Genetics of Cleft Lip and Cleft Palate

ELIZABETH J. LESLIE AND MARY L. MARAZITA

Orofacial clefts are common birth defects and can occur as isolated, nonsyndromic events or as part of Mendelian syndromes. There is substantial phenotypic diversity in individuals with these birth defects and their family members: from subclinical phenotypes to associated syndromic features that is mirrored by the many genes that contribute to the etiology of these disorders. Identification of these genes and loci has been the result of decades of research using multiple genetic approaches. Significant progress has been made recently due to advances in sequencing and genotyping technologies, primarily through the use of whole exome sequencing and genome-wide association studies. Future progress will hinge on identifying functional variants, investigation of pathway and other interactions, and inclusion of phenotypic and ethnic diversity in studies.

© 2013 Wiley Periodicals, Inc.

KEY WORDS: cleft lip; cleft palate; genetics; syndromes; nonsyndromic


INTRODUCTION

Orofacial clefts, notably cleft lip (CL) and cleft palate (CP), are the most common craniofacial birth defects in humans and represent a substantial personal and societal burden. Clefts affect approximately 1 in 700 individuals [Rahimov et al., 2012] with a lifetime cost of treatment estimated at $200,000 [Berk and Marazita, 2002; Wehby and Cassell, 2010]. Affected individuals initially face difficulties feeding and also experience speech, hearing, and dental problems. Although clefts can be surgically repaired, patients often undergo multiple craniofacial and dental surgeries, as well as speech and hearing therapy. Despite these interventions, patients can experience lifelong psychosocial effects from the malformation. In fact, individuals born with a cleft have increased incidence of mental health problems and higher mortality rates at all stages of life [Christensen et al., 2004; Wehby and Cassell, 2010]. Clefting is also associated with a higher risk of various cancer types, including breast, brain, and colon cancers, in the individual with a cleft as well as their family members [Zhu et al., 2002; Bille et al., 2005; Menezes et al., 2009; Dietz et al., 2012]. The complications of clefting in early life are particularly devastating in developing countries where access to medical care may be limited [Wehby et al., 2006]. Understanding the etiologies of clefting is important not only for our knowledge of developmental biology, but ultimately for improved prevention, treatment, and prognosis for individuals affected by orofacial clefting.

DEVELOPMENT OF OROFACIAL CLEFTS

Orofacial clefts arise from failure of normal craniofacial developmental processes. Proper development of the face requires coordination of a complex series of events and includes cell growth, migration, differentiation, and apoptosis. The development of the face begins in the 4th week of development when neural crest cells migrate to form the five cranial primordia: the frontonasal prominence, the paired mandibular processes, and the paired maxillary processes (MxP). After facial prominences are formed, the nasal placodes invaginate to form the medial (MNP) and lateral (LNP) nasal processes. During the 6th and 7th weeks of gestation, the LNP merge with the MxP and then fuse with the MNP, forming the upper lip and...
Orofacial clefts arise from failure of normal craniofacial developmental processes. Proper development of the face requires coordination of a complex series of events and includes cell growth, migration, differentiation, and apoptosis. The development of the face begins in the 4th week of development when neural crest cells migrate to form the five facial primordia: the frontonasal prominence, the paired mandibular processes, and the paired maxillary processes (MxP).

primary palate (reviewed by Jiang et al. [2006]). Failure in growth or fusion of these processes results in orofacial clefting involving the upper lip, alveolus, and/or primary palate. The secondary palate begins to develop in the 7th week of embryogenesis when the palatal shelves (PS) emerge as outgrowths from the MxP. The PS initially grows vertically along the sides of the developing tongue but later elevate into a horizontal position as the tongue flattens (reviewed by Gritli-Linde [2007]). Continued growth leads to the PS meeting at the midline followed by fusion along the medial edge epithelia (MEE). Successful fusion of the secondary palate results in complete separation of the nasal and oral cavities. Clefts of the palate can arise due to failure at any of several steps including PS elevation, migration, or fusion.

