704.1 4 exons) in 180 CL(P) patients (90 from Iowa and 90 from Philippines) and 180 controls (90 from Iowa and 90 from Philippines). An additional 630 cases from the Philippines were sequenced for the *PAX7* gene. We also sequenced 265 patients (171 Mongolians and 94 Japanese) and 92 Mongolian controls. Primers were designed using Primer 3 and optimized to the optimal annealing temperature with details available on request through the Murray lab website (see web resources). PCR products were sent for sequencing using an ABI 3730XL (Functional Biosciences, Inc., Madison, WI). 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).

Statistical Analyses

Transmission disequilibrium was tested using the Family Based Association Test (FBAT) under the additive model for the Philippine multiplex families. PLINK software was used for the TDT test in the Iowa, Japanese and Mongolian trios. The PLINK software [Purcell *et al*., 2007] was also used to estimate transmitted and un-transmitted allele counts, calculate odds ratios (affected transmitted allele count divided by unaffected transmitted allele count), and parental effects (separate maternal/paternal transmission counts).

We performed gene-gene (GxG) interaction analyses on the family data. We specifically targeted the *PAX7* and *VAX1* genes and their possible interaction with SNPs in the other cleft associated genes/regions: *ABCA4, IRF6,* 8q region*, FOXE1,* and *MAFB*. Our primary

gene-gene interaction evaluation was conducted using conditional logistic regression. A model containing one term for each SNP as a main effect and one interaction term was compared to another model without the interaction term. Assessment was based on a one degree of freedom Wald Chi-Square statistic contrasting these two models. Comparison of these models was implemented using R(trio) "full" option [Schwender *et al.*, 2011; Schwender *et al.*, 2012)] In addition, we considered the Cordell epistatic model [Cordell et al., 2002; Cordell et al., 2004] (with four interaction terms). However, there were convergence failures as a result of the sparseness in the data for specific SNP combinations. Therefore, the only the results from the one degree of freedom method are reported. The SNPs' linkage disequilibrium (LD) and minor allele frequencies were not problematic in the regression modeling.

RESULTS

Replication of markers in PAX7 and VAX1 genes in four populations

Mendelian errors were removed from the data by PLINK in executing the allelic TDT. Iowa patient-parent triads and Filipino multiplex families were genotyped for the *PAX7* SNP rs6659735. The threshold for significance was set as $p < 0.05$ and when we corrected for multiple testing, $p<0.001$ was significant.

The *PAX7* trinucleotide SNP (rs6659735) showed significant association ($p= 0.007$) in all the cleft groups combined in the Iowa population (Table I). There was strong evidence of linkage and association with markers in *VAX1* for CL(P) (rs7078160, p=2.7E-06 and rs4752028 p=0.0002 respectively) in both Mongolian and Japanese case-parent triads. The results in the CPO group were not significant (p=0.14 and p=0.87, respectively). For all clefts combined, the results were p=8.9E-06 and p=0.0006, respectively for Mongolian and Japanese case-parent triads) (Table I). A significant association with the CL(P) group was also observed among the Filipino multiplex families for rs70781860 ($p= 0.001$). No statistically significant findings were observed for *VAX1* in the Iowa sample (p=0.10 for all clefts combined and p=0.08 for the CL(P) group). There was strong evidence of association with SNPs in *VAX1* (rs7078160 and rs4752028) in the combined Asian samples (p=2.3E-08 and p=0.00004 respectively) (Table I). When data were analyzed from all four populations and all clefts combined, a significant association was also observed for markers in *VAX1* (rs7078160, p=7.4E-09 and rs4752028, p=9.8E-06). The CL(P) cleft group contributed the most to the association (p=1E-07 and p=0.00002). Interestingly, the CPO group also contributed to significant association findings in Asians, but for only rs7078160 (p=0.008) (Table I). The *VAX1* c.80130G>T marker (identified through sequencing) analyses suggested that the non-ancestral rare allele is not over-transmitted to patients (p=0.72 for all clefts combined and p-values ranging between 0.48 and 1 for the other cleft subgroups) (Supplementary table I).

