



Published in final edited form as:

Dev Biol. 2017 June 01; 426(1): 97–114. doi:10.1016/j.ydbio.2017.03.025.

Systems biology of facial development: contributions of ectoderm and mesenchyme

Joan E. Hooper^{1,2,*}, Weiguo Feng^{1,3}, Hong Li³, Sonia Leach⁴, Tzulip Phang^{2,5}, Charlotte Siska², Kenneth L. Jones⁶, Richard A. Spritz⁷, Lawrence E. Hunter^{2,8}, and Trevor Williams^{1,3}

¹Department of Cell and Developmental Biology, University of Colorado School of Medicine, 12801 E 17th Avenue, Aurora, CO 80045, USA

²Computational Bioscience Program, University of Colorado School of Medicine, 12801 E 17th Avenue, Aurora, CO 80045, USA

³Department of Craniofacial Biology, University of Colorado School of Dental Medicine, 12801 E 17th Avenue, Aurora, CO 80045, USA

⁴Department of Biomedical Research, National Jewish Health, 1400 Jackson Street, Denver, CO 80206, USA

⁵Department of Medicine, University of Colorado School of Medicine, 12801 E 17th Avenue, Aurora, CO 80045, USA

⁶Department of Pediatrics, University of Colorado School of Medicine, 12801 E 17th Avenue, Aurora, CO 80045, USA

⁷Human Medical Genetics and Genomics Program, University of Colorado School of Medicine, 12800 E 17th Avenue, Aurora, CO 80045, USA

⁸Department of Pharmacology, University of Colorado School of Medicine, 12801 E 17th Avenue, Aurora, CO 80045, USA

Abstract

The rapid increase in gene-centric biological knowledge coupled with analytic approaches for genomewide data integration provides an opportunity to develop systems-level understanding of facial development. Experimental analyses have demonstrated the importance of signaling between the surface ectoderm and the underlying mesenchyme in coordinating facial patterning. However, current transcriptome data from the developing vertebrate face is dominated by the mesenchymal component, and the contributions of the ectoderm are not easily identified. We have generated transcriptome datasets from critical periods of mouse face formation that enable gene expression to be analyzed with respect to time, prominence, and tissue layer. Notably, by

*Corresponding author. Joan Hooper. 12801 E 17th Avenue, rm 12103. University of Colorado School of Medicine. Aurora, CO 80045. Tel.: (303) 724-3417; FAX: (303) 724-3420; Joan.hooper@ucdenver.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

separating the ectoderm and mesenchyme we considerably improved the sensitivity compared to data obtained from whole prominences, with more genes detected over a wider dynamic range. From these data we generated a detailed description of ectoderm-specific developmental programs, including pan-ectodermal programs, prominence-specific programs and their temporal dynamics. The genes and pathways represented in these programs provide mechanistic insights into several aspects of ectodermal development. We also used these data to identify co-expression modules specific to facial development. We then used 14 co-expression modules enriched for genes involved in orofacial clefts to make specific mechanistic predictions about genes involved in tongue specification, in nasal process patterning and in jaw development. Our multidimensional gene expression dataset is a unique resource for systems analysis of the developing face; our co-expression modules are a resource for predicting functions of poorly annotated genes, or for predicting roles for genes that have yet to be studied in the context of facial development; and our analytic approaches provide a paradigm for analysis of other complex developmental programs.

Keywords

transcriptome; functional genomics; co-expression modules; epithelial-mesenchymal interaction; orofacial clefts

INTRODUCTION

Facial development requires finely choreographed growth and morphogenesis of bilaterally paired nasal, maxillary and mandibular prominences that converge to form the nose and the upper and lower jaws. In the mouse (reviewed in Depew et al., 2002), the discrete prominences arise by embryonic day (E) 10, originating as mesenchymal bulges encased in an overlying layer of ectoderm and surrounding the primitive oral cavity. By E13 these separate structures have fused to form an integrated unit (summarized in Figure S1). Facial mesenchyme is derived from both the mesoderm and the neural crest cell populations and eventually forms the bone, cartilage, connective tissues and muscles of the face. The cranial ectoderm gives rise to the epidermis as well as the lining of the oral and nasal cavities. Moreover, via placodal intermediates (reviewed in Singh and Groves, 2016), the ectoderm provides critical components of sensory organs, exocrine glands, and teeth.

Although mesenchymal tissues make up the bulk of the embryonic facial prominences, crucial patterning information is relayed to these cells by surrounding tissues including the ectoderm, endoderm and neural tube (Adameyko and Fried, 2016; Chai and Maxson, 2006; Singh and Groves, 2016). The facial ectoderm provides both permissive and instructive signals that are required for normal development of the underlying mesenchyme. Thus, manipulation of the embryonic facial ectoderm by either genetic or surgical means results in major developmental defects, including orofacial clefts (OFCs). The ectoderm also receives critical signaling input from mesenchyme to regulate its growth, to maintain its competence to drive face formation and to diversify its derivatives. Human clinical analyses and animal model studies have shown that normal face formation requires the integration of multiple signals between the ectoderm and mesenchyme (reviewed in Van Otterloo et al., 2016; Yuan et al., 2016). Some of these – including Fibroblast Growth Factors (Fgfs), Bone

Morphogenetic proteins (BMPs), Wnts, Hedgehogs (Hhs), Platelet Derived Growth Factors (PDGFs), Retinoic Acid (RA), and endothelin – are well known, but others remain to be identified. This intricate signaling crosstalk coordinates the convergent growth and morphogenesis of the facial prominences that is essential for aligning these structures prior to fusion (e.g. Geetha-Loganathan et al., 2014; Green et al., 2015; Hu et al., 2015; Linde-Medina et al., 2016; Suzuki et al., 2016). Subsequent interactions between the apposed epithelia then consummate lip and palate fusion (reviewed in Kousa and Schutte, 2016). The complex interplay of signaling, growth, morphogenesis and fusion that occur during face formation, presumably accounts for the susceptibility of this process to genetic and environmental insults, reflected by the prevalence of human craniofacial defects (reviewed in Dixon et al., 2011; Twigg and Wilkie, 2015).

Despite the importance of the ectoderm in facial morphogenesis, surprisingly little is known about the genetic programs that define the facial ectoderm and its derivatives, how those programs are influenced by mesenchymal or environmental cues or how the signals from the ectoderm influence the genetic programs of the adjacent mesenchyme. To understand the ectodermal genetic programs and how they coordinate and integrate with the mesenchymal programs to drive facial development, it is important to obtain a comprehensive understanding of the cellular and molecular changes that occur in each of these tissues during the early stages of facial development. Previous transcriptome studies of human, mouse or chick facial development have provided considerable information concerning gene expression in discrete tissues or regions at specific developmental stages or have generated gene expression profiles of whole prominences over a defined time course (Bhattacharjee et al., 2007; Brinkley et al., 2016; Brugmann et al., 2010; Brunskill et al., 2014; Buchtová et al., 2010; Cai et al., 2005; Ding et al., 2016; Feng et al., 2009; Garaffo et al., 2013; Han et al., 2014; Iwata et al., 2012; Mima et al., 2013; Musselmann et al., 2011; O’Connell et al., 2012; A. S. Potter and S. S. Potter, 2015; Warner et al., 2014). However none of these datasets have sufficient temporal breadth and spatial resolution to identify the contributions of the ectoderm to the genetic and signaling programs that are crucial for facial morphogenesis. To address this gap, we have generated a transcriptome resource from separated ectoderm and mesenchyme of the developing mouse face at 24 h intervals between E10.5 and E12.5 – a period that is critical for establishing overall facial shape as well as for the fusion of the lip and primary palate (Fig S1). This has allowed us to identify an ectodermal program that is radically different from the mesenchymal program that has dominated previous studies. In addition we have leveraged the spatial, temporal and tissue-specific dimensions of the data, using co-expression to generate functional modules specific to facial patterning. These modules can be used to predict functions of individual genes, to generate hypotheses about the developmental processes represented by the modules and to investigate how module member genes may act together to effect those processes.

Materials and Methods

Sample preparation, data capture and pre-processing

All animal experiments used inbred C57BL/6J mice (Jackson Labs) and were performed in accordance with protocols approved by the University of Colorado Denver (UCD) Animal

Care and Usage Committee. Animal husbandry, embryo staging by a combination of embryonic day and morphological criteria (Theiler staging), as well as sample preparation, was as described in Feng et al. (2009). Details of the dissections, morphological landmarks and tissue separation are further described in Li and Williams (2013) and its accompanying video presentation. Specifically, ectoderm and mesenchyme from dissected facial prominences were separated by ‘peeling’ following Dispase II digestion for 15 min at 37°C (Roche, cat no. 04 942 078 001). Figure S1 illustrates the tissues taken at each timepoint with the boundary of the maxillary and nasal prominences at the naso-lacrimal groove. The nasal prominences were processed intact, rather than being separated into lateral and medial nasal components to assist with direct comparison to data in Feng et al. (2009). Due to the intimate interdigitation of the olfactory epithelium and the mesenchyme, the nasal prominence mesenchyme samples included the olfactory epithelium. Finally, at E12.5, protruding tongue tissue was trimmed from the mandible samples to avoid excessive muscle gene expression signatures.

Dissected tissues from multiple embryos were pooled to generate at least 5 ug total RNA from each sample; this required an average of 50 embryos at E10.5, 18 embryos at E11.5 and 8 embryos at E12.5. Total RNA was purified using Trizol (Invitrogen, Carlsbad, CA) and the RNeasy MiniKit (Qiagen, Germantown, MD). 2–5 ug of total RNA was random primed to generate cDNA, from which biotin-labeled cRNA was derived and used to probe Affymetrix MoGene-1.0-st-v1 microarrays (Affymetrix, Santa Clara, CA). We chose this platform because it has been previously used for studies of the earliest stages of mouse face development (Brunskill et al., 2014, see Fig S1). Probe preparation and microarray analyses were carried out by the UCD Gene Expression Core Facility using standard procedures recommended by the manufacturer (Affymetrix). Three biologically independent replicates for each condition allowed for statistical analysis. Raw data were processed with Affymetrix Power Tools (apt-probeset-summarize, v 1.91, Robust Multi-array Average; RMA) to normalize and summarize probeset expression levels, and for DABG (detected above background at $p > 0.05$) analysis. MGI Gene symbols were mapped onto the probesets using MoGene-1_0-st-v1 Transcript Cluster Annotations, Release 34 (4/7/14). Statistics of microarray data are summarized in Table S1 and Fig S2. Gene expression data from this study, both as .cel files and as a probeset-by-sample expression spreadsheet, are available via GEO (GSE62214) and FaceBase (FB00000803) respectively. Plots of expression profiles for each gene across age, prominence and tissue, as well as for the earlier whole prominence data (Feng et al., 2009), are available in an accompanying Data in Brief (Leach et al., submitted).

Statistical analyses

All data and statistical analyses were performed using custom scripts in R, available upon request. After filtering out probesets that were not detected above background in at least one sample, and whose variance was less than the median of all probesets (median filtering), 9457 differentially expressed probesets were identified among the remaining 12072 probesets by 3-way analysis of variance (ANOVA), two-sample t-test for significance at $p = 0.01$ and correction for multiple testing ($p_{adj} < 0.01$; Benjamini and Hochberg, 1995). Hierarchical clustering used hclust (distance = correlation, cluster method = average, scaled

data) and was displayed with heatmap.2 (gplots package). The validity of the tree structure in Figure 1D was tested by bootstrapping using the Approximately Unbiased method ($n = 10,000$; Shimodaira, 2002). Biological replicates clustered together at better than 90% probability, with the exception of the Mandible mesenchyme samples: E10.5 ('bpm' in Fig 1D; >82%), E11.5 ('dpm'; >85%), and E12.5 ('fpm'; >72%). The colored blocks above the heatmap in Figure 1D indicate other groupings with probability greater than 99%. The clusters in Figure 4 were determined by tree-cutting ($n = 13$) and cluster V was further split to separate the Fn-enriched genes from the Md-enriched genes (Va and Vb).

Gene Ontology Biological Process (GO-BP) term enrichment used the Database for Annotation, Visualization and Integrated Discovery (DAVID6.7) functional annotation and clustering tools (D. W. Huang et al., 2009a; 2009b) with the RDAVIDWebService interface (Fresno and Fernández, 2013). GO-BP terms for Tables 1–3 were re-ranked for 'interestingness' using the sum of their p-value rank and number of genes associated with that term (Pop.Hits), a formula that performed better than a variety of alternatives. Overlapping/redundant terms were identified by DAVID functional annotation clusters and eliminated by manual selection. Reanalysis of functional annotations for Tables 1–3 with updated GO terms (DAVID6.8) identified about twice as many enriched terms, but the new terms do not have appreciable impact on the overall interpretation of the gene programs enriched in each age, layer or prominence.

Cross-platform comparisons with previous studies

Whole prominence expression values from .CEL files from Affymetrix mouse A430 arrays (Feng et al., 2009) were re-calculated using MAS5 normalization (Affymetrix GeneChip Operating Software) and MGI gene symbols extracted from Affymetrix transcript annotation files. Those data were aligned with data from this study by matching gene symbols, with 17,079 genes matched across platforms. For these cross-platform comparisons, where multiple probesets represented one gene, the mean of the probeset expression values was used.

Co-expression modules and their analysis

Co-expression modules were generated from 9337 differentially expressed probesets (from 1-way ANOVA after median filtering; $p_{adj} < 0.01$) corresponding to 8507 genes. The Weighted Gene Co-Expression Network Analysis (WGCNA) R package was used according to guidelines (Langfelder and Horvath, 2012) with a power of 8 for a signed hybrid-type adjacency matrix, a minimum module size of 5 probesets, eigengene correlation > 90% for module merging, and otherwise using default parameters. The signed hybrid adjacency matrix was chosen to optimize for genes that are up-regulated together. Other parameters were selected to maximize enrichment of GO-BP terms in the resulting modules. The resulting 79 modules were surveyed for chromosomally clustered genes that might not be transcriptionally independent, and four modules were eliminated for this reason. The result was 75 co-expression modules with membership ranging from 3 to 1285 genes (Table S10).

Functional enrichment analysis of the modules (Table S11, Figure S9) applied the Cluster Profiler R package (G. Yu et al., 2012) to our custom-curated list orofacial cleft genes and

29 signaling pathway genes (Table S12), GO-BP terms, GO-CC terms and non-coding RNAs (Entrez annotations). 'Background' for calculating enrichment of non-coding genes was the universe of genes detected in at least one sample; 'background' for signaling and orofacial cleft genes was the 18937 protein-coding genes that were also detected in at least one sample. Figure 8, illustrating GO term enrichment, was generated using the Cluster Profiler Plot function, after culling redundant terms from the term enrichment matrix. Functional networks were generated using the GeneMANIA server, incorporating protein-protein interactions, phenotype similarity, co-expression across GEO datasets, tissue co-localization and protein interactions predicted by cross-species homology (Warde-Farley et al., 2010; Zuberi et al., 2013).

Manual annotation and curation

Manual annotations in Table S9 focused on biological functions, expression patterns and mutant phenotypes relevant to the growth, morphogenesis, development and/or to the face. The broad categories were 'cell morphology and migration', 'transcription', 'roles in cancer' (including proliferation apoptosis and motility), 'expression' (where/when and subcellular location), 'mutant phenotype' and 'relation to orofacial clefting'. The annotations for the mouse genes and their human homologs were drawn from EntrezGene RIFs, PubMed, MGI gene expression and phenotype databases, and <http://locate.imb.uq.edu.au/cgi-bin/search.cgi> for subcellular localization.

A list of mouse orofacial clefting genes was compiled by searching MGI for the phenotypes 'cleft upper lip, palate, chin, bifid chin', then filtering the 307 entries for the 228 genes with Entrez ids. A list of human orofacial clefting genes was compiled from 379 genes returned from OMIM by searching for 'cleft' and from 124 genes compiled from the literature. The human genes were mapped onto their mouse homologs using the MGI human-mouse homology table and added to the mouse gene list, for a total of 569 mouse orofacial clefting genes. They are listed in Table S12 under the pathway 'clefting'.

Signaling pathway genes were compiled by pathway from existing databases and the literature as follows. Human canonical pathways and gene lists from MSigDB (<http://software.broadinstitute.org/gsea/msigdb/collections.jsp>; Liberzon et al., 2015) including Kegg, (c2.cp.kegg.v5.1.symbols.gmt) and Reactome (c2.cp.reactome.v5.1.symbols.gmt) and from NCI-PID (<http://public.ndexbio.org/>) were sorted into the 28 signaling pathways (or pathway groups) shown in Table S12.

For each pathway, pathway-specific genes were included while genes common to multiple pathways were eliminated. Thus most of the signaling pathway genes included in this compilation encode ligands, receptors or their modulators, while genes encoding transduction apparatus intermediates that are common to multiple pathways (e.g. Akt, Mapk, Ras, Rho), were eliminated. These classifications were supported by the literature, by pathway maps and compilations from 'cellsignal.com', 'lsresearch.com' and wikipathways. Additional pathway-specific components were added by literature searches, for a total of 1715 pathway-specific signaling genes (Table S12).

In situ hybridization

Whole mount in situ hybridization was performed on E10.5 embryos as described in (Feng et al., 2009). Hybridized probe was detected using an anti-digoxigenin antibody (Roche) and signal was developed in BM Purple (Roche). For sectioned material, fixed and dehydrated E10.5 and E12.5 embryos were rehydrated, embedded in OCT (Sakura Finetek, Torrance, CA), frozen and sectioned at 12 μ m. Sections affixed to slides were hybridized and developed as above, except that probe was diluted 50-fold, incubation and wash times were shortened, and slides were counterstained with Nuclear Fast Red (Gertsenstein et al., 2003). Details of probes are available on request.

RESULTS AND DISCUSSION

A transcriptome for the developing mouse face - dynamics and robustness

To elucidate genetic programs underlying facial development, we profiled the transcriptome of ectoderm and mesenchyme derived from the Mandibular, Maxillary, and nasal prominences (abbreviated MdP, MxP and FnP, respectively). Samples were derived from ages when the prominences first become distinct (E10.5/TS17), during their outgrowth and morphogenesis (E11.5/TS19) and at the end of fusion of the lip and primary palate (E12.5/TS20.3; see Fig 1A and Fig S1). 24 h intervals were chosen as optimal for the timecourse, based on a prior study of whole prominences at 12 h intervals (Feng et al., 2009). To minimize variability due to differences between individuals, each sample was comprised of pooled tissue from multiple embryos. The small size of the early MxP and FnP made it impractical to isolate sufficient E10.5 MxP or FnP ectoderm, or E10.5 MxP mesenchyme, so only 15 of the 18 possible conditions could be analyzed. Affymetrix mouse Gene 1.0 ST arrays were probed with random-primed total RNA. Three independent samples were analyzed for each condition, and these biological replicates showed excellent reproducibility (Figs 1B, S2). Table S1 presents an overview of the gene expression data. Of the 25779 genes represented on the microarray, we detected expression of 98% of these in at least one sample, and 64% in all 45 samples. 34% of detected genes (8575) were differentially expressed (multiple testing-corrected p-value < 0.01; Fig 1C). Of these 8575 differentially expressed genes, most differed between ectoderm and mesenchyme (7722; 90%), many differed by age (6048; 71%) and some differed by prominence (3635; 42%). Most of these differed by at least 1.5-fold (8239) and many by at least 2-fold (4267).