OROFACIAL CLEFT PHENOTYPES

Orofacial clefts are a heterogeneous group of disorders affecting the structure of the face and oral cavity that have been divided into three general categories: those that affect the lip only (CL, Fig. 1A), those affecting the lip and palate (CLP, Fig. 1B), and those affecting the palate alone (CP, Fig. 1C). Historically, CL and CLP have been considered variants of the same defect that only differ in severity [Marazita, 2012]. Although the primary palate and secondary palate have distinct developmental origins, CL and CLP share a defect of the primary palate, motivating the inclusion of CL and CLP into a common group—CL with or without CP (CL/P) [Fogh-Andersen, 1942; Fraser, 1955]. However, epidemiological [Groen et al., 2010] and biological [Rahimov et al., 2008; Ludwig et al., 2012] data suggest that CL and CLP may have separate genetic etiologies. Nonetheless, common pathways may underlie the etiologies of each group, as occasionally both CL/P and CP are present within the same pedigree. This event is often referred to as mixed clefting, and is most commonly noted in syndromic forms of clefting [Rahimov et al., 2012].

SUBCLINICAL PHENOTYPES IN OROFACIAL CLEFT FAMILIES

Despite the range of phenotypic presentations, orofacial clefts are typically thought of as a simple, qualitative trait (unaffected vs. affected). Recent evidence, however, suggests that individuals with clefts lie on a spectrum of overt phenotypes (CL, CLP, and CP) and a range of subclinical features, which may also be present in “unaffected” relatives of cases [Weinberg et al., 2009]. These phenotypes include craniofacial measures [Weinberg et al., 2006], dental anomalies (tooth agenesis, microdontia, and supernumerary teeth) [Vieira et al., 2008], brain structural differences [Nopoulos et al., 2002; Weinberg et al., 2013], and dermatoglyphic lip print whorls [Neiswanger et al., 2009].

Subclinical phenotypes of the lip and palate include microform clefts (also known as congenital healed CL), defects of the orbicularis oris muscle [Neiswanger et al., 2007; Rogers et al., 2008; Weinberg et al., 2008], bifid uvula, submucous CP, and velopharyngeal insufficiency. These subclinical phenotypes may help explain incomplete penetrance or apparent lack of

---

**Figure 1.** Examples of nonsyndromic cleft lip and cleft palate. A: Bilateral cleft lip alone, (B) unilateral cleft lip plus cleft palate, (C) cleft palate alone. Images courtesy of FaceBase (www.FaceBase.org).
Mendelian inheritance patterns observed in families with overt clefts. Subclinical phenotypes might also explain an interesting phenomenon in discordant monozygotic twin pairs. Despite a lack of concordance, the recurrence risk for offspring of the affected and unaffected twin in discordant monozygotic twin pairs is largely identical [Groen et al., 2011]. Subphenotypes and subclinical phenotypes are also important to consider in the design of genetic studies because power is reduced when diverse phenotypes of different etiologies are merged. Incorporation of such phenotypic distinctions allow more biologically relevant groupings.

SYNDROMIC OROFACIAL CLEFTS

The designation of orofacial clefts as syndromic is usually based on the presence of additional physical or cognitive abnormalities. At least 275 syndromes, in which clefting is a primary feature, have been identified (http://www.ncbi.nlm.nih.gov/OMIM) and these are caused by mutation of a single genetic locus, chromosomal abnormalities, or teratogens. Of the described syndromes, 75% have a known genetic cause, including hundreds of syndromes due to Mendelian inheritance at a single genetic locus (summarized in Table I).

Van der Woude syndrome (VWS; OMIM #119300) is the most common form of syndromic clefting, accounting for approximately 2% of all CL/P cases, with a prevalence of 1/34,000 live births [Burdick, 1986]. VWS and its allelic disorder popliteal pterygium syndrome (PPS; OMIM #119500) are caused by mutations in IRF6 [Kondo et al., 2002]. Hundreds of mutations have been reported to cause these disorders, and while they are enriched in the DNA-binding domain of this transcription factor, they are also found throughout the protein-binding domain [de Lima et al., 2009; Leslie et al., 2012b].

Most clefting syndromes are rare, affecting only one in several hundred thousand live births. Identification of genes causing these syndromes has been very successful [Dixon et al., 2011] and has been further facilitated by advances in sequencing technology. Exome sequencing recently identified genes causing Kabuki syndrome [Ng et al., 2010a], Miller syndrome [Ng et al., 2010b], and Bartosocas–Papas syndrome [Kalay et al., 2012; Mitchell et al., 2012].

