Cleft lip and palate: understanding genetic and environmental influences

Michael J. Dixon, Mary L. Marazita‡, Terri H. Beaty§ and Jeffrey C. Murray||*

Abstract | Clefts of the lip and/or palate (CLP) are common birth defects of complex aetiology. CLP can occur in isolation or as part of a broad range of chromosomal, Mendelian or teratogenic syndromes. Although there has been marked progress in identifying genetic and environmental triggers for syndromic CLP, the aetiology of the more common non-syndromic (isolated) forms remains poorly characterized. Recently, using a combination of epidemiology, careful phenotyping, genome-wide association studies and analysis of animal models, several distinct genetic and environmental risk factors have been identified and confirmed for non-syndromic CLP. These findings have advanced our understanding of developmental biology and created new opportunities for clinical translational research.

Nares
The nostrils or nasal passages

Primary palate
The anterior portion of the palate including the bony component in humans.

**Faculty of Medical and Human Sciences, Manchester Academic Health Sciences Centre, Michael Smith Building, University of Manchester, Oxford Road, Manchester M13 9PT, UK.*
‡Center for Craniofacial and Dental Genetics, Department of Oral Biology, School of Dental Medicine, Department of Human Genetics, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, Pennsylvania 14219, USA.
§Department of Epidemiology, School of Public Health, Johns Hopkins University, Baltimore, Maryland 21205, USA.
||Departments of Pediatrics, Biology and Epidemiology, University of Iowa, Iowa 52242, USA.
Correspondence to J.C.M.
e-mail: jeff-murray@uiowa.edu
doi:10.1038/nrg2933

Clefts of the lip and/or palate (CLP) are immediately recognizable disruptions of normal facial structure. Although not a major cause of mortality in developed countries, CLP does cause considerable morbidity to affected children and imposes a substantial financial risk for families with a concomitant societal burden[1]. Individuals with CLP may experience problems with feeding, speaking, hearing and social integration that can be corrected to varying degrees by surgery, dental treatment, speech therapy and psychosocial intervention. CLP is aetiologically heterogeneous, and this has crucial implications for understanding the biology of facial development, how environmental risks interact with genetic factors and how we can incorporate known aetiologic variables to improve clinical care. Recent successes in genome-wide linkage and association studies have identified novel loci that are significantly associated with CLP[2–6]. Researchers are currently striving to identify the aetiologic variants at these novel loci to understand the developmental disturbances leading to CLP. This knowledge should eventually result in improved prevention, treatment and prognosis for individuals with these conditions.

Development of the lip and palate is outlined in FIG. 1. The common forms of CLP involve disruption of tissue planes above the lip, extending into the nares and/or the palate (hard and/or soft) (FIG. 2). Fogh-Andersen and Fraser[7,8] noted that clefts involving the anterior structures (lip and primary palate) could be separated on both genetic and embryological grounds from those involving only the secondary palate. Although there are many disruptions affecting the craniofacial complex, the overwhelming majority involve only the upper lip and/or palate. Further, approximately 70% of cases of CLP occur as isolated entities with no other apparent cognitive or craniofacial structural abnormalities; this is commonly termed 'isolated, non-syndromic CLP'. Because the defects arise early in embryological development, have a complex aetiology (with both genetic and environmental contributions) and modest recurrence rates, it has proven difficult to identify specific aetiologic factors. A combination of epidemiologic, candidate gene and genome-wide studies, plus analysis of animal models, has recently provided deeper insights into the causes of non-syndromic CLP.

With the advent of the genomics era, there have been major advances in the identification of causative genetic mutations underlying syndromic forms of CLP (see the OMIM website for further information). By contrast, there has been less progress in advancing our understanding of the genetic aetiology of non-syndromic CLP owing to its genetic heterogeneity, departure from Mendelian inheritance patterns, the limited availability and expense of genomic tools and the necessity for very large data sets. However, the recent development of innovative approaches to phenotyping and powerful, cost-effective genomic tools, together with extrapolation from studies of syndromic forms of CLP, have increased our understanding of non-syndromic CLP. Because of its particular challenges, in this Review we focus on non-syndromic CLP and we summarize syndromic forms (which are genetically tractable) only briefly. We discuss important epidemiologic clues, environmental contributions, genetic architecture and issues of phenotyping,

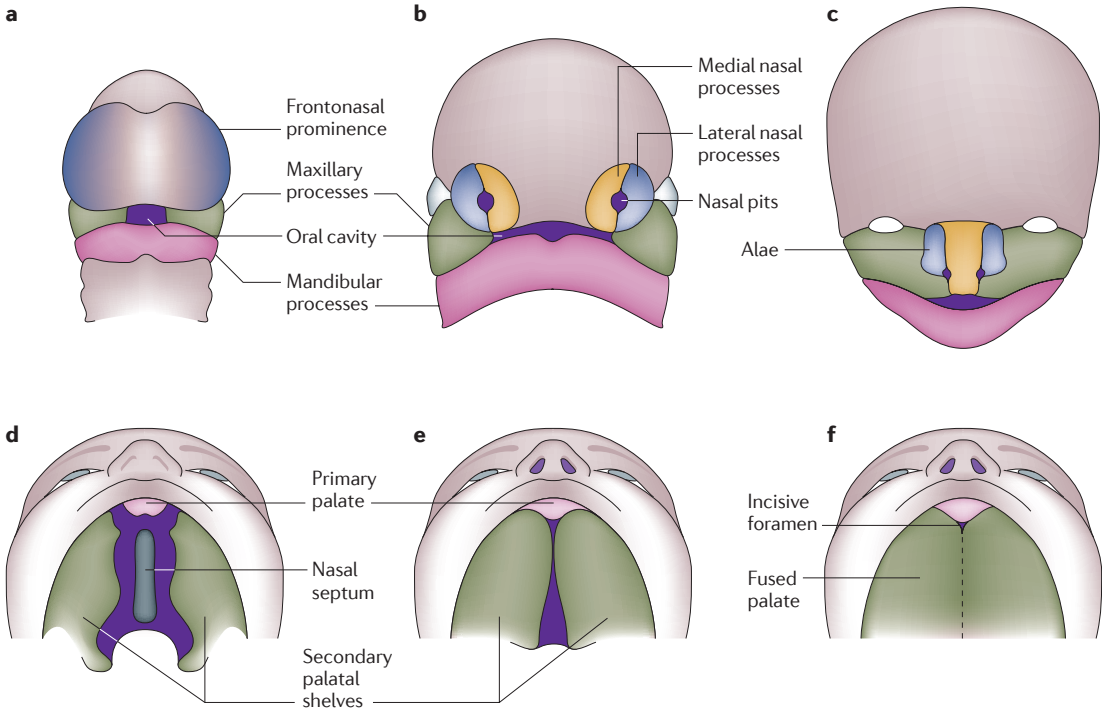

Figure 1 | **Development of the lip and palate.** Schematic diagrams of the development of the lip and palate in humans. **a** | The developing frontonasal prominence, paired maxillary processes and paired mandibular processes surround the primitive oral cavity by the fourth week of embryonic development. **b** | By the fifth week, the nasal pits have formed, which leads to the formation of the paired medial and lateral nasal processes. **c** | The medial nasal processes have merged with the maxillary processes to form the upper lip and primary palate by the end of the sixth week. The lateral nasal processes form the nasal alae. Similarly, the mandibular processes fuse to form the lower jaw. **d** | During the sixth week of embryogenesis, the secondary palate develops as bilateral outgrowths from the maxillary processes, which grow vertically down the side of the tongue. **e** | Subsequently, the palatal shelves elevate to a horizontal position above the tongue, contact one another and commence fusion. **f** | Fusion of the palatal shelves ultimately divides the oronasal space into separate oral and nasal cavities. Figure is modified, with permission, from REF. 137 © (2009) John Wiley and Sons Ltd.

gene discovery and insights into molecular pathogenesis. We also speculate about the implications of these findings for estimating recurrence, finding new clinical associations building on advances in imaging and using large databases to examine long-term outcomes.

Challenges in studying CLP

Epidemiology. CLP affects approximately 1 in 700 live births, with wide variability across geographic origin, racial and ethnic groups, as well as environmental exposures and socioeconomic status. In general, Asian and Native American populations have the highest reported birth prevalence rates, which are often as high as 1 in 500. European-derived populations have intermediate prevalence rates at approximately 1 in 1,000, and African-derived populations have the lowest prevalence rates at approximately 1 in 2,500. These observations suggest that the relative contribution of individual susceptibility genes may vary across different populations[6,9,10]. The frequency of CLP also differs by gender and laterality: there is a 2:1 male to female ratio for clefts involving the lip, approximately a 1:2 male to female ratio for clefts of the palate only and a 2:1 ratio of left to right sided clefts among unilateral cleft lip cases.

Historically, CLP has been divided into cleft palate only and cleft lip with or without cleft palate (CL/P)[7,8]. However, recent epidemiologic data suggest that cleft lip only may have unique aetiologic features, including strong genetic associations, whereas some individuals with cleft palate only show evidence of subclinical cleft lip[11–15]. Nevertheless, this broad sub-division of anatomical defects is consistent with the distinct developmental origins of the lip/primary palate versus the secondary palate. Furthermore, separate cellular and genetic aetiologies for CL/P and cleft palate only are consistent with the general observation that these two conditions do not segregate in the same family, although exceptions have been reported for families with aetiologic mutations in specific genes (for example, tumour protein p63 (*TP63*), msh homeobox 1 (*MSX1*), interferon regulatory factor 6 (*IRF6*) and fibroblast growth factor receptor 1 (*FGFR1*))[16–20]. Approximately 70% of all cases of CL/P and 50% of cases of cleft palate only are considered to be non-syndromic[21–23]. The remaining cases are composed of a wide range of malformation syndromes, including over 500 Mendelian syndromes (see the OMIM website for further information) as well as those arising secondary to chromosomal or teratogenic

Secondary palate
Posterior or soft palate in humans.

