ABSTRACT
Failure of facial prominence fusion causes cleft lip and palate (CL/P), a common human birth defect. Several potential mechanisms can be envisioned that would result in CL/P, including failure of prominence growth and/or alignment as well as a failure of fusion of the juxtaposed epithelial seams. Here, using geometric morphometrics, we analyzed facial outgrowth and shape change over time in a novel mouse model exhibiting fully penetrant bilateral CL/P. This robust model is based upon mutations in Tpap2a, the gene encoding transcription factor AP-2α, which has been implicated in both syndromic and non-syndromic human CL/P. Our findings indicate that aberrant morphology and subsequent misalignment of the facial prominences underlies the inability of the mutant prominences to fuse. Exencephaly also occurred in some of the Tpap2a mutants and we observed additional morphometric differences that indicate an influence of neural tube closure defects on facial shape. Molecular analysis of the CL/P model indicates that Fgf signaling is misregulated in the face, and that reducing Fgf8 gene dosage can attenuate the clefting pathology by generating compensatory changes. Furthermore, mutations in either Tpap2a or Fgf8 increase variance in facial shape, but the combination of these mutations restores variance to normal levels. The alterations in variance provide a potential mechanistic link between clefting and the evolution and diversity of facial morphology. Overall, our findings suggest that CL/P can result from small gene-expression changes that alter the shape of the facial prominences and uncouple their coordinated morphogenesis, which is necessary for normal fusion.

KEY WORDS: Craniofacial, TPAP2A, AP-2a, BOFS, Branchio-ocular-facial syndrome, Cleft lip/palate, Geometric morphometrics, Fgf signaling pathway

INTRODUCTION
Orofacial clefts are common human birth defects, with cleft lip with or without cleft palate (CL/P) occurring in about one in 700 births, and cleft palate only (CP) presenting about half as frequently (Mossey and Modell, 2012). Although many instances of orofacial clefts are syndromic and associated with a specific developmental syndrome, the majority of clefts are isolated and non-syndromic (Mossey and Modell, 2012; Rahimov et al., 2012).

Face shape has long been considered a factor in the development of CL/P. Studies of unaffected parents of children with CL/P using cephalometry, or more recently detailed three-dimensional (3D) morphometrics, show distinct differences in facial shape compared with unaffected families (McIntyre and Mossey, 2002; Nakasima and Ichinose, 1983; Wyszynski, 2002; Weinberg et al., 2006; Weinberg et al., 2009). These differences hold true for many different racial and ethnic backgrounds, and often include a more concave face.

In mouse, genetic manipulations tend to result in CP, rather than CL/P (which is more common in humans) (Juriloff and Harris, 2008). This difference presumably reflects a combination of alterations in facial shape, growth, morphogenesis and gene expression between these species. The greater prevalence of mutations leading to CP in mouse, combined with the relative ease of culturing isolated palatal shelves, has generated a better understanding of the mechanisms involved in the formation of the secondary palate. CP can result from a number of specific causes, including defects in growth and morphogenesis, defective fusion following apposition, mechanical defects caused by aberrant positioning of the tongue, or inappropriate fusion of the palatal epithelia to other oral epithelia (Bush and Jiang, 2012; Grilli-Linde, 2007). Studies on CL/P have not advanced as far, in part because, until recently, there were few models that displayed CL/P with high penetrance.

In both human and mouse, facial development begins with the appearance of discrete buds of tissue, termed the facial prominences (Brugmann et al., 2006; Chai and Maxson, 2006). Three pairs of prominences form the region of the upper jaw: the medial nasal, the lateral nasal and the maxillary. In mouse development, at embryonic day (E)9.5 the prominences have begun to form and, by E10.5, they are undergoing rapid outgrowth. The medial nasal prominences have met at the midline by E11.5, and the other prominences are in close apposition and are beginning to fuse. By E12.5 they have fused to create a continuous tissue band at the front of the face – the primary palate (Brugmann et al., 2006; Chai and Maxson, 2006; Ray and Niswander, 2012). Although the basic steps required for fusion are known, it is unclear which step or steps fail during the development of CL/P. There are two major and non-exclusive hypotheses for the etiology of CL/P: the first is that the epithelium at these locations is somehow incapable of fusing, whereas the second is that the shape...
TRANSLATIONAL IMPACT

Clinical issue

Orofacial clefting is one of the most frequent human birth defects, with cleft lip with or without cleft palate (CL/P) affecting between 1 in 500 to 1 in 1000 live births worldwide. Clefting of the secondary palate (CP; the roof of the mouth) occurs slightly less frequently. Clefting is treatable in humans, but even with treatment this condition can have serious consequences for speech, feeding and social interactions. CL/P or CP treatment requires a complex and multilayered approach, which includes successive surgeries, speech therapy and orthodontics; lifelong medical treatment is required for many individuals. Whereas, in humans, CL/P is more common than CP, most mouse models display CP. In addition, the classic mouse models of CL/P were hampered by partial penetrance of the phenotype.

Results

This study describes the generation of a new and a fully penetrant mouse model of CL/P caused by mutations in Tfap2a, a gene linked to human CL/P. With this model it is possible to study the gene expression changes underlying CL/P as well as to examine whether clefting results from altered facial shape or altered fusion events. The methodology of geometric morphometrics was employed to compare how facial shape changes between control and mutant embryos affects facial morphology. In the mutant model, clefting develops due to altered shape of the upper facial prominences, which changes prominence alignment, preventing fusion of the primary palate. Global gene expression analysis shows that alterations in gene expression in mutant mice are relatively minor despite the severity of the resulting pathology. However, alterations in the fibroblast growth factor (Fgf) signaling pathway are evident and decreasing the gene dosage of Fgf8 can generate a partial rescue of the phenotype.

Implications and future directions

This work suggests that the formation of the lip and primary palate is particularly sensitive to small changes in growth and gene expression. This finding might explain why this developmental process often goes awry in human development owing to genetic and/or environmental causes. Moreover, the finding that the clefting pathology in this model can be altered by manipulation of Fgf signaling during embryonic facial development could one day lead to directed early pharmacological interventions, rectifying facial growth and preventing the need for repetitive surgeries.

of the face and the direction of growth of the facial prominences are altered such that these structures are not aligned to allow for normal fusion (Ferretti et al., 2011; Trasler, 1968).

