



OPEN

Transcriptional analysis of cleft palate in TGF β 3 mutant mice

J. Liu^{1,4}, S. K. Chanumolu^{2,4}, K. M. White³, M. Albahrani², H. H. Otu² & A. Nawshad¹✉

Cleft palate (CP) is one of the most common craniofacial birth defects, impacting about 1 in 800 births in the USA. Tgf- β 3 plays a critical role in regulating murine palate development, and *Tgf- β 3* null mutants develop cleft palate with 100% penetrance. In this study, we compared global palatal transcriptomes of wild type (WT) and *Tgf- β 3* $-/-$ homozygous (HM) mouse embryos at the crucial palatogenesis stages of E14.5, and E16.5, using RNA-seq data. We found 1,809 and 2,127 differentially expressed genes at E16.5 vs. E14.5 in the WT and HM groups, respectively (adjusted $p < 0.05$; |fold change| > 2.0). We focused on the genes that were uniquely up/downregulated in WT or HM at E16.5 vs. E14.5 to identify genes associated with CP. Systems biology analysis relating to cell behaviors and function of WT and HM specific genes identified functional non-Smad pathways and preference of apoptosis to epithelial-mesenchymal transition. We identified 24 HM specific and 11 WT specific genes that are CP-related and/or involved in Tgf- β 3 signaling. We validated the expression of 29 of the 35 genes using qRT-PCR and the trend of mRNA expression is similar to that of RNA-seq data. Our results enrich our understanding of genes associated with CP that are directly or indirectly regulated via TGF- β .

Orofacial clefting is the most common craniofacial anomaly treated in pediatric hospitals and is the second most common birth defect, with a prevalence ranging from 1/500 to 1/2,500 in humans^{1,2}. Formation of a confluent palate is a precise orchestration of many palatal cellular processes, including cellular movement, cell death, and cell cycle progression^{3,4}. Many genes have been implicated in the etiology of cleft palate (CP). Transforming growth factor-beta (*Tgf- β*) isoforms (1, 2, and 3) are essential for proper development, including palate fusion^{5,6}. The roles of TGF- β in palatogenesis, participating in different phases of palate development, such as elevation, contact, and fusion, have been detailed previously³. A number of studies suggest that TGF- β 3 is a candidate gene for causing cleft palate in mice², chickens⁷, and humans⁸. The *Tgf- β 3* homozygous knockout (stated as homozygous, HM in this study) mouse model presents the phenotype of CP but no other major anomalies⁹.

The process of palatogenesis is remarkably similar among vertebrates. In mice, the palatal shelves grow out bilaterally from the internal surfaces of the maxillary processes (~E11.5). The shelves first vertically elongate on either side of the tongue (~E12.5) and then elevate, becoming horizontal above the tongue as the tongue descends (~E13.5). When the opposing shelves approach each other, the cells of the outer layer (periderm) covering the medial edge epithelia (MEE) slough off (~E14.0), exposing the lateral surfaces of the underlying basal MEE cells to close contact with each other (also known as “adherence” or “fusion”), promoting formation of the midline epithelial seam (MES) (~E14.5; Fig. 1A–C). The palatal epithelial seam subsequently disintegrates and reaches complete palatal mesenchymal confluency (by ~E16.5; Fig. 1G,H). This process is governed by epithelial mesenchymal transition (EMT) and/or apoptosis⁴, resulting in the mesenchymal portion of the two palatal shelves becoming continuous.

While both the wild type (WT) and HM palates proceed through the exact same stages of palatogenesis until E14.0 (i.e., growth, elongation, and elevation), the persistence of periderm on the *Tgf- β 3* HM palates at the final “adherence/fusion” stage at E14.5 hinders palatal MEE contact between opposite palatal shelves (~E14.5, Fig. 1D–F) resulting in palatal cleft (Fig. 1I, J). It is evident that the MEE contact is strictly regulated by *Tgf- β 3* in WT to ensure desquamation of the periderm for immaculate fusion^{10,11}. The absence of *Tgf- β 3* fails to facilitate desquamation of the periderm and hinders fusion resulting in subsequent palatal cleft in HM. Therefore, in this study, we aim to assess the difference in the temporal gene expression for the WT and *Tgf- β 3* HM samples from

¹Department of Oral Biology, College of Dentistry, University of Nebraska Medical Center, Lincoln, NE 68583, USA. ²Department of Electrical and Computer Engineering, University of Nebraska-Lincoln, Lincoln, NE 68588, USA. ³Department of Growth and Development, College of Dentistry, University of Nebraska Medical Center College of Dentistry, Lincoln, NE 68583, USA. ⁴These authors contributed equally: J. Liu and S. K. Chanumolu. ✉email: anawshad@unmc.edu