The largest differences in this dataset were between ectoderm and mesenchyme (Fig 1; Table S1). This is most clearly demonstrated by Principal Component Analysis, where the ectoderm and mesenchyme samples separated along the first principal component, and accounted for 45% of the variability across all samples (Fig 1B). The second principal component reflected embryonic age and accounted for 12% of the variability. Prominence identity did not map to any of the first four principal components. Hierarchical clustering (Fig 1D) confirmed that ectoderm-mesenchyme differences dominate in the developing face. Ectoderm samples clustered by age, while mesenchyme samples clustered by prominence. Within the mesenchyme, the FnP program was quite distinct, while the MxP/MdP programs were relatively similar to each other. Note that biological replicates clustered together (bootstrap probability > 95%), indicating that the dataset is statistically robust and biological

reproducible as well as demonstrating the distinct properties of each prominence and the clear changes associated with age.

A transcriptome for the developing mouse face - validation

We used four metrics to validate this dataset. First, to assess cross-contamination between ectoderm and mesenchyme, we used the fold-enrichment of known ectoderm- or mesenchyme-specific genes in the respective tissue layers. The ectodermal genes *Esrp1*, *Krt14* and *Perp* respectively showed 32, 26 and 23-fold average enrichment in MxP/MdP ectoderm or 4, 11 and 11-fold enrichment in FnP ectoderm (Table S2). The mesenchymal genes *Aldh1a2*, *Cdh5* and *Ebf1* respectively showed 12, 11 and 10-fold average enrichment in MxP/MdP mesenchyme and 9, 7 and 7-fold enrichment in FnP mesenchyme (Table S3). Therefore, we estimate the upper limit of contamination of MxP/MdP mesenchyme by ectoderm at less than 5% and of MxP/MdP ectoderm by mesenchyme at less than 10%. The cross-contamination in FnP was substantially higher: 9–25% of mesenchyme signal might be due to ectodermal contamination and 11–14% of ectodermal signal might be due to mesenchymal contamination. This reflects an ambiguous boundary between ectoderm and mesenchyme in the FnP, caused by invagination of the nasal pit and intermingling of ectodermal and neural crest cells in the ‘migratory mass’ that pioneers the olfactory nerve (e.g. Forni et al., 2011; Miller et al., 2010).

Second, to assess temporal and prominence-specificity and possible deleterious effects of the dispase treatment used for tissue separation, we compared published data for known prominence- and age-specific genes (Feng et al., 2009) with the expression profiles obtained in the current study. For instance, Feng et al. (2009) found 19 collagen-annotated genes that increased in expression from E10.5 to E12.5 (Fig S4 of Feng et al., (2009)). Similar expression profiles were observed in this study (Fig S3A), including shared subtleties such as lower expression of *Col8a2*, *Col9a2* and *Col9a3* in E12.5 MxP. In addition, we found that *Col17a1*, a component of epithelial hemidesmosomes (McGrath et al., 1995), is specifically expressed in the ectoderm, that five collagen genes (*Col5a1*, *Col7a1*, *Col12a1*, *Col8a1*, *Col14a1*) were expressed in both ectoderm and mesenchyme, and that the remaining genes were mesenchyme-specific. Expression profiles for MdP-specific genes (Fig S7A of Feng et al., (2009)) were also very similar between studies (Figure S3B), with most of these genes showing mesenchyme-specificity. Four genes, including *Dlx3* and *Foxa2*, were more highly expressed in the ectoderm or epithelial derivatives, consistent with previous findings (Magdaleno et al., 2006; Monaghan et al., 1993).

Third, we queried the mouse Gene Expression Database (Smith et al., 2014; Finger et al., 2017) and the literature for expression patterns corresponding to the top 10 genes (by fold-change) predicted to be enriched in ectoderm, mesenchyme, FnP, MxP or MdP (Tables S2–4). All 40 of the genes with relevant expression data were consistent with our microarray analysis (Table S6), providing confidence in the validity of this study.

Finally, we performed *in situ* hybridization to whole mount embryos at E10.5 and on sectioned tissue at E10.5 and E12.5 to validate the spatiotemporal gene expression profiles obtained from the microarray analysis (Figure 2, S4). We selected genes with limited published expression data that displayed profiles enriched in all ectoderm (*Trim29*, *Esrp1*,

and *Mpzl2*), E12.5 ectoderm (*Gabrp*), or mesenchyme (*Aldh1a2* and *AW551984*) – as illustrated by the plots shown in Figure 2 (left panels). Transcripts representing *Gabrp*, *Trim29*, *Esrp1*, and *Mpzl2* were detected only in the ectoderm and its derivatives, with distinct temporal and spatial distributions that mirrored their expression profiles. Thus, expression of *Trim29*, *Esrp1*, and *Mpzl2* was readily detected in the surface ectoderm of the facial prominences at both E10.5 and E12.5. In contrast, *Gabrp* showed limited, patchy expression in the surface ectoderm at E10.5, but robust expression throughout this tissue at E12.5. Further, *Esrp1* and *Mpzl2* transcripts were also detected in the nasal pit at E10.5 and the invaginating olfactory epithelium at E12.5. The fractionation of this convoluted ectoderm-derived layer with the nasal mesenchyme presumably accounts for the moderate levels of these two genes in the FnP mesenchyme expression profiles (Fig 2, left panel, red lines).

Similarly, transcripts corresponding to the two genes with mesenchymal expression profiles were detected in situ at E10.5 and E12.5 in patterns consistent with a subset of mesenchymal tissues, but were not apparent in the surface ectoderm (Figure 2). Additional in situ hybridizations for genes enriched in either ectoderm (*Cldn6*, *Fermt1*, *Lass3*, *Tpd52*) or mesenchyme (*Aff2*) confirmed the spatiotemporal dynamics of their microarray-derived expression profiles (Fig S4). Thus, in situ hybridization validated the gene expression distribution and intensity measured by microarray. All together these validations demonstrate that our gene expression dataset accurately describes the transcriptional profiles of individual genes during facial development with respect to time, prominence and tissue layer.

A transcriptome for the developing mouse face - effects of study design on data structure and quality

We next examined how this dataset overlaps, complements and extends other transcriptome data derived from the developing mouse face (Fig S1). In temporal terms, this study bridges the gap between an early gene expression atlas (up to E10.5) based on laser-capture samples (Brunskill et al., 2014) and several studies of later structures undergoing terminal differentiation including teeth, tongue, salivary glands and olfactory epithelium (Garaffo et al., 2013; Musselmann et al., 2011; O’Connell et al., 2012). In terms of resolution, this study delineates the ectoderm and mesenchyme components that are not resolved in whole prominence studies (Feng et al., 2009), while retaining the global profile of prominence expression that is not possible with the restricted sampling of laser capture (Brunskill et al., 2014). In technical terms, the tissue pooling strategy facilitates measuring gene expression in very small tissue samples while avoiding the quantification artifacts introduced by amplification steps (Russell et al., 2008). In practical terms, the present study identified many more differentially expressed genes than other studies of similar tissues at similar ages.

To better understand the power of this study, we compared our data to those from two similar studies that employed different sampling strategies to examine genome wide expression in the developing mouse face: Feng et al. (2009) used whole facial prominences; Brunskill (2014) used laser-capture of small groups of cells; whereas this study used micro-

dissection to separate ectoderm from mesenchyme. We first compared our tissue-separated data to the whole-prominence data by reanalyzing the whole prominence data using the same statistical parameters of the present study, and then matching 17084 genes across the two platforms (Fig S5). We found many fewer differentially expressed genes in whole prominence-derived tissue than in the separated ectoderm and mesenchyme (2076 vs. 3863 at $p_{\text{adj}} < 0.01$, fold change > 2 ; Figure S5B). This difference was not due to statistical considerations, as similar numbers of differential expressed genes were detected on purely statistical grounds in both datasets (7863 for whole prominences, 7502 in this study at $p_{\text{adj}} < 0.01$; Fig S5A). Nor was it due to overall data distribution, as the two studies were normalized to the same mean and quartile distributions. The increased sensitivity was largely due to the fold changes, which were considerably larger in the current study (Fig S5C; median 1.59 for whole prominences, 2.0 for this study).

We next asked whether the differences in fold-change between the two studies were driven by ectoderm- or mesenchyme-specific genes. To do so we compared whole-prominence expression values for 17084 genes common to the two platforms to either ectoderm or mesenchyme expression values, matching genes, ages and prominences. We found that ectoderm-enriched genes were consistently underestimated in whole prominence samples (blue vs. grey in Fig S5 D; $p < 0.001$; mean of 2.5-fold expression level), whereas mesenchyme-enriched genes distributed similarly to the overall population in both rank-order and expression values (red vs. grey in Fig S5 G; $p > 0.01$). This under-valuing of ectoderm but not mesenchyme genes in whole prominence data is consistent with ectoderm constituting the minority of tissue in whole prominences. We then asked whether the ectoderm and mesenchyme-enriched genes were over represented in the 2282 differentially expressed genes found only in the present study (Fig S5B). As shown in Fig S5E, almost half of these 2282 new genes (1027/2282) were differentially expressed by age or prominence, rather than by tissue layer. That is, the ectoderm- and mesenchyme-enriched genes (718 and 537 respectively) only accounted for about half of the increased sensitivity of this study. Apparently, sampling more homogeneous cell populations (ectoderm or mesenchyme) uncovers differences along additional dimensions (age and prominence) that are masked by sampling mixed populations.

Finally, we investigated the costs and benefits of sub-sampling the tissues within the facial prominences (laser-captured data from Brunskill et al., 2014). In a direct comparison of comparable samples (E10.5 mandibular ectoderm), we found considerably more sample-to-sample variability in the laser-capture samples than in the microdissected samples (Fig S5F), which reduces the statistical power to detect differentially expressed genes. For instance, laser-capture found 24 probesets differentially expressed between at E10.5 between FnP and 1st branchial arch mesenchyme, whereas 233 were identified from manually dissected tissue in this study. In contrast, the laser-capture data showed larger fold-differences; in the same comparison, 44 genes had at least 4-fold change in laser capture data but only 2 genes had this fold change in manual dissection data. In sum, the benefit of smaller cell populations is larger fold-changes, but the cost is more variability. The few differentially expressed genes that are detected have large expression differences and are likely to identify compartment-specific signatures. However, most differentially expressed genes go undetected.

By comparing these three studies we conclude that a balance between tissue resolution and reproducibility accounts for the excellent power of this study to detect differentially expressed genes. In addition, we can make general suggestions for optimal study designs in developing systems. First, it is crucial to separate tissues according to the principal difference(s) in genetic programs - in this case, ectoderm and mesenchyme lineages. This increases sensitivity by increasing measured fold changes, both for genes that differ along the first principal component and for genes that differ along other dimensions. The net effect is that more differentially expressed genes are detected with fewer biological replicates. Second, fold change cutoffs should take into consideration the cellular heterogeneity within the samples; with more heterogeneous samples, smaller fold changes can still be biologically significant. Third, spatial resolution often comes at the cost of reproducibility and hence statistical power to detect true positives. Therefore spatially high-resolution data from laser capture or single-cell sequencing might be most powerful when coupled with statistically high-resolution data such as this study, so that sub-populations and sub-programs are described within the context of the broader genetic program.

Functional annotations outline the facial developmental program

The overall motivation for this study was to delineate a facial development program- to understand its major sub-programs and their mechanistic underpinnings. As a first step, we used the genes that were expressed at least 2-fold higher in specific tissues, ages or prominences, and identified Gene Ontology Biological Process terms (GO-BP; (Ashburner et al., 2000) that were enriched in these genesets.

We first used ectoderm- or mesenchyme-enriched genes (Table S1, complete genelists in Tables S2 and S3) to characterize the gene expression programs active in these two tissue layers. Common developmental terms such as ‘cell adhesion’, ‘embryonic morphogenesis’, ‘cell fate commitment’, and ‘positive regulation of transcription’ dominated in both tissues, providing little insight into the relevant biological processes occurring in the ectoderm and mesenchyme (data not shown). To highlight the significant differences between conditions, we adjusted the term rankings so that the more specific terms, those attached to fewer genes, were favored in the rankings (see Materials and Methods for details). The resulting top-ranked terms highlight processes that differ between ectoderm and mesenchyme (Table 1). Terms enriched in the ectoderm geneset are associated with the development of the epidermis and its appendages, whereas those enriched in the mesenchyme geneset emphasize development of muscles, skeletal elements, nerves and blood vessels. Most notably, various signaling process terms were enriched in both genesets (Table 1, highlighted in purple). Both tissues shared ‘retinoic acid metabolic processes’; the ectoderm highlighted planar polarity, non-canonical Wnt signaling, FGF and BMP signaling - while VEGF, receptor tyrosine kinase, integrin-mediated signaling and regulation of Wnt signaling were prominent in the mesenchyme. This suggests that both shared and distinct signaling processes participate in patterning these tissues. Altogether, the differences in gene expression programs between ectoderm and mesenchyme reflect the very different tissues that they will generate, with extensive signatures for signaling both within (e.g. “planar polarity”) and between (e.g. “mesenchymal-epithelial”) tissue layers.

Similar analyses of the 1574 genes that vary in expression by age (Table 2, complete genelist in Table S4) found term numbers increasing with age, which parallels the increasing complexity that accompanies facial morphogenesis (Fig S1 and references therein) and is consistent with previous reports (Feng et al., 2009). The 342 genes that differed in expression at E10.5 were enriched in only 42 biological process terms. Various anabolic processes were highlighted, as well as ‘Vascular Endothelial Growth Factor signaling’; the latter process is consistent with the extensive vascular remodeling which occurs at this time (Hiruma et al., 2002). The 1122 genes that differed in expression at E12.5 relative to prominence and layer-matched samples encompassed 174 biological process terms, with top-ranked terms highlighting development, differentiation and morphogenesis - of bone, cartilage, muscle, axons, blood vessels, glands, epidermis, etc. Altogether, the temporal dynamics reflects the shift from early anabolism (growth) to later differentiation.

The fewest genes differed in expression by prominence (Table, 3, complete genelist in Table S5) consistent with previous reports (Feng et al., 2009). Of the 1026 genes that differed in expression by prominence relative to age and layer-matched samples, almost half (556) involved the FnP, and many (447) were unique to this region. Those FnP-specific genes were enriched for processes that highlight neurosensory development and epithelial differentiation, which is consistent with the extensive neurogenesis in the olfactory epithelium of the FnP at this age (Cuschieri and Bannister, 1975). 268 genes had expression specific to the MdP, and 114 to the MxP, with another 88 shared between just MxP and MdP. Terms enriched in MxP genes highlight signaling (‘Wnt’, ‘Retinoic Acid’, ‘Notch’) and synaptic function. The latter may reflect the extensive synaptogenesis in the MxP involving trigeminal nerve fibers (Stainier and Gilbert, 1990). MdP terms highlight muscle and skeletal development, consistent with the muscular character of the tongue (Mayo et al., 1992), the formation of Meckel’s cartilage, and dentary bone (Frommer and Margolies, 1971). Thus, the differences in gene expression programs between the facial prominences reflect their unique derivatives: ‘nose development’ and olfactory epithelium for the FnP; ‘palate development’ for the MxP; muscle (tongue) for the MdP; ‘odontogenesis’ and ‘skeletal system morphogenesis’ shared between MxP and MdP. Overall, the findings demonstrate that the datasets we have obtained reflect the known processes occurring during facial growth and fusion, and provide a strong rationale for further mining of specific gene expression programs to reveal the intricacies of the underlying molecular interactions.

Distinct ectoderm and mesenchyme programs in the developing face

As noted above, expression differences between tissue layers predominated in the dataset, consistent with the origin of the ectoderm and mesenchyme from distinct developmental lineages. Thus we re-defined the facial development program as an ectoderm program and a mesenchyme program (Fig 3) using the 1217 genes that were at least 2-fold enriched in ectoderm, relative to age- and prominence-matched mesenchyme, and the 1459 that were enriched in mesenchyme by the same criteria (Tables S1–S3). There was very little overlap between ectoderm- and mesenchyme-enriched genes; only 43 were enriched in ectoderm under some conditions and in mesenchyme under other conditions (Fig 3A). This lends strong support for structuring the facial development program around parallel ectoderm and mesenchyme programs.

To analyze these programs, we began by assessing which ectoderm- or mesenchyme-enriched genes were constitutively expressed in that tissue across all conditions, and which varied in expression by age and/or prominence by at least 2-fold (Fig 3). We found 858 pan-ectodermal genes and 628 pan-mesenchymal genes, and 143 ectodermal genes and 504 mesenchymal genes that varied in expression by prominence (detailed in Table S7). We also found 315 ectodermal genes and 691 mesenchymal genes that varied in expression by age. Note that many of these varied in expression by both age and prominence (Fig 3A). In both ectoderm and mesenchyme, temporal trajectories were dominated by genes whose expression increased with age (Fig 3C). This is consistent with the increasing tissue complexity (e.g. Fig S1). The most remarkable contrast in differential expression characteristics between the two tissue layers was the much larger proportion of genes varying by age and/or prominence within the mesenchyme versus the ectoderm (57% versus 29%; Fig 3A). This might indicate that mesenchyme is more heterogeneous than ectoderm, which is consistent with the multitude of differentiated types that derive from facial mesenchyme. Alternatively it might indicate that the heterogeneity within mesenchyme is better aligned with our sampling strategy (prominences), whereas ectodermal heterogeneity is at a smaller scale. The second notable difference was that expression of most ectodermal prominence-varying genes was specific to a single prominence (116/143 or 81%), whereas many of the mesenchymal prominence-varying genes were shared between the Mdp and MxP (Fig 3B).

The facial ectoderm developmental program

Though the facial mesenchyme gene expression programs have been studied previously (e.g. Feng et al., 2009; Tipney et al., 2009), the ectodermal program described here is novel and largely unexplored. To delineate the ectoderm-wide aspects of the facial developmental program we analyzed GO-BP terms that were enriched in the 858 pan-ectodermal genes (Table S8). Those terms highlight epithelial morphogenesis, cell adhesion, cell-cell signaling and neural development, which is similar to the picture from all ectoderm genes (compare to Table 1). One new feature was the term “sulfur metabolic process which was associated with genes involved in sulfated proteoglycan metabolism and glutathione-mediated catabolism. Of particular interest is an orofacial gene regulatory network based upon *Irf6* that is involved in epithelial fusion (Kousa and Schutte, 2016). Most of the genes in this network (*Irf6*, *Grhl2*, *Grhl3*, *Tfap2a*, *Tfap2b*, *Tfap2c*, *Tp63*, *Ovol1*, *Ripk4*, *Jag2*) belong to our pan-ectodermal gene class. Taken together, these studies define a pan-ectodermal genetic program specifying the maturing epithelia (oral, epidermal, dental, etc.) and associated basement membranes, with extensive signaling featuring protein kinase cascades, planar polarity and Wnt signaling. Moreover, specific sub-programs within this larger grouping may regulate important aspects of epithelial fusion that cause orofacial clefting when disrupted.

To investigate the more specialized aspects of the ectoderm developmental program, we clustered the 359 dynamically expressed ectoderm-enriched genes from Fig 3A by their spatio-temporal expression within the ectoderm and then used functional annotations to understand the developmental processes encoded/regulated by each cluster (Fig 4, Table S9). Many of the 12 clusters were not enriched for any GO-BP terms, so we manually annotated

each of the 359 dynamic ectodermal genes for biological functions and tissue specificity relevant to the developing face (Table S9). These included cell type/fate (e.g. epithelial, oral, dental, follicle, salivary gland, neural); stem/progenitor cell behaviors (proliferation, differentiation); contributions to tissue morphology (cell shape, motility and adhesion); differentiated phenotypes and signaling. These richer annotations then allowed us to interpret the more dynamic aspects of the facial ectoderm developmental program.