Geometric morphometrics provides a powerful means to evaluate the role of shape in the development of fusion defects of the upper lip and primary palate (Cooper and Albertson, 2008; Martínez-Abadias et al., 2012). Geometric morphometrics uses homologous-points-based data (landmarks) that are then scaled and superimposed for shape analysis by multivariate statistical methods (Klingenberg, 2010). In mouse, geometric morphometrics has been used to probe how facial shape might influence orofacial clefting in two strains for shape analysis by multivariate statistical methods (Klingenberg, 2010). In mouse, geometric morphometrics has been used to probe how facial shape might influence orofacial clefting in two strains. (Neo/Null). In this model, clefting develops due to altered shape of the upper facial prominences, which changes prominence alignment, preventing fusion of the primary palate. Global gene expression analysis shows that alterations in gene expression in mutant mice are relatively minor despite the severity of the resulting pathology. However, alterations in the fibroblast growth factor (Fgf) signaling pathway are evident and decreasing the gene dosage of Fgf8 can generate a partial rescue of the phenotype.

RESULTS

Derivation of the Neo/Null model of bilateral facial clefting

Loss of Tfap2a in the mouse results in major defects in development of the head, with widely separated mandibular and maxillary processes as well as exencephaly (Schorle et al., 1996; Zhang et al., 1996). More recently, several conditional Tfap2a alleles have been generated (Brewer et al., 2004), including the Ne allele (Fig. 1A), in which the neomycin resistance (neo) cassette is present between exons encoding part of the AP-2a DNA-binding domain (Williams and Tjian, 1991a; Williams and Tjian, 1991b). Because the neo insertion can create a hypomorphic allele in other instances (Hester et al., 2005; Meyers et al., 1998), we compared the phenotypes of mice that were heterozygous or homozygous for the Ne allele, as well as mice that were transheterozygous for the Ne and null alleles (Ne/Null). In this Tfap2a allelic series, Ne homozygotes and heterozygotes were normal and viable, whereas E18.5 Ne/Null mice had fully penetrant bilateral facial clefting with concomitant perinatal lethality (Fig. 1B,C). For further study, homozygous Neo mice were bred to heterozygous null (Tfap2a−/−) mice to generate 50% affected Neo/Null mice and 50% Neo/Wt mice. The latter had a normal appearance and served as the control group. Neo/Null embryos also showed additional gross external abnormalities, including polydactyly and mid/hindbrain exencephaly but at a lower frequency than CL/P (Fig. 1D and data not shown). With relevance for head formation, exencephaly occurred in ~30% of the E18.5 Neo/Null mice, compared with about a 12% incidence of neural tube closure defects in Tfap2a−/− mice (Kohlbeker et al., 2002) (R.M.G. and T.W., data not shown). Gross brain morphology in Neo/Wt and Neo/Null mice with normal neural tube closure was comparable, indicating that the mutants did not display a general disruption of brain patterning (data not shown). Although 30% of Neo/Null animals had exencephaly at E18.5, at E9.5 ~45% had abnormal neural tube closure, indicating that some embryos have a delay in this process that eventually resolves. Because more than 50% of Neo/Null embryos had normal neural tube closure at E9.5 but still developed bilateral CL/P, the CL/P phenotype could be studied either in isolation or in conjunction with the exencephaly.

Bone and cartilage staining of controls showed that the premaxilla formed a continuous arc from the midline to the maxilla, whereas, in the Neo/Null embryos (with closed neural tubes), the premaxilla was discontinuous (Fig. 1E-H). This resulted in the premaxilla protruding from the anterior of the face as a bulbous structure. Although both the primary and secondary palates were affected in
the Neo/Null mice (Fig. 1I,J), the secondary-palate defects were mainly associated with defects in the premaxilla and the premaxillary palatal processes. Thus, defects were generally not observed in the more posterior regions of the secondary palate, with the exception of a low penetrance of defects in the maxillary palatal processes (2/10 embryos).

Because Neo/Null mice are more severely affected than Tfp2a−− mice, we next used reverse transcription (RT)-PCR to determine whether the neo insertion significantly altered the expression and/or processing of the associated transcripts from the Neo allele. RNA was isolated at E10.5 from the faces of wild-type controls, Tfp2a−/− and Neo/Null mice, as well as Neo homozygotes and heterozygotes. Primer pairs located within exons 5 and 7 (Fig. 1A) detected a wild-type spliced product in all mice carrying the Neo allele, and provides a means to study the development of this pathology using geometric morphometrics during the critical period of normal facial fusion, between E9.5-E11.5. Because a significant fraction of the Neo/Null mice had exencephaly, we could also investigate how this pathology impacted facial shape. Thus, three groups were landmarked and analyzed using geometric morphometric methods.

**Independent shape differences associated with exencephaly and facial clefting**

The Neo/Null model exhibits completely penetrant bilateral CL/P and provides a means to study the development of this pathology using geometric morphometrics during the critical period of normal facial fusion, between E9.5-E11.5. Because a significant fraction of the Neo/Null mice had exencephaly, we could also investigate how this pathology impacted facial shape. Thus, three groups were studied: Neo/Wt; Neo/Null with closed neural tube; and Neo/Null with open neural tube. As an initial analysis, a total of ~100 embryos were studied: Neo/Wt; Neo/Null with closed neural tube; and Neo/Null with open neural tube. In contrast, primer pairs containing various mutant allele combinations, compared with wild type (Fig. 1A). The variants would generate protein truncations and frameshifts that would abolish AP-2α DNA-binding and dimerization. These findings demonstrate that the Neo allele generates aberrant transcripts that would not encode functional AP-2α, as well as reducing the quantities of wild-type product. In conjunction with the phenotypes observed from previous AP-2α studies as well as in the allelic series, these findings suggest that, in the Neo/Null mice, functional Tfp2a levels drop below a critical threshold required for normal facial fusion.