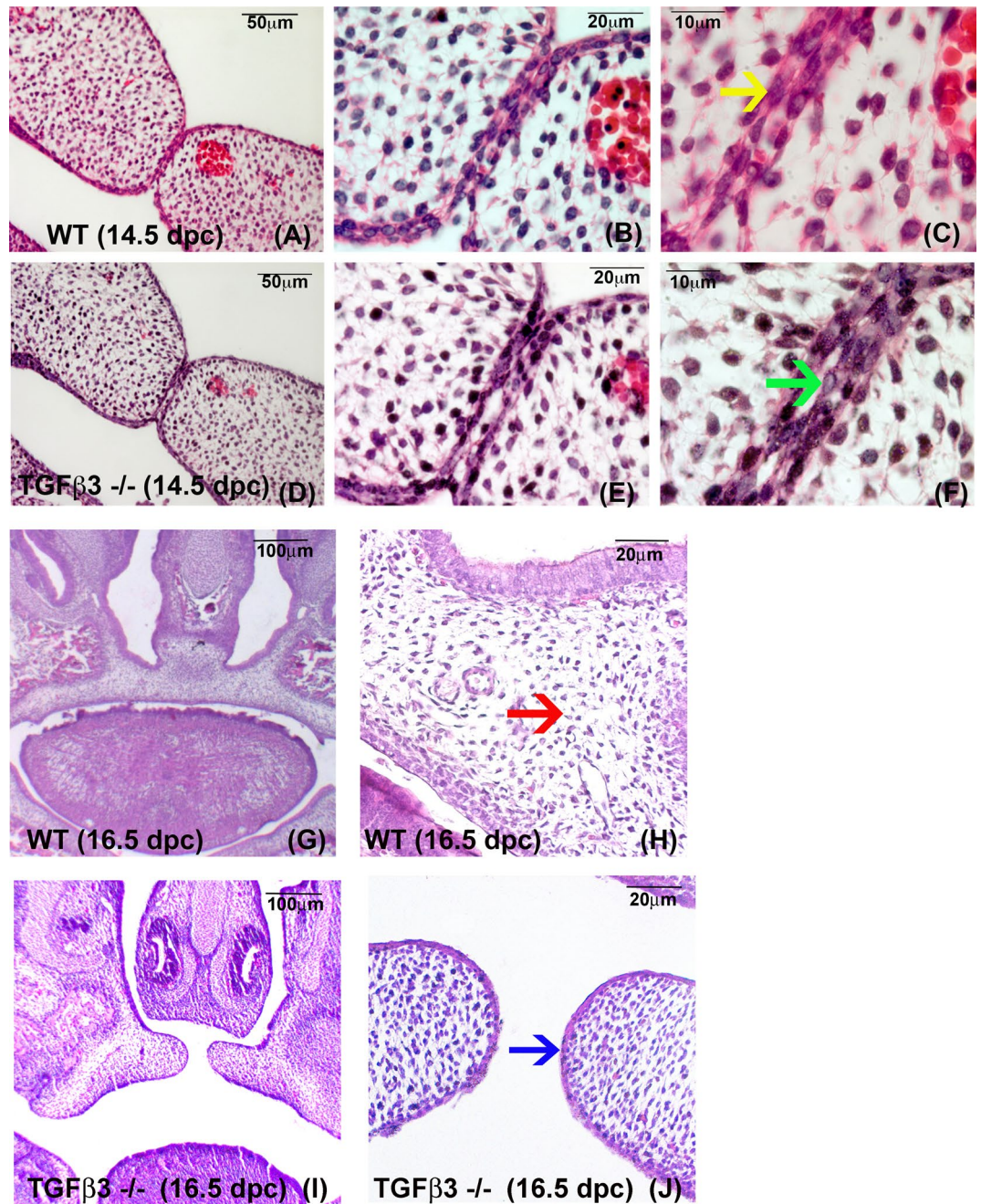


Figure 1. Histological illustration of murine palate formation in WT (A–C, G, H) and $TGF\beta 3^{-/-}$ (HM) mice (D–F, I, J). At 14.5 dpc, in WT palates the mid-palate region shows a complete union of a two-cell thick, basal medial edge epithelium (MEE), consisting of two opposite palates forming a tight epithelial seam in low (A) and higher magnifications (B); (C, yellow arrow). At 14.5 dpc, the HM ($TGF\beta 3^{-/-}$; D–F) palates demonstrate a trapped additional layer of flattened epithelium between the MEE of opposite palatal shelves (D); (E); (F, green arrow) [Figs. A–F are from *J. Cell. Physiol.* 230: 1,212–1,225, 2015. Wiley Periodicals, Inc., copyright John Wiley & Sons, Inc.]. At 16.5 dpc, the WT palates show complete disintegration of the seam and palatal mesenchymal confluence (G), higher magnification (H, red arrow). At 16.5 dpc, $TGF\beta 3^{-/-}$ (HM) palates fail to fuse (I), higher magnification (J), and drift apart as face continues to grow, resulting in cleft palate consequently. In the HM palates at 16.5 dpc, the palatal epithelia are stratified (J, blue arrow).

E14.5 to E16.5 to provide further understanding of genes functionally regulated by $TGF-\beta 3$ during murine palatal fusion and examine their contribution to the development of CP using our previously established RNA-seq data¹².

Several studies have explored CP-related genes using RNA-seq in murine and human CP samples^{13–15}, however, downstream molecular mechanisms directly and indirectly controlled by $Tgf-\beta 3$ signaling remain largely

unexplored. We previously used RNA-seq to obtain the transcriptome of WT and *Tgf-β3* HM mice at crucial stages of palatogenesis (E14.5 and E16.5)¹². However, our previous study was confined to only assessing the expression profile of the 322 known CP genes in our RNA-seq dataset and did not explore the genes differentially expressed between the WT and HM groups and/or between the E14.5 and E16.5 time points. In this study, using the same RNA-seq data established in our previous paper, we applied new and advanced data quality control, filtering, quantification, normalization, differential expression, and systems biology approaches that were not previously employed, and explored all of the differentially expressed genes across all of the genotypes and time points. This way, we could identify potential key molecular components and possible underlying mechanisms of palate formation during development, failure of which may result in CP.