The early facial ectoderm developmental program: E10.5-E11.5

The early program was defined by clusters of genes with a downward trajectory over time across all prominences; that is, with maximal expression at E10.5 in the MdP or at E11.5 for FnP and MxP as these latter two processes lacked E10.5 data (clusters I, III and Vb in Fig 4). Genes involved in signaling, cell cycle and cell shape dominated this early program (Fig 4B). To visualize this program, we mapped the corresponding gene products onto a model of an ectodermal cell, using color-coded functional categories (Fig 5). Signals known to regulate craniofacial patterning included Shh, Fgf8 and Bmp4, and associated regulators such as the Wnt activator Cpz (L. Wang et al., 2009) and the Nodal/Bmp modulator Nptx1 (Boles et al., 2014). Signals novel to early facial development included the Ret/Gfra ligand Artn and the Robo3 ligand Nell2, both of which may be involved in facial innervation (Jaworski et al., 2015). Th and Ddc in the catecholamine pathway were strongly expressed at E10.5. Catecholamines (CA) are synthesized by keratinocytes in response to wounding and CA signaling modulates wound healing (Schallreuter et al., 1995; Sivamani et al., 2014), but neither has previously been implicated in epidermal development. Two components of the Hippo signaling pathway, Amot and Vgll2, were also included in this early program. The Hippo signaling pathway transduces mechanical cues into changes in gene expression to control organ size or tissue homeostasis (e.g. (F.-X. Yu et al., 2015)). Amot is part of the tension-sensing junctional complexes, while Vgll2 is a transcriptional cofactor (Maeda et al., 2002). In the facial ectoderm the Hippo pathway promotes proliferation in maturing skin and oral epithelia (Liu et al., 2015) and signaling that is necessary for survival of the underlying mesenchyme (Johnson et al., 2011). This raises the possibility that tension within the ectoderm, sensed by the Hippo pathway, might be important for coordinating growth of the mesenchyme with that of the overlying ectoderm.

Cell morphology and cytoskeletal dynamics was the second major theme in the early program (orange and red in Figure 5), with multiple genes encoding proteins associated with the ECM and junctional complexes. Proliferation and its supporting metabolism (yellow and dark grey in Figure 5) was represented by the cell-cycle regulators Cdkn1a and Usp44. Genes involved in transcriptional regulation (laurel-green in Figure 5) included the chromatin remodeling factor Chd7 which is mutated in human CHARGE syndrome - a birth defect associated with craniofacial patterning abnormalities (Basson and van Ravenswaaij-Arts, 2015).

Prominence-specific aspects of the early program, which were not distinguished because our data lacked samples from FnP and MxP ectoderm at E10.5, were resolved by mining databases including GEO, Mouse Genome Informatics, FaceBase and PubMed. We found evidence for mandible-specific expression of some genes (e.g. Foxa2, Besnard et al., 2004)

and for co-expression of others in the MnP and MxP (e.g. *Emx1* and *Vgll2*, Briata et al., 1996; Gorski et al., 2002; Johnson et al., 2011; Maeda et al., 2002). Thus prominence-specific programs already exist in the E10.5 ectoderm. In addition, there is an extensive early program that is shared between the ectoderm of all three prominences, typified by genes involved in the cell cycle, junction complex formation, and cytoskeletal function, including *Krt8*, *Krt18*, *Cldn7*, *Ocln*, *Tubb4a*, *Cdkn1a* and the transcription factor *Sp6* (Gray et al., 2004; Magdaleno et al., 2006). Overall, these shared gene signatures are consistent with a proliferative and patterning phase of ectodermal development in all prominences, prior to differentiation and morphogenesis of the epithelia.

The later facial ectoderm developmental program: E11.5-E12.5

A later program shared by all prominences (Fig 6) was defined by clusters of genes with an upward trajectory over time and maximal expression at E12.5 across all prominences (clusters IX and XIII in Fig 4; Table S9). Genes involved in epidermal cell shape and differentiation dominated the later program (Fig 4B), encoding cytoskeletal/cell junction components, secretion/endocytosis proteins and transcription factors (orange, grey and green, respectively, in Fig 6). “Signaling” was the second major theme (Figure 4B, Fig 6) and two observations concerning this process were especially noteworthy. First, we detected expression of the voltage-gated potassium channel *Kcnh7* and the A-type current regulator *Kcni1* in this late ectoderm program. This finding is consistent with a role for potassium currents in craniofacial patterning (Adams et al., 2016; Dahal et al., 2012), and provides a specific temporal focus for electrical excitability in patterning the ectoderm. Second, the Gaba receptor A, Pi subunit (*Gabrp*) was highly expressed in ectoderm from E11.5 and E12.5 (e.g. Fig 2). In adults *Gabrp* is expressed across a broad array of undifferentiated epithelia (MGI Mouse Gene Expression Database, (Smith et al., 2014; Finger et al., 2017); in breast cancer, *Gabrp* is associated with undifferentiated phenotypes and acts through activation of Extracellular Regulated Kinase (ERK) 1/2 (Sizemore et al., 2014). Given that all other Gaba receptor subunits were absent from our facial ectodermal samples, these data suggest a role for *Gabrp* in normal epithelial differentiation, independent of GABA and possibly through ERK signaling. Altogether this later facial ectodermal program suggests a transition from proliferating uncommitted cells towards differentiation of an epithelium and its subtypes (e.g. epidermal versus mucosal).

Prominence-specific aspects of the facial ectoderm program

The remaining clusters defined prominence-specific aspects of the ectodermal program (Fig 4B clusters II, IV, Va, VI, VII, VIII, X, XI, XII, cellular models in Fig S6–8). Mid and mid-late clusters involving the FnP highlighted placode patterning or neural development (IV, Va, VII), while the late stage FnP cluster (VIII) was dominated by genes annotated for melanocytes. The former is consistent with the patterning and neurogenesis within the olfactory placode beginning around E 11 (e.g. Fig S1), while the latter coincides with pigmentation that appears at the tip of the nose around E12.5 (T.W. unpublished). The MxP-specific clusters (VI and X) showed axon guidance and synapse themes – presumably reflecting innervation of the maxilla by trigeminal nerve fibers. The MdP themes were more diverse. The mid Md cluster II featured cell fate specification and oral/odontogenic

epithelium while the late Md cluster XI included “secretion”, which might reflect the maturing oral mucosa.

Integrating these prominence-specific programs with the early-, late- and pan-ectodermal programs provides a detailed description of the various aspects of ectodermal behavior that underlie facial development. The approach of combining genes at least 2-fold enriched in ectoderm with manual functional annotations was remarkably sensitive, able to detect signatures of small cell populations such as the melanocytes in the nose, the odontogenic epithelium in the jaws and the immune-related dendritic/Langerhans cells in the E12.5 epidermis. It was also remarkably specific, able to detect shifts from neurogenesis to neuro-differentiation in the FnP between E11.5 (mid) and E12.5 (mid-late). This validates our analytic approach and illustrates the power of these multidimensional data for identifying detailed genetic programs in specific tissues.

Co-expression modules for facial development

One principle that has emerged from more than a decade of transcriptome-based systems biology is that genes sharing functions tend to be co-expressed. This has considerable value, both for inferring functions of unannotated genes and for identifying the functions that are most relevant in a particular system. This is borne out by our analysis of ectoderm-enriched genes; clustering by spatio-temporal expression generated groups of genes that tended to share functional annotations (Figs 4–6, S6–8), and the shared functions tended to be closely related to the known processes for the corresponding tissue/stage. Thus, our final step was to build co-expression modules encompassing the entire facial development program, including the mesenchyme. To take full advantage of our multi-dimensional gene expression data, and to avoid biases based on preconceived notions of the importance of tissue- or temporal-specific expression, we used all of the differentially expressed probesets across all conditions ($p_{\text{adj}} < 0.05$), and Weighted Gene Co-Expression Network Analysis (WGCNA), with parameters adjusted to maximize enrichment of functional annotations (GO-BP and GO-CC terms, signaling pathways, and OFC genes) in the resulting modules. The result was 79 co-expression modules with membership ranging from 5 to 1351 probesets representing 1 to 1285 genes (Table S10). Four modules were eliminated from further consideration because they represented a single gene or because they represented a chromosomal cluster whose co-expression might be driven by a shared cis-regulatory element rather than a shared functional relationship. Of the remaining 75 co-expression modules, 45 were enriched for various GO-BP terms, 40 for GO-CC terms, 26 for various cell-to-cell signaling pathways, 14 for OFC genes and 11 for ncRNAs (Fig. S9). Only 14 modules lacked enrichment in one of these categories at $p_{\text{adj}} < 0.05$ (32 failed at $p_{\text{adj}} < 0.01$). Thus these co-expression modules show highly significant enrichment for biological processes and functions that might be relevant to facial development.

To understand the relationships of these modules in terms of when and where their constituent genes are expressed, the consensus (eigengene) expression patterns for each module were grouped by similarity, the resulting cohorts were color-coded and interpreted in terms of tissue and stage, and plotted as a heatmap (Figure 7, Table S10). The four largest modules were tissue specific: 2 pan-ectoderm modules with 1285 and 927 genes; a pan-

mesenchyme module with 966 genes; and a mid-late mesenchyme module with 642 genes. Together they accounted for 44% of the differentially expressed genes. This aligns well with our previous analysis in which tissue differences dominated the data. In addition, the number of pan-mesenchyme genes was considerably smaller than the number of pan-ectoderm genes, and there were more mesenchyme modules than ectoderm modules, in good agreement with our conclusion from Figure 3A that there is more diversity in the mesenchyme than in the ectoderm.

Co-expression modules for facial morphogenesis

Because OFCs are the common outcome of many perturbations of facial morphogenesis, we reasoned that modules enriched in genes associated with OFCs might feature processes pertinent to facial morphogenesis. We therefore curated a list of genes implicated in OFC (Table S12), identified modules that were enriched in these OFC genes and focused on these 14 Putative Morphogenesis and/or oro-facial Clefting (PMC) modules for further analysis (arrows on dendrogram in Fig 7). To facilitate discussion of these 14 modules, we named them according to their consensus (eigengene) spatio-temporal expression characteristics (Figs 7–9). While the non-PMC modules were primarily expressed early (E10.5 or E10.5 + E11.5; orange and green cohorts in Fig 7), the PMC modules displayed great diversity in both the tissues and timepoints with which they were associated (Fig 7). Our interpretation of this distribution of PMC modules is that 1) both ectoderm and mesenchyme make substantial contributions to facial morphogenesis and/or the mechanisms underlying fusion of the prominences, and 2) important foci are: the FnP mesenchyme at E10.5 and ectoderm at E11.5, the Mx ectoderm + mesenchyme at E11.5, the Md mesenchyme at all ages and the commonalities between the Md+Mx mesenchyme at E11.5+E12.5. This is remarkably congruent with our current understanding of facial morphogenesis, in which early FnP growth aligns the nasal pits with the maxilla for primary palate fusion (Green et al., 2015; Hu et al., 2015), and where growth of the jaw allows the tongue to drop from between the palatal shelves, removing a barrier to their apposition and subsequent fusion (Parada et al., 2015; Z. Song et al., 2013). This congruence between expression dynamics of our PMC modules and the dynamics of processes known to underlie facial morphogenesis encouraged us to look deeper for specific processes and the underlying molecular machinery.

We first examined the functional annotations enriched in each of these PMC modules to look for functional groupings (Figure 8). There was surprisingly little overlap in the annotations associated with the different PMC modules, aside from the two large ectoderm modules; apparently our module definition parameters have effectively partitioned genes into functionally distinct groups. Next we used these functional annotations to ascertain how the genes in each module might contribute to facial development. The two large ‘ectoderm’ PMC modules (red cohort, $n = 1285$ and $n = 927$ in Figs 7 and 8) were enriched for many typical ectoderm/epithelial terms encompassing cell adhesion, epidermal development and cellular processes. Likewise the large ‘mesenchyme’ PMC modules (brown and purple cohorts, $n = 966$ and $n = 642$ in Figs 7 and 8) were enriched for many terms typifying mesenchyme behaviors including adhesion, cell motility, cellular processes and connective tissue development. However, the size and complexity of these four largest modules limits their utility in understanding details of facial morphogenesis and the pathogenesis of OFCs.

The smaller PMC modules provided more insight into potential cellular and molecular control of facial morphogenesis. The ‘late Md mes’ module (purple cohort, n = 122 in Figs 7 and 8, ‘lightyellow’ in Tables S10 and S11) was highly enriched for striated muscle functions, with at least two genes expressed in primordial tongue mesenchyme (Candia et al., 1992; Du et al., 2016). The combination of tongue/muscle features and later expression in MdP suggests that this is a tongue muscle specification module.

The ‘early Fn Mes’ PMC module (purple cohort, n = 44, ‘darkorange2’ in Tables S10 and S11) included 24 genes that are annotated as expressed in neural tissues (DAVID, D. W. Huang et al., 2009b; 2009a). It also included a number of transcription factors and signaling molecules, five of which are associated with orofacial clefting: *Alx3*, *Foxc1*, *Foxc2*, *Gdf11*, *Stra6* ((Bahuaui et al., 2002; Beverdam et al., 2001; McPherron et al., 1999; Seller et al., 1996; Winnier et al., 1997). This module fits well with the BMP, FGF and RA signals that integrate patterning of the olfactory placode and of the medial and lateral nasal processes (LaMantia et al., 2000; Maier et al., 2010; 2011; Sabado et al., 2012; Y. Song et al., 2004; Szabo-Rogers et al., 2008; 2009), patterning that is critical for organizing the olfactory epithelium and for correct apposition and fusion of the facial prominences.

The smallest PMC modules, with less than 30 genes, offered an opportunity to use network-based gene-by-gene interpretation of their functional themes (Figure 9). The ‘mid Mx +’ cleft module (yellow cohort, n = 20, ‘orangered3’ in Table S10 and S9) was enriched in miRNAs. Its 7 protein coding genes included the transcription factors *Sox11* and *Skida1*, the chromatin factors *Epc2* and *Hist1h1c*, and the RNA processing factor *Celf1*. A functional network built from this module (Figure 9A) reveals that these genes are co-expressed in a variety of tissues and experimental contexts, supporting the idea that they may constitute a functional unit. Their GO annotations do not suggest a common function, largely due to the sparse annotations for most of these genes. However, in the context of facial development, there is a clear theme; *Sox11* and seven of these miRNAs (in the miR-17-92 and miR-106b-25 clusters) are needed for mandibular growth and lip/palate fusion (H. Huang et al., 2016; J. Wang et al., 2013). Altogether this module suggests coordinated production of a cohort of miRNAs in the face at E11.5 that cooperate to promote mandible growth and lip/palate fusion.

The ‘Mx/Md mes’ module (purple cohort, n = 9, ‘yellow3’ in Tables S10 and S11) was constitutive in Mx and Md mesenchyme. Its functional network (Fig 9B) shows that 7 of its 9 genes are involved in gene expression, including four transcription factors involved in OFCs, *Barx1*, *Dlx1*, *Lhx8* and *Sim2*, as well as *Lhx6* and the histone deacetylase co-repressor complex members *Mab2111* and *Mab2112*. Four of these 9 genes are involved in establishing the dental lamina (Denaxa et al., 2009; Grigoriou et al., 1998; Mitsiadis and Drouin, 2008), while *Ndnf*, *Mab2112*, *Lhx6*, *Lhx8b*, *Barx1* and probably also *Dlx1a* are regulated by the interaction of the *Edn1* and Notch signaling pathways that restricts cartilage primordia in the upper face in zebrafish (Barske et al., 2016). Because there is no obvious role of the dental lamina in earlier facial morphogenesis, we interpret this module as a jaw patterning program, possibly specific to the maxilla. We further suggest that the *Mab211* co-repressor complex cooperates with the other transcription factors to initiate chondrogenic

condensations, and that *Ndnf*, and perhaps also glutamate signaling might participate in this process.

The ‘early Md mes’ PMC module (purple cohort, $n = 25$, ‘mediumorchid’ in Tables S10 and S11) was specific to Md mesenchyme with highest expression at E10.5. The functional network built from this module (Fig 9C) shows a core network of six genes (*Hand1*, *Gbx2*, *Gpr50*, *Cited1*, *Dgkk*, *Magel2*) that are involved in mesodermal patterning, based on its enriched GO-BP annotations ‘blood vessel morphogenesis’ and ‘embryonic organ morphogenesis’; it also includes the four OFC genes (*Gbx2*, *Hand1*, *Tcf21*, *Dlx1*). In addition, this module includes a secondary network involving cell-cell junctions (*Jph2*, *Des*, *Popdc2*, *Mtus2*) and outliers suggesting neural function (*Syt6*, *Chrna6*, *Trpm1*). Because five genes in this module are involved in proximal-distal patterning of the mandible (*Hand1*, *Nkx3-2*; Lettice et al., 1999; Firulli et al., 2014; reviewed in Clouthier et al., 2013) and/or are regulated by the distal mandible patterning genes *Dlx5/6* (*Dlx1as*, *Hand1*, *Cited1*, *Rgs5*; Jeong et al., 2008), we interpret the core network as patterning the mandible proximal-distal axis. If co-expression indeed reflects shared function, then this module also predicts that mandible patterning involves a cell-cell junction sub-network, the orphan GPCR *Gpr50*, the signal transduction/hypospadias (urethral fusion) risk factor *Dgkk* (van der Zanden et al., 2011), and the endosome recycling factor *Magel2* (Hao et al., 2013).

In this analysis of co-expression modules we used the 14 modules enriched in OFC genes for proof-of-principle that our co-expression modules could be used to develop hypotheses about cohorts of genes with specific roles in facial morphogenesis. We found that 1) the consensus spatiotemporal expression pattern of the PMC modules aligned with the dynamics of the processes known to underlie facial morphogenesis and cleft pathogenesis; 2) the larger modules were of limited use for understanding specific mechanisms in facial development; they tended to be enriched for general terms that lacked sufficient specificity to suggest novel mechanisms of development. This is likely because the larger modules encompass several distinct functional groups, which were not separated by our spatio-temporal sampling strategies; 3) five of the seven smaller modules (>125 genes) showed clear functional themes that align with our knowledge of cleft pathogenesis. Importantly, the functional themes within these modules suggested novel mechanisms for facial development, while the specific genes provide tools to test those novel hypotheses. They offer an exciting opportunity to mine CLP GWAS datasets for signals at loci associated with the OFC modules. While we focused on modules enriched in OFC genes for the purpose of facial morphogenesis, the module membership data in Table S10 could easily be queried for other functional enrichments (e.g. cell-cell signaling) to identify modules pertinent to other aspects of facial development and OFC pathogenesis.

CONCLUSIONS

We present a multidimensional gene expression dataset from developing mouse face, including three ages and three facial prominences, with ectoderm separated from mesenchyme for each condition. The transcriptome data are statistically and biologically robust, with minimal variability between replicates and minimal cross-contamination between tissues. It was validated *in silico* and by *in situ* hybridization. Separating ectoderm

from mesenchyme, that is, separating tissues by lineage, was crucial both for increasing the dynamic range and sensitivity, and for detecting differential expression.

The ages, structures, and tissues analyzed in this dataset provide a foundational gene expression resource encompassing central aspects of facial morphogenesis. It is particularly suited to understanding the genes involved in the growth, morphogenesis, and prominence-specific patterning and differentiation processes that occur in the face up until the fusion of the lip and primary palate. It is also pertinent to early secondary palate development, specifically the initiation of the palatal shelves from the maxillary prominences. It does not cover later stages of secondary palate development (elevation, apposition and fusion of the palatal shelves), which have been extensively addressed in previous studies (e.g. Han et al., 2014; Iwata et al., 2012; Mima et al., 2013; Ozturk et al., 2013; Pelikan et al., 2013; A. S. Potter and S. S. Potter, 2015; reviewed in Lane and Kaartinen, 2014). Our dataset is relevant to orofacial clefting, both to the growth and morphogenesis that aligns the facial prominences, and to the fusion process itself. Although it lacks the spatial resolution to examine the dynamics of gene expression at the site of fusion of the lip and primary palate, its statistical robustness makes it well suited to layer with other more spatially refined datasets (e.g. Brunskill et al., 2014; Garaffo et al., 2013; O'Connell et al., 2012; A. S. Potter and S. S. Potter, 2015) to address roles of specific tissues and pathways in facial development and palate fusion.