**Fig. 1. The Neo/Null allele and the bilateral cleft phenotype.** (A) Schematic diagram of the Tfp2a alleles. The Neo/Null model is a combination of the Neo (Brewer et al., 2004) and Tfp2a null (Zhang et al., 1996) alleles. The positions of primers for RT-PCR and qRT-PCR are shown, along with the intron:exon structure of the gene, and the position of sequences introduced by gene targeting. See Materials and Methods for details of primers. Frt, flippase recombination target. (B-D) E18.5 embryos showing the control phenotype (B), and the bilateral cleft phenotype in Neo/Null with normal neural tube closure (C) and Neo/Null with exencephaly (D). (E-H) Bone and cartilage staining from P1 Neo/Wt (E,G) and Neo/Null (F,H) after mandible removal showing norma basilaris (E,F) and norma lateralis (G,H). An asterisk after the name of the bone represents an area where differences were observed between Neo/Wt control and Neo/Null embryos. (I,J) Images of the palate from Neo/Wt (I) and Neo/Null (J) E18.5 embryos. Note the cleft in J extending from the primary palate into the secondary. (K) RT-PCR analysis of the Tfp2a transcripts that are present in the E10.5 face using the primer pairs shown in A. ‘−RT’ is a no reverse transcriptase control, and ‘Wt’ is wild type. (L) qRT-PCR data showing the ratio of Tfp2a transcripts containing exons 6-7 to those containing exons 2-3. Transcripts lacking exon 6-7 would lack the DNA-binding domain. *P<0.05 to all other genotypes. The two genotypes marked by a bar are not statistically different from each other, but they are different from all other alleles.
All data were regressed against tail somite number and centroid size to remove small differences within the age groups.

E9.5 was the earliest time point investigated, because landmarking requires sufficient surface detail to mark consistent morphological features. At E9.5, 20 landmarks could be assigned and these could be correlated across all samples of all ages (Materials and Methods, and supplementary material Table S1). When scanning later time points, E10.5 and E11.5, there was sufficient surface detail to assign additional landmarks to the developing face. To compare groups and examine the between-group differences, a Procrustes permutation test was used, followed by a canonical variates analysis (CVA) to interpret the nature of the differences. At E9.5, there were statistically significant differences between all groups (Fig. 2A; supplementary material Table S2). Three groupings had large Procrustes distances, corresponding to embryos with cranial neural tubes that were closed, delayed or everted/presumptive exencephalic (supplementary material Fig. S2A shows representative raw scans of each grouping). In contrast, the two groups with a closed neural tube clustered more tightly together, despite their different genotypes. Therefore, at this time point, the most significant differences in facial shape correspond to the state of neural tube closure, rather than the *Tfap2a* genotype. Notably, embryos with an open neural tube had a narrower telencephalon that is likely to impact the associated frontonasal tissue. By E11.5, the groups separated more strongly by genotype, in addition to neural tube closure, creating three statistically distinct groups: Neo/Wt; Neo/Null open neural tube; and Neo/Null closed neural tube (Fig. 2A). The separation between the two Neo/Null groups indicates that the failure of neural tube closure independently affects facial shape.

Facial clefting correlates with altered morphology of the nasal pit and maxilla

Next, we focused on the development of Neo/Wt (control) and Neo/Null (CL/P) embryos with appropriate neural tube closure to...
address specifically the shape changes that were involved in the development of the orofacial cleft, especially in the area around the newly forming nasal pit. The earliest time point examined was ~E10.25, equivalent to 23 or greater somites (Som) from tail tip to the forelimb, when the nasal pits first become visible and can be landmarked. At this stage, an established set of 48 landmarks that focus on the developing facial prominences and nasal pit could be used. The data were then analyzed using a principal component analysis (PCA) to determine the principal axes of variation. At this early developmental stage, 23-35 Som, the three first principal components (PC1-3) did not distinguish between the different genotypes (supplementary material Fig. S3). However, examination of PC4 indicates that alterations in facial morphology specific to the Neo/Null mutation can already be detected (Fig. 2B). PC4 accounts for ~8% of the overall shape variance (supplementary material Fig. S3). The morphological significance of PC4 can be observed through the deformation of the wireframe diagram at various points along the PC. The Neo/Wt embryos are clustered along the negative end and the Neo/Null along the positive. The values ~0.05 and +0.05 were chosen to represent Neo/Wt and Neo/Null, respectively. Based on the wireframe deformations, at this stage, genotype correlated with changes in the shape of the nasal pits and maxilla (Fig. 2B,D). Because PC4 (genotype) accounts for only a small percentage of the overall morphological variation, the remaining variation might be due to rapid and major shape changes occurring during this period of facial development that are not fully removed by regressing on somite stage. In addition, some variation might be attributable to genetic background from the use of an outbred mouse strain (Black Swiss). In addition to PCA, we used CVA to visualize differences between the E10.25 Neo/Null and Neo/Wt embryos. The shape differences between embryos of these two genotypes at this stage, although relatively small, were highly significant ($P<0.005$; Procrustes permutation test; data not shown).

By E11.5 (40-50 Som), the percentage of the shape change that can be attributed to genotype increased, and was now seen in the first PC, which accounted for about 30% of the total shape variance. Based on the wireframe deformation graph for PC1, drastic changes could be seen in both the relative size and shape of the nasal pit and the maxilla (Fig. 2C). Furthermore, it seems that, when a typical control embryo is warped to fit the landmark coordinates that would resemble a typical Neo/Null embryo, a notch begins to form underneath the nasal pit in the area where a cleft will form (Fig. 2E, arrow), suggesting that a quantitative deformation of the control facial shape to match the Neo/Null points is sufficient to develop a cleft.