NONSYNDROMIC OROFACIAL CLEFTS

Epidemiology

The majority of orofacial cleft cases lack additional features and are categorized as “non-syndromic,” that is, 70% of all CL/P cases and 50% of all CPO cases [Jugessur et al., 2009a]. Although many studies have reported the prevalence of CL/P, those that have distinguished CLP from CL observed that CLP is twice as common as CL [Jensen et al., 1988].

The majority of orofacial cleft cases lack additional features and are categorized as “non-syndromic,” that is, 70% of all CL/P cases and 50% of all CPO cases [Jugessur et al., 2009a].

Although many studies have reported the prevalence of CL/P, those that have distinguished CLP from CL observed that CLP is twice as common as CL.

Among nonsyndromic clefts, CL/P is twice as frequent in males than in females, while CP is twice as frequent in females [Mossey et al., 2009]. Approximately 75% of clefts involving the lip are unilateral. Among unilateral clefts, those affecting the left side are twice as common as right-sided clefts [Gundlach and Mau, 2006].

Interestingly, the prevalence of nonsyndromic CL with or without CP (NSCL/P) also varies by ancestry. NSCL/P most commonly affects those of Asian or Amerindian descent (1/500 live births) and least commonly those of African descent (1/2,500) [Dixon et al., 2011]. Caucasian populations have an intermediate prevalence rate at approximately 1/1,000 [Mossey et al., 2009]. The influence of socioeconomic status on prevalence of orofacial clefts has not been conclusively determined [Carmichael et al., 2003; Clark et al., 2003; Yang et al., 2008]. Possible explanations for the differences in prevalence among geographic origins and socioeconomic statuses include environmental factors such as vitamin use, nutrition, access to medical care, and lifestyle risk factors such as smoking.

Evidence for Genetic Etiology in NSCL/P

NSCL/P is a genetically complex disorder caused by the interaction of multiple genetic and environmental risk factors. Although the familiality of orofacial clefts has long been noted (see, e.g., Trew [1757], Sproule [1863], and Darwin [1875]), formal genetic studies did not begin until Fogh-Andersen [1942] proposed that genetic factors contribute to NSCL/P after observing an increased frequency of clefting in relatives of a patient with a cleft [Fogh-Andersen, 1942]. Segregation analysis [Marazita et al., 1984] and twin studies [Mitchell, 2002] later supported a genetic component to NSCL/P. NSCL/P has a high rate of family recurrence; the risk for CL in first degree relatives is estimated to be 32 times the risk for individuals without a family history of CL [Sivertsen et al., 2008]. The concordance rate of 40–60% in monozygotic twins is higher than the 3–5% rate in dizygotic twins and also suggests a strong, but not purely, genetic etiology [Little and Bryan, 1986].

Environmental Risk Factors

Epidemiological data support a role for environmental risk factors in the development of orofacial clefts. Maternal
smoking has been consistently associated with an increased risk of clefting, with a population-attributable risk estimated as high as 20% and an odds ratio of 1.3 for CLP [Shi et al., 2008]. While alcohol is an established teratogen [West and Blake, 2005], evidence supporting a role for maternal alcohol use in clefting has been inconsistent [Murray, 2002]. However, some support for maternal alcohol consumption comes from an association between clefting and genetic variants in the alcohol dehydrogenase gene ADH1C [Jugessur et al., 2009b]. Moreover, a recent study demonstrated