In particular, we identified the differentially expressed genes between genotypes (at a given time point) as well as the differentially expressed genes that changed temporally from E14.5 to E16.5 (for a given genotype). We further applied a comparative temporal transcriptome analysis of *Tgf-β3* WT and HM mice and identified genes that were uniquely regulated by each genotype, which may contribute to CP formation in HM mice. Our systems biology analysis based on the Gene Ontology (GO) and Ingenuity Pathway Analysis (IPA; QIAGEN Inc., <https://www.qiagenbioinformatics.com/products/ingenuity-pathway-analysis>) platforms identified the biological functions, molecular networks, and regulatory pathways (especially in relation to TGF-β signaling) for the differentially expressed genes. Finally, we validated the RNA-seq analysis results with real-time quantitative polymerase chain reaction (qRT-PCR) for 29 genes that are either known CP-related genes and/or genes that are regulated by TGF-β and fall under its downstream signaling pathways. For all of the 29 genes, the degree and direction of differential expression based on RNA-seq analysis agreed with the qRT-PCR results. Collectively, this data will enrich our understanding of TGF-β signaling during palate development and provide insight into the temporal regulation of downstream TGF-β-regulated genetic modulators that control cell morphology, cellular differentiation, apoptosis, and morphogenesis of the embryo.

Results

Our previously acquired RNA-seq data¹² involving wild type and *Tgf-β3* homozygous knockout mouse models on E14.5 and E16.5 represented four sample groups (denoted as “WT14.5,” “WT16.5,” “HM14.5,” and “HM16.5”), each represented by two biological replicates (denoted as “a,” or “b”) for a total of eight samples. The average read count for the raw RNA-seq data was ~65.5 M paired-end reads (i.e., ~130.1 M total reads) per sample providing a high coverage of the transcriptome. The re-analysis of this data involved data trimming and filtering, normalization, expression quantification, clustering, differential analysis, comparative analysis, systems biology analysis, and qRT-PCR validation.

RNA-seq data processing and clustering. After trimming and filtering, the number of average total reads per sample came down to ~128.1 M (Supplementary Fig. 1A). The average read length was 101 bp in the raw data, which decreased to 95.86 bp following trimming and filtering (Supplementary Fig. 1B). On the other hand, the average Phred read quality score increased to 36.79 from 35.84 (Supplementary Fig. 1C); and the percentage of high-quality bases (bases with a quality score > 20) per sample increased to 99.20% from 96.28% (Supplementary Fig. 1D) following trimming and filtering. Therefore, both the total number of reads and the average read length parameters showed small changes in quantity after trimming and filtering; but this resulted in significant data quality improvement.