Analyses using parent-of-origin effects showed significant excess transmission of the minor allele from both parents with the effect in the mothers (p=6.5E-05, OR (transmission) =1.91) more striking than in the fathers (p=0.004, OR (transmission) =1.67) for *VAX1* marker rs7078160 in the combined Mongolian and Japanese samples when all cleft types were combined. Affected males were mostly responsible for the rs7078160 SNP association with all clefts in the combined Japanese and Mongolian case-parent trios (male cases: p=6.4E-06 and female cases: p=0.02). In males with clefts, parental transmission effects were also different: (rs7078160; maternal (p=3.4E-05), OR=2.65 for and paternal (p=0.03), OR=1.69). No statistically significant parental transmission effects were seen in the affected female analysis (paternal (p=0.08) and maternal (p=0.14)).

Sequencing

Identification of variants in PAX7—We found 19 new variants, eight of which were missense mutations and three were synonymous mutations (Table II). The other newly identified variants were in introns. For cases with the missense mutations, we sequenced the samples for the available parent (in this case the mothers) but did not find the variants in them. Since we did not have samples from the fathers in these families, we cannot make any conclusions concerning the mode of inheritance (i.e., whether the missense variant segregated in the family or it was a *de novo* variant).

Identification of variants in VAX1—We found 12 new variants in *VAX1* (Table III). One of the SNPs rs14874160G>T showed significant association *(p*= 0.007) with CL(P) using the Fisher's exact test) when analyzed in 270 cases and 270 controls in the Philippines samples (Supplementary Table II).

Two of the new *VAX1* variants were missense mutations, (pMet117Arg) c. 96026T>G and (p.Ala233Ser) c.93824G>T, found only in Mongolian cases. A canonical splice acceptor variant c.96171G>T at position-1 was found in a Mongolian case. Two more variants were synonymous, found in a single Japanese case and in a Mongolian case) in coding regions (Table III). A Proline in-frame insertion (p.A233_P234insP) c.698_699insCCC was found in a single Filipino case and both parents.

Gene-gene interactions (GxG)—None of the GxG results reached the threshold for significance under the multiple testing done in this project (i.e., *p*-values ≤ 0.001), however there were a few GxG results that were nominally significant and warrant future investigations. For example, in the Asian group the *MAFB* rs17820943 SNP interacted with the *PAX7* rs4920520 SNP (*p*=0.03) and with *VAX1* rs7078160 (*p*=0.05) and *VAX1* rs4752028 (*p*=0.03). Also in Asians, *VAX1* rs7078160 interacts with *FOXE1* rs3758249 (*p*=0.01). In the combined sample across all populations, there was evidence for interaction of *VAX1* rs4752028 with 8q rs987525 (*p*=0.03) and with *MAFB*rs17820943 (*p*=0.05).

DISCUSSION

Significant genetic associations have been identified for 11 genes or loci with non-syndromic cleft lip with or without cleft palate (CL(P)) in five GWAS reports including a recent meta-analysis [Ludwig *et al.*, 2012]. To replicate two of these GWAS signals, we investigated the role of common and rare variants in the *PAX7* and *VAX1* genes.

Pax7 is important in neural crest development and required for the expression of the neural crest markers Slug, Sox9, Sox10 and HNK-1 in vivo [Basch *et al.*, 2006]. Evidence from animal models suggests *Vax1* has biological functions in the craniofacial region [Hallonet *et al.*, 1998]. Since the older SNP arrays were less comprehensive for non-Caucasian ethnicities and do not include rare variants [Pritchard, 2001; McCarthy *et al.*, 2008], we sequenced *PAX7 and VAX1* in individuals from Iowa, Japan, Mongolia and Philippines and identified eight rare missense variants in *PAX7* and two in *VAX1*. None of the *PAX7* rare variants in the cases were found in over 6500 individuals in the 1000 genome and whole exome variant sequence databases [Leslie and Murray, 2012]. However, the rare variant (p.Pro937Leu) c.89576C>T found in a control individual from Iowa, was also found in five out of 2188 chromosomes in the 1000 Genomes database. In *VAX1*, one of the missense mutations (p.Met117Arg) c.96062T>G is predicted to be probably damaging by Polyphen. These variants may affect protein structure and function. None of these missense variants were found in the 1000 genome and whole exome sequence databases. A recent study reported a missense mutation in *VAX1* in an individual with bilateral microphthalmia, cleft

lip and palate, small optic nerves, hippocampal malformations, corpus callosum agenesis, and absence of the pineal gland [Slavotinek *et al.*, 2011].