<table>
<thead>
<tr>
<th>Syndrome</th>
<th>Cleft type observed</th>
<th>Gene</th>
<th>Refs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ankyloblepharon-ectodermal dysplasia-clefting</td>
<td>CL/P</td>
<td>TP63</td>
<td>McGrath et al. [2001]</td>
</tr>
<tr>
<td>Apert</td>
<td>CP</td>
<td>FGFR2</td>
<td>Wilkie et al. [1995]</td>
</tr>
<tr>
<td>Bamforth–Lazarus</td>
<td>CP</td>
<td>FOXE1</td>
<td>Bamforth et al. [1989]</td>
</tr>
<tr>
<td>Bartoscas-Papas</td>
<td>CL/P</td>
<td>RIK4</td>
<td>Kalay et al. [2012], Mitchell et al. [2012]</td>
</tr>
<tr>
<td>Branchio-oculo-facial</td>
<td>CL/P</td>
<td>TFAP2A</td>
<td>Milunsky et al. [2008]</td>
</tr>
<tr>
<td>Campomelic dysplasia</td>
<td>CP</td>
<td>SOX9</td>
<td>Foster et al. [1994], Wagner et al. [1994]</td>
</tr>
<tr>
<td>CHARGE</td>
<td>CP</td>
<td>CHD7</td>
<td>Vissers et al. [2004]</td>
</tr>
<tr>
<td>CLP ectodermal dysplasia</td>
<td>CL/P</td>
<td>PVRL1</td>
<td>Suzuki et al. [2000]</td>
</tr>
<tr>
<td>Cornelia de Lange</td>
<td>CP</td>
<td>NIPBL</td>
<td>Krantz et al. [2004], Tonkin et al. [2004]</td>
</tr>
<tr>
<td>Crouzon</td>
<td>CP</td>
<td>FGFR2</td>
<td>Reardon et al. [1994]</td>
</tr>
<tr>
<td>DiGeorge</td>
<td>CP</td>
<td>TBX1</td>
<td>Packham and Brook [2003]</td>
</tr>
<tr>
<td>Eccodontally-ectodermal dysplasia-clefting</td>
<td>CL/P</td>
<td>TP63</td>
<td>Celli et al. [1999]</td>
</tr>
<tr>
<td>Familial gastric cancer and CLP</td>
<td>CL/P</td>
<td>CDH1</td>
<td>Frebourg et al. [2006]</td>
</tr>
<tr>
<td>Gorlin</td>
<td>CL/P</td>
<td>PTCH1</td>
<td>Hahn et al. [1996], Johnson et al. [1996]</td>
</tr>
<tr>
<td>Holoprosencephaly</td>
<td>CL/P</td>
<td>GLI2</td>
<td>Roessler et al. [2003]</td>
</tr>
<tr>
<td>Holoprosencephaly</td>
<td>CL/P</td>
<td>SHH</td>
<td>Roessler et al. [1996]</td>
</tr>
<tr>
<td>Holoprosencephaly</td>
<td>CL/P</td>
<td>SIX3</td>
<td>Wallis et al. [1999]</td>
</tr>
<tr>
<td>Holoprosencephaly</td>
<td>CL/P</td>
<td>TGF</td>
<td>Gripp et al. [2000]</td>
</tr>
<tr>
<td>Isolated cleft palate</td>
<td>CP</td>
<td>SATB2</td>
<td>FitzPatrick et al. [2003]</td>
</tr>
<tr>
<td>Kabuki</td>
<td>CL/P</td>
<td>MLL2, KDM6A</td>
<td>Lederer et al. [2012], Ng et al. [2010a]</td>
</tr>
<tr>
<td>Kallmann</td>
<td>CL/P</td>
<td>FGFR1</td>
<td>Dode et al. [2003]</td>
</tr>
<tr>
<td>Lethal and Escobar multiple pterygium</td>
<td>CP</td>
<td>CHRG</td>
<td>Morgan et al. [2006]</td>
</tr>
<tr>
<td>Loey-Dietz</td>
<td>CP</td>
<td>TFGBR1, TFGBR2</td>
<td>Loey et al. [2005]</td>
</tr>
<tr>
<td>Miller</td>
<td>CP</td>
<td>DHO</td>
<td>Ng et al. [2010b]</td>
</tr>
<tr>
<td>Oculo-otospondi-vertebral</td>
<td>CP</td>
<td>BCO2</td>
<td>Ng et al. [2004]</td>
</tr>
<tr>
<td>Optiz G/BBB</td>
<td>CL/P</td>
<td>MID1</td>
<td>Quadri et al. [1997]</td>
</tr>
<tr>
<td>Oro-facial-digital</td>
<td>CL/P</td>
<td>GLI3</td>
<td>Johnston et al. [2010]</td>
</tr>
<tr>
<td>Oro-facial-digital type 1</td>
<td>CL/P</td>
<td>OFD1</td>
<td>Ferrante et al. [2001]</td>
</tr>
<tr>
<td>Otopalatodigital types 1 and 2</td>
<td>CP</td>
<td>FLNA</td>
<td>Robertson et al. [2003]</td>
</tr>
<tr>
<td>Pierre Robin</td>
<td>CP</td>
<td>SOX9</td>
<td>Benko et al. [2009]</td>
</tr>
<tr>
<td>Popliteal pterygium</td>
<td>CL/P</td>
<td>IRF6</td>
<td>Kondo et al. [2002]</td>
</tr>
<tr>
<td>Saethre-Chotzen</td>
<td>CL/P</td>
<td>TWIST1</td>
<td>El Ghouzzi et al. [1997], Howard et al. [1997]</td>
</tr>
<tr>
<td>Stickler type 1</td>
<td>CP</td>
<td>COL2A1</td>
<td>Snead and Yates [1999]</td>
</tr>
<tr>
<td>Stickler type 2</td>
<td>CP</td>
<td>COL1A1, COL1A2</td>
<td>Snead and Yates [1999]</td>
</tr>
<tr>
<td>Tetra-amelia with CLP</td>
<td>CL/P</td>
<td>WNT3</td>
<td>Niemann et al. [2004]</td>
</tr>
<tr>
<td>Tooth agenesis with or without cleft</td>
<td>CL/P</td>
<td>MSX1</td>
<td>van den Boogaard et al. [2000]</td>
</tr>
<tr>
<td>Treacher Collins</td>
<td>CP</td>
<td>TCOF1</td>
<td>Group [1996]</td>
</tr>
<tr>
<td>Van der Woude</td>
<td>CL/P</td>
<td>IRF</td>
<td>Kondo et al. [2002]</td>
</tr>
<tr>
<td>X-Linked cleft palate and ankyloglossia</td>
<td>CP</td>
<td>TBX22</td>
<td>Braybrook et al. [2001]</td>
</tr>
<tr>
<td>Siderius X-linked mental retardation</td>
<td>CL/P</td>
<td>PHF8</td>
<td>Laumonnier et al. [2005]</td>
</tr>
</tbody>
</table>