RNA-seq analysis generated expression data for 103,215 transcripts. Transcripts that showed a TPM value of less than 1 in both samples in all of the four groups were eliminated from downstream analysis leaving 52,475 transcripts. Biological replicates showed high degree of correlation ($r > 0.98$). In Supplementary Fig. 2, we show the hierarchical clustering of the samples using all 52,475 transcripts. This global unsupervised grouping reveals that the samples were separated clearly by time as there are two main clades where one clade only consists of E14.5 samples and the other clade only consists of E16.5 samples. The effect of genotype on the transcriptional profiling was subtle because in the clustering tree we do not see a grouping based on the genotype neither globally, nor within a time point. Hence, there is a need for supervised analysis methods to identify the differences in gene expression due to genotypic variance. Furthermore, the height of the branching points in Supplementary Fig. 2 both for the E14.5 and E16.5 sample groups implied that the similarity between the WT and HM groups was higher at E14.5 than it was at E16.5. This is because of the shorter branch lengths among samples in the E14.5 clade compared to the branch lengths among samples in the E16.5 clade. Therefore, the samples at E16.5 showed a more divergent transcriptional profile and thus the effects of *Tgf-β3* knockout were more pronounced at E16.5.

Differential expression analysis. Differential expression analysis also showed the stark temporal difference in gene expression between samples (Table 1). There were 4,115 and 5,304 significantly differentially expressed genes (SDEG) between the E16.5 and E14.5 time points for the WT and HM samples, respectively indicating that the temporal transcriptional change in HM samples was more widespread than in normal controls. Furthermore, in both genotypes, the number of upregulated genes with time (2,421 for WT and 3,153 for HM) was more than the number of downregulated genes (1,694 for WT and 2,151 for HM) showing that transcriptional induction overshadowed transcriptional silencing. This difference in up/downregulation was further emphasized when we applied a 2.0-fold change (FC) cutoff (on top of the adjusted $p < 0.05$ cutoff) to define SDEGs. Most of the upregulated genes survived this FC cutoff (1,675 out of 2,421 in WT and 1,936 out of 3,153 in HM had an $|FC| > 2.0$); but the number of SDEG that were downregulated decreased dramatically (134 out of 1,694 in WT and 191 out of 2,151 in HM had an $|FC| > 2.0$). This further underlined the trend of significant transcriptional induction with time.