Rare coding variants have been reported in re-sequencing of genes implicated by GWAS [Johansen *et al.*, 2010] and significant GWAS SNPs may actually be reflecting signals from multiple associated functional rare variants that have large effects on risks [Dickson *et al.*, 2010]. Therefore, it is possible that these rare variants that have profound effects on the disease risks are responsible for the missing heritability even if there is no evidence of Mendelian segregation [Keinan and Clark, 2012; Tennessen *et al.*, 2012].

In the present study, we found statistically significant association with all clefts types and *VAX1* GWAS SNPs in the three Asian populations combined (rs7078160 p=2.3E-08 and rs4752028 p=0.00004) and all four populations combined (rs7078160 p=7.4E-09 and rs4752028 p=9.8E-06).

Significant association signals were observed using the same reported *VAX1* GWAS markers in the Asian case-parent trios. The findings of this study support previous GWAS [Beaty *et al.*, 2010, Mangold *et al.*, 2010] and is consistent with the replication studies reporting significant association with *VAX1* [Nikopensius *et al.*, 2010, Rojas-Martinez *et al.*, 2010; Ludwig *et al.*, 2012]. Our findings suggest the GWAS signals are in LD with the unobservable functional variants associated with NSCL(P). This assumption is supported by the evidence which suggest that GWAS signals can extend across multiple LD blocks [Dickson *et al.*, 2010].

In conclusion, our study has replicated previous findings from GWAS for markers in *VAX1* across three independent Asian populations, and identified rare variants in *PAX7* that may contribute to the etiology of CL(P). Determining the specific role of these rare variants in the etiology of CL(P) will require further investigations. We also found some evidence of interactions between SNPs in different genes suggesting they may act in the same pathway. These findings provide solid support for *PAX7* and *VAX1* as genes playing a role in human CLP.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We are particularly grateful to the many patients and their families who participated in this study. We were greatly assisted by many people including those in the Division of Research and Treatment for Oral and Maxillofacial Congenital Anomalies, Aichi-Gakuin University, Japanese Cleft Palate Foundation (Aichi-Gakuin University, Nagoya, Japan), a medical organization that provides clinical care for underserved populations in Japan, Vietnam, Mongolia, and Cambodia, and physicians and volunteers in Mongolia. We sincerely appreciate the outstanding support in Iowa of Sandy Daack Hirsh, Susie McConnell, Nancy Davin, Erin Brothers Smith, Raul Petrin, and Tamara Busch. The project was supported by the National Institute for Dental and Craniofacial Research (NIH) grants K99-DE022378-01, R37-DE08559 and R01-DE016148, and U01-DE-018993 to the consortium for GWAS genotyping and analysis. KAKENHI-Grant-in-Aid for Young Scientist (B) No. 20791560 (Aichi-Gakuin), Face Base grant U01 DE-20057.

References

Avila JR, Jezewski PA, Vieira AR, Orioli IM, Castilla EE, Christensen K, Daack-Hirsch S, Romitti PA, Murray JC. PVRL1 variants contribute to non-syndromic cleft lip and palate in multiple populations. Am J Med Genet A. 2006; 140:2562–2570. [PubMed: 17089422]

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 / without cleft palate using case-parent trios of European and Asian ancestry identifies *MAFB* and *ABCA4* as novel candidate genes. Nat Genet. 2010; 42:525–529. [PubMed: 20436469]