CL, cleft lip; CP, cleft palate; CL/P, cleft lip with or without cleft palate.
that the combination of ADH1C variants with reduced enzymatic activity and heavy maternal alcohol use increased the risk for orofacial clefts [Boyles et al., 2010]. However, a role for alcohol may be confounded by other risk factors such as nutrition, smoking, or stress that can be associated with alcohol consumption in some contexts.

Nutrition during pregnancy has been suggested as another contributing factor based on observational and interventional studies using folic supplements as a preventive measure [Wehby and Murray, 2010]. The beneficial effect of folate use, however, remains controversial and has not been consistently replicated [Wilcox et al., 2007; Wehby and Murray, 2010]. Other nutrients, including cholesterol [Porter, 2006], zinc [Munger et al., 2009], and general multivitamin use [Johnson and Little, 2010], have also been studied, but need to be expanded to larger populations. Finally, other exposures to teratogens and environmental toxins have also been associated with increased risk of clefting [Abbott, 2010] such as retinoic acid, valproic acid, and phenytoin. A more controversial and has not been consistently replicated [W ehby and Murray, 2010]. The beneficial effect of folate use, however, remains controversial and has not been consistently replicated [Wilcox et al., 2007; Wehby and Murray, 2010]. Other nutrients, including cholesterol [Porter, 2006], zinc [Munger et al., 2009], and general multivitamin use [Johnson and Little, 2008], have also been studied, but need to be expanded to larger populations.

Genomic rearrangements and copy number variants
Genomic rearrangements are caused by improper recombination of chromosomes and include deletions, duplications, translocations, and inversions that can occur within or between chromosomes. Analysis of breakpoints in patients with balanced rearrangements has identified CLPTM1 [Yoshii rura et al., 1998], SATB2 [FitzPatrick et al., 2003], SUMO1 [Alkuraya et al., 2006], and FGFR1 [Kim et al., 2005] as candidate genes for CL/P, and implicated 9q and 17q as potential risk loci [Machida et al., 2009]. In contrast, copy number variants (CNVs) and microdeletions are submicroscopic gains or losses of DNA segments ranging from kilobases to megabases. Although most CNVs have been found in syndromic forms of clefting, such as DiGeorge syndrome and VWS, recent studies have focused on the role of CNVs and microdeletions in nonsyndromic forms of clefting [Osoegawa et al., 2008; Menezes et al., 2009]. Genes implicated in NSCL/P from these approaches include FGFR2 [Osoegawa et al., 2008], TTFAP2A [Shi et al., 2009], and SUMO1 [Shi et al., 2009].