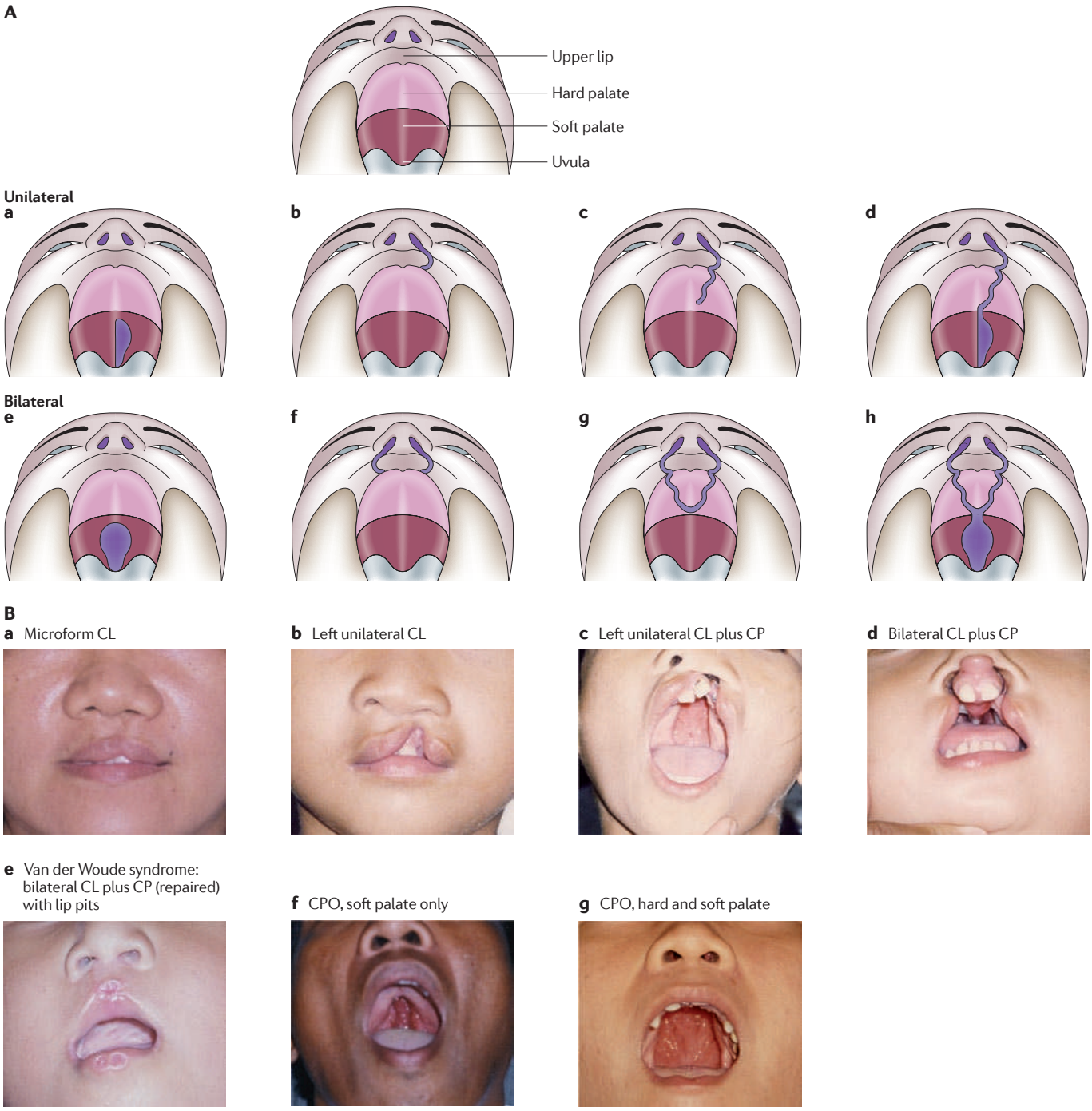

Figure 2 | **Types of cleft. A** | Illustrative drawings of types of cleft lip and/or palate (CLP)[114]. **a** and **e** show unilateral and bilateral clefts of the soft palate; **b**, **c** and **d** show degrees of unilateral cleft lip and palate; **f**, **g** and **h** show degrees of bilateral cleft lip and palate. Clefts are indicated in purple. **B** | A collection of images of different types of clefts, some with associated anomalies such as lip pits. Descriptions are given above the images. CL, cleft lip; CP, cleft palate; CPO, cleft palate only. Images collected during J.C.M.'s research. Part **A** is modified, with permission, from REF. 114

effects. These syndromic forms are more tractable to genetic analysis, and BOX 1 provides a summary of a subset of syndromes in which the underlying genetic mutation has been identified (see also Supplementary information S1 (table)).

Genetic architecture and phenotyping. Whereas twin studies and familial clustering studies have provided compelling evidence for a genetic component to non-syndromic CLP[24], few pedigrees show clear-cut Mendelian inheritance and most cases appear to be

Box 1 | **Clefting syndromes in which the mutated gene has been identified**

Cleft lip ± cleft palate

Autosomal-dominant developmental malformations, deafness and dystonia — *ACTB*
Familial gastric cancer and CLP — *CDH1*
Craniofrontonasal — *EFNB1*
Roberts — *ESCO2*
Holoprosencephaly — *GLI2*
'Oro-facial-digital' — *GLI3*
Hydrolethalus — *HYLS1*
Van der Woude/popliteal pterygium — *IRF6*
X-linked mental retardation and CL/P — *PHF8*
Gorlin — *PTCH1*
CLP, ectodermal dysplasia — *PVRL1*
Holoprosencephaly — *SHH*
Holoprosencephaly — *SIX3*
Branchio-oculo-facial — *TFAP2A*
Holoprosencephaly — *TGIF1*
Ectrodactyly-ectodermal dysplasia-clefting — *TP63*
Ankyloblepharon-ectodermal dysplasia-clefting — *TP63*
Tetra-amelia with CLP — *WNT3*

Cleft palate only

Oculofaciocardiodental — *BCOR*
CHARGE — *CHD7*
Lethal and Escobar multiple pterygium — *CHRNG*
Stickler type 1 — *COL2A1*
Stickler type 2 — *COL11A1*
Stickler type 3 — *COL11A2*
Desmosterolosis — *DHCR24*
Smith–Lemli–Opitz — *DHCR7*
Miller — *DHODH*
Craniofrontonasal — *EFNB1*
Kallmann — *FGFR1*
Crouzon — *FGFR2*
Apert — *FGFR2*
Otopalatodigital types 1 and 2 — *FLNA*
Larsen syndrome; atelosteogenesis — *FLNB*
Hereditary lymphedema-distichiasis — *FOXC2*
Bamforth–Lazarus — *FOXE1*
'Oro-facial-digital' — *GLI3*
Van der Woude/popliteal pterygium — *IRF6*
Andersen — *KCNJ2*
Kabuki — *MLL2*
Cornelia de Lange — *NIPBL*
X-linked mental retardation — *PQBP1*
Isolated cleft palate — *SATB2*
Diastrophic dysplasia — *SLC26A2*
Campomelic dysplasia — *SOX9*
Pierre Robin — *SOX9*
DiGeorge — *TBX1*
X-linked cleft palate and ankyloglossia — *TBX22*
Treacher Collins — *TCOF1*
Loeys–Dietz — *TGFBR1*
Loeys–Dietz — *TGFBR2*
Saethre-Chotzen — *TWIST1*

Midline cleft lip

Opitz G/BBB — *MID1*
Oro-facial-digital type I — *OFD1*

An expanded version of these data with a full reference list is provided as Supplementary information S1 (table). *ACTB*, actin, β; *BCOR*, BCL6 corepressor; *CDH1*, cadherin 1; CHARGE, coloboma, heart defect, atresia choanae, retarded growth and development, genital abnormality, and ear abnormality; *CHD7*, chromodomain helicase DNA binding protein 7; *CHRNG*, cholinergic receptor, nicotinic, γ; CLP, clefts of the lip and/or palate; *COL*, collagen; *DHCR*, dehydrocholesterol reductase; *DHODH*, dihydroorotate dehydrogenase; *EFNB1*, ephrin-B1; *ESCO2*, establishment of cohesion 1 homologue 2; *FGFR*, fibroblast growth factor receptor; *FLN*, filamin; *FOX*, forkhead box; *HYLS1*, hydrolethalus syndrome 1; *IRF6*, interferon regulatory factor 6; *KCNJ2*, potassium inwardly-rectifying channel, subfamily J, member 2; *MID1*, midline 1 (Opitz/BBB syndrome); *MLL2*, myeloid/lymphoid or mixed-lineage leukaemia 2; *NIPBL*, Nipped-B homologue; *OFD1*, oral-facial-digital syndrome 1; *PHF8*, PHD finger protein 8; *PQBP1*, polyglutamine binding protein 1; *PTCH1*, patched 1; *PVRL1*, poliovirus receptor-related 1 (herpesvirus entry mediator C); *SATB2*, SATB homeobox 2; *SHH*, sonic hedgehog; *SLC26A2*, solute carrier family 26 (sulfate transporter), member 2; *SIX3*, SIX homeobox 3; *SOX9*, SRY (sex determining region Y)-box 9; *TBX*, T-box; *TCOF1*, Treacher Collins-Franceschetti syndrome 1; *TFAP2A*, transcription factor AP2α (activating enhancer binding protein 2α); *TGFBR*, transforming growth factor, β receptor; *TGIF1*, TGFB-induced factor homeobox 1; *TP63*, tumour protein p63; *WNT3*, wingless-type MMTV integration site family, member 3.

sporadic[25]. Moreover, CLP is known to be influenced by environmental risk factors[26,27]; consequently, a multifactorial model of inheritance is favoured in which genetic risk factors of small, individual impact may interact with environmental covariates[12]. These combined factors complicate genetic analysis of non-syndromic forms of CLP.

Accurate phenotyping is crucial to understanding both the epidemiology and aetiology of any congenital malformation because the power to detect effects is weakened when heterogeneous groups are treated as a single entity. Although clefts of the lip and palate show a range of phenotypic expression (FIG. 2), they are generally defined as qualitative traits (that is, affected or unaffected). Dividing CLP in this simplistic way could potentially result in important information being lost. For example, different patterns of genome-wide linkage are observed when multiplex families are divided into subgroups depending on the overt CLP phenotypes present in affected individuals. This observation suggests that careful attention to phenotypes will be an important tool for furthering our understanding of the genetic heterogeneity underlying non-syndromic CLP[2]. Furthermore, numerous lines of evidence now suggest that the phenotypic spectrum of non-syndromic CLP is more complex than previously realized and should include a variety of subclinical phenotypic features observed in either an individual with CLP and/or their 'unaffected' relatives[28].