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, Bergé SJ, Reich RH, Schiefke F, Hemprich A, Pötzsch S, Steegers-Theunissen RP, Pötzsch B, Moebus S, Horsthemke B, Kramer FJ, Wienker TF, Mossey PA, Propping P, Cichon S, Hoffmann P, Knapp M, Nöthen MM, Mangold E. Key susceptibility locus for nonsyndromic cleft lip with or withgout cleft palate on chromosome 8q24. Nature Genet. 2009; 41:473–477. [PubMed: 19270707]

Chiquet BT, Lidral AC, Stal S, Mulliken JB, Moreno LM, Arcos-Burgos M, Valencia-Ramirez C, Blanton SH, Hecht JT. CRISPLD2: a novel NSCLP candidate gene. Hum Mol Genet. 2007; 16:2241–2248. [PubMed: 17616516]

Cordell HJ. Epistasis: what it means, what it doesn't mean, and statistical methods to detect it in humans. Hum Molec Genet. 2002; 11:2463–2468. [PubMed: 12351582]

Cordell HJ, Barratt BJ, Clayton DG. Case/Pseudocontrol Analysis in Genetic Association Studies: A Unified Framework for Detection of Genotype and Haplotype Associations, Gene-Gene and Gene-Environment Interactions, and Parent-of-Origin Effects. Genet Epi. 2004; 26:167–185.

Dickson SP, Wang K, Krantz I, Hakonarson H, Goldstein DB. Rare Variants Create Synthetic Genome-Wide Associations. PLoS Biol. 2010; 8:e1000294. [PubMed: 20126254]

Dixon MJ, Marazita ML, Beaty TH, Murray JC. Cleft lip and palate: understanding genetic and environmental influences. Nat Rev Genet. 2011; 12:167–178. [PubMed: 21331089]

Grant SFA, Wang K, Zhang H, Glaberson W, Annaiah K, Kim CE, Bradfield PJ, 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 non-syndromic cleft lip with or without cleft palate on 8q. 24. Journal of Pediatr. 2009; 155:909–913.

Hallonet M, Hollemann T, Wehr R, Jenkins NA, Copeland NG, Pieler T, Gruss P. Vax1 is a novel homeobox-containing gene expressed in the developing anterior ventral forebrain. Development. 1998; 125:2599–2610. [PubMed: 9636075]

Jezewski PA, Vieira AR, Nishimura C, Ludwig B, Johnson M, O'Brien SE. Complete sequencing shows a role for *MSX1* in nonsyndromic cleft lip and palate. J Med Genet. 2003; 40:399–407. [PubMed: 12807959]

Johansen CT, Wang J, Lanktree MB, Cao H, McIntyre AD, Ban MR, Martins RA, Kennedy BA, Hassell RG, Visser ME. Excess of rare variants in genes identified by genome-wide association study of hypertriglyceridemia. Nat Genet. 2010; 42:684–687. [PubMed: 20657596]

Jugessur A, Farlie PG, Kilpatrick N. The genetics of isolated orofacial clefts: from genotypes to subphenotypes. Oral Dis. 2009; 15:437–453. [PubMed: 19583827]

Leslie EJ, Murray JC. Evaluating rare coding variants as contributing causes to nonsyndromic cleft lip and palate. Clin Genet. 2012 In Press.

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

Keinan A, Clark AG. Recent explosive human population growth has resulted in an excess of rare genetic variants. Science. 2012; 336:740–743. [PubMed: 22582263]

Mangold E, Ludwig KU, Birnbaum S, Baluardo C, Ferrian M, Herms S, Reutter H, Almeida de Assis N, Al Chawa T, Mattheisen M. Genome-wide association study identifies two susceptibility loci for nonsyndromic cleft lip with or without cleft palate. Nat Genet. 2010; 42:24–26. [PubMed: 20023658]

Mansouri A, Stoykova A, Torres M, Gruss P. Dysgenesis of cephalic neural crest derivatives in Pax7−/− mutant mice. Development. 1996; 122:831–838. [PubMed: 8631261]