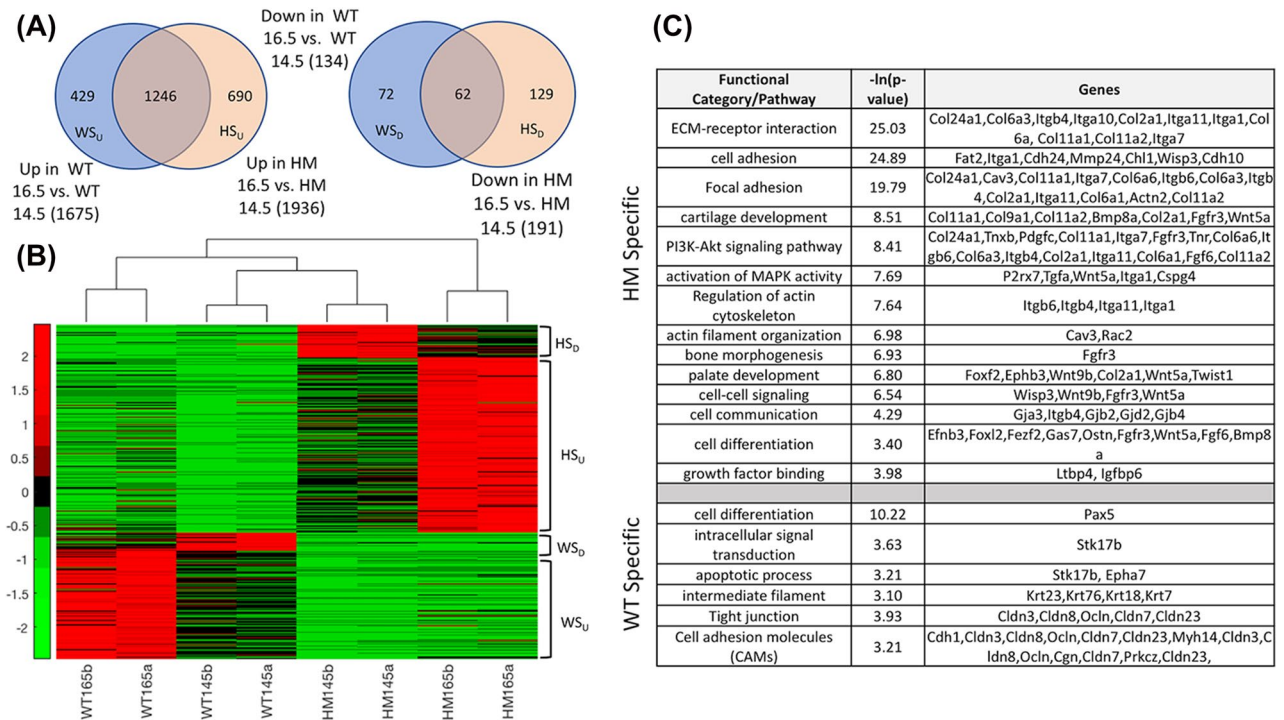


Figure 2. Hierarchical clustering of 501 (429 + 72) WT-specific and 819 (690 + 129) HM specific genes with the color bar showing the row z-score. **(A)** Comparison of significantly differentially expressed genes (SDEGs) between E16.5 and E14.5 in the WT and HM groups. WT-Specific Up (WS_U): genes uniquely upregulated in the WT group at E16.5; WT-Specific Down (WS_D): genes uniquely downregulated in the WT group at E16.5; HM-Specific Up (HM_U): genes uniquely upregulated in the HM group at E16.5; HM-Specific Down (HM_D): genes uniquely downregulated in the HM group at E16.5. **(B)** Hierarchical clustering of all four groups (eight samples) using the 501 (429 + 72) WT-specific and 819 (690 + 129) HM-specific genes identified in **(A)**. **(C)** Significantly enriched Gene Ontology (GO) categories in the WT-specific and HM-specific gene lists (sample genes in the groups are shown), fold change of relevant WT- and HM-specific genes.

We observed a very subtle difference between the two genotypes at a given time point. At days E14.5 and E16.5, there were only 13 and 38 SDEGs between the WT and HM samples, respectively. In concordance with the unsupervised hierarchical clustering results, we saw a greater difference between the genotypes at E16.5. This was further strengthened by observing 13 SDEGs between the two genotypes with an $|FC| > 2.0$ at E16.5, while there were no SDEGs between the WT and HM samples with an $|FC| > 2.0$ at E14.5.

WT specific and HM specific gene identification. Since a direct comparison between the HM and WT groups yielded a very subtle difference, we defined the effect due to *Tgf-β3* knockout by comparing the temporal SDEGs in the two groups. Following our adjusted $p < 0.05$ and $|FC| > 2.0$ cutoffs, we compared the 1,675 and 1,936 SDEGs that were upregulated in E16.5 vs. E14.5 in the WT and HM groups, respectively. We identified 429 genes that were upregulated in the WT group, but not in the HM group; similarly, we identified 690 genes that were upregulated in the HM group, but not in the WT group. Conversely, we compared the 134 SDEGs downregulated in E16.5 vs. E14.5 in the WT group with the 191 SDEGs downregulated in E16.5 vs. E14.5 in the HM group. We identified 72 genes that were downregulated in the WT group, but not in the HM group and similarly we identified 129 genes that were downregulated in the HM group, but not in the WT group. These results are summarized in Fig. 2A. We called the 501 (429 + 72) SDEGs uniquely up/downregulated in the WT group “WT specific”; and similarly, we called the 819 (690 + 129) SDEGs uniquely up/downregulated in the HM group “HM specific.”

In Fig. 2B, we show the hierarchical clustering of the 1,320 (501 + 819) WT and HM specific SDEGs across the four sample groups. This visual representation of expression values (heatmap) clearly shows how WT specific genes are significantly up/downregulated between WT E14.5 and WT E16.5 samples, whereas the expression of these genes remains “not significantly altered” between the HM E14.5 and HM E16.5 samples. A similar (but complementary) observation holds for the HM-specific genes. Figure 2B also shows the sample similarity based on genes specific to the genotypes, which clustered the biological replicates together and grouped the samples in E14.5 under the same clade. This implies that the temporal difference in the dataset was more dominant than the genotypic difference as we do not see samples from the same genotype in the two different time points group together. The HM samples at E16.5 stood out as a distinct group, which is reasonable as the phenotypic difference between the genotypes was not visible at E14.5 but emerged clearly at E16.5, making the HM samples at E16.5 a distinctly separate phenotypic group among the four.