Multiplex family
A family in which multiple members are affected by an inherited disease.

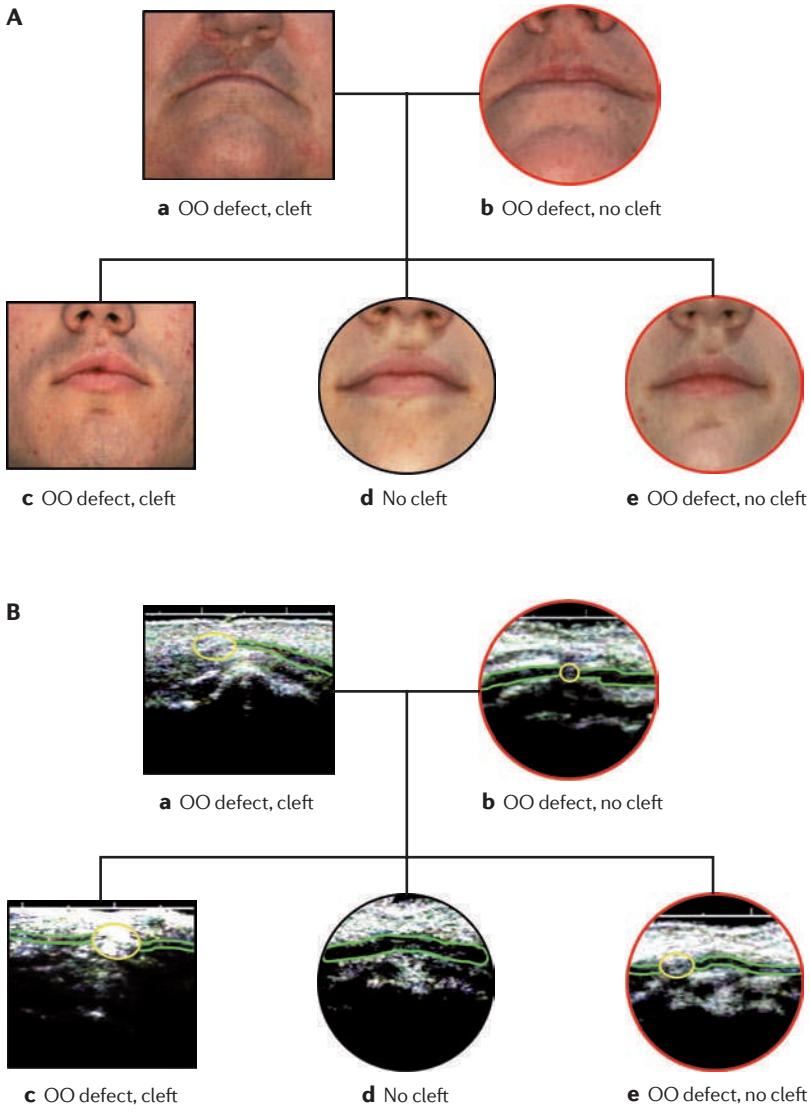

Figure 3 | **Subclinical phenotypes. A** | Photographs of the upper lip region for each member of a nuclear family with two family members affected with nonsyndromic cleft lip and/or palate (CLP) (surgically repaired). The other three family members do not have externally visible defects, but two of them have subclinical defects of the orbicularis oris (OO) muscle (pedigree symbols circled in red). **B** | The upper lip ultrasounds of each member of the family shown in panel **A**. Note the disruptions in the orbicularis oris muscle in the two people with CLP in the family, plus in two people with no external manifestation (pedigree symbols circled in red). Images courtesy of M.L.M.

Subclinical phenotypes can include minor structural variants, including lip pits/prints[29], dental anomalies[30], defects of the orbicularis oris muscle[31,32], three-dimensional facial image measurement[28], brain variants as assessed by MRI[33,34] or by surrogate measures[35,36], and speech or cognitive differences such as velopharyngeal insufficiency, reading disability and IQ. Palatal subphenotypes have been less explored but also include bifid uvula, submucous cleft palate, the differentiation of clefts of the hard and soft palate and possibly ankyloglossia. In the future, our understanding of palatal subdivisions by phenotype and pathway will benefit from both human and mouse models[37,38]. Defects of the orbicularis oris muscle show particular promise for enhancing the search for causative genetic variants and for contributing to clinical risk assessment[31,39–42]. Orbicularis oris defects can be assessed using high-resolution ultrasound of the upper lip (FIG. 3). Subclinical phenotyping therefore holds great promise to enhance the power of family studies and may lead to opportunities for translational research that is relevant for both clinical care of patients and clinical genetics as a science.

Lip pit
An indentation and/or sinus tract in the lower lip that is usually located to one or both sides of the midline. Lip prints are visual impressions taken of the upper and lower lip that can indicate the presence of pits.

Orbicularis oris
The muscle surrounding the mouth.

Gene discovery in non-syndromic CLP

To date, genetic approaches to non-syndromic CLP have included: linkage analysis using large, multiplex families or smaller but inbred families, or analysis of affected relative pairs; association studies using case–parent trios or case–control samples; identification of chromosomal anomalies or microdeletions in cases; and direct sequencing of DNA samples from affected individuals. These methods can be applied to candidate genes or genome-wide strategies can be used. Each approach has its own advantages and disadvantages, some of which will depend on the underlying genetic architecture of the disease, as well as the realities of economics and technology. We briefly summarize successes using a range of approaches, followed by further details on the results of recent genome-wide association (GWA) studies. Most studies of non-syndromic clefts to date have focused on CL/P rather than isolated cleft palate. This has been biased perhaps by the larger numbers of cases, easier ascertainment and less confusion from confounding syndromes. Future studies will need to address this gap and also the somewhat counter-intuitive observation that more mouse models are available for cleft palate than for cleft lip.

Candidate genes, chromosomal anomalies, linkage and sequencing. Candidate gene studies have been at the core of cleft research since Ardinger and colleagues[43] suggested a role for transforming growth factor-α (*TGFA*) variants in risk for non-syndromic CL/P. The identification of candidate genes has traditionally relied on gene expression and developmental analyses performed in model organisms, particularly the mouse, either to first identify the candidate genes or to provide biological plausibility for the association. More recently, extrapolation from the study of syndromic forms of CL/P has proven to be a useful adjunct to this approach. As with candidate gene studies of many complex disorders, rigorous confirmatory replication is not common, with only variants in interferon regulatory factor 6 *(IRF6)* yielding consistent evidence of association across multiple studies[12,44–47] (discussed further below). Analysis of chromosomal anomalies in patients has proven to be a productive route for the identification or confirmation of CL/P loci, with recent successes for fibroblast growth factor receptor 2 (*FGFR2*)[48] and *SUMO1* (a member of the small ubiquitin-like modifier family)[49–51]. Candidate gene-based association studies and analyses of chromosomal anomalies have recently been reviewed in detail[27,52].

There have been many attempts to use linkage analysis to identify regions of the genome that are likely to carry genes controlling pathogenesis of CLP, and the region surrounding the *FOXE1* (forkhead box E1) gene reached genome-wide levels of significance with subsequent fine-mapping and replication[2,53]. There have been several resequencing studies of candidate genes to identify specific variants that might underlie statistical associations with clefting, and the strongest current evidence is for mutations in *MSX1* (REFS 18,54), *FGFR1* and *FGF8* (REF. 55), and bone morphogenetic protein 4 (*BMP4*)[39]. Whole-exome sequencing has recently been successful in identifying causative genetic variants for Mendelian traits[56,57], including Miller syndrome[58] (an autosomal-recessive syndrome that can include cleft palate) and Kabuki syndrome[59] (a dominant disorder than can include cleft palate), but is yet to be successful for complex and heterogeneous traits such as non-syndromic CLP.

Genome-wide association studies. As is now apparent for many common complex disorders, GWA studies have provided recent major advances in our understanding of genes and pathways that have a role in the aetiology of CLP. To date, there are three published GWA studies for CL/P using the case–control design[3–5] and one case–parent trio study from an international consortium that is part of GENEVA (the gene environment association studies consortium)[6,60]. These studies have mostly excluded cases with cleft palate only, based on likely aetiologic heterogeneity. Birnbaum and colleagues[3] confirmed the impact of *IRF6*, which had previously been identified in candidate gene studies[12,44], and discovered a new region on chromosome 8q24 that gave extremely strong evidence of association in their European case–control sample. Grant and colleagues independently confirmed that this 'gene desert' region on chromosome 8q24 was strongly associated with CL/P in a sample of European-American cases and controls[4]. Mangold and colleagues subsequently used an expanded data set from Europe and identified additional loci at chromosomes 10q25 (with peak values closest to ventral anterior homeobox 1 (*VAX1*)) and 17q22 (with peak values closest to noggin (*NOG*)) that achieved genome-wide significance[5].

The GENEVA Cleft Consortium study used case–parent trios from multiple populations and reconfirmed the *IRF6* findings, as well as replicating the chromosome 8q24 and 10q25 (*VAX1*) associations[6]. Interestingly, in this consortium study, the level of statistical evidence from markers within chromosome 8q24 was much stronger among case–parent trios of European ancestry than among those of Asian ancestry, whereas the evidence for linkage and association for markers in *IRF6* was much stronger in trios of Asian ancestry. This GENEVA study identified at least two new loci (near *MAFB* and *ABCA4*) that were not previously associated with CL/P that reached genome-wide significance with stronger signals in Asian compared to European populations[6]. The signals and this population difference were replicated using independent families from multiple populations (see further details below).