We also ran simulations by warping an E10.5 control embryo either along the ‘Neo/Wt’ or ‘Neo/Null’ side of the E11.5 PC1 to visualize how changes in shape over this period affect development (supplementary material Movies 1, 2). These simulations reinforce the idea that the overall growth trajectories of the facial prominences in the two types of embryos were very different. In the embryo developing along the Neo/Wt phenotype, growth and morphogenesis resulted in the whole face moving inward toward the midline. In the Neo/Null example, the maxilla seemed to grow away from the midface, and the tissues that were marked by the landmarks along the base of the nasal pit, which come together in the control, maintained their distance. The abnormal growth of the maxilla in the Neo/Null embryos is also apparent in the raw scans (supplementary material Fig. S2B,C). Specifically, at E11.5, the Neo/Null maxilla was located in a more rostrolateral position, such that, in lateral view, it covered part of the developing eye. Taken together, these data support the idea that the failure of the prominences to fuse and form a normal primary palate is a direct effect of the change in shape.

The Neo/Null model exhibits subtle changes in growth and Fgf pathway gene expression

It is notable that small differences in local growth rates can produce significant effects on shape (Boehm et al., 2010). Therefore, one explanation for the difference in facial shape and growth trajectory would be a subtle difference in cell proliferation levels between the mutant and the control embryos. To examine proliferation, we utilized EdU, a thymidine analog, to label newly synthesized DNA, as well as an antibody against phospho-histone H3 to identify proliferating cells. Labeled cells were then examined in either the ectoderm or mesenchyme of the prominences. In the mesenchyme, both methods identified consistent decreases in proliferation in the Neo/Null samples for the medial and lateral nasal prominences as well as the maxilla, compared with Neo/Wt controls (Fig. 3A-H). In contrast, proliferation in the ectoderm was not significantly different between Neo/Wt and Neo/Null mutants, with the exception of phospho-histone H3 staining in the maxilla (compare Fig. 3F,H). TUNEL staining did not reveal any significant differences in apoptosis between mutant and control animals, with the exception of the nasal pit ectoderm. In wild-type samples, apoptotic cells were detected where the medial and lateral edges of the nasal pit ectoderm were beginning to fuse. This region of staining, presumably due to apoptosis of the epithelial seam at the fusion point, was not observed in Neo/Null sections, consistent with the failure of fusion between these nasal prominences (supplementary material Fig. S4A,B).

We next compared gene expression profiles in Neo/Null and Neo/Wt nasal and maxillary prominences at E10.5 using microarray analysis. Although differences in shape and proliferation were already apparent at this stage, gene expression changes were moderate (Fig. 4A,B), with less than 300 genes detected as altered by >1.25-fold and few of these altered by >twofold (supplementary material Tables S3, S4). However, analysis of these gene lists using gene set enrichment analysis (GSEA) highlighted enrichment of Fgf pathway members specifically in the nasal prominence (Fig. 4C). Because Fgf signaling has major roles in facial development, this association was further investigated using qRT-PCR and in situ hybridization. First, RT-PCR was used on Neo/Wt controls and Neo/Null mutants (Fig. 4N) to examine expression of Fgf8 as well as Wnt9b, an additional gene associated with bilateral CL/P (Jin et al., 2012). In agreement with the microarray data, a consistent ~1.3-fold increase was detected in Fgf8 expression in the Neo/Null mutants, whereas Wnt9b levels were relatively unchanged from controls (Fig. 4N). A slight increase in Fgf8 expression in the ectoderm of the facial prominences was also revealed by in situ analysis coupled with optical projection tomography (OPT) imaging (Fig. 4D-G,L-M). This increase was most notable in the margin of the lateral nasal prominence bordering the nasal pit. The potential upregulation of Fgf signaling was further investigated by examining the expression of pathway targets Dusp6 and Mapk1. Dusp6, a negative feedback regulator of the pathway, showed increased expression in the developing facial prominences of Neo/Null mice compared with controls, in contrast to the more comparable expression in the limbs (Fig. 4H-K). However, we did not detect significant changes in the levels of Mapk1-encoded total Erk or phospho-Erk by western blot analysis of whole E10.5 facial tissue (supplementary material Fig. S5).

Reduction of Fgf8 gene dosage leads to a partial rescue of Neo/Null bilateral clefting

To determine whether changes in Fgf signaling could impact the Neo/Null phenotype, the Neo/Null allelic combination was crossed onto an Fgf8 heterozygous background (Meyers et al., 1998). This
generated four genotypes with various combinations of the Tfap2a and Fgf8 alleles (Fig. 5A). Interestingly, examination of the Neo/Null neonates heterozygous for the Fgf8 null allele indicated that 8/18 presented with a unilateral cleft primary palate as opposed to the fully penetrant bilateral cleft characteristic of Neo/Null animals with wild-type Fgf8 gene dosage (Fig. 5A-F). Skeletal organization of ‘rescued’ Neo/Null;Fgf8het mice was assessed by μCT and confirmed a unilateral premaxillary fusion defect (Fig. 5F, red arrow). Note that there was no laterality preference for the ‘rescue’ in Neo/Null;Fgf8het mice, with remaining clefts being detected on either the left or right side of the face. The altered susceptibility to the bilateral facial clefting caused by the reduction of Fgf8 gene dosage in Neo/Null mice prompted a morphometric analysis of E10.5 embryos to identify early associated shape changes. CVA demonstrated significant between-group differences, with CV1 reflecting differences caused by the Fgf8 mutation and CV2 reflecting the differences caused by the Tfap2a genotype (Fig. 6A). Increases along CV1 represent greater width of the frontal nasal prominence and a more external orientation of the nasal pits. Intriguingly, the combination of Neo/Null and Fgf8 null alleles that partially rescued the orofacial clefting did not return the phenotype nearer to the control shape, but resulted in a distinct grouping (Fig. 6A). PCA was next utilized to examine these differences in greater detail (Fig. 6B,C) and indicated that, when compared with the Neo/Wt;Fgf8wt control group, separation along PC1 indicates that the Neo/Null;Fgf8het has a slight widening of the midface but more obviously the nasal pits are somewhat larger and are angled slightly downwards toward the maxilla (Fig. 6B, lower images). Therefore, the frequent rescue of the bilateral clefting in the Neo/Null background caused by reduction of Fgf8 gene dosage does not occur by simply rectifying the changes observed in the Neo/Null embryos at a similar stage (PC4; Fig. 2B,D and supplementary material Fig. S3). Instead, the reduction of Fgf8 causes an additional set of changes in the shape of the midface and nasal pits. Furthermore, when these Neo/Null;Fgf8het compound mutants are compared with the Neo/Null;Fgf8wt, we noted a subtle increase in the width of the midface, a lateral pivot of the nasal pit toward the maxilla, and an increase in both the size and width of the nasal pit (Fig. 6C, lower images). These changes would push the nasal pit more into alignment with the maxilla in the Neo/Null;Fgf8het embryos.