We have used this dataset, along with manual curation of annotations relevant to facial development, to delineate a program for developing facial ectoderm. It includes an early program of structures, functions and cellular machinery characteristic of proliferating uncommitted cells (Fig 5), a later program transitioning towards differentiation of the epidermis (Fig 6), and prominence-specific programs that reflect the diverse roles of the FnP, MxP and MdP in generating ectoderm derivatives like whiskers, teeth, salivary glands and other tissues of the mature face (Fig S6–8). All this is built on top of a pervasive pan-ectoderm program of cell adhesion, cell-cell signaling, epithelial morphogenesis, and neural development (Table S8). The remarkable concordance between the functional themes represented by the various programs (Figs 5, 6, S6–8) and the underlying developmental processes (Fig S1) validates these programs; the novelties offer insights into the roles of unannotated or under-annotated genes, as well as genes and pathways (e.g. catecholamines, Gabrp) that are not part of our current understanding of the mechanisms underlying facial development.

Finally, we used co-expression across this dataset to derive modules of genes that may be functionally related in the context of facial development. We validated these modules by focusing on those that were enriched for genes involved in the pathogenesis of OFCs, and by using functional annotations and functional networks to analyze their roles in facial morphogenesis. The developmental and regulatory themes that emerged aligned well with functions known to be important in the tissues and stages represented by each module's consensus expression characteristics. These insights remain to be experimentally demonstrated, but provide specific hypotheses and tools to test these hypotheses. Altogether, this transcriptome dataset, the ectoderm genetic programs and the co-expression modules provide an important new resource for exploring the systems biology of facial development.

They offer a means of identifying candidates for new human cleft lip/palate loci, for mining facial genome-wide association studies, and for interpreting mechanisms underlying facial development and morphology.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

TW, LH, RS and WF conceived and initiated this study. WF prepared all samples. Microarray probe preparation and analyses were performed by the UCD Gene Expression Core Facility. TP and CS pre-processed the data. KJ analyzed data for and generated Fig S5F. HL mined public domain data supporting prominence-specific expression at E10.5. JH, HL and TW performed the RNA in situ analyses. RS curated human orofacial cleft genes. SL contributed data visualization and advice on analytics. JH performed all other analyses, annotations and interpretations. JH prepared the manuscript in consultation with TW. Special thanks to Irene Choi for expert technical assistance with the in situ analysis; to Dr. Bethany Reid for SEM images; to members of the Hunter lab, the Computational Biosciences Program and the Biostatistics Program for advice, guidance and encouragement on data analysis; to our colleagues in the Williams lab and the Cells, Stem cells and Development Program for provocative discussions and feedback on data interpretation; to Dr. Eric Van Otterloo for critical comments on the manuscript. Funding was provided by NIDCR via DE012728 (TW), DE15191 (RS), DE024429 (TW, JH, KJ) and NLM R01LM008111 and R01LM009254 (LH). There are no competing interests.

References

- Adameyko I, Fried K. The Nervous System Orchestrates and Integrates Craniofacial Development: A Review. *Front Physiol.* 2016; 7:49.doi: 10.3389/fphys.2016.00049 [PubMed: 26924989]
- Adams DS, Uzel SGM, Akagi J, Wlodkowic D, Andreeva V, Yelick PC, Devitt-Lee A, Pare JF, Levin M. Bioelectric signalling via potassium channels: a mechanism for craniofacial dysmorphogenesis in KCNJ2-associated Andersen-Tawil Syndrome. *J Physiol (Lond).* 2016; 594:3245–3270. DOI: 10.1113/JP271930 [PubMed: 26864374]
- Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet.* 2000; 25:25–29. DOI: 10.1038/75556 [PubMed: 10802651]
- Bahuau M, Houdayer C, Tredano M, Soupre V, Couderc R, Vazquez MP. FOXC2 truncating mutation in distichiasis, lymphedema, and cleft palate. *Clin Genet.* 2002; 62:470–473. [PubMed: 12485195]
- Barske L, Askary A, Zuniga E, Balczerski B, Bump P, Nichols JT, Crump JG. Competition between Jagged-Notch and Endothelin1 Signaling Selectively Restricts Cartilage Formation in the Zebrafish Upper Face. *PLoS Genet.* 2016; 12:e1005967.doi: 10.1371/journal.pgen.1005967 [PubMed: 27058748]
- Basson MA, van Ravenswaaij-Arts C. Functional Insights into Chromatin Remodelling from Studies on CHARGE Syndrome. *Trends Genet.* 2015; 31:600–611. DOI: 10.1016/j.tig.2015.05.009 [PubMed: 26411921]
- Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B* 1995
- Besnard V, Wert SE, Hull WM, Whitsett JA. Immunohistochemical localization of Foxa1 and Foxa2 in mouse embryos and adult tissues. *Gene Expr Patterns.* 2004; 5:193–208. DOI: 10.1016/j.modgep.2004.08.006 [PubMed: 15567715]
- Beverdam A, Brouwer A, Reijnen M, Korving J, Meijlink F. Severe nasal clefting and abnormal embryonic apoptosis in Alx3/Alx4 double mutant mice. *Development.* 2001; 128:3975–3986. [PubMed: 11641221]
- Bhattacharjee V, Mukhopadhyay P, Singh S, Johnson C, Philipose JT, Warner CP, Greene RM, Pisano MM. Neural crest and mesoderm lineage-dependent gene expression in orofacial development. *Differentiation.* 2007; 75:463–477. DOI: 10.1111/j.1432-0436.2006.00145.x [PubMed: 17286603]

- Boles NC, Hirsch SE, Le S, Corneo B, Najm F, Minotti AP, Wang Q, Lotz S, Tesar PJ, Fasano CA. NPTX1 regulates neural lineage specification from human pluripotent stem cells. *Cell Rep.* 2014; 6:724–736. DOI: 10.1016/j.celrep.2014.01.026 [PubMed: 24529709]
- Briata P, Di Blas E, Gulisano M, Mallamaci A, Iannone R, Boncinelli E, Corte G. EMX1 homeoprotein is expressed in cell nuclei of the developing cerebral cortex and in the axons of the olfactory sensory neurons. *Mech Dev.* 1996; 57:169–180. [PubMed: 8843394]
- Brinkley JF, Fisher S, Harris MP, Holmes G, Hooper JE, Jabs EW, Jones KL, Kesselman C, Klein OD, Maas RL, Marazita ML, Selleri L, Spritz RA, van Bakel H, Visel A, Williams TJ, Wysocka J, FaceBase C, Chai Y. The FaceBase Consortium: a comprehensive resource for craniofacial researchers. *Development.* 2016; 143:2677–2688. DOI: 10.1242/dev.135434 [PubMed: 27287806]
- Brugmann SA, Powder KE, Young NM, Goodnough LH, Hahn SM, James AW, Helms JA, Lovett M. Comparative gene expression analysis of avian embryonic facial structures reveals new candidates for human craniofacial disorders. *Hum Mol Genet.* 2010; 19:920–930. DOI: 10.1093/hmg/ddp559 [PubMed: 20015954]
- Brunskill EW, Potter AS, Distasio A, Dexheimer P, Plassard A, Aronow BJ, Potter SS. A gene expression atlas of early craniofacial development. *Dev Biol.* 2014; 391:133–146. DOI: 10.1016/j.ydbio.2014.04.016 [PubMed: 24780627]
- Buchtová M, Kuo WP, Nimmagadda S, Benson SL, Geetha-Loganathan P, Logan C, Au-Yeung T, Chiang E, Fu K, Richman JM. Whole genome microarray analysis of chicken embryo facial prominences. *Developmental Dynamics.* 2010; 239:574–591. DOI: 10.1002/dvdy.22135 [PubMed: 19941351]
- Cai J, Ash D, Kotch LE, Jabs EW, Attie-Bitach T, Auge J, Mattei G, Etchevers H, Vekemans M, Korshunova Y, Tidwell R, Messina DN, Winston JB, Lovett M. Gene expression in pharyngeal arch 1 during human embryonic development. *Hum Mol Genet.* 2005; 14:903–912. DOI: 10.1093/hmg/ddi083 [PubMed: 15703188]
- Candia AF, Hu J, Crosby J, Lalley PA, Noden D, Nadeau JH, Wright CV. Mox-1 and Mox-2 define a novel homeobox gene subfamily and are differentially expressed during early mesodermal patterning in mouse embryos. *Development.* 1992; 116:1123–1136. [PubMed: 1363541]
- Chai Y, Jiang X, Ito Y, Bringas P, Han J, Rowitch DH, Soriano P, McMahon AP, Sucov HM. Fate of the mammalian cranial neural crest during tooth and mandibular morphogenesis. *Development.* 2000; 127:1671–1679. [PubMed: 10725243]
- Chai Y, Maxson RE. Recent advances in craniofacial morphogenesis. *Dev Dyn.* 2006; 235:2353–2375. DOI: 10.1002/dvdy.20833 [PubMed: 16680722]
- Clouthier DE, Passos-Bueno MR, Tavares ALP, Lyonnet S, Amiel J, Gordon CT. Understanding the basis of auriculocondylar syndrome: Insights from human, mouse and zebrafish genetic studies. *Am J Med Genet C Semin Med Genet.* 2013; 163C:306–317. DOI: 10.1002/ajmg.c.31376 [PubMed: 24123988]
- Cuschieri A, Bannister LH. The development of the olfactory mucosa in the mouse: light microscopy. *J Anat.* 1975; 119:277–286. [PubMed: 1133096]
- Dahal GR, Rawson J, Gassaway B, Kwok B, Tong Y, Ptáček LJ, Bates E. An inwardly rectifying K⁺ channel is required for patterning. *Development.* 2012; 139:3653–3664. DOI: 10.1242/dev.078592 [PubMed: 22949619]
- Denaxa M, Sharpe PT, Pachnis V. The LIM homeodomain transcription factors Lhx6 and Lhx7 are key regulators of mammalian dentition. *Dev Biol.* 2009; 333:324–336. DOI: 10.1016/j.ydbio.2009.07.001 [PubMed: 19591819]
- Depew MJ, Tucker AS, Sharpe PT. Craniofacial development 2002
- Ding HL, Hooper JE, Batzel P, Eames BF, Postlethwait JH, Artinger KB, Clouthier DE. MicroRNA Profiling during Craniofacial Development: Potential Roles for Mir23b and Mir133b. *Front Physiol.* 2016; 7:281.doi: 10.3389/fphys.2016.00281 [PubMed: 27471470]
- Dixon MJ, Marazita ML, Beaty TH, Murray JC. Cleft lip and palate: understanding genetic and environmental influences. *Nat Rev Genet.* 2011; 12:167–178. DOI: 10.1038/nrg2933 [PubMed: 21331089]

- Du W, Prochazka J, Prochazkova M, Klein OD. Expression of FGFs during early mouse tongue development. *Gene Expr Patterns*. 2016; 20:81–87. DOI: 10.1016/j.gep.2015.12.003 [PubMed: 26748348]
- Feng W, Leach SM, Tipney H, Phang T, Geraci M, Spritz RA, Hunter LE, Williams T. Spatial and temporal analysis of gene expression during growth and fusion of the mouse facial prominences. *PLoS ONE*. 2009; 4:e8066.doi: 10.1371/journal.pone.0008066 [PubMed: 20016822]
- Finger JH, Smith CM, Hayamizu TF, McCright IJ, Xu J, Law M, Shaw DR, Baldarelli RM, Beal JS, Blodgett O, Campbell JW, Corbani LE, Lewis JR, Forthofer KL, Frost PJ, Giannatto SC, Hutchins LN, Miers DB, Motenko H, Stone KR, Eppig JT, Kadin JA, Richardson JE, Ringwald M. The mouse Gene Expression Database (GXD): 2017 update. *Nucleic Acids Res*. 2017; 45:D730–D736. DOI: 10.1093/nar/gkw1073 [PubMed: 27899677]
- Fink JL, Aturaliya RN, Davis MJ, Zhang F, Hanson K, Teasdale MS, Kai C, Kawai J, Carninci P, Hayashizaki Y, Teasdale RD. LOCATE: a mouse protein subcellular localization database. *Nucleic Acids Res*. 2006; 34:D213–7. DOI: 10.1093/nar/gkj069 [PubMed: 16381849]
- Firulli BA, Fuchs RK, Vincentz JW, Clouthier DE, Firulli AB. Hand1 phosphoregulation within the distal arch neural crest is essential for craniofacial morphogenesis. *Development*. 2014; 141:3050–3061. DOI: 10.1242/dev.107680 [PubMed: 25053435]
- Forni PE, Taylor-Burds C, Melvin VS, Williams T, Williams T, Wray S. Neural crest and ectodermal cells intermix in the nasal placode to give rise to GnRH-1 neurons, sensory neurons, and olfactory ensheathing cells. *J Neurosci*. 2011; 31:6915–6927. DOI: 10.1523/JNEUROSCI.6087-10.2011 [PubMed: 21543621]
- Fresno C, Fernández EA. RDAVIDWebService: a versatile R interface to DAVID. *Bioinformatics*. 2013; 29:2810–2811. DOI: 10.1093/bioinformatics/btt487 [PubMed: 23958726]
- Frommer J, Margolies MR. Contribution of Meckel's Cartilage to Ossification of the Mandible in Mice. *J Dent Res*. 1971; 50:1260–1267. DOI: 10.1177/00220345710500052801 [PubMed: 5285785]
- Garaffo G, Provero P, Molineris I, Pinciroli P, Peano C, Battaglia C, Tomaiuolo D, Etzion T, Gothilf Y, Santoro M, Merlo GR. Profiling, Bioinformatic, and Functional Data on the Developing Olfactory/GnRH System Reveal Cellular and Molecular Pathways Essential for This Process and Potentially Relevant for the Kallmann Syndrome. *Frontiers in Endocrinology*. 2013; 4:203.doi: 10.3389/fendo.2013.00203 [PubMed: 24427155]
- Geetha-Loganathan P, Nimmagadda S, Fu K, Richman JM. Avian facial morphogenesis is regulated by c-Jun N-terminal kinase/planar cell polarity (JNK/PCP) wingless-related (WNT) signaling. *J Biol Chem*. 2014; 289:24153–24167. DOI: 10.1074/jbc.M113.522003 [PubMed: 25008326]
- Gertsenstein M, Vintersten K, Behringer R. *Manipulating the mouse embryo: a laboratory manual*. 2003
- Gorski JA, Talley T, Qiu M, Puelles L, Rubenstein JLR, Jones KR. Cortical excitatory neurons and glia, but not GABAergic neurons, are produced in the *Emx1*-expressing lineage. *J Neurosci*. 2002; 22:6309–6314. [PubMed: 12151506]
- Gray PA, Fu H, Luo P, Zhao Q, Yu J, Ferrari A, Tenzen T, Yuk DI, Tsung EF, Cai Z, Alberta JA, Cheng LP, Liu Y, Stenman JM, Valerius MT, Billings N, Kim HA, Greenberg ME, McMahon AP, Rowitch DH, Stiles CD, Ma Q. Mouse brain organization revealed through direct genome-scale TF expression analysis. *Science*. 2004; 306:2255–2257. DOI: 10.1126/science.1104935 [PubMed: 15618518]
- Green RM, Feng W, Phang T, Fish JL, Li H, Spritz RA, Marcucio RS, Hooper J, Jamniczky H, Hallgrímsson B, Williams T. *Tfap2a*-dependent changes in mouse facial morphology result in clefting that can be ameliorated by a reduction in *Fgf8* gene dosage. *Dis Model Mech*. 2015; 8:31–43. DOI: 10.1242/dmm.017616 [PubMed: 25381013]
- Grigoriou M, Tucker AS, Sharpe PT, Pachnis V. Expression and regulation of *Lhx6* and *Lhx7*, a novel subfamily of LIM homeodomain encoding genes, suggests a role in mammalian head development. *Development*. 1998; 125:2063–2074. [PubMed: 9570771]
- Hall JMH, Bell ML, Finger TE. Disruption of sonic hedgehog signaling alters growth and patterning of lingual taste papillae. *Dev Biol*. 2003; 255:263–277. [PubMed: 12648489]

- Han A, Zhao H, Li J, Pelikan R, Chai Y. ALK5-mediated transforming growth factor β signaling in neural crest cells controls craniofacial muscle development via tissue-tissue interactions. *Mol Cell Biol*. 2014; 34:3120–3131. DOI: 10.1128/MCB.00623-14 [PubMed: 24912677]
- Hao YH, Doyle JM, Ramanathan S, Gomez TS, Jia D, Xu M, Chen ZJ, Billadeau DD, Rosen MK, Potts PR. Regulation of WASH-dependent actin polymerization and protein trafficking by ubiquitination. *Cell*. 2013; 152:1051–1064. DOI: 10.1016/j.cell.2013.01.051 [PubMed: 23452853]
- Harada H, Toyono T, Toyoshima K, Yamasaki M, Itoh N, Kato S, Sekine K, Ohuchi H. FGF10 maintains stem cell compartment in developing mouse incisors. *Development*. 2002; 129:1533–1541. [PubMed: 11880361]
- Hiruma T, Nakajima Y, Nakamura H. Development of pharyngeal arch arteries in early mouse embryo. *J Anat*. 2002; 201:15–29. [PubMed: 12171473]
- Hu D, Young NM, Li X, Xu Y, Hallgrímsson B, Marcucio RS. A dynamic Shh expression pattern, regulated by SHH and BMP signaling, coordinates fusion of primordia in the amniote face. *Development*. 2015; 142:567–574. DOI: 10.1242/dev.114835 [PubMed: 25605783]
- Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc*. 2009a; 4:44–57. DOI: 10.1038/nprot.2008.211 [PubMed: 19131956]
- Huang DW, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res*. 2009b; 37:1–13. DOI: 10.1093/nar/gkn923 [PubMed: 19033363]
- Huang H, Yang X, Bao M, Cao H, Miao X, Zhang X, Gan L, Qiu M, Zhang Z. Ablation of the Sox11 Gene Results in Clefting of the Secondary Palate Resembling the Pierre Robin Sequence. *J Biol Chem*. 2016; 291:7107–7118. DOI: 10.1074/jbc.M115.690875 [PubMed: 26826126]
- Iwata J, Hacia JG, Suzuki A, Sanchez-Lara PA, Urata M, Chai Y. Modulation of noncanonical TGF-beta signaling prevents cleft palate in Tgfb2 mutant mice. *The Journal of Clinical Investigation*. 2012; 122:873–885. DOI: 10.1172/JCI61498 [PubMed: 22326956]
- Jaworski A, Tom I, Tong RK, Gildea HK, Koch AW, Gonzalez LC, Tessier-Lavigne M. Operational redundancy in axon guidance through the multifunctional receptor Robo3 and its ligand NELL2. *Science*. 2015; 350:961–965. DOI: 10.1126/science.aad2615 [PubMed: 26586761]
- Jeong J, Li X, McEvelly RJ, Rosenfeld MG, Lufkin T, Rubenstein JLR. Dlx genes pattern mammalian jaw primordium by regulating both lower jaw-specific and upper jaw-specific genetic programs. *Development*. 2008; 135:2905–2916. DOI: 10.1242/dev.019778 [PubMed: 18697905]
- Johnson CW, Hernandez-Lagunas L, Feng W, Melvin VS, Williams T, Artinger KB. Vgl2a is required for neural crest cell survival during zebrafish craniofacial development. *Dev Biol*. 2011; 357:269–281. DOI: 10.1016/j.ydbio.2011.06.034 [PubMed: 21741961]
- Kousa YA, Schutte BC. Toward an orofacial gene regulatory network. *Developmental Dynamics*. 2016; 245:220–232. DOI: 10.1002/dvdy.24341 [PubMed: 26332872]
- LaMantia AS, Bhasin N, Rhodes K, Heemskerk J. Mesenchymal/epithelial induction mediates olfactory pathway formation. *Neuron*. 2000; 28:411–425. [PubMed: 11144352]
- Lane J, Kaartinen V. Signaling networks in palate development. *Wiley Interdiscip Rev Syst Biol Med*. 2014; 6:271–278. DOI: 10.1002/wsbm.1265 [PubMed: 24644145]
- Langfelder P, Horvath S. Fast R Functions for Robust Correlations and Hierarchical Clustering. *J Stat Softw*. 2012:46.
- Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics*. 2008; 9:559. doi: 10.1186/1471-2105-9-559 [PubMed: 19114008]
- Lettice LA, Purdie LA, Carlson GJ, Kilanowski F, Dorin J, Hill RE. The mouse bagpipe gene controls development of axial skeleton, skull, and spleen. *Proc Natl Acad Sci USA*. 1999; 96:9695–9700. [PubMed: 10449756]
- Li H, Williams T. Separation of mouse embryonic facial ectoderm and mesenchyme. *J Vis Exp*. 2013; doi: 10.3791/50248
- Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov JP, Tamayo P. The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst*. 2015; 1:417–425. DOI: 10.1016/j.cels.2015.12.004 [PubMed: 26771021]