Candidate gene approaches
In contrast to genome-wide approaches, such as linkage analysis, selection of candidate genes for study typically relies on a priori knowledge about the biological processes involved and can utilize a variety of sources, from animal models to Mendelian clefting syndromes. Mouse models with spontaneous clefts, those induced by chemical mutagenesis, or by knockout experiments have all been used for gene discovery. Expression analysis is another powerful tool for identifying and focusing lists of candidate genes. The Craniofacial and Oral Gene Expression Network database (COGENE, now available through the FaceBase online resource, www.F aceBase.org) catalogs gene expression patterns from early stage human embryos, while the EMAGE database catalogs extensive gene expression information for the developing mouse embryo [Armit et al., 2012].

Candidate genes can also be selected based on their role in syndromes that include clefting as part of the phenotype (see Table I). The rationale for this approach is that while rare, damaging mutations in a gene can cause a Mendelian clefting syndrome, the common, less deleterious variants of the same gene can contribute to a similarly less severe, isolated cleft [Stanier and Moore, 2004]. VWS is an excellent model for this as approximately 15% of VWS patients present as phenocopies of isolated, nonsyndromic clefts. While mutations in IRF6 are causative for VWS and PPS, variants in and around IRF6 have been associated with NSCL/P [Kondo et al., 2002; Zuccher o et al., 2004; Rahimov et al., 2008]. Resequencing of candidate genes has successfully identified specific variants that may contribute to the statistical associations of several candidate genes including MSX1, FGFR1, FGFR8, and BMP4 [Leslie and Murray, 2012].
Genome-wide association studies

GWAS have become widely used for their unbiased approach for identifying candidate genes or loci associated with complex traits such as NSCL/P. To date, four independent GWAS and a meta-analysis have been published for NSCL/P [Birnbaum et al., 2009; Grant et al., 2009; Beaty et al., 2010; Mangold et al., 2010; Ludwig et al., 2012] (positive results summarized in Table II). Birnbaum et al. [2009] found an extremely strong association between markers on 8q24 and NSCL/P in a German population which was subsequently replicated in a population from the United States [Grant et al., 2009]. In the third study, Mangold et al. [2010] identified additional significant signals near VAX1 on chromosome 10q25 and NOG on chromosome 17q22. The fourth GWAS of NSCL/P was performed by the GENEVA Cleft Consortium study [Beaty et al., 2010]. Novel features of this study were the use of case-parent trios for the first time in a NSCL/P GWAS and the inclusion of a variety of families of European, Asian, and mixed ancestry. In the combined analysis for all populations, this study confirmed the previous associations with 1q32 and 8q24 and identified novel loci on 1p22 and 20q12. When stratified by population, markers near 1q32, 1p22, and 20q12 reached genome-wide significance in Asians, while only the 8q24 signal was formally significant in Europeans. An examination of marker information content and haplotype diversity for the 8q24 locus demonstrated low heterozygosity and low haplotype diversity in Asian populations [Murray et al., 2012]. The inclusion of additional SNPs by imputation revealed a strong, but not genome-wide significant, signal in Asian populations. Thus, the absence of a signal in Asian populations in the GWAS may not be due to a population effect, but lack of power due to low minor allele frequency.

To identify additional susceptibility loci, Ludwig et al. [2012] performed a meta-analysis by combining the GENEVA Cleft Consortium and Mangold et al. studies, which are the largest of the published GWAS. Combining the European case-control data with the European–American trios resulted in six loci reaching genome-wide significance (8q24, 10q25, 17q22, 2p21, 13q31, and 15q22). Six additional regions reached genome-wide significance (1p36, 1p22, 1q32, 3p11, 8q21, and 20q12). Mangold et al. also performed separate analyses for CL and CLP, demonstrating that the 13q31 locus was exclusively associated with CLP.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1p22</td>
<td>ARHGAP29</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1p36</td>
<td>PAX7</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1q32</td>
<td>IRF6</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2p21</td>
<td>THADA</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3p11</td>
<td>EPHA3</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8q21</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8q24</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10q25</td>
<td>VAX1</td>
<td>X</td>
<td>x</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13q31</td>
<td>SPRY2</td>
<td>x</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15q22</td>
<td>TPM1</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17q22</td>
<td>NOG</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20q12</td>
<td>MAFB</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