Marazita ML, Murray JC, Lidral AC, Arcos-Burgos M, Cooper ME, Maher BS, Daack-Hirsch S, Schultz R, Mansilla MA, Field LL, Liu YE, Prescott N, Malcolm S, Winter R, Ray A, Moreno L, Valencia C, Neiswanger K, Wyszynski DF, Bailey-Wilson JE, Albacha-Hejazi H, Beaty TH, McIntosh I, Hetmanski JB, Tunçbilek G, Edwards M, Harkin L, Scott R, Roddick LG. Meta-analysis of 13 gene scans reveals multiple cleft lip / palate genes with novel loci on 9q21 and 2q32–35. Am J Hum Genet. 2004; 75:161–173. [PubMed: 15185170]

McCarthy MI, Hirschhorn JN. Genome-wide association studies: potential next steps on a genetic journey. Hum Mol Genet. 2008; 17:156–165.

Moreno LM, Mansilla MA, Bullard SA, Cooper ME, Busch TD, Machida J, Johnson MK, Brauer D, Krahn K, Daack-hirsch S, L'heureux J, Valencia-Ramirez C, Rivera D, López AM, Moreno MA, Hing A, Lammer EJ, Jones M, Christensen K, Lie RT, Jugessur A, Wilcox AJ, Chines P, Pugh E, Doheny K, Arcos-Burgos M, Marazita ML, Murray JC, Lidral AC. *FOXE1* association with isolated cleft lip with or without cleft palate; and isolated cleft palate. Human Mol Genet. 2009; 18:4879–4896. [PubMed: 19779022]

Mossey, PA.; Little, J. Epidemiology of oral clefts: an international perspective. In: Wyszynski, DF., editor. Cleft Lip and Palate: From Origin to Treatment. Oxford University Press; 2002. p. 127-144.

Murray JC, Daack-Hirsch S, Buetow KH, Munger R, Espina L, Paglinawan N, Villanueva E, Rary J, Magee K, Magee W. Clinical and epidemiologic studies of cleft lip and palate in the Philippines. Cleft Palate Craniofac J. 1997; 34:7–10. [PubMed: 9003905]

Nikopensius T, Birnbaum S, Ludwig KU, Jagomägi T, Saag M, Herms S, Knapp M, Hoffmann P, Nöthen MM, Metspalu A, Mangold E. Susceptibility locus for non-syndromic cleft lip with or without cleft palate on chromosome 10q25 confers risk in Estonian patients. Eur J Oral Sci. 2010; 18:317–319. [PubMed: 20572868]

Pritchard JK. Are rare variants responsible for susceptibility to common diseases? Am J Hum Genet. 2001; 69:124–137. [PubMed: 11404818]

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklar P, de Bakker PIW, Daly MJ, Sham PC. PLINK: a toolset for whole-genome association and population-based linkage analysis. Am J Hum Genet. 2007; 81:559–575. [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-2α binding site in an *IRF6* enhancer is strongly associated with cleft lip. Nat Genet. 2008; 40:1341–1347. [PubMed: 18836445]

Riley BM, Mansilla MA, Ma J, Daack-Hirsch S, Maher BS, Raffensperger LM, Russo ET, Vieira AR, Dodé C, Mohammadi M, Marazita ML, Murray JC. Impaired FGF signaling contributes to cleft lip and palate. Proc Natl Acad Sci U S A. 2007; 104:4512–4517. [PubMed: 17360555]

Rojas-Martinez A, Reutter H, Chacon-Camacho O, Leon-Cachon RBR, Munoz-Jimenez SG, Nowak S, Becker J, Herberz R, Ludwig KU, Paredes-Zenteno M, Arizpe-Cantú A, Raeder S, Herms S, Ortiz-Lopez R, Knapp M, Hoffmann P, Nöthen MM, Mangold E. Genetic Risk Factors for Non-syndromic Cleft Lip with or without Cleft Palate in a Mesoamerican Population: Evidence for *IRF6* and Variants at 8q24 and 10q25. Birth Defects Research (Part A). 2010; 88:535–537.

Scapoli L, Martinelli M, Arlotti M, Palmieri A, Masiero E, Pezzetti F, Carinci F. Genes causing clefting syndromes as candidates for non-syndromic cleft lip with or without cleft palate: a family-based association study. Eur J Oral Sci. 2008; 116(6):507–511. [PubMed: 19049519]

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. 2011 In Press.