- Linde-Medina M, Hallgrimsson B, Marcucio R. Beyond cell proliferation in avian facial morphogenesis. *Developmental Dynamics*. 2016; 245:190–196. DOI: 10.1002/dvdy.24374 [PubMed: 26637960]
- Liu M, Zhao S, Lin Q, Wang XP. YAP regulates the expression of Hoxa1 and Hoxc13 in mouse and human oral and skin epithelial tissues. *Mol Cell Biol*. 2015; 35:1449–1461. DOI: 10.1128/MCB.00765-14 [PubMed: 25691658]
- Maeda T, Chapman DL, Stewart AFR. Mammalian vestigial-like 2, a cofactor of TEF-1 and MEF2 transcription factors that promotes skeletal muscle differentiation. *J Biol Chem*. 2002; 277:48889–48898. DOI: 10.1074/jbc.M206858200 [PubMed: 12376544]
- Magdaleno S, Jensen P, Brumwell CL, Seal A, Lehman K, Asbury A, Cheung T, Cornelius T, Batten DM, Eden C, Norland SM, Rice DS, Dosooye N, Shakya S, Mehta P, Curran T. BGEM: an in situ hybridization database of gene expression in the embryonic and adult mouse nervous system. *PLoS Biol*. 2006; 4:e86.doi: 10.1371/journal.pbio.0040086 [PubMed: 16602821]
- Maier E, von Hofsten J, Nord H, Fernandes M, Paek H, Hébert JM, Gunhaga L. Opposing Fgf and Bmp activities regulate the specification of olfactory sensory and respiratory epithelial cell fates. *Development*. 2010; 137:1601–1611. DOI: 10.1242/dev.051219 [PubMed: 20392740]
- Maier E, Nord H, von Hofsten J, Gunhaga L. A balance of BMP and notch activity regulates neurogenesis and olfactory nerve formation. *PLoS ONE*. 2011; 6:e17379.doi: 10.1371/journal.pone.0017379 [PubMed: 21383851]
- Maklad A, Conway M, Hodges C, Hansen LA. Development of Innervation to Maxillary Whiskers in Mice. *The Anatomical Record*. 2010; 293:1553–1567. DOI: 10.1002/ar.21194 [PubMed: 20648571]
- Shi L, Reid LH, Jones WD, Shippy R, Warrington JA, Baker SC, Collins PJ, de Longueville F, Kawasaki ES, Lee KY, Luo Y, Sun YA, Willey JC, Setterquist RA, Fischer GM, Tong W, Dragan YP, Dix DJ, Frueh FW, Goodsaid FM, Herman D, Jensen RV, Johnson CD, Lobenhofer EK, Puri RK, Schrf U, Thierry-Mieg J, Wang C, Wilson M, Wolber PK, Zhang L, Amur S, Bao W, Barbacioru CC, Lucas AB, Bertholet V, Boysen C, Bromley B, Brown D, Brunner A, Canales R, Cao XM, Cebula TA, Chen JJ, Cheng J, Chu T-M, Chudin E, Corson J, Corton JC, Croner LJ, Davies C, Davison TS, Delenstarr G, Deng X, Dorris D, Eklund AC, Fan X-H, Fang H, Fulmer-Smentek S, Fuscoe JC, Gallagher K, Ge W, Guo L, Guo X, Hager J, Haje PK, Han J, Han T, Harbottle HC, Harris SC, Hatchwell E, Hauser CA, Hester S, Hong H, Hurban P, Jackson SA, Ji H, Knight CR, Kuo WP, LeClerc JE, Levy S, Li Q-Z, Liu C, Liu Y, Lombardi MJ, Ma Y, Magnuson SR, Maqsoodi B, McDaniel T, Mei N, Myklebost O, Ning B, Novoradovskaya N, Orr MS, Osborn TW, Papallo A, Patterson TA, Perkins RG, Peters EH, Peterson R, Philips KL, Pine PS, Pusztai L, Qian F, Ren H, Rosen M, Rosenzweig BA, Samaha RR, Schena M, Schroth GP, Shchegrova S, Smith DD, Staedtler F, Su Z, Sun H, Szallasi Z, Tezak Z, Thierry-Mieg D, Thompson KL, Tikhonova I, Turpaz Y, Vallanat B, Van C, Walker SJ, Wang SJ, Wang Y, Wolfinger R, Wong A, Wu J, Xiao C, Xie Q, Xu J, Yang W, Zhang L, Zhong S, Zong Y, Slikker W. MAQC Consortium. The MicroArray Quality Control (MAQC) project shows inter- and intraplatform reproducibility of gene expression measurements. *Nat Biotechnol*. 2006; 24:1151–1161. DOI: 10.1038/nbt1239 [PubMed: 16964229]
- Mayo ML, Bringas P Jr, Santos V, Shum L. Desmin expression during early mouse tongue morphogenesis. *Int J Dev Biol*. 1992
- McGrath JA, Gatalica B, Christiano AM, Li K, Owaribe K, McMillan JR, Eady RA, Uitto J. Mutations in the 180-kD bullous pemphigoid antigen (BPAG2), a hemidesmosomal transmembrane collagen (COL17A1), in generalized atrophic benign epidermolysis bullosa. *Nat Genet*. 1995; 11:83–86. DOI: 10.1038/ng0995-83 [PubMed: 7550320]
- McPherron AC, Lawler AM, Lee SJ. Regulation of anterior/posterior patterning of the axial skeleton by growth/differentiation factor 11. *Nat Genet*. 1999; 22:260–264. DOI: 10.1038/10320 [PubMed: 10391213]
- Miller AM, Treloar HB, Greer CA. Composition of the migratory mass during development of the olfactory nerve. *J Comp Neurol*. 2010; 518:4825–4841. DOI: 10.1002/cne.22497 [PubMed: 21031554]

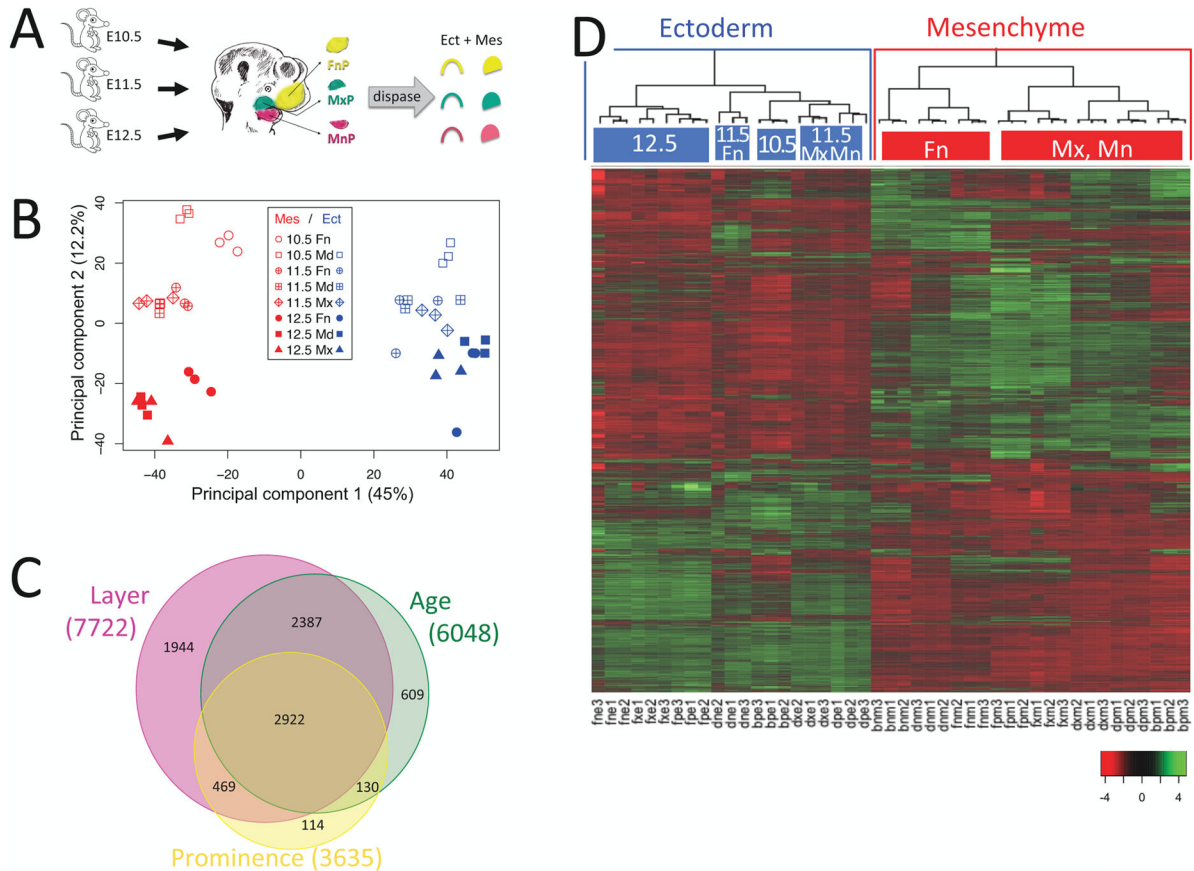
- Mima J, Koshino A, Oka K, Uchida H, Hieda Y, Nohara K, Kogo M, Chai Y, Sakai T. Regulation of the epithelial adhesion molecule CEACAM1 is important for palate formation. *PLoS ONE*. 2013; 8:e61653.doi: 10.1371/journal.pone.0061653 [PubMed: 23613893]
- Mitsiadis TA, Drouin J. Deletion of the Pitx1 genomic locus affects mandibular tooth morphogenesis and expression of the Barx1 and Tbx1 genes. *Dev Biol*. 2008; 313:887–896. DOI: 10.1016/j.ydbio.2007.10.055 [PubMed: 18082678]
- Monaghan AP, Kaestner KH, Grau E, Schütz G. Postimplantation expression patterns indicate a role for the mouse forkhead/HNF-3 alpha, beta and gamma genes in determination of the definitive endoderm, chordamesoderm and neuroectoderm. *Development*. 1993; 119:567–578. [PubMed: 8187630]
- Musselmann K, Green JA, Sone K, Hsu JC, Bothwell IR, Johnson SA, Harunaga JS, Wei Z, Yamada KM. Salivary gland gene expression atlas identifies a new regulator of branching morphogenesis. *J Dent Res*. 2011; 90:1078–1084. DOI: 10.1177/0022034511413131 [PubMed: 21709141]
- O’Connell DJ, Ho JWK, Mammoto T, Turbe-Doan A, O’Connell JT, Haseley PS, Koo S, Kamiya N, Ingber DE, Park PJ, Maas RL. A Wnt-bmp feedback circuit controls intertissue signaling dynamics in tooth organogenesis. *Sci Signal*. 2012; 5:ra4.doi: 10.1126/scisignal.2002414 [PubMed: 22234613]
- Ozturk F, Li Y, Zhu X, Guda C, Nawshad A. Systematic analysis of palatal transcriptome to identify cleft palate genes within TGFβ3-knockout mice alleles: RNA-Seq analysis of TGFβ3 Mice. *BMC Genomics*. 2013; 14:113.doi: 10.1186/1471-2164-14-113 [PubMed: 23421592]
- Parada C, Han D, Grimaldi A, Sarrion P, Park SS, Pelikan R, Sanchez-Lara PA, Chai Y. Disruption of the ERK/MAPK pathway in neural crest cells as a potential cause of Pierre Robin sequence. *Development*. 2015; 142:3734–3745. DOI: 10.1242/dev.125328 [PubMed: 26395480]
- Pelikan RC, Iwata J, Suzuki A, Chai Y, Hacia JG. Identification of candidate downstream targets of TGFβ signaling during palate development by genome-wide transcript profiling. *J Cell Biochem*. 2013; 114:796–807. DOI: 10.1002/jcb.24417 [PubMed: 23060211]
- Potter AS, Potter SS. Molecular Anatomy of Palate Development. *PLoS ONE*. 2015; 10:e0132662.doi: 10.1371/journal.pone.0132662 [PubMed: 26168040]
- Russell, S., Meadows, LA., Russell, RR. *Microarray Technology in Practice*. Academic Press; 2008.
- Sabado V, Barraud P, Baker CVH, Streit A. Specification of GnRH-1 neurons by antagonistic FGF and retinoic acid signaling. *Dev Biol*. 2012; 362:254–262. DOI: 10.1016/j.ydbio.2011.12.016 [PubMed: 22200593]
- Schallreuter KU, Lemke KR, Pittelkow MR, Wood JM, Körner C, Malik R. Catecholamines in human keratinocyte differentiation. *J Invest Dermatol*. 1995; 104:953–957. [PubMed: 7769265]
- Schwind JL. The development of the hypophysis cerebri of the albino rat. *American Journal of Anatomy*. 1928; 41:295–319. DOI: 10.1002/aja.1000410206
- Seller MJ, Davis TB, Fear CN, Flinter FA, Ellis I, Gibson AG. Two sibs with anophthalmia and pulmonary hypoplasia (the Matthew-Wood syndrome). *Am J Med Genet*. 1996; 62:227–229. DOI: 10.1002/(SICI)1096-8628(19960329)62:3<227::AID-AJMG5>3.0.CO;2-Q [PubMed: 8882778]
- Shimodaira H. An approximately unbiased test of phylogenetic tree selection. *Syst Biol*. 2002; 51:492–508. DOI: 10.1080/10635150290069913 [PubMed: 12079646]
- Singh S, Groves AK. The molecular basis of craniofacial placode development. *Wiley Interdiscip Rev Dev Biol*. 2016; 5:363–376. DOI: 10.1002/wdev.226 [PubMed: 26952139]
- Sivamani RK, Shi B, Griffiths E, Vu SM, Lev-Tov HA, Dahle S, Chigbrow M, La TD, Mashburn C, Peavy TR, Isseroff RR. Acute wounding alters the beta2-adrenergic signaling and catecholamine synthetic pathways in keratinocytes. *J Invest Dermatol*. 2014; 134:2258–2266. DOI: 10.1038/jid.2014.137 [PubMed: 24614156]
- Sizemore GM, Sizemore ST, Seachrist DD, Keri RA. GABA(A) receptor pi (GABRP) stimulates basal-like breast cancer cell migration through activation of extracellular-regulated kinase 1/2 (ERK1/2). *J Biol Chem*. 2014; 289:24102–24113. DOI: 10.1074/jbc.M114.593582 [PubMed: 25012653]
- Smith CM, Finger JH, Hayamizu TF, McCright IJ, Xu J, Berghout J, Campbell J, Corbani LE, Forthofer KL, Frost PJ, Miers D, Shaw DR, Stone KR, Eppig JT, Kadin JA, Richardson JE,

- Ringwald M. The mouse Gene Expression Database (GXD): 2014 update. *Nucleic Acids Res.* 2014; 42:D818–24. DOI: 10.1093/nar/gkt954 [PubMed: 24163257]
- Song Y, Hui JN, Fu KK, Richman JM. Control of retinoic acid synthesis and FGF expression in the nasal pit is required to pattern the craniofacial skeleton. *Dev Biol.* 2004; 276:313–329. DOI: 10.1016/j.ydbio.2004.08.035 [PubMed: 15581867]
- Song Z, Liu C, Iwata J, Gu S, Suzuki A, Sun C, He W, Shu R, Li L, Chai Y, Chen Y. Mice with Tak1 deficiency in neural crest lineage exhibit cleft palate associated with abnormal tongue development. *J Biol Chem.* 2013; 288:10440–10450. DOI: 10.1074/jbc.M112.432286 [PubMed: 23460641]
- Sprenger J, Lynn Fink J, Karunaratne S, Hanson K, Hamilton NA, Teasdale RD. LOCATE: a mammalian protein subcellular localization database. *Nucleic Acids Res.* 2008; 36:D230–3. DOI: 10.1093/nar/gkm950 [PubMed: 17986452]
- Stainier, DY., Gilbert, W. Pioneer neurons in the mouse trigeminal sensory system. Presented at the Proceedings of the National ...; 1990.
- Suzuki A, Sangani DR, Ansari A, Iwata J. Molecular mechanisms of midfacial developmental defects. *Developmental Dynamics.* 2016; 245:276–293. DOI: 10.1002/dvdy.24368 [PubMed: 26562615]
- Szabo-Rogers HL, Geetha-Loganathan P, Nimmagadda S, Fu KK, Richman JM. FGF signals from the nasal pit are necessary for normal facial morphogenesis. *Dev Biol.* 2008; 318:289–302. [PubMed: 18455717]
- Szabo-Rogers HL, Geetha-Loganathan P, Whiting CJ, Nimmagadda S, Fu K, Richman JM. Novel skeletogenic patterning roles for the olfactory pit. *Development.* 2009; 136:219–229. DOI: 10.1242/dev.023978 [PubMed: 19056832]
- Thesleff I. Epithelial-mesenchymal signalling regulating tooth morphogenesis. *J Cell Sci.* 2003; 116:1647–1648. [PubMed: 12665545]
- Tipney HJ, Leach SM, Feng W, Spritz R, Williams T, Hunter L. Leveraging existing biological knowledge in the identification of candidate genes for facial dysmorphology. *BMC Bioinformatics.* 2009; 10(Suppl 2):S12.doi: 10.1186/1471-2105-10-S2-S12
- Tucker AS. Salivary gland development. *Semin Cell Dev Biol.* 2007; 18:237–244. DOI: 10.1016/j.semcdb.2007.01.006 [PubMed: 17336109]
- Twigg SRF, Wilkie AOM. New insights into craniofacial malformations. *Hum Mol Genet.* 2015; 24:R50–9. DOI: 10.1093/hmg/ddv228 [PubMed: 26085576]
- van der Zanden LFM, van Rooij IALM, Feitz WFJ, Knight J, Donders ART, Renkema KY, Bongers EMHF, Vermeulen SHHM, Kiemeny LALM, Veltman JA, Arias-Vásquez A, Zhang X, Markljung E, Qiao L, Baskin LS, Nordenskjöld A, Roeleveld N, Franke B, Knoers NVAM. Common variants in DGKK are strongly associated with risk of hypospadias. *Nat Genet.* 2011; 43:48–50. DOI: 10.1038/ng.721 [PubMed: 21113153]
- Van Otterloo E, Williams T, Artinger KB. The old and new face of craniofacial research: How animal models inform human craniofacial genetic and clinical data. *Dev Biol.* 2016; 415:171–187. DOI: 10.1016/j.ydbio.2016.01.017 [PubMed: 26808208]
- Wang J, Bai Y, Li H, Greene SB, Klysik E, Yu W, Schwartz RJ, Williams TJ, Martin JF. MicroRNA-17-92, a Direct Ap-2a Transcriptional Target, Modulates T-Box Factor Activity in Orofacial Clefting. *PLoS Genet.* 2013; 9:e1003785.doi: 10.1371/journal.pgen.1003785 [PubMed: 24068957]
- Wang L, Shao YY, Ballock RT. Carboxypeptidase Z (CPZ) links thyroid hormone and Wnt signaling pathways in growth plate chondrocytes. *J Bone Miner Res.* 2009; 24:265–273. DOI: 10.1359/jbmr.081014 [PubMed: 18847325]
- Warde-Farley D, Donaldson SL, Comes O, Zuberi K, Badrawi R, Chao P, Franz M, Grouios C, Kazi F, Lopes CT, Maitland A, Mostafavi S, Montojo J, Shao Q, Wright G, Bader GD, Morris Q. The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. *Nucleic Acids Res.* 2010; 38:W214–20. DOI: 10.1093/nar/gkq537 [PubMed: 20576703]
- Warner DR, Mukhopadhyay P, Brock G, Webb CL, Michele Pisano M, Greene RM. MicroRNA expression profiling of the developing murine upper lip. *Dev Growth Differ.* 2014; 56:434–447. DOI: 10.1111/dgd.12140 [PubMed: 24849136]