X, results reach formal genome-wide significance; x, results approach genome-wide significance.
Unlike the multiple studies of NSCL/P, there are few genetic studies of nonsyndromic CP. SNPs from the four European NSCL/P GWAS loci (1q32, 8q24, 10q25, and 17q22) were tested in CP trios, but were not significant [Mangold et al., 2010]. A GWAS for CP was recently completed but lacked statistically significant SNP main effects. However, with the addition of environmental exposures, several significant gene–environment interaction effects emerged [Beaty et al., 2011]. There was little or no overlap between loci associated with NSCL/P and those associated with CP, suggesting that these clefting phenotypes have distinct etiologies and supporting the historical separation of NSCL/P from CP.

Genes and Genetic Loci Best-Supported as Involved in NSCL/P

There are now multiple genes and genetic loci with support for involvement in NSCL/P from multiple lines of evidence. Many of these also have biological plausibility, as will now be reviewed.

IRF6

Interferon regulatory factor 6 (IRF6) was first implicated in orofacial clefting when it was discovered to be the cause for two autosomal-dominant clefting syndromes: VWS and PPS [Kondo et al., 2002]. 15% of VWS patients lack lip pits, making them clinically indistinguishable from individuals with nonsyndromic clefts. This led to the hypothesis that hypomorphic alleles in IRF6 might contribute to the etiology of NSCL/P, which was confirmed in a large study [Zucchero et al., 2004] involving several populations and subsequently replicated in genome–wide linkage studies [Marazita et al., 2004, 2009]. GWAS [Birnbaum et al., 2009; Grant et al., 2009; Beaty et al., 2010] and numerous candidate gene studies [Ghassibe et al., 2005; Park et al., 2007; Jugesur et al., 2008]. A study later identified a common SNP in the MCS-9.7 regulatory element that is over transmitted in nonsyndromic CL [Rahimov et al., 2008]. The risk allele disrupts a highly conserved binding site for transcription factor AP-2α, which is mutated in the autosomal dominant clefting syndrome branchio-oculo-facial syndrome, which includes pits of the upper lip among its phenotypic features [Milunsky et al., 2008].

MAFB

The MAFB gene encodes a basic leucine zipper transcription factor that is involved in development and differentiation of keratinocytes [Borrelli et al., 2010] and at least in the mouse has been shown to be expressed in the PS and the MEE during palatal fusion [Beaty et al., 2010]. Several markers on chromosome 20q12 near MAFB achieved genome-wide significance in the GENEVA Cleft Consortium GWAS in European and Asian populations, although the signal was much stronger in Asians [Beaty et al., 2010]. Several independent studies using diverse populations have replicated this association [Beaty et al., 2010; Pan et al., 2011; Lennon et al., 2012]. Sequencing of the single MAFB exon identified a missense mutation, H131Q (rs121912307), present in 3.5% of Filipinos with CL/P but only 0.7% of controls (P < 0.0001). This variant is located in a poly-histidine tract, a motif that has been implicated in subcellular localization in other proteins. Although the function of this variant (and motif) is currently unknown, it is bioinformatically predicted to damage protein structure and/or function. Mutations in the transactivation domain cause multicentric carpotarsal osteolysis (MCTO), an autosomal dominant skeletal dysplasia characterized by progressive bone resorption [Zankl et al., 2012; Dworschak et al., 2013]. Individuals with MCTO can have craniofacial abnormalities such as triangular faces, micrognathia, and exophthalmos, but have not been reported to have CL/P.

ARHGAP29

Multiple markers on chromosome 1p22, located in and around the ABCA4 gene, which encodes an ATP-binding cassette transporter, reached genome-wide significance by GWAS [Beaty et al., 2010; Ludwig et al., 2012]. The most strongly associated SNPs were replicated in independent populations [Beaty et al., 2010; Pan et al., 2011; Lennon et al., 2012; Yildirim et al., 2012], which also showed a stronger signal in Asian families than in European ones. Mutations in ABCA4 cause several autosomal recessive retinal diseases [Tsibovsky et al., 2010]. Despite identifying numerous missense mutations in this gene in individuals with NSCL/P, expression of mouse Abca4 was restricted to the retina at the RNA and protein level [Beaty et al., 2010]. Leslie et al. [2012a] demonstrated craniofacial expression of the adjacent gene Arhgap29, which was decreased in If6 knockout mice. Mutation screening of ARHGAP29 identified multiple rare coding variants in families with NSCL/P. ARHGAP29 encodes Rho GTPase activating protein 29, a protein that mediates the cyclical regulation of small GTP binding proteins [Saras et al., 1997] which are involved in many functions critical for craniofacial development related to cellular shape, movement, cell–cell interactions, and proliferation [Mosey et al., 2009].