Schwender H, Li Q, Ruczinski I. Trio: Detection of disease-associated SNP interactions in case-parent trio data. R package version 1. 4. 23. Biometrics. 2012 In Press.

Slavotinek AM, Chao R, Vacik T, Yahyavi M, Vacik T, Abouzeid H, Bardakjian T, Schneider A, Shaw G, Sherr EH, Lemke G, Youssef M, Schorderet DF. *VAX1* mutation associated with microphthalmia, corpus callosum agenesis and orofacial clefting-the first description of a *VAX1* phenotype in humans. Hum Mutat. 2011; 33:364–8. [PubMed: 22095910]

Suzuki S, Marazita ML, Cooper ME, Miwa N, Hing A, Jugessur A, Natsume N, Shimozato K, Ohbayashi N, Suzuki Y, Niimi T, Minami K, Yamamoto M, Altannamar TJ, Erkhembaatar T, Furukawa H, Daack-Hirsch S, L'heureux J, Brandon CA, Weinberg SM, Neiswanger K, Deleyiannis FW, de Salamanca JE, Vieira AR, Lidral AC, Martin JF, Murray JC. Mutations in *BMP4* Are Associated with Sub-epithelial, Microform, and Overt Cleft Lip. Am J Hum Genet. 2009; 84:406–411. [PubMed: 19249007]

Tennessen JA, Bigham AW, O'Connor TD, Fu W, Kenny EE, Gravel S, McGee S, Do R, Liu X, Jun G, Kang HM, Jordan D, Leal SM, Gabriel S, Rieder MJ, Abecasis G, Altshuler D, Nickerson DA, Boerwinkle E, Sunyaev S, Bustamante CD, Bamshad MJ, Akey JM, Broad GO, Seattle GO. NHLBI Exome Sequencing Project. Evolution and functional impact of rare coding variation from deep sequencing of human exomes. Science. 2012; 337:64–69. [PubMed: 22604720]

Turhani D, Item CB, Watzinger E, Sinko K, Watzinger F, Lauer G, Ewers R. Mutation analysis of *CLPTM 1* and *PVRL1* genes in patients with non-syndromic clefts of lip, alveolus and palate. J Craniomaxillofac Surg. 2005; 33:301–306. [PubMed: 16122939]

Underhill DA. Genetic and biochemical diversity in the *Pax* gene family. Biochem Cell Biol. 2000; 78:629–638. [PubMed: 11103953]

Vieira AR, Avila JR, Daack-Hirsch S, Dragan E, Felix TM, Rahimov F, Harrington J, Schultz RR, Watanabe Y, Johnson M, Fang J, O'Brien SE, Orioli IM, Castilla EE, Fitzpatrick DR, Jiang R, Marazita ML, Murray JC. Medical sequencing of candidate genes for nonsyndromic cleft lip and palate. PLoS Genet. 2005; 1:64.

Am J Med Genet A. Author manuscript; available in PMC 2013 May 01.

Table I

Family based association analysis for Iowa case triads for all cleft sub-types using markers for *PAX7* and *VAX1.*