Functional analysis. In order to understand the genes and corresponding signaling network and functional mechanisms that would explain the differences observed between HM and WT samples across E16.5 and E14.5 time points, we highlighted the GO functional categories and KEGG pathways that were statistically significantly enriched in the WT and HM specific gene lists (Fig. 2C). A complete list of SDEGs, enriched GO categories and KEGG pathways can be found in the Supplementary Data. In Table 2, we list 24 HM-specific and 11 WT-specific genes based on their functional relevance by any one or more of the following criteria: (1) falling under the TGF- β signaling pathway, (2) contributing to palatogenesis or cleft palate in mice, (3) being involved in palatal cellular processes, such as EMT, apoptosis, differentiation, proliferation that are functional in palatal and craniofacial morphogenesis, and (4) being directly or indirectly regulated by TGF- β signaling.

Systems biology analysis. Using IPA, we analyzed the signaling networks and crosstalks amongst the downstream molecules that could be uniquely functional in WT and HM palates. Our network analysis results suggest that the TGF- β , ERK/MAPK, p38MAPK and PI3K/AKT pathways directly regulate *p38MAPK* in WT (Fig. 3A and Supplementary Fig. 3), which is not the case in HM (Fig. 3B and Supplementary Fig. 7). The canonical pathway analysis using IPA further strengthened this observation when we specifically considered the TGF- β signaling pathway separately for WT specific genes (Fig. 3A) and for HM specific genes (Fig. 3B). Furthermore, we showed that *Gooseoid* (*Gsc*) is downregulated by TGF- β signaling in WT (Fig. 3A), whereas, in the HM, TGF- β signaling downregulates transcription factor *Tlx2* (Fig. 3B). Finally, our results suggest that, in both in WT and HM, TGF- β signaling regulates expression of neither Smads nor of transcripts encoding proteins that act upstream or downstream of Smads (Fig. 3A,B; Supplementary Figs. 3 and 7).

We identified WT specific genes that are also TGF- β downstream target genes. When we overlaid the Epithelial Mesenchymal Transition (EMT) functional category on these genes, we observed that only two genes were selected (Supplementary Fig. 4a). When we similarly overlaid the “cell death” functional category on the same gene list, we observed nine selected genes (Supplementary Fig. 5a). These results potentially demonstrate that in WT, TGF- β signals to regulate “cell morphology” and “differentiation” that facilitate apoptosis over EMT as the number of TGF- β downstream target genes are predominantly pro-apoptotic genes but not EMT. We further analyzed the mechanistic networks generated by IPA that involve TGF- β signaling. These networks show the interaction among upstream regulators that best explain the changes observed in the SDEG list. We performed this analysis for WT specific genes and found that these upstream regulators were predominantly pro apoptotic, (12 out of 14, Supplementary Fig. 5b) not EMT (none out of 14, Supplementary Fig. 4b) genes.

qRT-PCR validation. We further assessed the expression levels of the 29/35 differentially expressed genes listed in Table 2 using real-time qRT-PCR. Our qRT-PCR results were all in agreement with our RNA-seq results (Table 2), which demonstrated that in WT, *Cdh1*, *Ocln*, *F2rl1*, and *Tnfrsf11b* were uniquely upregulated (folds of increase: 4.19, 5.84, 4.97, and 3.39, respectively) while *Fndc3c1*, *Dlx1*, and *Gas2* were uniquely downregulated (folds of increase: -8.99, -6.89, and -7.18, respectively) at E16.5 when compared to E14.5 (Fig. 4A). Similarly, in HM palates, genes that uniquely showed upregulation at E16.5 in RNA-seq data (Table 2), including *Chrng*, *Col2a1*, *Col11a1*, *Col11a2*, *Llcam*, *Adam12*, *Fas*, *Hspg2*, *Lox*, *Itgb4*, *Klf5*, *Cldn1*, and *Nrcam*, exhibited significantly increased mRNA expression in qRT-PCR (folds of increase > 4.02), when compared to E14.5 (Fig. 4B). Similarly, qRT-PCR results verified that *Wnt9b*, *Alx4*, *Pax1*, *Kcp*, *Msx1*, *Ppp1r17*, *Twist1*, and *Wnt5a* were also in concordance with RNA-seq data, presenting downregulated mRNA expression (folds of decrease < -4.99) uniquely in HM (Fig. 4B).