8q24

8q24 is a gene desert identified with striking significance in several complex diseases including NSCL/P [Birnbaum et al., 2009; Grant et al., 2009; Beaty et al., 2010; Mangold et al., 2010], prostate cancer [Haiman et al., 2007], colorectal cancer [Zanke et al., 2007; Haerian et al., 2011], bladder cancer [Kremenek et al., 2008], and breast cancer [Easton et al., 2007]. The 8q24 region was strongly associated with NSCL/P in Caucasian populations and was independently identified by all GWAS. The closest gene to this locus is the oncogene MYC, which directly interacts with at least one part of the 8q24 locus [Sotelo et al., 2010]. Myc is strongly expressed in the mandible and maxilla at E14.5 [Richardson et al., 2010] and is also required for neural crest cell formation [Bellmeyer et al., 2003; Wei et al., 2007]. Several regions demonstrating enhancer activity in craniofacial tissue have been identified at the 8q24 locus (www.FaceBase.org), leading to the hypothesis
that this region contains multiple regulatory elements critical for proper craniofacial development.

**VAX1**

VAX1, ventral anterior homeobox 1, is a transcriptional regulator containing a DNA-binding homeobox domain. Markers in or near VAX1 approached genome-wide significance in studies by Mangold et al. [2010] and the GENEVA Cleft Consortium [Beaty et al., 2010]; this association has been replicated in three independent Asian populations [Butali et al., 2013]. Resequencing of VAX1 failed to identify an excess of rare variants in NSCL/P [Nasser et al., 2012]. Vax1 is expressed in several craniofacial structures and mice deficient for Vax1 develop CP [Hallonet et al., 1999]. Recently a homozygous missense mutation was described in a child from a consanguineous family with bilateral microphthalmia, bilateral CLP, and corpus callosum agenesis [Slavotinek et al., 2013], mimicking the phenotype of the Vax1-/- mouse.

**PAX7**

Paired box protein Pax-7 (PAX7) is a transcription factor that in the mouse has been shown to have a role in neural crest development by regulating the expression of neural crest markers Slug, Sox9, and Sox10 [Basch et al., 2006]. Pax7 is expressed in the PS, Meckel's cartilage, and various nasal structures including the nasal epithelium. Mutant Pax7 mice have malformations of the maxilla and nose, confirming its role in craniofacial development [Mansouri et al., 1996]. In humans, several markers around PAX7 approached genome-wide significance by GWAS [Beaty et al., 2010] and meta-analysis [Ludwig et al., 2012], suggesting a role for common variants of PAX7 in the etiology of NSCL/P. Notably, PAX7 was previously associated with NSCL/P in four populations in a candidate gene association study [Sull et al., 2009].

**SUMMARY AND RECOMMENDATIONS**

Clefts of the lip and palate have been of interest in the scientific literature since at least the 1700s, and currently represent one of the major success stories in applying modern molecular genetic techniques to a common, complex disorder. There have been major successes in sequencing Mendelian forms to identify causative variants (see Table I) and in performing genome-wide linkage and association approaches to identify genes and regions involved in NSCL/P (see Table II). Given the strong and consistent results obtained from GWAS, meta-analysis, and other genetic studies of NSCL/P, a number of research groups are now conducting sequencing studies to identify specific functional variants in the genes and regions identified, as well as their regulatory regions.

**ACKNOWLEDGMENTS**

The authors have no conflicts of interest to declare. Many thanks to our many colleagues and mentors in the field of cleft lip/palate genetics. This review was partially supported by NIH grants T15-LM007059 (E.J.L.), DE016148 and DE0200570 (M.L.M.).

**REFERENCES**


Letra A, Menezes R, Govil M, Fonseca RF, McHenry T, Granjeiro JM, Castilla EE, Orioli IM, Marazita ML, Vieira AR. 2010. Follow-up association studies of


Vitale Palace Craniofac J 45:518–524.