Previous studies have indicated that a reduction in Fgf8 dosage alters the mean facial shape relative to wild-type populations and might also lead to an increase in phenotypic variation (Griffin et al., 2013). Such phenotypic variation might influence how an individual mutation affects a population or leads to partial penetrance (Hallgrimsson et al., 2002; Jamniczky and Hallgrimsson, 2009). We therefore investigated whether alterations in Tfap2a could also influence phenotypic variance as a potential mechanism for interactions with Fgf8 gene dosage in modifying orofacial clefting. Specifically, we assessed the variance within the E10.5 dataset, using
the trace of the variance covariance matrix, a multivariate measure of variance (Schaefer and Lauder, 1996). In both the Neo/Null;Fgf8wt and Neo/Wt;Fgf8het embryos, the phenotypic variance for shape was significantly increased (Fig. 7A). Surprisingly, however, the combination of these mutant alleles in Neo/Null;Fgf8het embryos did not produce a further increase. In fact, the shape variance in this latter allelic combination was not significantly increased over the controls. Thus, the increase in variance seen in Neo/Null;Fgf8wt and Neo/Wt;Fgf8het seems to be moderated in Neo/Null;Fgf8het. These findings argue against the Neo/Null and Fgf8 heterozygous changes acting in an additive or synergistic manner on the variance, and instead indicate that the combination of these mutations unexpectedly dampens the amount of variance.

Another method for observing the interaction between shape changes is to examine how different parts of the face are integrated with one another, or simply to identify how much changes in one region correlate with changes in another. Observations of integration can also determine how underlying gene expression changes can affect shape. Therefore, a two-block partial least squares regression (PLS) analysis was used to examine correlation of shape changes between the nasal processes and the maxilla between the various genotypes (Fig. 7B and supplementary material Table S5). This analysis indicated a very high degree of integration of shape differences between the prominences in the various models. The vectors describing the directions of landmark displacements that correspond to the covariances between the nasal and maxillary processes were also significantly correlated across the samples (P-values in supplementary material Table S5). This indicates that both the pattern and magnitude of integration among the processes are similar across genotypes and further suggests that the same processes that increase variance are also linked to overall structural changes within the face.

**DISCUSSION**

In this study, we describe the derivation of a new mouse model displaying fully penetrant bilateral CL/P that is caused by an insertional mutation in an intron of Tfap2a (Tfap2a(neo/neo) × Tfap2a(+/-);Fgf8(+/-) crosses (L and R are left- and right-sided cleft, respectively). (B,C) Frontal images of the heads of Neo/Null;Fgf8wt (B) and Neo/Null;Fgf8het (C) P0 pups. Black arrow shows the unilateral cleft line. (D-F) Ventral view of skulls using μCT for Neo/Wt;Fgf8wt (D), Neo/Null;Fgf8wt (E) and Neo/Null;Fgf8het (F) P0 pups. Red arrows show clefting at points where the premaxilla and the palatal process of the premaxilla should be fused.
genes, including p63 and IRF6 (Ferretti et al., 2011; Rahimov et al., 2008; Wang et al., 2013). Here, we have shown that, as the level of functional Tfap2a transcripts are decreased, presumably below a certain threshold, there are changes in Fgf signaling, prominence growth and facial shape that are associated with orofacial clefting.

The primary conclusion from our study is that the failure of facial fusion in the Neo/Null model is due to changes in the growth and morphogenesis of the facial prominences. Using μCT and geometric morphometrics, differences in facial shape between Neo/Null and control embryos were apparent by E10.25, and the morphologies of the facial prominences became highly divergent by E11.5. Analysis of the relative movement of the components of the face over time in control embryos indicates that all the facial prominences normally wrap around the mediolateral axis and converge towards the midline. However, in Neo/Null embryos, the lateral edge of the nasal pit flares outward instead of growing in toward the center and the maxillary prominence is also projected laterally. These alterations from the normal morphogenetic pathway prevent apposition and fusion of these facial prominences. Therefore, in common with some forms of CP, CL/P develops in this model because the prominences are not aligned properly for fusion. Previous studies in chick have indicated that directional differences in cell proliferation can affect shape and overall growth trajectories of the facial prominences (Li et al., 2013b). Similarly, in the Neo/Null mouse model, we show that one consequence of decreased Tfap2a levels is a reduced rate of proliferation in the mesenchyme of the nasal and maxillary prominences, which would presumably lead to a change in the overall growth and, potentially, the shape of the nasal pit and mid-face.

A secondary conclusion from these analyses is that neural tube closure defects correlate with differences in the shape of the face from very early time points. A subset of the Neo/Null mice developed exencephaly and, by E9.5, this group could be clearly distinguished on the basis of facial shape from other embryos with closed or delayed neural tube closure. There are several possible explanations for the effect of aberrant neural tube closure on face development. One hypothesis suggested by the μCT imagery is that facial defects are linked with defective morphogenesis of the
telencephalon. In particular, failure to close the neural tube leads to a much narrower telencephalon and, based on data indicating that the forebrain acts as a major signaling center for midfacial development (Gongal et al., 2011; Hu and Marcucio, 2009), this would alter regulatory crosstalk between these two regions. Alternatively, the neural tube defect might lead to abnormal formation or migration of the neural crest that could limit the population of mesenchymal cells available for face formation. Whatever mechanism is responsible for the distinct facial development of the forebrain (Inman et al., 2013; Tabler et al., 2013). Crucially, we have shown that modulating Fgf signaling through altering Fgf8 gene dosage can either exacerbate or ameliorate mouse craniofacial defects (Inman et al., 2013; Tabler et al., 2013). These findings indicate that reduced Tfap2a levels do not cause drastic changes in the expression of a specific set of gene targets, but instead might cause more subtle differences in the overall gene regulatory network controlling facial development. We note that similar conclusions were obtained from a study regarding the influence of the Myc transcription factor on mouse facial development (Uslu et al., 2014). However, we did find evidence of alterations in Fgf signaling in the Neo/Null model, including a modest upregulation of Fgf8, Fgf12 and Fgf17 expression as well as of the Fgf target Dusp6. Fgf8 has proven roles in the regulation of facial shape, facial development and midfacial integration in multiple species (reviewed in Dorey and Amaya, 2010; Griffin et al., 2013; Inman et al., 2013). Previous studies have also shown that modulation of Fgf signaling through altering Fgf8 gene dosage can either exacerbate or ameliorate mouse craniofacial defects (Inman et al., 2013; Tabler et al., 2013). Crucially, we have shown that modulating Fgf signaling through altering Fgf8 gene dosage can either exacerbate or ameliorate mouse craniofacial defects (Inman et al., 2013; Tabler et al., 2013). The Neo/Null strain, which has been studied using 2D morphometric methods, shows shape differences primarily in the nasal pit (Parsons et al., 2008; Young et al., 2007).