These observations suggest not only that there are multiple genetic variants influencing risk of CL/P but also that some of these genes may be differentially tagged by polymorphic markers in a population-specific manner. For example, in the chromosome 8q24 region, the most significant SNP (rs987525) showed similar patterns of over-transmission to the affected child but had a higher minor allele frequency among parents of European ancestry compared to parents of Asian ancestry (0.26 versus 0.07)[6]. In fact, the entire region of signal on chromosome 8q24 showed higher rates of heterozygosity among parents of European ancestry compared to those of Asian ancestry, which means that European trios would be far more informative than Asian trios for this region. Therefore, it may be more difficult to identify causal genetic variants in some populations compared to others. Some putative causal genes or loci have been identified through polymorphic markers in most populations (for example, *IRF6*), whereas others (for example, 8q24, *MAFB* and *ABCA4*) seem to be more population-specific, which could reflect variable coverage by available marker panels or true allelic heterogeneity. True allelic heterogeneity, in which multiple mutations occurred on different background haplotypes, would make it much more difficult to identify causal genes through association studies. However, Dickson and colleagues[61] noted that there may be mixtures of multiple rare alleles on common haplotypes within a single causal gene for complex and heterogeneous disorders such as CLP.

Below, we provide a short summary of each of the genes confirmed or identified through GWA studies together with insights into the molecular pathogenesis derived from analysis of animal models. In TABLE 1, we summarize genes with a confirmed role in non-syndromic CLP, those that seem likely to be involved and those that have been intensively studied but have less-convincing supporting data.

Insights into molecular pathogenesis

Although GWA studies will increase the number of CLP loci identified, the move from a GWA study signal to a causative variant will still be challenging. Animal models and gene expression data are powerful tools for identifying candidate genes for complex traits; importantly, they also contribute to our knowledge of normal facial development and the molecular pathogenesis of CLP. The mouse is the pre-eminent model organism for studies of this type, as facial development mirrors human craniofacial development, and mouse strains with high rates of CLP are available. A number of excellent reviews have described the cellular and molecular mechanisms underlying normal and abnormal development[62,63]; here we provide examples of how the mouse has influenced our understanding of the molecular pathogenesis of CLP in humans.

IRF6. Mutations in *IRF6* were first identified as aetiologic in the autosomal-dominant Van der Woude syndrome, which can include CL/P and/or cleft palate only along with dental anomalies and lip fistulas[19]. Subsequent research showed that common alleles in *IRF6* were

Velopharyngeal insufficiency
Incomplete closing of the velopharyngeal sphincter (soft palate muscle) during speech.

associated with non-syndromic CL/P[44]. This association has been independently replicated in GWA studies as well as in many candidate gene studies[3–6,13,44–47,64]; some failures of replication were possibly due to population differences[65]. Recently, an approach that integrated the identification of *cis*-regulatory elements using sequence conservation across multiple species, analysis of animal models and biochemical analyses resulted in the identification of one specific sequence variant (rs642961, located within an enhancer ~10 kb upstream of the *IRF6* transcription start site) that is significantly over-transmitted in non-syndromic cleft lip only[12]. Importantly, this apparent risk allele was found to disrupt a binding site for the transcription factor AP2α, which is mutated in the autosomal-dominant CLP disorder branchio-oculo-facial syndrome[66], therefore strongly suggesting that this SNP is a contributory variant[12].

A role of *IRF6* in CLP is further supported by analysis of animal models. Recent research has shown that *Irf6* mutant mice exhibit a hyper-proliferative epidermis that fails to undergo terminal differentiation, which leads to multiple epithelial adhesions that can occlude the oral cavity and result in cleft palate[67,68]. These results demonstrated that IRF6 is a key determinant of the keratinocyte proliferation–differentiation switch, and subsequent research indicated that IRF6 also has a key role in the formation of oral periderm, spatiotemporal regulation of which is essential for ensuring appropriate palatal adhesion[69]. Recently, a combination of mouse genetics, gene expression analyses, chromatin immunoprecipitation studies and luciferase reporter assays has shown that *IRF6* is a direct target of p63, which underlies several malformation syndromes that include CLP as a hallmark feature[16,17]. p63 activates *IRF6* transcription through the *IRF6* enhancer element, variation within which increases susceptibility to cleft lip only[70].

MAFB. The *MAFB* gene encodes a basic leucine zipper transcription factor. Markers near *MAFB* achieved genome-wide significance in the GENEVA Cleft Consortium study[6], with trios of Asian descent providing much stronger statistical evidence than trios of European descent. In independent replication samples, 1,149 pedigrees of European ancestry showed evidence of linkage and association with a SNP (rs13041247; $p = 0.0007$) located 260 bp from the SNP yielding the strongest signal among Asian families (rs11696257; $p = 0.0009$ in 331 independent pedigrees). A missense mutation, H131Q, in *MAFB* was found in 3.5% of Filipinos with CL/P but only 0.7% of controls ($p < 0.0001$). This variant occurs in a region of strongly conserved sequence, suggesting that there may be a rare variant in *MAFB* that contributes to the observed signal in the GWA study. It is noteworthy that the gene-poor regions either side of *MAFB* include numerous binding sites for transcription factors that are known to have a role in palate development (including transcription factors in the MSX, IRF, SRY-box containing (SOX) and BTB and CNC homology (BACH) gene families). In the mouse, *Mafb* is highly expressed in the epithelium of the palatal shelves and in the medial edge epithelium during palatal fusion[6].

Oral periderm
A superficial layer of flattened cells which develops from the single-cell-layered ectoderm to form a transient covering for the oral epithelia.

Table 1 | **Genes with a role in non-syndromic CLP**

Class/gene	Evidence	Refs
Confirmed*		
IRF6	GWA, LD, L, M	3,12,44
VAX1	GWA, LD	5,6
8q24 locus	GWA, LD	3,4,6
Likely‡		
ABCA4 (locus only)	GWA	6
BMP4	M	39,115
FGFR2	M	48,55,116
FOXE1	L, LD, M	53,117,118
MAFB	GWA	6
MSX1	LD, M	18,54,119–121
MYH9	LD	3,122–124
17q22 locus	GWA	5,6
Intensively studied§		
CRISPLD2	LD	125,126
FGF8	M	55,116
GSTT1	LD	83
MTHFR	LD	127,128
PDGFC	LD, M	25,129,130
PVRL1	M, LD	131–133
SUMO1	M	49,50,51,134
TGFA	LD	43,120,134
TGFB3	LD, M	119,120,135,136

*At least two independent studies reaching conservative levels of significance. ‡At least one study with conservation/compelling data and other supportive studies. §Multiple studies, no consensus or convincing meta-analysis. *BMP4*, bone morphogenetic protein 4; CLP, clefts of the lip and/or palate; *CRISPLD2*, cysteine-rich secretory protein LCCL domain containing 2; *FGFR*, fibroblast growth factor receptor; *FOXE1*, forkhead box E1 (thyroid transcription factor 2); *GSTT1*, glutathione S-transferase-θ1; GWA, genome-wide association; *IRF6*, interferon regulatory factor 6; L, linkage; LD, candidate gene association; M, mutation detection; *MAFB*, v-maf musculoaponeurotic fibrosarcoma oncogene homologue B; *MSX1*, msh homeobox 1; *MTHFR*, methylenetetrahydrofolate reductase (NAD(P)H); *MYH9*, myosin, heavy chain 9, non-muscle; *PDGFC*, platelet-derived growth factor C; *PVRL1*, poliovirus receptor-related 1 (herpesvirus entry mediator C); *TGFA*, transforming growth factor-α; *TGFB3*, transforming growth factor-β3; *VAX1*, ventral anterior homeobox 1.

ABCA4. *ABCA4* encodes an ATP-binding cassette transporter. Multiple markers in *ABCA4* (within and 5′ to the transcribed region) gave evidence of linkage and association at the genome-wide significance level in the GENEVA Cleft Consortium GWA study[6], again with stronger evidence among Asian samples. Two of the SNPs with the strongest signals were replicated in independent family samples, and one of these SNPs (rs560426) gave a far stronger signal in Asian families ($p = 0.0003$ in 331 pedigrees) compared to European families ($p = 0.005$ in 1149 pedigrees). This difference in the strength of statistical evidence again raises the possibility of either an allele common to both groups but with differing frequencies, or multiple risk alleles

occurring on different haplotype backgrounds. *ABCA4* is known to cause the autosomal-recessive retinal degenerative disease Stargardt's disease, and sequencing of the 50 exons of *ABCA4* in 190 CL/P cases identified 27 different missense mutations, many of which have been previously reported in Stargardt's or other ocular disorders (see the OMIM website for details). As *ABCA4* is surrounded by many other genes, the peak signal in *ABCA4* may be a surrogate for aetiologic variants in another gene nearby. Furthermore, no *Abca4* expression has been seen in mouse palatal shelves around the time of palatal fusion[6].

VAX1. In the studies by Mangold *et al.*[5] and the GENEVA Cleft Consortium[6], markers in or near the *VAX1* gene at chromosome 10q25 yielded evidence approaching genome-wide significance; the same two alleles of SNPs in *VAX1* (rs7078160 and rs4752028) were overrepresented in CL/P cases in both studies. *VAX1* encodes a transcriptional regulator with a DNA-binding homeobox domain. Mouse knockouts for *Vax1* develop cleft palate, and this gene is expressed widely in developing craniofacial structures[71]; thus, variants in *VAX1* itself are strong candidates for contributing to CLP.

WNT signalling. Although not yet implicated by GWA studies, variants within WNT genes have been reported to be associated with non-syndromic CL/P[72], and mutations in *WNT3* underlie autosomal-recessive tetra-amelia with cleft lip and palate[73]. Although the evidence for the involvement of WNT signalling in non-syndromic CL/P is not strong, these findings have led to further analyses of genes in the WNT signalling pathway as candidates for normal development of the lip and palate. Targeted mutation of *Wnt9b* in mice leads to CLP, and the A/WySn strain of mice, which have increased incidence of spontaneous CLP, have a retrotransposon inserted 6.6 kb downstream of the *Wnt9b* gene (a site known as the *clf1* locus)[74]. These findings suggest that WNT9B has a key role in the development of the lip[74–76]. Further support for this hypothesis arises from the observation that canonical WNT signalling is activated during midfacial morphogenesis in mice[77]. Additionally, genetic inactivation of low density lipoprotein receptor-related protein 6 (*Lrp6*), a co-receptor of the WNT–β-catenin signalling pathway, causes CLP[78]. Intriguingly, *Msx1* and *Msx2* (see below) are downstream targets of this WNT–β-catenin signalling pathway during lip formation and fusion[78].