Discussion

All isoforms of TGF- β are involved in both palatal epithelial and mesenchymal cell proliferation, differentiation, transformation, and apoptosis¹⁶. TGF- β 3 plays a distinctive role in periderm removal, palatal fusion and seam disintegration chronologically that involves crosstalk and reciprocal signaling between the two cell populations by epithelial-mesenchymal interaction (EMI)^{17,18}, which cannot be achieved by other isoforms^{10,11}. Therefore, in this study, we analyzed *Tgf- β 3* HM whole palates to accurately explain the global TGF- β 3 signaling during palate development. Usage of both epithelial and mesenchymal cells is a standard practice as previously done^{19,20}. In one of the most elegant studies of the palatal transcriptome analysis by NIH Facebase consortium, the authors also used regions of palate with a combination of both palatal epithelia and mesenchyme¹³.

Our differential expression analysis showed a very subtle difference between the WT and HM palates—there were only 13 and 38 SDEGs between the WT and HM samples, respectively, at days E14.5 and E16.5 (see Supplementary Data). In agreement with the unsupervised hierarchical clustering data, we noticed a greater difference between the genotypes at E16.5. This difference was not only in numbers but also in the degree of differential expression. When the $|FC| > 2.0$ cutoff was employed, there were no SDEGs at E14.5 and only 13 SDEGs at E16.5 between the HM and WT groups (Table 1). This observation may be attributable to the obvious phenotypical difference, a palatal cleft at E16.5 and no difference at E14.5. Interestingly, 9 out of the 13 SDEGs in HM vs. WT at E16.5 were epithelial-specific genes with a majority of them being involved in palate deformities and cleft palate (Table 3). Henceforth, the differences between these two genotypes mainly impacted palatal epithelial cell-specific function, which is in accordance with the expression of TGF- β 3 in WT or lack of it in HM.

The subtle difference in gene expression at a given time point is likely to trigger larger downstream changes through direct and/or indirect regulatory signaling that results in a phenotypic difference. Such a large effect would be observed through variations in the temporal gene expression patterns for the two genotypes, which also accounts for any subtle cell population differences as direct comparisons between the genotypes are not involved. Therefore, we focused on genes that show temporal changes in WT (from E14.5 to E16.5) due to the presence of *Tgf- β 3*, which are not identified as changed in HM (from E14.5 to E16.5) due to the absence of *Tgf- β 3*

| Comparison | Group | Gene list | No. of SDEG | No. of SDEG ($ FC > 2.0$) |
|---------------------|-------|----------------------------------|-------------|------------------------------|
| Between time points | WT | Up in WT E16.5 versus WT E14.5 | 2,421 | 1,675 |
| | | Down in WT E16.5 versus WT E14.5 | 1694 | 134 |
| | HM | Up in HM E16.5 versus HM E14.5 | 3,153 | 1,936 |
| | | Down in HM E16.5 versus HM E14.5 | 2,151 | 191 |
| Between genotype | E14.5 | Up in HM E14.5 versus WT E14.5 | 8 | 0 |
| | | Down in HM E14.5 versus WT E14.5 | 5 | 0 |
| | E16.5 | Up in HM E16.5 versus WT E16.5 | 30 | 10 |
| | | Down in HM E16.5 versus WT E16.5 | 8 | 3 |

Table 1. Significantly differentially expressed genes (SDEG), multiple hypothesis testing corrected p value < 0.05 , across time and genotype points listed separately for up/downregulation (FC, fold change).

and vice versa. In order to identify genes that show temporal difference uniquely in the WT or HM groups, we identified WT specific and HM specific gene lists (Fig. 2A). The functional analysis of these gene lists (Fig. 2C, Supplementary Data) render GO categories and KEGG pathways that are statistically significantly enriched in the WT or HM, such as the “apoptotic process,” “cell adhesion molecules (CAMs),” and “focal adhesion.” These lists show specific signaling pathways and TGF- β downstream molecules modulating both the epithelial and the mesenchymal cellular functions that are unique to WT and HM, resulting in immaculate palatogenesis in WT but cleft in HM.