Neo/Null is the first fully penetrant bilateral clefting model examined with morphometric methods and combines the separate aspects of these other two models, with simultaneous morphological differences in both the maxilla and nasal pit. We propose that the combination of these two potentially independent shape changes distinguishes the fully penetrant CL/P phenotype from the two previously studied partially penetrant models. Taken together, though, the morphometric data obtained from these three models of CL/P indicate that they each exhibit modest changes in the shape and growth direction of the facial prominences. Such alterations would prevent these structures from being in the correct positions at the correct times, thus precluding normal fusion.

Despite the similarities with the A/WySn model noted above, we did not detect differences in the expression of Wnt9b in the Neo/Null mice. Indeed, overall gene expression changes between Neo/Null and controls were limited in both number and fold change despite the resulting severe facial pathology. These findings indicate that the frequency of 50-100% depending on background (Jin et al., 2012; Juriolfo and Harris, 2008; Juriolfo et al., 2006). C/Fr and A/WySn strains have considerable within-group variation, which might correlate with the partially penetrant nature of their mutant facial phenotypes, and it has not been possible to identify definitive early changes that will result in CL/P. Nevertheless, these studies are pertinent to the shape changes observed in Neo/Null mice. In the A/WySn model, the facial prominences are less well developed overall compared with controls at a similar embryonic stage, and the shape differences seem to be concentrated in the maxilla. In common with the Neo/Null model, the hypoplastic development might be due to decreased cell proliferation in the mesenchyme, because this also occurs in the Wnt9b-null mouse (Jin et al., 2012). The C/Fr strain, which has been studied using 2D morphometric methods, shows shape differences primarily in the nasal pit (Parsons et al., 2008; Young et al., 2007).

As noted above, in the Neo/Null model, we observe changes in the shape of both the nasal and maxillary prominences that places them in more lateral positions. The shape change in the Fgf8 heterozygous mice instead affects only the nasal pit, which is however also tilted outwards. It seems, therefore, that the Neo/Null clefting defect could
be rescued when combined with an Fgf8 mutation because the nasal pit is positioned more laterally and proximally, allowing an increased area for interaction, and therefore fusion, with the maxilla. Thus, the alteration in shape when Fgf8 gene dosage is reduced does not seem to directly ‘reverse’ or ‘undo’ the shape change in the Neo/Null mutant; rather, it provides a compensatory change that facilitates fusion. The observation that manipulation of Fgf signaling can rescue aspects of bilateral CL/P in the Neo/Null model might also have wider significance for understanding and treating orofacial clefting in both mouse and human. In future, it would valuable to determine if genetic or pharmaceutical alteration of Fgf signaling could alter the severity and frequency of clefting in other mouse models that exhibit this pathology.

In addition to changes in mean shape, we also observed changes in shape variance in our model, which could be pertinent to the evolutionary constraints that act upon facial morphology. Recent studies have indicated that only a limited number of facial shapes are developmentally tolerated across a wide array of species (Young et al., 2014). In support of this theory, human studies demonstrate differences in the facial shape of control groups compared with those of unaffected relatives of individuals with CL/P, suggesting that morphology might predispose individuals to clefting (Weinberg et al., 2009; Weinberg et al., 2008). Phenotypic heterogeneity is a common feature of human birth defects, including CL/P (Milunsky et al., 2011). The extent to which increased shape variance, caused by underlying mutations, contributes to this heterogeneity remains virtually unknown. In mouse, CL/P has been hypothesized to occur at extremes of normal variation of facial shape (Juriloff and Trasler, 1976; Trasler, 1968). Therefore, increased variance might also contribute to both penetrance and expressivity of human CL/P, because this might push individual, or sides within individuals, over a theoretical threshold beyond which normal lip fusion fails (Parsons et al., 2008). It is therefore intriguing that alteration of Tcap2a and Fgf8 can influence variance (this study) (Griffin et al., 2013). Moreover, the observation that the combination of mutations in these two genes dampens the variance caused by mutations affecting only a single gene is intriguing and suggests interacting pathways. Further studies are warranted to determine the mechanistic links between Tcap2a and Fgf signaling as well as a possible wider role for variance in facial shape and orofacial clefting.

In summary, analysis of the robust Neo/Null orofacial clefting model provides the first direct evidence in mouse demonstrating that facial shape changes directly interfere with prominence fusion. We also note that the overall shape and gene expression changes in the Neo/Null model are relatively modest and yet lead to drastic morphological consequences. Based on these observations, we speculate that the process of facial fusion is exquisitely sensitive to subtle changes in the shape of the facial prominences, which would render this developmental process prone to genetic and environmental perturbation consistent with the frequent occurrence of human CL/P birth defects. Our data also suggest that combinations of effects across multiple pathways can influence facial shape and the generation of dysmorphology in ways that are complex and not easily predicted. Recently, additional mouse models with fully penetrant CL/P have been identified, including mutations involving Lrp6, Pbx and Bmpr1a (Das and Crump, 2012; Ferretti et al., 2011; Liu et al., 2005; Song et al., 2009; Wang et al., 2013). Morphometric analysis of these additional models would determine whether they also involve similar subtle changes in growth trajectories of the prominences, or whether further mechanisms, such as a failure of fusion following normal apposition of the prominences, can also account for this widespread human pathology.