- Winnier GE, Hargett L, Hogan BL. The winged helix transcription factor MFH1 is required for proliferation and patterning of paraxial mesoderm in the mouse embryo. *Genes Dev.* 1997; 11:926–940. [PubMed: 9106663]
- Wrenn JT, Wessells NK. The early development of mystacial vibrissae in the mouse. *J Embryol Exp Morphol.* 1984; 83:137–156.
- Yu FX, Zhao B, Guan KL. Hippo Pathway in Organ Size Control, Tissue Homeostasis, and Cancer. *Cell.* 2015; 163:811–828. DOI: 10.1016/j.cell.2015.10.044 [PubMed: 26544935]
- Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS.* 2012; 16:284–287. DOI: 10.1089/omi.2011.0118 [PubMed: 22455463]
- Yuan G, Singh G, Chen S, Perez KC, Wu Y, Liu B, Helms JA. Cleft Palate and Aglossia Result From Perturbations in Wnt and Hedgehog Signaling. *Cleft Palate Craniofac J.* 2016; :15–178. DOI: 10.1597/15-178
- Zuberi K, Franz M, Rodriguez H, Montojo J, Lopes CT, Bader GD, Morris Q. GeneMANIA prediction server 2013 update. *Nucleic Acids Res.* 2013; 41:W115–22. DOI: 10.1093/nar/gkt533 [PubMed: 23794635]

Highlights

- A statistically robust and extensively verified multidimensional gene expression resource for mouse facial development.
- Data derived from separated ectoderm and mesenchyme from each facial prominence at E10.5 through E12.5
- A facial ectoderm developmental program highlights cell-cell signaling, epithelial diversification and maturation.
- Facial development co-expression modules identify sub-programs relevant to facial patterning, morphogenesis and the pathogenesis of orofacial clefts.
- A resource for interpreting clinical genetic data pertaining to facial development, morphology and orofacial cleft pathogenesis

**Figure 1.**

Ectoderm-mesenchyme differences dominate the developing mouse face. **A.** Experimental design. Ectoderm and mesenchyme were separated by dispase treatment of microdissected facial prominences from ages E10.5, E11.5 and E12.5. Affymetrix MoGene-1.0-st-v1 microarrays were probed with random primed RNA from triplicate tissue samples. **B.** Principal Component Analysis of facial gene expression. The first component separated samples by tissue layer (ectoderm versus mesenchyme) and accounted for 44.95% of the variation between samples. The second component separated by age and accounted for 12.22% of the variation. The third component, accounting for 8.52% of the variation, did not separate by prominence. **C.** Differentially expressed genes. 9457 probesets in 8575 genes showed differential expression with p -value < 0.01 after multiple testing correction (Benjamini and Hochberg, 1995) by three-way Analysis of Variance (ANOVA). 90% (7722 genes) varied between ectoderm and mesenchyme, 71% (6048 genes) varied by age, 42% (3635 genes) varied by prominence, and 34% (2922 genes) varied by all three. **D.** Ectoderm and mesenchyme gene expression programs. Heatmap of differentially expressed genes after scaling (mean = 0 and standard deviation = 1, color key at lower right) and agglomerative hierarchical clustering using average distance and Pearson correlation coefficient. The array dendrogram (top) shows two major divisions (highlighted with blue and red boxes) corresponding to the ectodermal and mesenchymal tissues. Statistically significant subgroups ($p > 99\%$) are indicated with colored top-bars, and labeled according to their

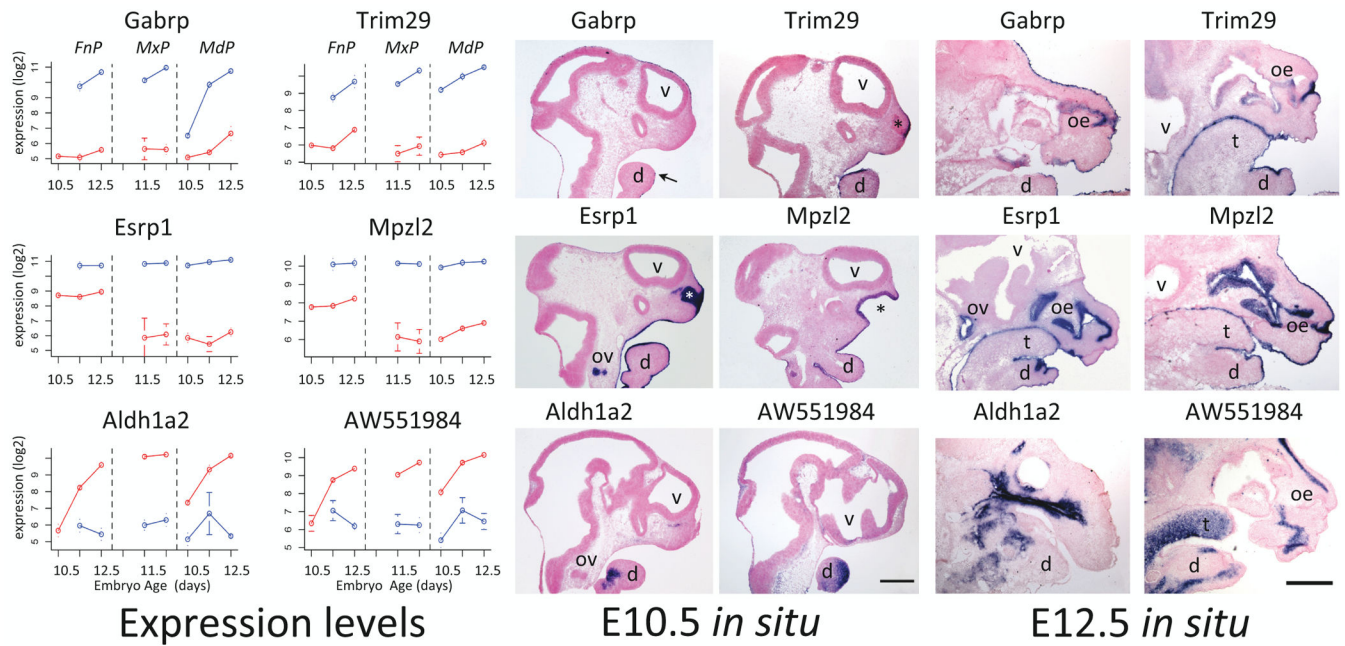
unifying characteristics. The bootstrap probabilities used to derive statistical significance are described in Materials and Methods. Sample labels: b, E10.5; d, E11.5; f, E12.5; n, FnP; p, MdP; x, MxP; e, Ect; m, Mes; number refers to which biological replicate.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Figure 2.**

Validation by in situ hybridization. Left Panel: Mean microarray expression levels in ectoderm (blue) and mesenchyme (red) with error bars indicating standard deviation for four genes with ectoderm-specific expression (Gabrp, Trim29, Esrp1, Mpzl2) and two with mesenchyme-specific expression (Aldh1a2, AW551984). Center Panel: E10.5 in situ hybridization to sagittal sections of the head. Right Panel: E12.5 in situ hybridization to sagittal sections of the face, with mandible (d), tongue (t), olfactory epithelium (oe), otic vesicle (ov) and ventricles (v). Asterisk (*) indicates the nasal pit and the arrow patchy expression of Gabrp in the E10.5 mandible. Ventral is right, rostral is up. Scale bars, 500 microns.

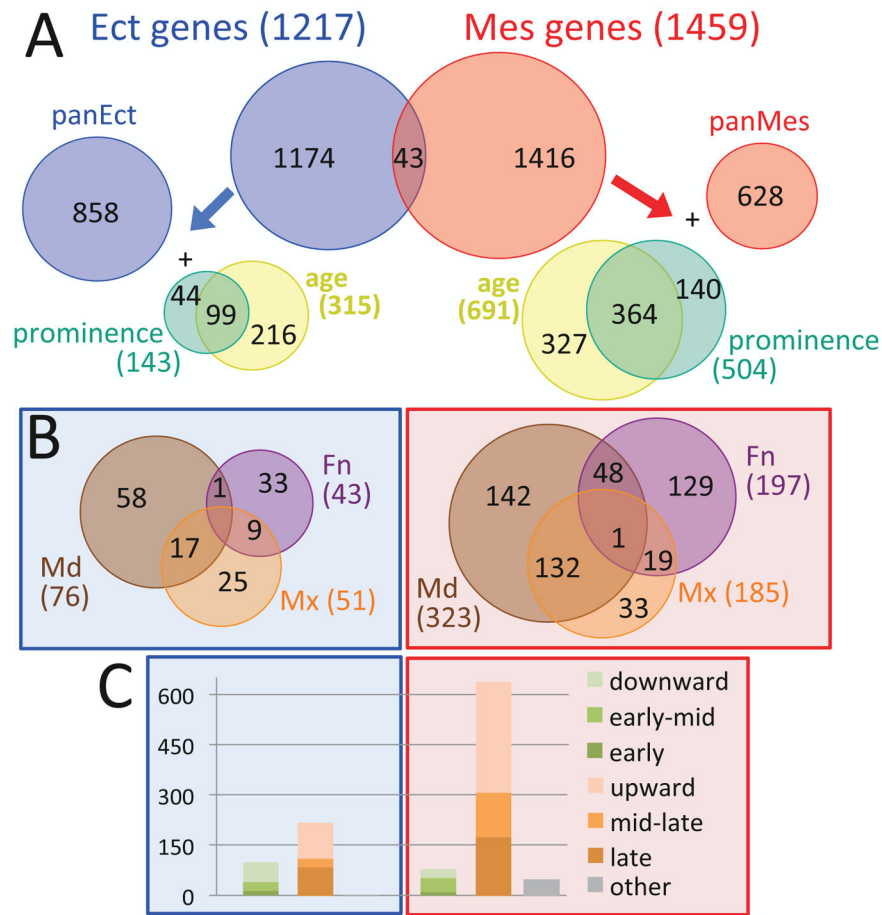


Figure 3. Spatio-temporal dynamics of ectoderm and mesenchyme programs. 2633 genes varied between ectoderm (blue) and mesenchyme (red) in age/prominence-matched conditions ($p_{adj} < 0.01$, fold-change > 2). These defined an ectoderm program of 1217 genes (Table S2) and a mesenchyme program of 1459 genes (Table S3). The Euler diagrams in the top panel (A) show the distribution of constitutively expressed (panEct or panMes), age- and/or prominence-enriched genes in the two programs. The Euler diagrams in the middle panel (B) show the distribution of prominence-enriched genes in the two programs. The stacked barplots in the bottom panel (C) show the temporal trajectories of the age-varying genes in the two programs, with those decreasing over time (shades of green), increasing over time (shades of orange), or more complex (grey). The y-axis is number of genes.

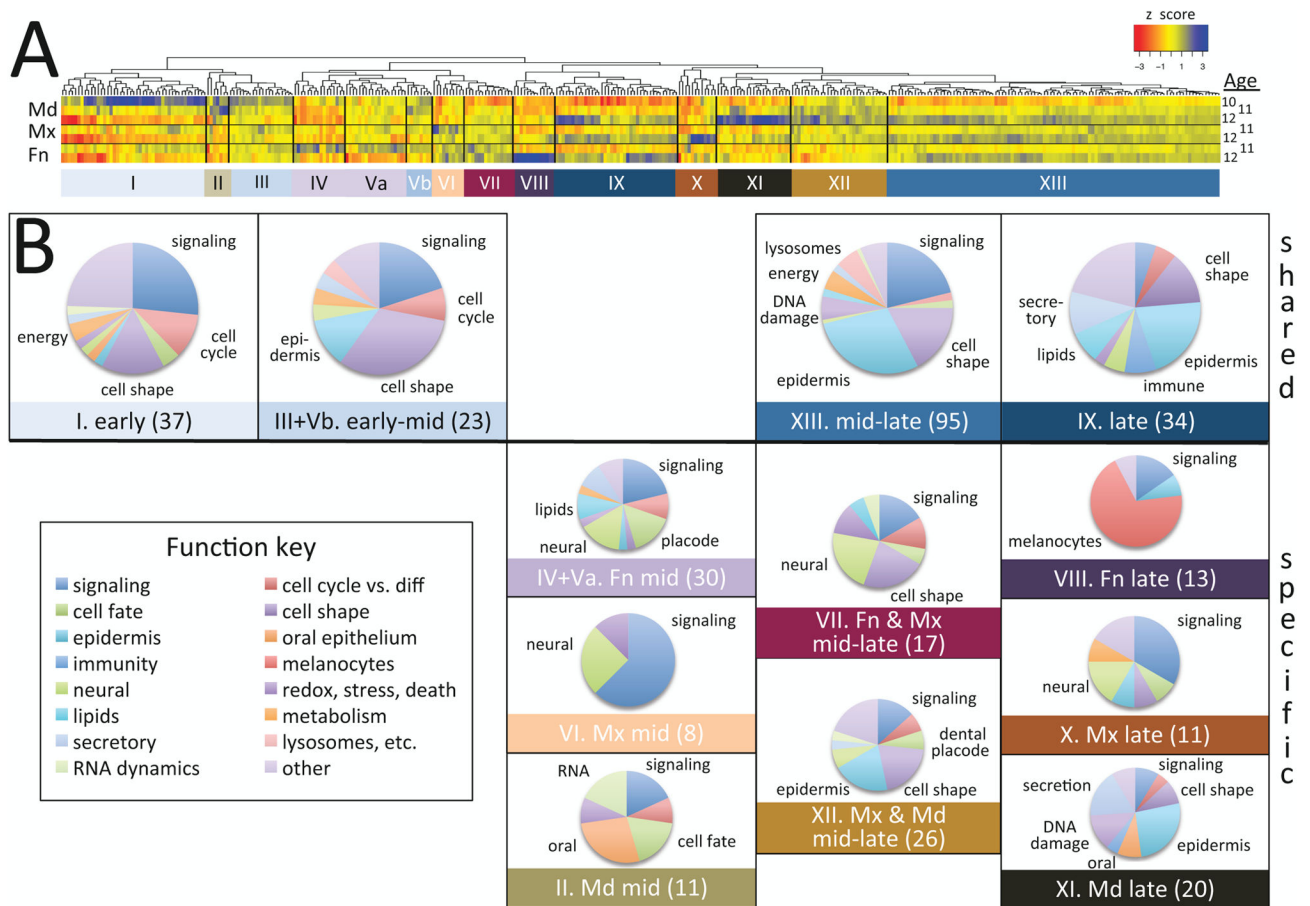


Figure 4. Spatio-temporal aspects of the facial ectodermal program. 359 ectoderm-enriched genes were dynamically expressed. (A) Hierarchical clustering (euclidean distance, complete agglomeration) of the 359 dynamically expressed ectoderm-enriched genes identified 13 spatio-temporal sub-programs. Cluster V, which encompassed two distinct spatio-temporal expression groups, was split into two sub-clusters (Va and Vb). The heatmap displays expression levels for each condition (mean of three samples) across the ectoderm. All values are scaled to mean = 0 and variation from the mean by z score. The dendrogram above the heatmap displays the clustering and the colored bars below the heatmap indicate the clusters based on the dendrogram. (B) Each gene was manually annotated (Table S9) for themes relevant to facial development (Function Key), and the representation of those themes within each cluster is displayed as pie charts. For each cluster, its name (roman numerals), expression characteristic and number of annotated genes is displayed below its pie chart. Clusters that vary only by age (shared between prominences) are in the top row. The lower groups (specific) are prominence- and age-specific. ‘early’ is E10.5; ‘early-mid’ is E10.5 & E11.5; ‘mid’ is E11.5; ‘mid-late’ is E11.5 & E12.5; ‘late’ is E12.5. Note that cluster Vb and was grouped with cluster III, and Va with VI, based on similarity of their spatio-temporal expression patterns.