Gene	SNP	CLP (N[a]=65)	CLO (N=46)	CL(P)[b] (N=114)	CPO (N=40)	Total[c] (N=157)
VAX1	rs7078160	0.10	0.39	0.08	0.81	0.10
	rs4752028	0.13	1.00	0.26	0.32	0.10
PAX7	rs4920520	0.56	0.75	0.92	0.27	0.66
	rs766325	0.31	0.34	0.68	0.72	0.94
	rs6659735	0.05	0.34	0.02	0.37	0.007
Family based association analysis for all Asian populations combined for all cleft sub-types using markers for *PAX7* and *VAX1.*						
Gene	**SNP**	**CLP (N[a]=270)**	**CLO (N=103)**	**CL(P)[b] (N=440)**	**CPO (N=54)**	**Total[c] (N=494)**
VAX1	rs7078160	6.00E-07	0.02	4.6E-07	0.008	2.3E-08
	rs4752028	0.0003	0.01	0.00002	0.78	0.00004
PAX7	rs4920520	0.91	0.47	0.70	0.78	0.75
	rs766325	0.91	0.37	0.42	0.70	0.48
Family based association analysis for all four populations combined for all cleft sub-types using markers for *PAX7* and *VAX1.*						
Gene	**Location**	**CLP (N[a]=335)**	**CLO (N=149)**	**CL(P)[b] (N=554)**	**CPO (N=94)**	**Total[c] (N=651)**
VAX1	rs7078160	1.5E-07	0.01	1.1E-07	0.02	7.4E-09
	rs4752028	0.0001	0.02	0.00001	0.34	9.8E-06
PAX7	rs4920520	0.87	0.45	0.70	0.54	0.61
	rs766325	0.60	0.84	0.61	0.70	0.58

CLO is cleft lip only, CPO is cleft palate only, CLP is cleft lip and palate and CL(P) is cleft lip with or without cleft palate.

[a] Number of informative families

[b] CL(P) includes families with both CL and CLP in their cleft history

[c] Total includes cleft families where cleft type is unknown.

Table II

Newly described variants in *PAX7* and nucleotide changes following sequence analysis of Iowan, Philippine, Mongolian and Japanese individuals.

GRCh37/hg19	Amino Acid Change	Iowa		Filipinos		Polyphen
		Cases	Controls	Cases	Controls	
c.189,614G>C	G411R	0/81	0/81	1/92	0/87	Possibly damaging
c.189,617G>A	G412S	1/81	0/81	0/92	0/87	Benign
c.062,366G>A	G466S	0/81	0/81	1/810	0/87	Benign
c.088,401C>T		0/81	0/81	1/92	0/87	
c.018,438C>A		0/81	0/81	1/92	0/87	
c.957,713C>T		2/81	2/82	11/86	4/89	
c.958,083C>A		3/81	0/82	0/86	0/89	
c.957,933C>T		0/81	0/82	1/86	0/89	
c.072014C>A		2/83	0/82	0/86	0/86	
c.072317G>A		1/83	0/82	0/86	0/86	
c.189576C>T	P397L	0/81	1/81	0/92	0/87	Probably damaging
GRCh37/hg19		Mongolian		Japanese		
		Cases	Controls	Cases	Controls	
c.958,104G>A	A3T	1/87	ND[b]	0/88	ND[b]	Benign
c.958,140G>A	A15T	0/87	ND	1/87	ND	Benign
c.958,168G>A	G24E	0/87	ND	2/88	ND	Possibly damaging
c.961,014T>C		1/87	ND	0/88	ND	
c.961,088G>T		2/85	ND	0/88	ND	
c.961,515C>-		1/87	ND	0/88	ND	
c.029895A>T		8/87	ND	0/88	ND	
c.062,129A>C	M387L	1/87	ND	0/88	ND	Benign

[b]Not determined

Table III

Newly described variants in *VAX1* and nucleotide changes following sequence analysis of Iowan, Philippine, Mongolian and Japanese individuals.

GRCh37/hg19	Amino acid change	Iowans		Philippines		Polyphen
		Cases	Controls	Cases	Controls	
c.898,982>C		0/92	0/81	1/91	1/88	
rs148741605C>T		0/92	0/78	13/270	2/270	
c.840,949G>A		6/91	6/72	0/90	0/82	
c.840,439C>A		0/91	2/78	7/86	6/85	
GRCh37/hg19		Mongolian		Japanese		
		Cases	Controls	Cases	Controls	
c.897,839C>T		8/171	ND[b]	0/94	ND[b]	
c.896,171G>T		1/171	ND	0/94	ND	
c.896,062T>G	M117R	1/171	0/92	0/94	ND	Probably damaging
c.893,947C>T		0/169	0/92	1/72	ND	
c.893,882G>T		1/169	0/92	1/72	ND	
c.893,824G>T	A233S	1/169	0/92	0/72	ND	Benign

[b]Not determined