**MATERIALS AND METHODS**

**Breeding and genotyping**

Mice were maintained on an outbred Black Swiss background (Charles River, Wilmington, MA). To obtain Neo/Null embryos, female Neo homozygote animals (Neofox as described previously (Brewer et al., 2004)) were crossed with Tcap2a+/− male mice (Zhang et al., 1996) to yield 50% Neo/Null and 50% Neo/Wt. To obtain Neo/Null embryos that were also heterozygous for Fgf8, female Neo homozygote animals were crossed with Tcap2a+/−; Fgf8+/− male mice. The Fgf8 null allele was derived by crossing the Fgf8lox line, Fgf8lox+12.4dmt, with β-actin Cre transgenic mice (Meyers et al., 1998), and then crossing back to Black Swiss mice to remove the β-actin Cre transgene. Females were examined in the morning for presence of a vaginal plug, and the presence of a plug was designated E0.5. Genomic DNA for genotyping was derived from tail biopsies or from the embryonic yolk sac and genotyping was performed as previously reported (Brewer et al., 2004; Meyers et al., 1998; Zhang et al., 1996). All animal experiments were performed in accordance with protocols approved by the University of Colorado Denver Animal Care and Usage Committee.

**Microarray**

E10.5 facial tissue was microdissected into separate nasal and maxillary prominence fractions. Samples were pooled to generate sufficient tissue for three independent biological replicates. Microarray analyses were performed in triplicate using these pooled samples as described (Feng et al., 2009). Briefly, total RNA was extracted using Trizol (Life Technologies, Carlsbad, CA), then purified using the RNAeasy MiniKit (Qiagen, Venlo, The Netherlands). Microarray analyses were carried out by the UCD Gene Expression Core Facility. In vitro transcription (IVT) was performed to generate biotin-labeled cRNA using an RNA Transcript Labeling Kit (Enzo Inc., Farmingdale, NY). Biotin-labeled cRNA was purified using an RNAeasy affinity column (Qiagen) and the cRNA was fragmented and analyzed on the Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA). cRNA was hybridized to Affymetrix GeneChip Mouse430 2.0 microarrays. Processing was performed in the GeneChipH Operating Software (Affymetrix) as discussed previously (Feng et al., 2009) using a scaling factor to bring the average intensity for all probes on the array to the same target intensity value, allowing samples to be compared across arrays. All data from our analysis are available via GEO (GSE60058; http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE60058). Analysis of gene representations was performed using GSEA (Subramanian et al., 2005).

**Reverse-transcription PCR and real-time PCR**

RNA was isolated from microdissected E10.5 faces through a standard Trizol protocol (Life Technologies). cDNA was made using the SuperScript III First-Strand Synthesis Kit for RT-PCR (Life Technologies) using 1 μg RNA/10 μl reaction. Three biological replicates were used for all analyses. For RT-PCR, the following primers were used: 5′-GAGACGTAAAAGC-TGCCAAACGTTACCCTCCTC-3′; 7R: 5′-CATGGGAGTAGGTTGA-AGTTGGTGTCAGGC-3′; PGKup: 5′-GGCCTCCACTACCCGGTAACTGTTGCCGGA-3′. For real-time PCR, pre-validated probes and primers were used to examine Tcap2a levels targeting exons 2-3 and 6-7 (IDT PrimerTime assay Mm.PT.56a.158069 and 12995135, Integrated DNA Technologies, Coralville, IA) as well as Fgf8 (TaQMan gene expression assay Mm00438922_m1, Applied Biosystems, San Francisco, CA) and Wnt9b (TaQMan gene expression assay Mm00457102_m1). Gapdh was used as a housekeeping control for all assays (ABI, TaqMan gene expression assay Mm99999915_g1). The TaqMan Gene Expression Master Mix was used and samples were run on a BioRad MyCycler. Each biological replicate was run in duplicate. Quantification was performed by comparison to a standard curve.

**In situ hybridization and optical projection tomography (OPT)**

In situ hybridization was performed as previously described (Feng et al., 2009). Briefly, dissected embryos were fixed overnight in 4% paraformaldehyde in phosphate buffered saline (PBS), dehydrated through a PBS/methanol gradient series (75% PBS/25% methanol, 50/50, 25/75) and
then stored in 100% methanol at −20°C. At the start of the protocol embryos were taken back through the series into PBST (0.01% Tween 20). Embryos were incubated in 10 μg/ml proteinase K for 20 minutes, re-fixed in 4% paraformaldehyde + 0.1% glutaraldehyde for 20 minutes, then pre-hybridized in hybridization buffer (50% Formamide, 1.3-SSC, 50 μg/ml Yeast tRNA, 100 μg/ml Heparin, 0.2% Tween 20, 0.5% CHAPS, 5 mM EDTA) for 3 hours at 70°C. Embryos were then incubated in fresh hybridization buffer containing dig-UTP-labeled probe overnight at 70°C. Next, embryos were washed in hybridization buffer then in Maleic acid buffer saline with 0.05% Tween 20 (MABT). Embryos were then blocked in 20% sheep serum with 2% blocking reagent (Roche, Penzberg, Germany). Probe was detected using an anti-digoxigenin antibody (Roche) diluted 1:2000 in the blocking reagent and incubated overnight at 4°C. Following 2 days of rinsing in MABT and alkaline phosphatase buffer, signal was developed in BM Purple (Roche).

For OPT, after completion of the whole-mount in situ hybridization procedure, embryos were embedded in 1% low-melt agarose (Life Technologies) then cleared in methanol for 24 hours followed by overnight in a mixture of 1 benzyl alcohol:2 benzyl benzoate. Cleared embryos were imaged on the Bioptonics 3000 scanner and images reconstructed using the SkyScan software.