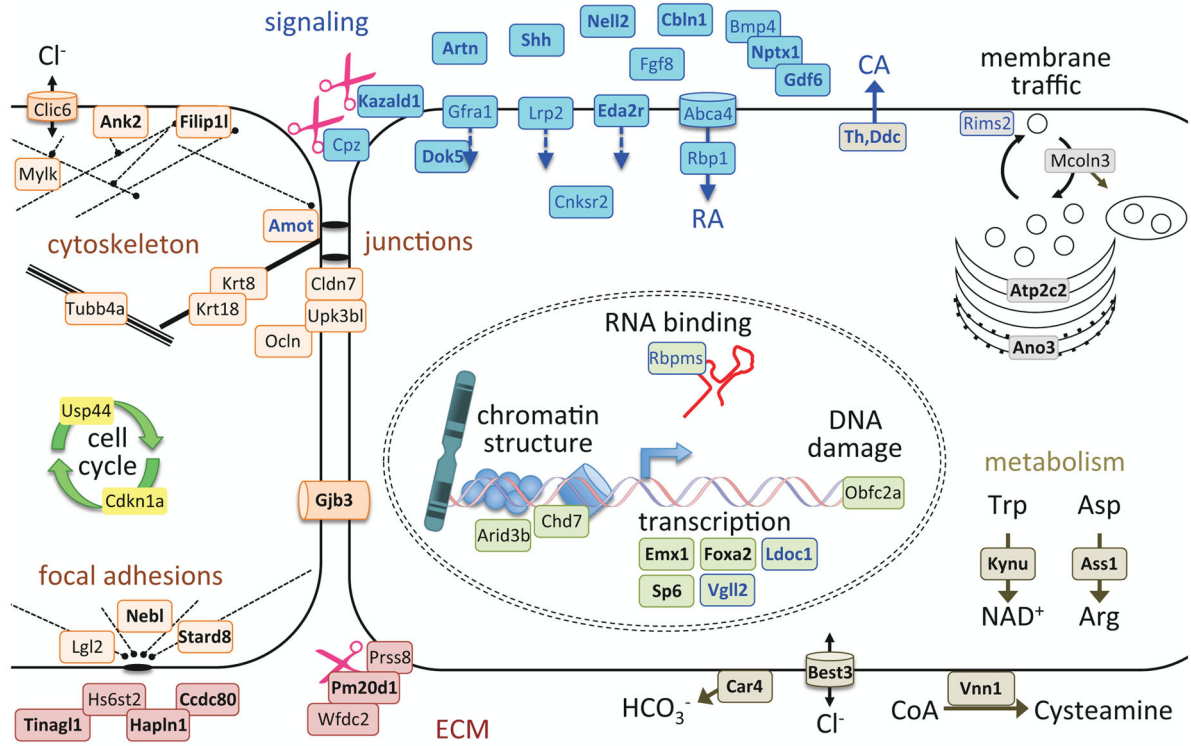


Figure 5. Cellular model of the early ectoderm program. The proteins encoded by genes from the early and early-mid shared clusters (I, III, Vb in Figure 4) were mapped onto this cellular model, based on their annotations (Table S9) and predicted subcellular localization (Fink et al., 2006; Sprenger et al., 2008). Encoded proteins are color-coded according to cellular compartment/function; cell-cell signaling (blue), extracellular matrix and cell-matrix adhesion (red), cell junction and cytoskeletal (peach-orange), cell cycle (yellow), nuclear (laurel-green), metabolic/homeostasis (taupe) and membrane trafficking (light grey). Proteins annotated to two categories are shaded with one color and outlined with another color. Scissors represent protease activity or its modifiers. Dashed segments represent actin fibers; solid segments indicate intermediate filaments; triple segments indicate microtubules. Bold face indicates the genes whose expression is likely to be prominence-specific. Eight early-mid genes were not placed in this model because they lack sufficient annotations: 2310007B03Rik, C330021F23Rik, Ccdc68, Dapl1, Necab1, Tecr1, Tmc4, probeset ids 10412380 and 10508907.

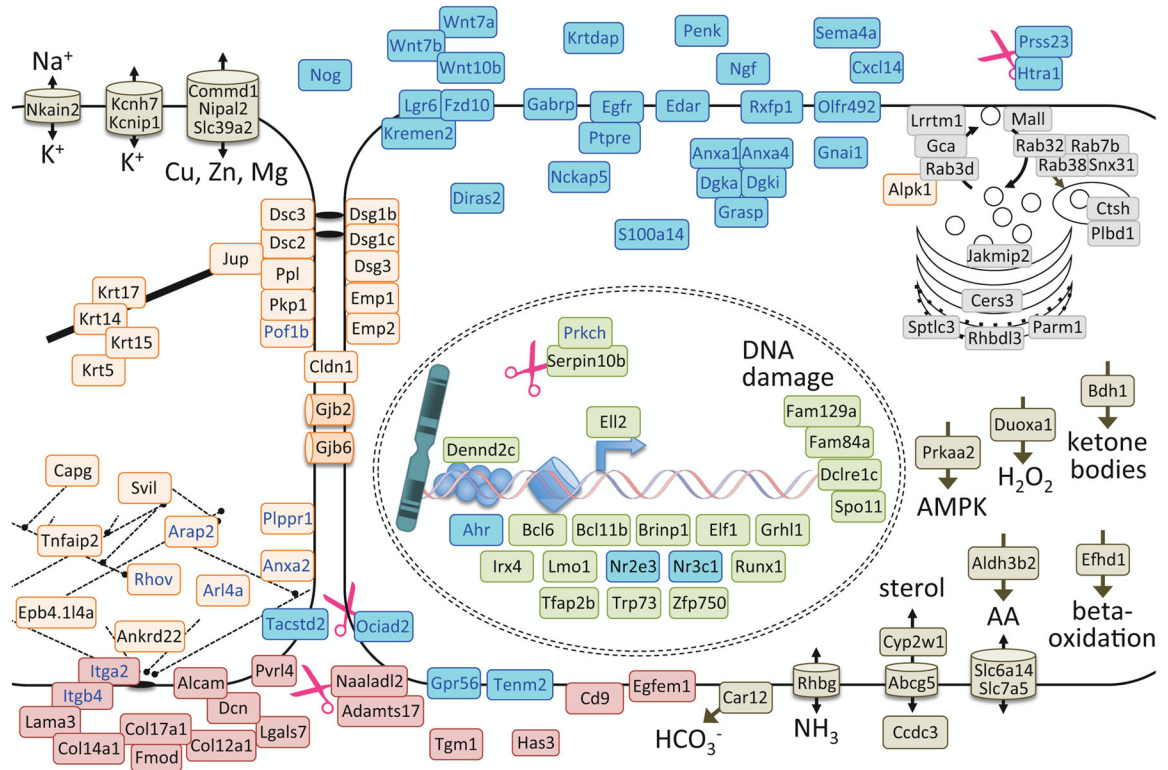


Figure 6. Cellular model of the later ectoderm program. The genes from the mid-late and late clusters (XIII and IX in Figure 4) were mapped onto this cellular model as described in Figure 5. Eight mid-late genes were not placed in this model because they lack sufficient annotations: A730017C20Rik, Bnpl, Fam63b, Gm1087, Gatsl3, Mpped1, probeset Ids 10497501 and 10444589.

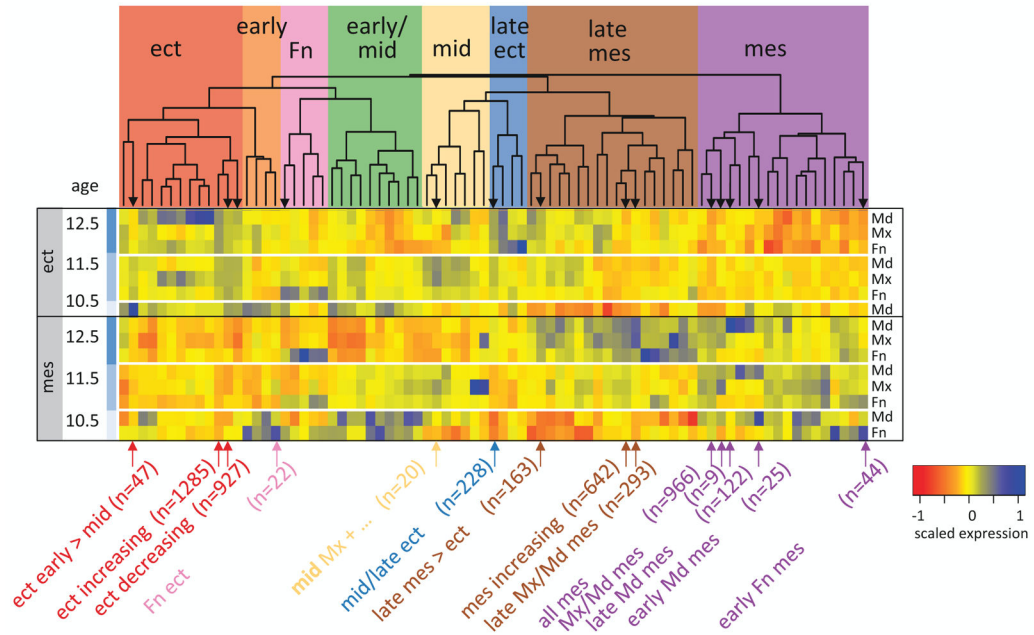


Figure 7. Co-expression modules: spatio-temporal dynamics. The differentially expressed probesets ($p_{adj} < 0.01$) were allocated into 79 co-expression modules using Weighted Gene Co-expression Network Analysis (Langfelder and Horvath, 2012; 2008) with minimum module size 5, eigengene correlation $> 90\%$ for module merging (module gene membership in Table S10). The heatmap displays the consensus expression pattern for each module- plotting the module eigengene mean of three samples for each condition. The modules are grouped by similarity of expression patterns, using hierarchical clustering of the eigengenes, with the dendrogram displayed above the heatmap. The modules with similar expression patterns, based on the dendrogram, were grouped into cohorts and are color-coded and labeled according to the dominant theme. The 14 modules that were enriched in cleft genes are indicated by arrows at the tips of the dendrogram and labeled under the heatmap by its eigengene spatio-temporal dynamics, with number of member genes (n).

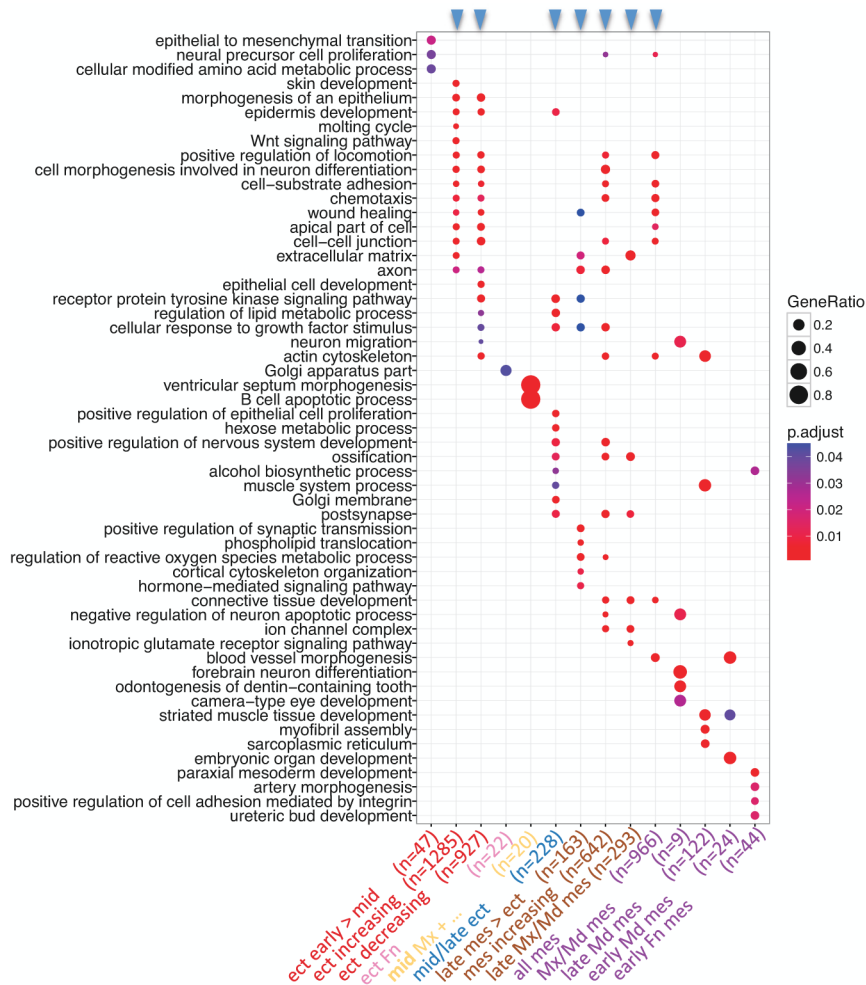


Figure 8. PMC module-associated functions. Each of the 14 PMC modules, which were defined by enrichment of genes associated with OFCs ($p < 0.05$), was tested for enrichment of GO-BP terms and GO-CC terms. The top GO terms from each module (ranked by statistical significance) were then used to compare functions across all PMC modules. Modules are plotted against GO terms with the dot color indicating statistical significance for enrichment, and the dot size indicating fraction of module genes associated with that term. Module order is the same as in Figure 7, with the color of the module label indicating its spatio-temporal expression cohort (from Figure 7), the text label describing its expression characteristics and n indicating the number of genes in the module. The blue arrowheads at the top indicate the larger modules, each of which were enriched for many functional annotations - likely reflecting several functional groups within each of these modules.

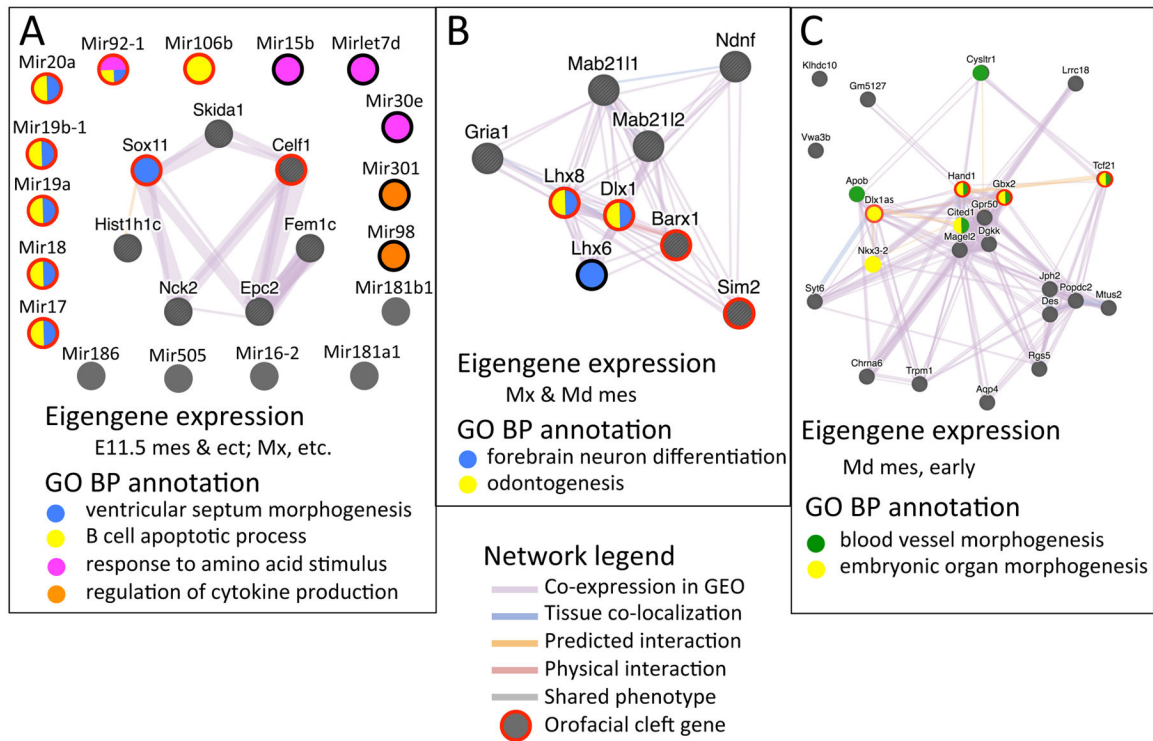


Figure 9.

Network-based Interpretation of functional themes in PMC modules. Gene networks incorporated protein-protein interactions, phenotype similarity, co-expression across GEO datasets, tissue co-localization and protein interactions predicted by cross-species homology, with genes (nodes) color coded according to enriched GO annotations or for OFC genes. Distance between nodes reflects composite similarity of all connections. (A) The ‘mid Mx +’ clefting module (‘orangered3’ in Table S10) included 7 protein coding genes and 17 miRNA genes including three co-transcribed clusters (Mir17/Mir18/Mir19b-1/Mir19a/Mir20a/Mir92-1, Mir15b/Mir16-2 and Mir181a1/Mir181b1). The emergent theme was miRNAs. (B) The ‘Mx/Md mes’ clefting module (‘yellow3’ in Table S10) included 9 genes. The emergent theme was a jaw chondrogenic program, possibly specific to the maxilla. (C) The ‘early Md mes’ clefting module (‘mediumorchid’ in Table S10) included 25 genes. The core network suggested proximo-distal Md patterning while the sub-network suggested cell-cell junctions.

Gene Ontology Biological Process Terms Enriched in genes differentially expressed by Layer. Of 2633 genes that differed by layer relative to age and layer-matched samples (Fig 3; $p\text{-adj} < 0.01$; fold change > 2), 1217 were enriched in ectoderm and 1459 were enriched in mesenchyme. Gene Ontology Biological Process (GO-BP) terms; Ashburner et al., 2000) that were enriched in genes differentially expressed at each age were identified using Database for Annotation, Visualization and Integrated Discovery (DAVID) functional annotation and clustering tools (D. W. Huang et al., 2009b; 2009a). Enriched terms were ranked by information content and p-value and filtered for redundancy (see Materials and Methods for details). The resulting top-ranked terms represented by genes in each layer are shown (blue for ectoderm, red for mesenchyme). Signaling-associated terms are highlighted in purple. Pop.Hits, number of mouse genes associated with the GO term; Count, number of differentially expressed genes associated with the GO term; Fold.Enrich, ratio of differentially expressed genes to all genes associated with GO term; $p\text{-adj}$, p-value after multiple testing correction (Benjamini and Hochberg, 1995); NA, terms were not enriched at this layer.

Ectoderm and mesenchyme gene functions based on enrichment in Gene Ontology Biological Process terms

Gene Ontology Biological Process Term	Ectoderm				Mesenchyme				Best Layer
	Pop.hits	count	fold.enrich	p.adjust	count	fold.enrich	p.adjust	Best Layer	
GO:0001763~morphogenesis of a branching structure	125	30	4.00	3.42EH08	33	3.58	2.85EH08	both	
GO:0045165~cell fate commitment	147	29	3.29	3.50EH06	33	3.04	1.27EH06	both	
GO:0030334~regulation of cell migration	92	17	3.08	2.68EH03	18	2.65	7.76EH03	both	
GO:0042573~retinoic acid metabolic process	17	7	6.87	6.95EH03	6	4.78	8.27EH02	both	
GO:0031128~developmental induction	21	5	3.97	2.94EH01	8	5.16	1.06EH02	both	
GO:0002009~morphogenesis of an epithelium	173	43	4.14	1.91EH12	33	2.58	4.17EH05	Ect	
GO:0022612~gland morphogenesis	84	28	5.56	7.22EH11	17	2.74	7.97EH03	Ect	
GO:0007398~ectoderm development	133	32	4.01	7.61EH09	NA	NA	NA	Ect	
GO:0042476~odontogenesis	45	17	6.30	4.09EH07	NA	NA	NA	Ect	
GO:0030855~epithelial cell differentiation	123	28	3.80	4.23EH07	16	1.76	2.92EH01	Ect	
GO:0045664~regulation of neuron differentiation	102	25	4.09	6.12EH07	15	1.99	1.69EH01	Ect	
GO:0007242~intracellular signaling cascade	915	99	1.80	9.35EH07	88	1.30	9.48EH02	Ect	
GO:0007431~salivary gland development	31	14	7.53	9.63EH07	7	3.06	2.14EH01	Ect	
GO:0043583~ear development	104	24	3.85	3.42EH06	20	2.61	4.46EH03	Ect	
GO:0007156~homophilic cell adhesion	117	23	3.28	8.06EH05	21	2.43	7.28EH03	Ect	
GO:0060562~epithelial tube morphogenesis	111	22	3.30	1.16EH04	17	2.07	9.17EH02	Ect	
GO:0016053~Wnt receptor signaling pathway	130	24	3.08	1.25EH04	21	2.19	2.33EH02	Ect	