MSX1 and BMP signalling. As in humans, loss-of-function mutations in the homeobox gene *Msx1* result in cleft palate in mice[79]. *Msx1* is a downstream target of BMP signalling in a number of embryonic tissues and *Msx1* is necessary for expression of *Bmp4* and/or *Bmp2* (REF. 80). In mice, loss-of-function of type I BMP receptor (*Bmpr1a*) in the craniofacial primordia resulted in CL/P, whereas deficiency of *Bmp4* resulted in cleft lip only[81]; this shows that BMP signalling has distinct functions in development of the lip versus the secondary palate. In the context of *Bmp4* deficiency, all *Bmp4* mutant embryos exhibited bilateral cleft lip at embryonic day 12 (E12), but only 22% still displayed cleft lip at E14, which suggests that there is some kind of *in utero* repair mechanism[81].These observations parallel the findings that mutations in *BMP4* may underlie a subset of cases of subepithelial, microform and overt cleft lip in humans[39].

Environment and gene–environment interaction

The identification of environmental components of clefting and studies of gene by environment interaction require large (ideally prospective) cohort studies and access to genetic material to be optimally effective. Although a few such resources are available (in Denmark, Norway, and the United States)[11,14,15], they are still primarily in the analysis phase. Nonetheless, there are a few studies that have begun to provide data on environmental risks. Because the environment is more malleable, the identification of environmental risks, particularly if they can be personalized with genetic covariates, provides the best short-term opportunities to be applied to prevention.

Maternal smoking has been associated repeatedly with increased risk of CLP, and meta-analysis strongly supports an overall odds ratio (OR) for having CLP of ~1.3 among offspring of mothers who smoke[82–84]. The increased risk resulting from exposure to maternal smoking during the peri-conceptual period raises the possibility that genes in certain metabolic pathways may have a role in the development of CLP. Specifically, markers in the glutathione S-transferase-θ1 (*GSTT1*) or nitric oxide synthase 3 (*NOS3*) genes appear to influence risk of CL/P in the presence of maternal smoking[83,85–87]. The *GSTT1* markers are gene deletion variants, which suggests that deficiencies in detoxification pathways may underlie some of this susceptibility. Smoking has also been recently associated with a joint risk with variants in *IRF6,* and the same study reported interactions between multivitamins and *IRF6* variants[88]. These findings provide evidence that gene–environment interactions are important in CLP. In addition, some specific teratogens[26,27,89] — for example, valproic acid — have yielded evidence of association with cleft palate[90].

Exposure to maternal alcohol consumption has also been suggested as a risk factor, but the evidence has been more inconsistent[27]. Studies also suggest that 'binge' drinking patterns (high doses of alcohol in short periods of time) increase risk[91], and this is supported by associations with variation in the *ADH1C* alcohol dehydrogenase gene[92]. However, these links to alcohol consumption remain to be confirmed. Nutritional factors, such as folate deficiency, have also been suggested to influence risk of CL/P, based on both observational studies and interventional trials using folate supplementation to prevent recurrences of CL/P in families[93]. However, the studies of vitamin supplementation with folate remain controversial[1,94] and recent studies of levels of folate receptor antibodies did not find an association with CL/P[95]. Furthermore, food fortification programmes using folic acid have shown

Odds ratio
A measurement of association that is commonly used in case–control studies. It is defined as the odds of exposure to the susceptible genetic variant in individuals with disease compared with that in controls. If the odds ratio is significantly greater than one, the genetic variant is associated with the disease.

detectable decreases in the rates of clefting in some[96,97] but not all[98,99] studies. In the future, other nutrient and micronutrient studies will need to be expanded to look for evidence of effects. For example, there are some data to support roles for zinc deficiency in risk of oral clefts in populations in which zinc status is highly compromised[100], for cholesterol deficiency in facial clefting[101], and for multivitamins in general in cleft prevention[97].

Besides nutrients and toxins, other environmental exposures have been, and should continue to be, assessed for possible roles in clefting. These exposures include hyperthermia[102], stress, maternal obesity, occupational exposures, ionizing radiation and infection[10]. Pregnancy planning has been shown to have a protective effect, and the basis of this observation needs to be more deeply explored[103]. Nonetheless, there is no consensus yet on the harmful effects of these factors, and prospective cohort studies large enough to measure effects on a relatively rare disorder such as clefting may be required. A particular challenge will be to determine the specificity of the role of an exposure in contributing to clefting, as many exposures will have both identifiable and unidentifiable coincident risks. Analytic approaches such as Mendelian randomization will be helpful in making these determinations[104]. A new, developing database (FaceBase) is providing a common source for human and animal model data on genes and gene expression relevant to facial clefting.

Integrating evidence into clinical care

Despite the recent identification of genes that are likely to influence the risk of non-syndromic CLP, these results have yet to have any direct impact on genetic counselling or clinical management. Improved epidemiologic information does, however, allow for better point estimates for familial recurrence risks[14]. Furthermore, it seems likely that genotypic information for apparent risk alleles associated with higher risk of oral clefts could be useful in clinical assessment (once we have a better definition of the full number of causal genes and their potential interactions with one another and with environmental risk factors). The next critical phase of statistical analyses will be to examine the heterogeneity underlying the aetiology of oral clefts and to investigate the gene–gene and gene–environment interactions that control risk. A range of study designs will be needed to achieve this level of documentation, including family studies, case–control studies and eventually prospective cohort data. Importantly, incorporating information from subclinical phenotypes, such as orbicularis oris defects or dental anomalies, may also allow us to identify aetiologically homogeneous subgroups of cleft cases, and thus should enhance family studies and estimates of recurrence risk[42]. New array-based copy-number-variant analysis and whole-exome or even whole-genome resequencing could also provide future opportunities for improved molecular diagnostics, and the continually improving ultrasound analysis of the fetus may allow earlier identification of the presence and severity of cleft type before birth.

Gene expression in time and space. Global approaches to expression analyses of genes in craniofacial structures have already provided a broad view of gene expression. For example, the Craniofacial and Oral Gene Expression Network (COGENE) project provides public web access to human gene expression data for 24 craniofacial-specific human tissues isolated from day 26 to day 60 human embryos. In zebrafish, mRNA sequencing and microRNA analysis have been informative for understanding palate development, so it would be useful to build on this knowledge[105]. Similarly, the ability to analyse tissues in their correct three-dimensional orientation is central to understanding biological processes, particularly when tissues undergo a complex and intricate series of movements relative to each other, as occurs in the developing craniofacial region. The mapping of gene and protein expression patterns within these complex shapes can provide important clues about their biological functions and also indicates which genes and/or proteins may interact with one another. The expression of genes relative to each other in both time and space can be visually represented using optical projection tomography (OPT)[106], and an atlas of craniofacial gene expression patterns is available online in the EMAGE database.

Cis-*regulatory element identification.* Much of the genetic variation underlying complex disorders (such as non-syndromic CLP) is likely to occur in regulatory elements outside coding sequences of genes. These elements are challenging to identify as they often regulate genes across substantial genomic distances. Although evolutionary sequence conservation can facilitate the discovery of regulatory elements, this technique does not predict their spatiotemporal pattern of activity *in vivo*[107]. Recently, chromatin immunoprecipitation followed by next-generation sequencing analysis (ChIP–seq) for the enhancer-associated protein p300 has been demonstrated to be a highly sensitive method to accurately identify enhancer elements and their associated activities[108]. Clearly, detailed mapping of regulatory elements will provide additional (and functionally relevant) targets for sequence analysis, particularly where they fall within regions of the genome implicated by GWA studies or other approaches. The power of integrating association studies in well-characterized patient populations with identification of *cis*-regulatory elements, analysis of animal models and biochemical analyses is amply illustrated by the example of *IRF6* noted above.

Wider implications. Biological roles outside the craniofacial complex are known for some of the candidate genes associated with CLP, increasing the importance of CLP gene-finding endeavours. One recent publication on a small data set suggests a role for *IRF6* in wound healing, at least in the autosomal-dominant Van der Woude syndrome[109]. Long-term outcomes of individuals born with clefts may include risks for higher overall mortality rates, mental health problems[110], a higher risk of cancer (particularly breast cancer) in affected individuals[111] and

Mendelian randomization
The random assignment of alleles from parents to offspring that occurs during gamete formation. It is the underlying concept of a method to genetically stratify individuals in a large population sample and then to evaluate phenotypic differences based on a pre-specified genotype.

their family members[112] and alterations in child bearing patterns[113]. Identifying long-term adverse outcomes (for example, cancer and psychiatric disorders) that are seemingly unrelated to a common birth defect may eventually result in decreasing an individual's lifelong health burden by recognizing risks at their early, pre-symptomatic stages. Studies into the aetiology of clefts may well enhance our understanding of other common, complex traits and allow us to move beyond the attitude that CLP is only a structural birth defect, but instead is a lifelong disorder for which therapies and prevention can promise a fuller and healthier lifespan.

Future approaches. Future advances in our understanding of the molecular pathogenesis of CLP will require strategies that increasingly integrate genetic analysis of precisely phenotyped cohorts of patients, global approaches for the identification and ranking of candidate genes, and improved methods for delineating and analysing functional elements controlling gene expression. Integration of genetic and environmental risk using epigenetics, systems biology, gene expression and epidemiology will all be required to generate a synthesis that will more completely characterize aetiologies, as well as provide access to better clinical care and prevention.