**Bone and cartilage staining**

E18.5 fetuses were sacrificed and stained as previously described (Brewer and Williams, 2004). Briefly, fetuses were incubated 2 days in 100% ethanol, 2 days in acetone, 1 week in dye solution (0.1% Alizarin red and 0.3% Alcian blue in 70% ethanol) then cleared through a series of potassium hydroxide (2%) and glycerol.

**Western blot**

Phosphatase inhibitor (Pierce cat # 8868, Rockford, IL)-treated lysates from E10.5 dissected faces were separated by 10% denaturing PAGE then transferred to a nitrocellulose membrane, blocked 1 hour in 5% BSA, then probed for phospho-Erk (Cell Signaling Technology #4370, Danvers, MA), total Erk1 (Santa Cruz Biotech SC-94, Santa Cruz, CA) and Erk2 (Santa Cruz SC-153).

**Cell proliferation and apoptosis**

Pregnant females were injected with 200 μg of Click-it EdU (BrdU analog, Life Technologies) 30 minutes prior to sacrifice. Embryos were dissected at E10.5, immediately fixed in 4% paraformaldehyde in PBS overnight, dehydrated through an ethanol series (25%, 50%, 70% overnight, 100% 1 hour), then processed through a xylene and paraffin series (100% xylene: 50% xylene + 50% paraffin; 100% paraffin), embedded in paraffin and cut as 5-μm sections. Only embryos with closed neural tubes were selected for analysis. Sections were dewaxed and rehydrated and then imaged following the manufacturer’s instructions omitting the Hoechst blue. Following the Click-it chemistry using the 594 fluorophore, slides were blocked for 1 hour in Cell Signaling block buffer, and incubated overnight with a rabbit anti-phospho histone H3 antibody (Cell Signaling Technology #9701). Cells were imaged following detection using goat anti-rabbit Alexa Fluor 488 secondary antibody (Cell Signaling Technology #4412). For TUNEL, alternate sections were selected and stained using the ApopTag fluorescein kit following the manufacturer’s instructions, except that sections were permeabilized according to the Click-it kit (0.1% Triton-X for 10 minutes) (Millipore #7160, Millipore, Billerica, MA). All sections were counterstained with Hoechst blue and Draq5 (Cell Signaling Technology #4084). Images were generated on a Leica SP5 confocal microscope and nuclei were hand counted on ImageJ. Three embryos per condition were counted and three sections were analyzed per animal.

**Morphometrics**

Embryos were dissected between E9.5 and E11.5, and immediately fixed in 4% paraformaldehyde and 5% glutaraldehyde as described previously (Schmidt et al., 2010). These time points were chosen because they cover the majority of craniofacial development prior to fusion of the primary palate. Fixed embryos were immersed in Cysto ConRay II® (isothalamate meglumine) contrast agent for 1 hour then scanned on a μCT35 scanner to a 3.5 or 7.5 μm resolution. For the analysis of skeletons, neonates were fixed in 4% paraformaldehyde and scanned at 7.5 μm resolution. Landmarking was performed in MeshLab (http://meshlab.sourceforge.net). Landmarks represent Bookstein type 1 and 2 landmarks (Bookstein, 1997; Zelditch et al., 2012) and are similar to the lab-established landmarks (Parsons et al., 2011) with additional landmarks around the maxillary prominence. Additionally, about 20 previously generated E10.5 C57Bl/6J scans from Schmidt et al. (Schmidt et al., 2010) were re-landmarked and used as an additional control for background or strain effects. All analyses, except scaled variance of the eigenvectors and trace of the eigenvectors, were performed in MorphoJ 1.05b (Klingenberg, 2011). Procrustes superimposition was used to remove the effects of alignment, rotation or scale. All embryos were divided into groups based on tail somite numbers. Tail somites (Som) were counted from the tail tip to the base of the forelimb. 12-21 Som were classified as E9.25-9.75, 23-35 as E10.25-10.75, and 40-50 as E11.25-11.75. Data were regressed against tail somites using group centered regressions and only within each age-based group, because it is easier to regress out age if the age-related changes are smaller. All further analysis was performed on the residuals from the regression. PCA was used to view spread of data and identify major axes of variance among individuals. CVA was used to determine major axes describing among-group variation. Procrustes distance and Procrustes permutation tests were used to determine P-values for between-group differences. The two-block PLS was run using block 1 as the nasal prominence landmarks and block 2 as the maxillary prominence landmarks, and the compare vector test was used to examine differences between groups. Movies and mesh warp images were created using landmark IDAV (Wiley et al., 2005), which uses 3D splines to warp one 3D image to a new set of 3D coordinates. Screen shots were then taken of the new 3D surface at various angles. For movies, a screen shot was taken for every 2% difference and then images were loaded into iMovie (Apple) to create the animations. Scaled variance of the eigenvalues and trace of the eigenvectors were performed as reported previously using the statistical software R (http://www.r-project.org) (Hallgrímsson et al., 2009).

**Acknowledgements**

The authors thank Irene Choi and Cvett Trpkov for technical assistance on the project as well as Denise Liberton for advice on statistics and help with R.

**Competing interests**

The authors declare no competing financial interests.

**Author contributions**

W.F., H.L. and T.W. developed the Neo/Null mouse model. R.M.G. performed all morphometric analyses with the assistance of H.J. and B.H. J.L.F. and R.S.M. were involved in the interpretation of the Fgfl data. T.P. and J.H. assisted with the analysis of the expression time course data. Studies were conceived by R.A.S., R.M.G., B.H. and T.W. Manuscript was prepared by R.M.G. with assistance from B.H. and T.W. All authors were involved in editing and approving the version for submission.

**Funding**

This research was funded by National Institutes of Health grants DE012728 (T.W.), DE15191 (R.A.S.), DE019638 and DE021708 (R.S.M. and B.H.), and DE022214 (R.M.G.)-a grant from the American Cleft Palate Association (T.W.) and Natural Sciences and Engineering Research Council Grant #238992-12 (B.H.).

**Supplementary material**

Supplementary material available online at http://dmm.biologists.org/lookup/suppl; doi:10.1242/dmm.017616; jdc1

**References**


Teratology 14, 213-218.


Development 136, 1401-1407.


Teratology 14, 213-218.