Gene Ontology Biological Process Term	Ectoderm				Mesenchyme				Best Layer
	Pop.hits	count	fold.enrich	p.adjust	count	fold.enrich	p.adjust	Best Layer	
GO:0001942~hair follicle development	50	14	4.67	2.47EH04	NA	NA	NA	Ect	
GO:0007223~Wnt receptor signaling pathway, calcium modulating pathway	20	9	7.50	4.41EH04	NA	NA	NA	Ect	
GO:0016331~morphogenesis of embryonic epithelium	78	17	3.63	4.89EH04	12	2.08	2.39EH01	Ect	
GO:0060512~prostate gland morphogenesis	27	10	6.17	6.21EH04	6	3.01	3.35EH01	Ect	
GO:0048638~regulation of developmental growth	37	11	4.96	1.39EH03	8	2.93	1.67EH01	Ect	
GO:0031076~embryonic camera-type eye development	15	7	7.78	3.45EH03	5	4.52	1.95EH01	Ect	
GO:0001738~morphogenesis of a polarized epithelium	11	6	9.09	5.89EH03	NA	NA	NA	Ect	
GO:0001736~establishment of planar polarity	5	4	13.34	3.35EH02	NA	NA	NA	Ect	
GO:0030509~BMP signaling pathway	24	7	4.86	3.98EH02	NA	NA	NA	Ect	
GO:0040036~regulation of fibroblast growth factor receptor signaling pathway	11	5	7.58	4.78EH02	NA	NA	NA	Ect	
GO:0048705~skeletal system morphogenesis	130	15	1.92	2.33EH01	39	4.06	2.02EH11	Mes	
GO:0030198~extracellular matrix organization	101	14	2.31	9.50EH02	34	4.56	2.43EH11	Mes	
GO:0051216~cartilage development	78	15	3.21	4.59EH03	29	5.04	1.19EH10	Mes	
GO:0001525~angiogenesis	133	23	2.88	4.98EH04	37	3.77	6.96EH10	Mes	
GO:0016477~cell migration	240	29	2.01	1.16EH02	51	2.88	1.90EH09	Mes	
GO:0035107~appendage morphogenesis	115	16	2.32	5.47EH02	33	3.89	3.38EH09	Mes	
GO:0007409~axonogenesis	163	26	2.66	5.05EH04	39	3.24	1.33EH08	Mes	
GO:0008284~positive regulation of cell proliferation	284	38	2.23	3.10EH04	54	2.58	2.51EH08	Mes	
GO:0060537~muscle tissue development	136	17	2.08	9.66EH02	33	3.29	2.39EH07	Mes	
GO:0048010~vascular endothelial growth factor receptor signaling pathway	14	NA	NA	NA	11	10.64	2.78EH07	Mes	
GO:0010811~positive regulation of cell-substrate adhesion	27	NA	NA	NA	14	7.02	8.93EH07	Mes	
GO:0007169~transmembrane receptor protein tyrosine kinase signaling pathway	192	26	2.26	4.73EH03	38	2.68	2.57EH06	Mes	
GO:0014032~neural crest cell development	33	10	5.05	2.65EH03	13	5.34	7.13EH05	Mes	
GO:0002062~chondrocyte differentiation	21	NA	NA	NA	10	6.45	2.50EH04	Mes	
GO:0030111~regulation of Wnt receptor signaling pathway	39	7	2.99	2.58EH01	13	4.52	4.19EH04	Mes	
GO:0050921~positive regulation of chemotaxis	11	NA	NA	NA	7	8.62	1.42EH03	Mes	
GO:0001756~somitogenesis	41	NA	NA	NA	12	3.97	3.26EH03	Mes	
GO:0010464~regulation of mesenchymal cell proliferation	23	6	4.35	1.25EH01	9	5.30	3.70EH03	Mes	
GO:0048634~regulation of muscle development	43	7	2.71	3.36EH01	12	3.78	4.92EH03	Mes	

Gene Ontology Biological Process Term	Ectoderm				Mesenchyme				Best Layer
	Pop.hits	count	fold.enrich	p.adjust	count	fold.enrich	p.adjust	Best Layer	
GO:0060638~mesenchymal-epithelial cell signaling	6	NA	NA	NA	5	11.29	8.34EH03	Mes	
GO:0048565~gut development	39	6	2.56	5.04EH01	11	3.82	8.38EH03	Mes	
GO:0007229~integrin-mediated signaling pathway	76	NA	NA	NA	15	2.67	2.12EH02	Mes	
GO:0050679~positive regulation of epithelial cell proliferation	31	NA	NA	NA	9	3.93	2.47EH02	Mes	

Gene Ontology Biological Process Terms Enriched in genes differentially expressed by Age. 1574 genes were differentially expressed by age (Fig 3; p.adj < 0.01; fold change > 2) relative to prominence and layer-matched samples. 342 differed at E10.5, 598 differed at E11.5 & 1122 differed at E12.5. The top-ranked GO Biological Process terms for E10.5, E11.5 and E12.5 are respectively highlighted with light green, teal and dark green. Analysis and abbreviations are described in Table 1.

Age+dependent gene functions based on enrichment in Gene Ontology Biological Process terms

GO Biological Process Term	E 10.5					E 11.5					E 12.5				
	Pop.Hits	Count	Fold.Enrich	p.adj	Count	Pop.Hits	Count	Fold.Enrich	p.adj	Count	Pop.Hits	Count	Fold.Enrich	p.adj	Count
GO:0001763~morphogenesis of a branching structure	125	14	6.56	1.62E104	17	4.22	6.11E105	24	3.19	5.88E105	all				
GO:0009069~serine family amino acid metabolic process	22	5	13.31	2.56E102	NA	NA	NA	NA	NA	NA	10.5				
GO:0044271~nitrogen compound biosynthetic process	302	15	2.91	3.11E102	NA	NA	NA	NA	NA	NA	10.5				
GO:0006544~glycine metabolic process	12	4	19.52	4.17E102	NA	NA	NA	NA	NA	NA	10.5				
GO:0045165~cell fate commitment	147	15	5.98	1.10E104	25	5.28	1.07E108	17	1.92	1.77E101	10.5, 11.5				
GO:0048010~vascular endothelial growth factor receptor signaling pathway	14	3	12.55	2.71E101	5	11.08	1.18E102	NA	NA	NA	10.5, 11.5				
GO:0007423~sensory organ development	257	13	2.96	5.75E102	37	4.47	7.84E111	29	1.87	3.09E102	11.5				
GO:0001656~metanephros development	58	7	7.07	2.49E102	16	8.56	3.01E108	15	4.29	2.26E104	11.5				
GO:0035107~appendage morphogenesis	115	11	5.60	5.22E103	21	5.67	6.37E108	17	2.45	2.64E102	11.5				
GO:0045664~regulation of neuron differentiation	102	5	2.87	5.80E101	18	5.47	1.05E106	12	1.95	3.55E101	11.5				
GO:0050678~regulation of epithelial cell proliferation	63	4	3.72	5.69E101	13	6.40	1.78E105	12	3.16	2.34E102	11.5				
GO:0014033~neural crest cell differentiation	33	4	7.10	2.30E101	10	9.40	1.82E105	7	3.52	1.53E101	11.5				
GO:0021537~telencephalon development	72	NA	NA	NA	12	5.17	3.45E104	11	2.53	1.31E101	11.5				
GO:0001756~somitogenesis	41	NA	NA	NA	9	6.81	7.52E104	9	3.64	4.53E102	11.5				
GO:0042573~retinoic acid metabolic process	17	NA	NA	NA	6	10.95	2.53E103	4	3.90	5.00E101	11.5				
GO:0030111~regulation of Wnt receptor signaling pathway	39	NA	NA	NA	8	6.36	3.52E103	9	3.83	3.44E102	11.5				
GO:0021983~pituitary gland development	29	NA	NA	NA	7	7.49	4.39E103	5	2.86	5.48E101	11.5				
GO:0045445~myoblast differentiation	12	3	14.64	2.20E101	5	12.93	6.67E103	NA	NA	NA	11.5				
GO:0016337~cell-cell adhesion	236	NA	NA	NA	32	4.21	8.19E109	48	3.37	9.61E111	11.5, 12.5				
GO:0007409~axonogenesis	163	10	3.59	6.40E102	26	4.95	1.28E108	38	3.87	3.22E110	11.5, 12.5				

GO Biological Process Term	E 10.5					E 11.5					E 12.5						
	Pop.Hits	Count	Fold.Enrich	p.adj	Count	Fold.Enrich	p.adj	Count	Fold.Enrich	p.adj	Count	Fold.Enrich	p.adj	Count	Fold.Enrich	p.adj	best age
GO:0045944~positive regulation of transcription from RNA polymerase II promoter	358	15	2.45	8.89E102	37	3.21	9.45E108	53	2.46	2.76E107	53	2.46	2.76E107	53	2.46	2.76E107	11.5, 12.5
GO:0030198~extracellular matrix organization	101	6	3.48	3.15E101	15	4.61	9.38E105	26	4.27	1.04E107	26	4.27	1.04E107	26	4.27	1.04E107	11.5, 12.5
GO:0007169~transmembrane receptor protein tyrosine kinase signaling pathway	192	8	2.44	4.04E101	23	3.72	8.06E106	32	2.77	2.01E105	32	2.77	2.01E105	32	2.77	2.01E105	11.5, 12.5
GO:0007267~cell-cell signaling	290	12	2.42	1.69E101	26	2.78	1.59E104	43	2.46	5.21E106	43	2.46	5.21E106	43	2.46	5.21E106	11.5, 12.5
GO:0048514~blood vessel morphogenesis	198	11	3.25	6.89E102	21	3.29	1.31E104	29	2.43	7.07E104	29	2.43	7.07E104	29	2.43	7.07E104	11.5, 12.5
GO:0014706~striated muscle tissue development	127	NA	NA	NA	16	3.91	2.87E104	24	3.14	7.58E105	24	3.14	7.58E105	24	3.14	7.58E105	11.5, 12.5
GO:0022612~gland morphogenesis	84	8	5.58	2.86E102	13	4.80	3.00E104	20	3.95	1.98E105	20	3.95	1.98E105	20	3.95	1.98E105	11.5, 12.5
GO:0048565~gut development	39	4	6.01	3.11E101	8	6.36	3.52E103	10	4.25	9.80E103	10	4.25	9.80E103	10	4.25	9.80E103	11.5, 12.5
GO:0016055~Wnt receptor signaling pathway	130	NA	NA	NA	14	3.34	4.51E103	22	2.81	9.97E104	22	2.81	9.97E104	22	2.81	9.97E104	11.5, 12.5
GO:0030509~BMP signaling pathway	24	3	7.32	4.70E101	6	7.76	1.22E102	7	4.84	4.22E102	7	4.84	4.22E102	7	4.84	4.22E102	11.5, 12.5
GO:0060324~face development	17	NA	NA	NA	5	9.12	2.34E102	7	6.83	8.10E103	7	6.83	8.10E103	7	6.83	8.10E103	11.5, 12.5
GO:0042476~odontogenesis	45	4	5.21	3.76E101	7	4.83	3.79E102	12	4.42	1.68E103	12	4.42	1.68E103	12	4.42	1.68E103	11.5, 12.5
GO:0051145~smooth muscle cell differentiation	10	NA	NA	NA	4	12.41	4.13E102	5	8.30	3.56E102	5	8.30	3.56E102	5	8.30	3.56E102	11.5, 12.5
GO:0060348~bone development	118	6	2.98	4.21E101	16	4.21	1.22E104	29	4.08	3.22E108	29	4.08	3.22E108	29	4.08	3.22E108	12.5
GO:0051216~cartilage development	78	NA	NA	NA	13	5.17	1.43E104	22	4.68	3.81E107	22	4.68	3.81E107	22	4.68	3.81E107	12.5
GO:0042692~muscle cell differentiation	117	NA	NA	NA	15	3.98	4.50E104	25	3.55	5.14E106	25	3.55	5.14E106	25	3.55	5.14E106	12.5
GO:0008544~epidermis development	125	NA	NA	NA	10	2.48	1.68E101	24	3.19	5.88E105	24	3.19	5.88E105	24	3.19	5.88E105	12.5
GO:0007178~transmembrane receptor protein serine/threonine kinase signaling pathway	78	5	3.75	3.94E101	10	3.98	1.22E102	18	3.83	1.15E104	18	3.83	1.15E104	18	3.83	1.15E104	12.5
GO:0048634~regulation of muscle development	43	NA	NA	NA	NA	NA	NA	13	5.02	2.15E104	13	5.02	2.15E104	13	5.02	2.15E104	12.5
GO:0048638~regulation of developmental growth	37	NA	NA	NA	5	4.19	2.32E101	12	5.38	2.54E104	12	5.38	2.54E104	12	5.38	2.54E104	12.5
GO:0051797~regulation of hair follicle development	7	NA	NA	NA	3	13.30	1.69E101	5	11.85	9.61E103	5	11.85	9.61E103	5	11.85	9.61E103	12.5
GO:0008593~regulation of Notch signaling pathway	8	NA	NA	NA	3	11.63	2.05E101	5	10.37	1.57E102	5	10.37	1.57E102	5	10.37	1.57E102	12.5

Table 3

Gene Ontology Biological Process Terms Enriched in genes differentially expressed by Prominence. 1026 genes were differentially expressed by prominence (Fig 3; $p_{\text{adj}} < 0.01$; fold change > 2) relative to age and layer-matched samples. 556 differed in FnP, 265 differed in MxP, and 424 differed in MdP. The top-ranked GO-BP terms represented by genes enriched in each prominence layer are shown (fuchsia for FnP, orange for MxP, tan for MdP). Analysis and abbreviations are described in Table 1.

Prominence0specific gene functions based on enrichment in Gene Ontology Biological Process terms

GO Biological Process Term	FnP			MxP			MdP			bestProm	
	Pop.Hits	Count	Fold.Enrich	p.adj	Count	Fold.Enrich	p.adj	Count	Fold.Enrich		p.adj
GO:0045165~cell fate commitment	147	26	5.89	1.61EJ10	17	7.74	4.80EJ08	20	6.14	2.93EJ08	all
GO:0016477~cell migration	240	28	3.89	1.93EJ07	16	4.46	7.02EJ05	18	3.39	4.82EJ04	all
GO:0045944~positive regulation of transcription from RNA polymerase II promoter	358	31	2.88	1.16EJ05	26	4.86	1.53EJ08	35	4.41	1.69EJ10	all
GO:0050767~regulation of neurogenesis	132	19	4.79	3.02EJ06	12	6.09	9.94EJ05	13	4.45	6.75EJ04	all
GO:0051216~cartilage development	78	14	5.98	1.40EJ05	10	8.58	5.42EJ05	11	6.37	1.83EJ04	all
GO:0022612~gland morphogenesis	84	10	3.96	1.43EJ02	9	7.17	5.96EJ04	7	3.76	8.79EJ02	all
GO:0030178~negative regulation of Wnt receptor signaling pathway	26	6	7.69	1.45EJ02	5	12.87	6.32EJ03	6	10.42	3.05EJ03	all
GO:0007423~sensory organ development	257	42	5.44	2.85EJ16	19	4.95	2.00EJ06	23	4.04	2.13EJ06	Fn
GO:0007409~axonogenesis	163	31	6.33	1.85EJ13	17	6.98	1.48EJ07	16	4.43	8.32EJ05	Fn
GO:0060429~epithelium development	271	33	4.06	2.40EJ09	18	4.45	1.56EJ05	22	3.66	1.76EJ05	Fn
GO:0014033~neural crest cell differentiation	33	13	13.12	6.67EJ09	7	14.20	1.67EJ04	9	12.31	1.28EJ05	Fn
GO:0035107~appendage morphogenesis	115	19	5.50	4.33EJ07	8	4.66	1.60EJ02	13	5.10	1.96EJ04	Fn
GO:0045664~regulation of neuron differentiation	102	18	5.88	4.56EJ07	9	5.91	1.97EJ03	11	4.87	1.40EJ03	Fn
GO:0001763~morphogenesis of a branching structure	125	18	4.80	6.56EJ06	11	5.89	3.31EJ04	12	4.33	1.63EJ03	Fn
GO:0001656~metanephros development	58	12	6.89	2.72EJ05	7	8.08	2.89EJ03	8	6.23	3.56EJ03	Fn
GO:0043584~nose development	9	5	18.50	1.76EJ03	NA	NA	NA	NA	NA	NA	Fn
GO:0042063~gliogenesis	52	9	5.76	2.65EJ03	5	6.44	5.91EJ02	NA	NA	NA	Fn
GO:0008038~neuron recognition	12	5	13.88	5.60EJ03	NA	NA	NA	NA	NA	NA	Fn
GO:0042438~melanin biosynthetic process	10	4	13.32	3.64EJ02	NA	NA	NA	NA	NA	NA	Fn
GO:0021537~telencephalon development	72	13	6.01	3.69EJ05	11	10.23	3.27EJ06	6	3.76	1.55EJ01	Fn, Mx
GO:0021983~pituitary gland development	29	7	8.04	3.34EJ03	5	11.54	9.07EJ03	4	6.23	1.80EJ01	Fn, Mx

GO Biological Process Term	FnP			MxP			MdP			bestProm	
	Pop.Hits	Count	Fold.Enrich	p.adj	Count	Fold.Enrich	p.adj	Count	Fold.Enrich		p.adj
GO:0021761~limbic system development	33	4	4.04	4.57EJ01	7	14.20	1.67EJ04	NA	NA	NA	Mx
GO:0030111~regulation of Wnt receptor signaling pathway	39	7	5.98	1.53EJ02	7	12.01	4.12EJ04	6	6.95	1.67EJ02	Mx
GO:0009954~proximal/distal pattern formation	25	NA	NA	NA	6	16.06	5.25EJ04	NA	NA	NA	Mx
GO:0042573~retinoic acid metabolic process	17	3	5.88	5.10EJ01	5	19.69	1.48EJ03	NA	NA	NA	Mx
GO:0050804~regulation of synaptic transmission	100	9	3.00	1.05EJ01	9	6.02	1.77EJ03	NA	NA	NA	Mx
GO:0060021~palate development	35	4	3.81	4.95EJ01	5	9.56	1.68EJ02	4	5.16	2.63EJ01	Mx
GO:0008593~regulation of Notch signaling pathway	8	3	12.49	1.91EJ01	3	25.10	4.86EJ02	NA	NA	NA	Mx
GO:0048705~skeletal system morphogenesis	130	14	3.59	2.63EJ03	14	7.21	2.32EJ06	17	5.90	1.02EJ06	Mx, Md
GO:0042476~odontogenesis	45	5	3.70	3.34EJ01	7	10.41	8.30EJ04	8	8.03	9.48EJ04	Mx, Md
GO:0001756~somitogenesis	41	NA	NA	NA	6	9.80	4.29EJ03	8	8.81	5.43EJ04	Mx, Md
GO:0014706~striated muscle tissue development	127	10	2.62	1.35EJ01	11	5.80	3.75EJ04	27	9.60	2.39EJ15	Md
GO:0048706~embryonic skeletal system development	83	12	4.82	7.87EJ04	9	7.26	5.60EJ04	16	8.70	1.86EJ08	Md
GO:0048514~blood vessel morphogenesis	198	14	2.35	7.25EJ02	9	3.04	7.37EJ02	21	4.79	6.18EJ07	Md
GO:0001503~ossification	106	8	2.51	3.07EJ01	8	5.05	1.06EJ02	15	6.39	2.84EJ06	Md
GO:0030198~extracellular matrix organization	101	7	2.31	4.84EJ01	9	5.96	1.88EJ03	14	6.26	9.82EJ06	Md
GO:0048638~regulation of developmental growth	37	NA	NA	NA	NA	NA	NA	7	8.54	2.14EJ03	Md
GO:0045109~intermediate filament organization	13	3	7.69	3.83EJ01	NA	NA	NA	5	17.36	2.15EJ03	Md
GO:0048565~gut development	39	NA	NA	NA	4	6.87	1.36EJ01	7	8.10	2.74EJ03	Md
GO:0060638~mesenchymal-epithelial cell signaling	6	NA	NA	NA	NA	NA	NA	4	30.10	2.78EJ03	Md
GO:0048634~regulation of muscle development	43	NA	NA	NA	NA	NA	NA	7	7.35	4.23EJ03	Md