1. Wehby, G. & Cassell, C. H. The impact of orofacial clefts on quality of life and healthcare use and costs. *Oral Dis.* **16**, 3–10 (2010).
2. Marazita, M. L. *et al.* Genome scan, fine-mapping, and candidate gene analysis of non-syndromic cleft lip with or without cleft palate reveals phenotype specific differences in linkage and association results. *Hum. Hered.* **68**, 151–170 (2009).
3. Birnbaum, S. *et al.* Key susceptibility locus for nonsyndromic cleft lip with or without cleft palate on chromosome 8q24. *Nature Genet.* **41**, 473–477 (2009).
This was the first successful GWA study in clefting and identified a significant and previously unsuspected locus for clefts at 8q24 in a large gene desert.
4. Grant, S. F. *et al.* A genome-wide association study identifies a locus for nonsyndromic cleft lip with or without cleft palate on 8q24. *J. Pediatr.* **155**, 909–913 (2009).
5. Mangold, E. *et al.* Genome-wide association study identifies two susceptibility loci for nonsyndromic cleft lip with or without cleft palate. *Nature Genet.* **42**, 24–26 (2010).
6. Beaty, T. H. *et al.* A genome-wide association study of cleft lip with and without cleft palate identifies risk variants near MAFB and ABCA4. *Nature Genet.* **42**, 525–529 (2010).
This GWA study used the case–parent trio approach and identified at least two new loci (near the *ABCA4* and *MAFB* genes) for clefts. It further demonstrated that population heterogeneity is an important consideration in GWA studies.
7. Fogh-Andersen, P. *Inheritance of Harelip and Cleft Palate* (Munksgaard, Copenhagen,1942).
8. Fraser, F. C. Thoughts on the etiology of clefts of the palate and lip. *Acta Genet. Stat. Med.* **5**, 358–369 (1955).
9. Christensen, K. & Mitchell, L. E. Familial recurrence-pattern analysis of nonsyndromic isolated cleft palate — a Danish Registry study. *Am. J. Hum. Genet.* **58**, 182–190 (1996).
10. Mossey, P., Little, J., Munger, R. G., Dixon, M. J. & Shaw, W. C. Cleft lip and palate. *Lancet.* **374**, 1773–1785 (2009).
11. Harville, E. W., Wilcox, A. J., Lie, R. T., Vindenes, H. & Abyholm, F. Cleft lip and palate versus cleft lip only: are they distinct defects? *Am. J. Epidemiol.* **162**, 448–453, (2005).
12. Rahimov, F. *et al.* Disruption of an AP-2α binding site in an IRF6 enhancer is associated with cleft lip. *Nature Genet.* **40**, 1341–1347 (2008).
A still rare demonstration of moving from an associated SNP to finding one of the likely aetiologic SNPs for clefts. It also brought a new gene (*TFAP2A*) and pathway into cleft studies.
13. Weinberg, S. *et al.* Rethinking isolated cleft palate: evidence of occult lip defects in a subset of cases. *Am. J. Med. Genet. A* **146A**, 1670–1675 (2008).
14. Grosen, D. *et al.* A cohort study of recurrence patterns among more than 54,000 relatives of oral cleft cases in Denmark: support for the multifactorial threshold model of inheritance. *J. Med. Genet.* **47**, 162–168 (2010).
The most extensive study to date of recurrence risks for clefts in first, second and third degree relatives.
15. Genisca, A. E. *et al.* Orofacial clefts in the National Birth Defects Prevention Study, 1997–2004. *Am. J. Med. Genet. A* **149A** 1149–1158 (2009).
16. Celli, J. *et al.* Heterozygous germline mutations in the p53 homolog p63 are the cause of EEC syndrome. *Cell* **99**, 143–153 (1999).
17. McGrath, J. A. *et al.* Hay-Wells syndrome is caused by heterozygous missense mutations in the SAM domain of p63. *Hum. Mol. Genet.* **10**, 221–229 (2001).
18. van den Boogaard, M. J., Dorland, M., Beemer F. A. & van Amstel, H. K. *MSX1* mutation is associated with orofacial clefting and tooth agenesis in humans. *Nature Genet.* **24**, 342–343 (2000).
19. Kondo, S. *et al.* Mutations in *IRF6* cause Van der Woude and popliteal pterygium syndromes. *Nature Genet.* **32**, 285–289 (2002).
20. Dodé, C. *et al.* Loss-of-function mutations in *FGFR1* cause autosomal dominant Kallmann syndrome. *Nature Genet.* **33**, 463–465 (2003).
21. Jones, M. C. Etiology of facial clefts: prospective evaluation of 428 patients. *Cleft Palate J.* **25**, 16–20 (1988).
22. FitzPatrick, D. R., Raine, P. A. & Boorman, J. G. Facial clefts in the west of Scotland in the period 1980–1984: epidemiology and genetic diagnoses. *J. Med. Genet.* **31**,126–129 (1994).
23. Marazita, M. L. *et al.* Nonsyndromic cleft lip with or without cleft palate in China: assessment of candidate regions. *Cleft Palate Craniofac. J.* **39**, 149–156 (2002).
24. Mitchell, L. E. in *Cleft Lip and Palate: From Origin to Treatment* (ed.Wyszyski, D. F.) 234–239 (Oxford Univ. Press, 2002).
25. Jugessur, A. *et al.* Genetic determinants of facial clefting: analysis of 357 candidate genes using two national cleft studies from Scandinavia. *PLoS ONE* **4**, e5358 (2009).
26. Murray, J. C. Gene/environment causes of cleft lip and/or palate. *Clin. Genet.* **61**, 248–256 (2002).
27. Mossey P & Little, J. Addressing the challenges of cleft lip and palate research in India. *Indian J. Plast. Surg.* **42**, 9–18 (2009).
28. Weinberg, S. *et al.* Face shape of unaffected parents with cleft affected offspring: combining three-dimensional surface imaging and geometric morphometrics. *Orthod. Craniofac. Res.* **12**, 271–281 (2009).
29. Neiswanger, K. *et al.* Whorl patterns on the lower lip are associated with nonsyndromic cleft lip with or without cleft palate. *Am. J. Med. Genet. A* **149A**, 2673–2679 (2009).
30. Vieira, A. R., McHenry, T. G., Daack-Hirsch, S., Murray, J. C. & Marazita, M. L. Candidate gene/loci studies in cleft lip/palate and dental anomalies finds novel susceptibility genes for clefts. *Genet. Med.* **10**, 668–674 (2008).
31. Neiswanger, K. *et al.* Orbicularis oris muscle defects as an expanded phenotypic feature in nonsyndromic cleft lip with or without cleft palate. *Am. J. Med. Genet. A* **143A**, 1143–1149 (2007).
This study opened the door for subphenotyping as a crucial variable in cleft studies. It also provided an opportunity to use a clinical test in determining risks for recurrence of clefts in families.
32. Weinberg, S. M. *et al.* Three-dimensional morphometric analysis of craniofacial shape in the unaffected relatives of individuals with nonsyndromic orofacial clefts: a possible marker for genetic susceptibility. *Am. J. Med. Genet. A* **146A**, 409–420 (2008).
33. Nopoulos, P. *et al.* Structural brain abnormalities in adult males with clefts of the lip and/or palate. *Genet. Med.* **4**,1–9 (2002).
34. Conrad, A. L. *et al.* Cerebellum structure differences and relationship to speech in boys and girls with non-syndromic cleft of the lip and/or palate. *Cleft Palate-Cran. J.* **47**, 469–475 (2010).
35. Wentzlaff, K. *et al.* Association between non-righthandedness and cleft lip with or without cleft palate in a Chinese population. *J. Craniofac. Genet. Dev. Bio.* **17**, 141–147 (1997).
36. Scott, N. M., Weinberg, S. M., Neiswanger, K., Brandon, C. A. & Marazita, M. L. Hair whorls and handedness: informative phenotypic markers in nonsyndromic cleft lip with or without cleft palate (NS CL/P) cases and their unaffected relatives. *Am. J. Med. Genet. A* **136**,158–161 (2005).
37. Pauws, E. *et al. Tbx22* null mice have a submucous cleft palate due to reduced palatal bone formation and also display ankyloglossia and choanal atresia phenotypes. *Hum. Mol. Genet.* **18**, 4171–4179 (2009).
38. Baek, J. A. *et al.* Bmpr1a signaling plays critical roles in palatal shelf growth and palatal bone formation *Dev. Biol.* 23 Dec 2010 (doi:10.1016/j. ydbio.2010.12.028).
39. Suzuki, S. *et al.* Mutations in *BMP4* are associated with subepithelial, microform, and overt cleft lip. *Am. J. Hum. Genet.* **84**, 406–411 (2009).
40. Marazita, M. Subclinical features in non-syndromic cleft lip with or without cleft palate (CL/P): review of the evidence that subepithelial orbicularis oris muscle defects are part of an expanded phenotype for CL/P. *Orthod. Craniofac. Res.* **10**, 82–87 (2007).
41. Rogers, C.R. *et al.* Anatomical basis for apparent subepithelial cleft lip: a histological and ultrasonographic survey of the orbicularis oris muscle. *Cleft Palate-Cran. J.* **45**, 518–524 (2008).
42. Klotz, C. M. *et al.* Revisiting the recurrence risk of nonsyndromic cleft lip with or without cleft palate. *Am. J. Med. Genet. A* **152A**, 2697–2702 (2010).
43. Ardinger, H. H. *et al.* Association of genetic variation of the transforming growth factor-α gene with cleft lip and palate. *Am. J. Hum. Genet.* **45**, 348–353 (1989).
44. Zucchero, T. M. *et al.* Interferon regulatory factor 6 (*IRF6*) gene variants and the risk of isolated cleft lip or palate. *N. Engl. J. Med.* **351**, 769–780 (2004).
45. Ghassibé, M. *et al.* Interferon regulatory factor-6: a gene predisposing to isolated cleft lip with or without cleft palate in the Belgian population. *Eur. J. Hum. Genet.* **13**, 1239–1242 (2005).
46. Park, J. *et al.* Association between *IRF6* and nonsyndromic cleft lip with or without cleft palate in four populations. *Genet. Med.* **9**, 219–227 (2007).
47. Scapoli, L. *et al.* Strong evidence of linkage disequilibrium between polymorphisms at the *IRF6* locus and nonsyndromic cleft lip with or without cleft palate, in an Italian population. *Am. J. Hum. Genet.* **76**, 180–183 (2005).
48. Osoegawa, K. *et al.* Identification of novel candidate genes associated with cleft lip and palate using array comparative genomic hybridization. *J. Med. Genet.* **45**, 81–86 (2008).
49. Alkuraya, F. S. *et al.* SUMO1 haploinsufficiency leads to cleft lip and palate. *Science* **313**, 1751 (2006).
50. Shi, M. *et al.* Identification of microdeletions in candidate genes for cleft lip and/or palate. *Birth Defects Res. A Clin. Mol. Teratol.* **85**, 42–51 (2009).

51. Mostowska, A. *et al.* Association between genetic variants of reported candidate genes or regions and risk of cleft lip with or without cleft palate in the Polish population. *Birth Defects Res. A Clin. Mol. Teratol.* **88**, 538–545 (2010).
52. Jugessur, A., Farlie, P. G. & Kilpatrick, N. The genetics of isolated orofacial clefts: from genotypes to subphenotypes. *Oral Dis.* **15**, 437–453 (2009).
53. Moreno, L. *et al. FOXE1* association with both isolated cleft lip with or without cleft palate; and isolated cleft palate. *Hum. Mol. Gen.* **18**, 4879–4896 (2009).
This study moved from a linkage localization for clefts as a complex trait to finding the specific gene (*FOXE1*) that is likely to be involved. The path from linkage to gene identification has been relatively unsuccessful, but this study showed that large populations can be used to identify both rare and common variants contributing to a phenotype.
54. Jezewski, P. A. *et al.* Complete sequencing shows a role for MSX1 in non-syndromic cleft lip and palate. *J. Med. Genet.* **40**, 399–407 (2003).
55. Riley, B. M. & Murray, J. C. Sequence evaluation of *FGF* and *FGFR* gene conserved non-coding elements in non-syndromic cleft lip and palate cases. *Am. J. Med. Genet. A* **143A**, 3228–3234 (2007).
56. Lupski, J. R. *et al.* Whole-genome sequencing in a patient with Charcot-Marie-Tooth neuropathy. *N. Eng. J. Med.* **362**, 1181–1911 (2010).
57. Roach, J. C. *et al.* Analysis of genetic inheritance in a family quartet by whole-genome sequencing. *Science* **328**, 636–639 (2010).
58 Ng, S. B. *et al.* Exome sequencing identifies the cause of a Mendelian disorder. *Nature Genet.* **42**, 30–35 (2010).
The first application of whole-exome sequencing to a dominant disorder that includes craniofacial features. This success opens up a whole new approach to Mendelian craniofacial disorders and suggests that it may be useful in complex traits as well.
59. Ng, S. B. *et al.* Exome sequencing identifies *MLL2* mutations as a cause of Kabuki syndrome. *Nature Genet.* **42**, 790–793 (2010).
60. Cornelis, M. C. *et al.* The Gene, Environment Association Studies consortium (GENEVA): maximizing the knowledge obtained from GWAS by collaboration across studies of multiple conditions. *Genet. Epidemiol.* **34**, 364–372 (2010).
61. Dickson, S., Wang, K., Krantz, I., Hakonarson, H. & Goldstein, D. B. Rare variants create synthetic genome-wide associations. *PLoS Biol.* **8**, e1000294 (2010).
62. Jiang, R., Bush, J. O. & Lidral, A. C. Development of the upper lip: morphogenetic and molecular mechanisms. *Dev. Dyn.* **235**, 1152–1166 (2006).
63. Gritli-Linde, A. Molecular control of secondary palate development. *Dev. Biol.* **301**, 309–326 (2007).
64. Blanton, S. H. *et al.* Variation in *IRF6* contributes to nonsyndromic cleft lip and palate. *Am. J. Med. Genet. A* **137A**, 259–262 (2005).
65. Blanton, S. H., Garcia, E., Mulliken, J. B., Stal, S. & Hecht, J. T. Ethnic heterogeneity of *IRF6* AP-2a binding site promoter SNP association with nonsyndromic cleft lip and palate. *Cleft Palate-Cran. J.* **47**, 574–577 (2010).
66. Milunsky, J. M. *et al. TFAP2A* mutations result in branchio-oculo-facial syndrome. *Am. J. Hum. Genet.* **82**, 1171–1177 (2008).
67. Richardson, R. J. *et al. Irf6* is a key determinant of the keratinocyte proliferation-differentiation switch. *Nature Genet.* **38**, 1329–1334 (2006).
68. Ingraham, C. R. *et al.* Abnormal skin, limb and craniofacial morphogenesis in mice deficient for interferon regulatory factor 6 (*Irf6*). *Nature Genet.* **38**, 1335–1340 (2006).
References 67 and 68 established a crucial mouse model for isolated clefts. These papers also demonstrate the role of the first gene associated with clefting with certainty (*IRF6*) in keratinocyte differentiation.
69. Richardson, R., Dixon, J., Jiang, R. & Dixon, M. J. Integration of IRF6 and Jagged2 signalling is essential for controlling palatal adhesion and fusion competence. *Hum. Mol. Gen.* **18**, 2632–2642 (2009).
70. Thomason, H. A. *et al.* Cooperation between the transcription factors p63 and IRF6 is essential to prevent cleft palate in mice. *J. Clin. Invest.* **120**, 1561–1569 (2010).
71. Hallonet, M., Hollemann, T., Pieler, T. & Gruss, P. *Vax1*, a novel homeobox-containing gene, directs development of the basal forebrain and visual system. *Genes Dev.* **13**, 3106–3114 (1999).
72. Chiquet, B. *et al.* Variation in WNT genes is associated with non-syndromic cleft lip with or without cleft palate. *Hum. Mol. Gen.* **17**, 2212–2218 (2008).
73. Niemann, S. *et al.* Homozygous *WNT3* mutation causes tetra-amelia in a large consanguineous family. *Am. J. Hum. Genet.* **74**, 558–563 (2004).
74. Juriloff, D. M., Harris, M. J., McMahon, A. P., Carroll, T. J. & Lidral, A. C. *Wnt9b* is the mutated gene involved in multifactorial nonsyndromic cleft lip with or without cleft palate in A/WySn mice, as confirmed by a genetic complementation test. *Birth Defects Res. A Clin. Mol. Teratol.* **76**, 574–579 (2006).
75. Carroll, T., Park, J. S., Hayashi, S., Majumdar, A. & McMahon, A. P. Wnt9b plays a central role in the regulation of mesenchymal to epithelial transitions underlying organogenesis of the mammalian urogenital system. *Dev. Cell* **9**, 283–292 (2005).
76. Juriloff, D. M. *et al.* Investigations of the genomic region that contains the clf1 mutation, a causal gene in multifactorial cleft lip and palate in mice. *Birth Defects Res. A Clin. Mol. Teratol.* **73**, 103–113 (2005).
77. Lan, Y. *et al.* Expression of *Wnt9b* and activation of canonical Wnt signaling during midfacial morphogenesis in mice. *Dev. Dyn.* **235**, 1448–1454 (2006).
78. Song, L. *et al.* Lrp6-mediated canonical Wnt signaling is required for lip formation and fusion. *Development.* **136**, 3161–3171 (2009).
79. Satokata, I. & Maas, R. *Msx1* deficient mice exhibit cleft palate and abnormalities of craniofacial and tooth development. *Nature Genet.* **6**, 348–356 (1994).
80. Zhang, Z. *et al.* Rescue of cleft palate in *Msx1*-deficient mice by transgenic *Bmp4* reveals a network of BMP and Shh signaling in the regulation of mammalian palatogenesis. *Development* **129**, 4135–4146 (2002).
81. Liu, W. *et al.* Distinct functions for Bmp signaling in lip and palate fusion in mice. *Development* **132**, 1453–1461 (2005).
82. Little, J., Cardy, A. & Munger, R. G. Tobacco smoking and oral clefts: a meta-analysis. *Bull. World Health Organ.* **82**, 213–218 (2002).
An excellent meta-analysis providing overwhelming support for a role of maternal tobacco use in contributing to clefting.
83. Shi, M. *et al.* Orofacial cleft risk is increased with maternal smoking and specific detoxification-gene variants. *Am. J. Hum. Genet.* **80**, 76–90 (2007).
84. Shi, M., Wehby, G. L. & Murray, J. C. Review on genetic variants and maternal smoking in the etiology of oral clefts and other birth defects. *Birth Defects Res. C Embryo Today* **84**, 16–29 (2008).
85. van Rooij, I. A. *et al.* Smoking, genetic polymorphisms in biotransformation enzymes, and nonsyndromic oral clefting: a gene-environment interaction. *Epidemiology* **12**, 502–507 (2001).
86. Lammer, E. J., Shaw, G. M., Iovannisci, D. M., Van Waes, J. & Finnell, R. H. Maternal smoking and the risk of orofacial clefts: susceptibility with *NAT1* and *NAT2* polymorphisms. *Epidemiology* **15**, 150–156 (2004).
87. Zhu, H., Kartiko, S. & Finnell, R. H. Importance of gene-environment interactions in the etiology of selected birth defects. *Clin. Genet.* **75**, 409–423 (2009).
88. Wu, T. *et al.* Evidence of gene-environment interaction for the *IRF6* gene and maternal multivitamin supplementation in controlling the risk of cleft lip with/without cleft palate. *Hum. Genet.* **128**, 401–410 (2010).
89. Abbott, B. D. The etiology of cleft palate: a 50-year search for mechanistic and molecular understanding. *Birth Defects Res. B Dev. Reprod. Toxicol.* **89**, 266–274 (2010).
90. Jentink, J. *et al.* Valproic acid monotherapy in pregnancy and major congenital malformations. *N. Engl. J. Med.* **362**, 2185–2193 (2010).
91. Deroo, L. A. & Wilcox, A. J. First-trimester maternal alcohol consumption and the risk of infant oral clefts in Norway: a population-based case–control study. *Am. J. Epidemiol.* **168**, 638–646 (2008).
92. Boyles, A. L. *et al.* Maternal alcohol consumption, alcohol metabolism genes, and the risk of oral clefts: a population-based case-control study in Norway, 1996–2001. *Am. J. Epidemiol.* **172**, 924–931 (2010).
93. Wehby, G. L. & Murray, J. C. Folic acid and orofacial clefts: a review of the evidence. *Oral Dis.* **16**, 11–19 (2010).
94. Wilcox, A. *et al.* Folic acid supplements and risk of facial clefts: national population based case-control study. *BMJ* **334**, 464 (2007).
95. Bille, C. *et al.* Autoantibodies to folate receptor α during early pregnancy and risk of oral clefts in Denmark. *Pediatr. Res.* **67**, 274–279 (2010).
96. Yazdy, M., Honein, M. A. & Xing, J. Reduction in orofacial clefts following folic acid fortification of the, U. S. grain supply. *Birth Defects Res. A Clin. Mol. Teratol.* **79**, 16–23 (2007).
97. Johnson, C. & Little, J. Folate intake, markers of folate status and oral clefts: is the evidence converging? *Int. J. Epidemiol.* **37**, 1041–1058 (2008).
98. Ray, J. G., Vermeulen, M. J., Wyatt, P. R. & Cole, D. E. Association between folic acid food fortification and congenital orofacial clefts. *J. Pediatr.* **143**, 805–807 (2003).
99. López-Camelo, J. S., Castilla E. E. & Orioli, I. M. Folic acid flour fortification: impact on the frequencies of 52 congenital anomaly types in three South American countries. *Am. J. Hum. Genet. A* **152A**, 2444–2458 (2010).
100. Munger, R. G. *et al.* Plasma zinc concentrations of mothers and the risk of oral clefts in their children in Utah. *Birth Defects Res. A Clin. Mol. Teratol.* **85**, 151–155 (2009).
101. Porter, F. Cholesterol precursors and facial clefting. *J. Clin. Invest.* **116**, 2322–2325 (2006).
102. Shahrukh Hashmi, S., Gallaway, M. S., Waller, D. K., Langlois, P. H. & Hecht, J. T. National Birth Defects Prevention Study. Maternal fever during early pregnancy and the risk of oral clefts. *Birth Defects Res. A Clin. Mol. Teratol.* **288**, 186–194 (2010).
103. Mossey, P., Davies, J. A. & Little, J. Prevention of orofacial clefts: does pregnancy planning have a role? *Cleft Palate-Cran. J.* **434**, 244–250 (2007).
104. Wehby, G. L., Ohsfeldt, R. L. & Murray, J. C. 'Mendelian randomization' equals instrumental variable analysis with genetic instruments. *Stat. Med.* **27**, 2745–2749 (2008).
105. Eberhart, J. *et al.* MicroRNA Mirn140 modulates Pdgf signaling during palatogenesis. *Nature Genet.* **40**, 290–298 (2008).
106. Sharpe, J. *et al.* Optical projection tomography as a tool for 3D microscopy and gene expression studies. *Science* **296**, 541–545 (2002).
107. Visel, A. *et al.* Ultraconservation identifies a small subset of extremely constrained developmental enhancers. *Nature Genet.* **40**, 158–160 (2008).
108. Visel, A., Rubin, E. M. & Pennacchio, L. A. Genomic views of distant-acting enhancers. *Nature* **461**, 199–205 (2009).
109. Jones, J. L. *et al.* Wound complications following cleft repair in children with Van der Woude syndrome. *J. Craniofac. Surg.* **21**, 1350–1353 (2010).
110. Christensen, K., Juel, K., Herskind, A. M. & Murray, J. C. Long term follow up study of survival associated with cleft lip and palate at birth. *BMJ* **328**, 1405 (2004).
111. Bille, C. & Knudsen, B. Changing lifestyles and oral clefts occurerence in Denmark. *Cleft Palate-Cran J.* **42**, 255–259 (2005).
112. Menezes, R. *et al.* AXIS inhibition protein 2, orofacial clefts and family history for cancer. *J. Am. Dent. Assoc.* **140**, 80–84 (2009).
113. Yttri, J. E., Christensen, K., Knudsen. L. & Bille, C. Reproductive patterns among Danish women with oral clefts. *Cleft Palate-Cran. J.* 8 Sep 2010 (doi:10.1597/09-245).
114. Muenke, M. The pit, the cleft and the web. *Nature Genet.* **32**, 219–220 (2002).
115. Jianyan, L. *et al.* Analysis of interactions between genetic variants of *BMP4* and environmental factors with nonsyndromic cleft lip with or without cleft palate susceptibility. *Int. J. Oral Maxillofac. Surg.* **39**, 50–56 (2010).
116. Riley, B. M. *et al.* Impaired FGF signaling contributes to cleft lip and palate. *Proc. Natl Acad. Sci. USA* **104**, 4512–4517 (2007).
117. Vieira, A. R. *et al.* Medical sequencing of candidate genes for nonsyndromic cleft lip and palate. *PLoS Genet.* **1**, e64 (2005).

118. Venza, M. *et al. FOXE1* gene mutation screening by multiplex PCR/DHPLC in CHARGE syndrome and syndromic and non-syndromic cleft palate. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* **836**, 39–46 (2006).
119. Lidral, A. C. *et al.* Association of *MSX1* and *TGFB3* with nonsyndromic clefting in humans. *Am. J. Hum. Genet.* **63**, 557–568 (1998).
120. Vieira, A. R. *et al. MSX1* and *TGFB3* contribute to clefting in South America. *J. Dent. Res.* **82**, 289–292 (2003).
121. Suzuki, Y. *et al.* In a Vietnamese population, *MSX1* variants contribute to cleft lip and palate. *Genet. Med.* **6**, 117–125 (2004).
122. Martinelli, M. *et al.* Cleft lip with or without cleft palate: implication of the heavy chain of non-muscle myosin IIA. *J. Med. Genet.* **44**, 387–389 (2007).
123. Chiquet, B. T. *et al.* Genomic screening identifies novel linkages and provides further evidence for a role of *MYH9* in nonsyndromic cleft lip and palate. *Eur. J. Hum. Genet.* **17**, 195–204 (2009).
124. Jia, Z. L. *et al.* Association among polymorphisms at *MYH9*, environmental factors, and nonsyndromic orofacial clefts in western China. *DNA Cell Biol.* **29**, 25–32 (2010).
125. Chiquet, B. T. *et al. CRISPLD2*: a novel NSCLP candidate gene. *Hum. Mol. Genet.* **16**, 2241–2248 (2007).
126. Letra, A. *et al.* CRISPLD2 variants including a C471T silent mutation may contribute to nonsyndromic cleft lip with or without cleft palate. *Cleft Palate Craniofac. J.* 1 Jul 2010 (doi:10.1597/09-227).
127. Mills, J. L. *et al.* Folate-related gene polymorphisms as risk factors for cleft lip and cleft palate. *Birth Defects Res. A Clin. Mol. Teratol.* **82**, 636–643 (2008).
128. Jagomägi, T. *et al. MTHFR* and *MSX1* contribute to the risk of nonsyndromic cleft lip/palate. *Eur. J. Oral Sci.* **118**, 213–220 (2010).
129. Ding, H. *et al.* A specific requirement for PDGF-C in palate formation and PDGFR-α signaling. *Nature Genet.* **36**, 1111–1116 (2004).
130. Choi, S. J. *et al.* The PDGF-C regulatory region SNP rs28999109 decreases promoter transcriptional activity and is associated with CL/P. *Eur. J. Hum. Genet.* **17**, 774–784 (2009).
131. Sözen, M. A. *et al.* Mutation of *PVRL1* is associated with sporadic, non-syndromic cleft lip/palate in northern Venezuela. *Nature Genet.* **29**, 141–142 (2001).
132. Avila, J. R. *et al. PVRL1* variants contribute to non-syndromic cleft lip and palate in multiple populations. *Am. J. Med. Genet. A* **140**, 2562–2570 (2006).
133. Sözen, M. A., Hecht, J. T. & Spritz, R. A. Mutation analysis of the *PVRL1* gene in caucasians with nonsyndromic cleft lip/palate. *Genet. Test. Mol. Biomarkers* **13**, 617–621 (2009).
134. Carter, T. C. *et al.* Testing reported associations of genetic risk factors for oral clefts in a large Irish study population. *Birth Defects Res. A Clin. Mol. Teratol.* **88**, 84–93 (2010).
135. Beaty, T. H. *et al.* Testing candidate genes for non-syndromic oral clefts using a case-parent trio design. *Genet. Epidemiol.* **22**, 1–11 (2002).
136. Suazo, J., Santos, J. L., Scapoli, L., Jara, L. & Blanco, R. Association between *TGFB3* and nonsyndromic cleft lip with or without cleft palate in a Chilean population. *Cleft Palate Craniofac. J.* **47**, 513–517 (2010).
137. Thomason, H. A. & Dixon, M. J. Craniofacial defects and cleft lip palete. *Enc. Life Sci.* 15 Mar 2009 (doi:10.1002/9780470015902.a0020915).

Acknowledgements

We should like to thank many colleagues who have collaborated with us over the years and in particular K. Christensen, R. Lie, A. Jugessur, A. Lidral, J. Hecht, A. Vieira, M. Shi, P. Jezewski, D. Fitzpatrick, R. Munger, P. Trainor, J. Dixon, P. Romitti, P. Nopoulos, J. Canady, B. Schutte, K. Buetow, A. Sander, G. Wehby, S. Daack-Hirsch and S. Weinberg, as well as many students. We apologize for being unable to cite all of the relevant papers. We gratefully acknowledge generous funding sources including the Medical Research Council (G0901539), Wellcome Trust (082,868), US National Institutes of Health (P50-DE016215, R01-DE08559, R01-DE016148, R01-DE014581, U01-DE018993 and U01-DE20057), the Healing Foundation and the Manchester NIHR Biomedical Research Centre.

Competing interests statement

The authors declare no competing financial interests.

FURTHER INFORMATION

Jeffrey C. Murray's homepage: http://genetics.uiowa.edu/
COGENE: http://humgen.wustl.edu/COGENE
EMAGE database of *in situ* gene expression patterns in the mouse embryo: http://genex.hgu.mrc.ac.uk/emage/home.php
FaceBase: https://www.facebase.org
Online Mendelian Inheritance in Man (OMIM): http://www.ncbi.nlm.nih.gov/omim

SUPPLEMENTARY INFORMATION

See online article: S1 (table)

ALL LINKS ARE ACTIVE IN THE ONLINE PDF