In the Supplementary Data, we also list the 1,308 genes commonly up/down regulated between E16.5 and E14.5 both in WT and HM groups. The rationale behind the table is to describe the common genes that are responsible for phases that are common in both WT and HM, such as palatal cell proliferation, differentiation, transformation and apoptosis. These cellular changes are ongoing in both genotypes, hence, the genes listed in this Supplementary Data represent a common palatal cellular behavior and function that may not necessarily imply genes that can trigger cleft palate. Of note among commonly upregulated genes are *Sntn*, which participates in cell growth/maintenance and is involved in cell communication process in the nasopharynx²¹, and *Tmem212*, which codes for a transmembrane protein known to interact with transcription factor *Tcf12*, an important paralog of *Tcf4*²² and *Forkhead box k2 (Frk2)*. *Frk2* is an important regulator palatal cell proliferation via activation of the phosphoinositide 3-kinase/AKT pathway²³. Other genes to names are, *Ppp1r32*, which is a substrate for cGMP-dependent protein kinase and is involved in central nervous system development and intracellular signal transduction²⁴, is upregulated in both WT and HM palates. It implements protein serine/threonine phosphatase inhibitor activity and inhibits phosphatase activities of *protein phosphatase 1 (Pp1)* and *protein phosphatase 2A (Pp2A)* complexes. Pp1 and Pp2A were reported by Weston et al.²⁵ to account for virtually all detectable serine/threonine protein phosphatase activity during the development of embryonic palate. We also observed common upregulation of *Tspan1* and *33* in both WT and HM going from E14.5 to E16.5. *Tspan1* have been shown to play crucial roles in biologic processes including cell adhesion, proliferation, differentiation, and migration²⁶. Our data suggest that the role of *Tspans* may be limited to cellular proliferation and differentiation via Smad pathways. Another commonly upregulated transcript is *Tnfrsf11a (RANK)*, which is a key regulator of bone homeostasis²⁷. As early as E14.5, mesenchymal condensations undergo chondrogenesis initially and ultimately membranous ossification that give rise to the hard palate²⁸. We expect the palatal cells express RANK to regulate osteoclast function in palatal bone formation in both genotypes. *In terms of epithelial markers*, *Krt13*, which has been shown to be expressed in the suprabasal layer of stratified palatal keratinized epithelia²⁹, is also among the commonly upregulated transcripts. *Krt4*, a type II cyokeratin, is specifically expressed in differentiated layers of all of oral mucosal epithelia along with family member *Krt13*³⁰. Similar to *Krt13*, *Krt4* is also upregulated in both WT and HM. Showing a similar expression pattern to *Krt13* among commonly upregulated transcripts is *Spr3* (aka Loricrin), a marker for terminally differentiated keratinized and non-keratinized oral mucosa³¹. This is potentially due to the fact that palatal epithelia undergo significant stratification with terminally differentiated keratinized epithelia covering the oral side of the palatal lining in both WT and HM palates.

Using IPA, we identified several novel findings which may suggest, and potentially imply, new signaling network molecules as well as cellular functions. Our data suggests that TGF- β signaling may induce palatogenesis through regulating *p38MAPK* in WT palates (Fig. 3A,B, Supplementary Figs. 3 and 7). These findings are in agreement with previous work^{32–34} showing *p38MAPK* activation by TGF- β signaling during palatogenesis. It has been previously demonstrated that in palatogenesis TGF- β signals through the SMAD pathway^{35–38}. Our results indicate no change in SMADs and its down- and up-stream genes at the mRNA levels both for WT and HM palates (Fig. 3A,B, Supplementary Figs. 3 and 7). This observation does not necessarily imply an inactivation of the SMAD pathway as transcriptional level activity does not always imply protein level functionality. On the other hand, these results may imply *non-Smad* pathways to be also at play both in the WT and HM while in HM, unlike WT, this signaling cascade does not include *p38MAPK* (Supplementary Figs. 3 and 7).

Uniquely in the WT palate, *Gooseoid (Gsc)*, a homeobox-containing gene, is downregulated by TGF- β signaling in palatogenesis (Fig. 3A). *Gsc* mutant mice display defects in the pharyngeal muscles and the pharyngeal mucosa³⁹. TGF- β signaling is known to promote EMT by regulating the *Gsc* gene during embryonic Spemann's organizer formation⁴⁰ as well as breast cancer metastasis⁴¹. Since GSC is known to be a homeobox transcription factor that promotes EMT, it is likely that apoptosis is favored for seam disintegration since an EMT gene, such

| ENSEMBLE gene ID | Gene name | Adjusted <i>p</i> value | Fold change (HM E16.5/HM E14.5) | Function |
|--------------------|------------------|-------------------------|---------------------------------|--|
| HM specific | | | | |
| 24,330 | <i>Coll1a2*</i> | 1.72E-27 | 6.16 | Variants involved in CP ^{58,59} |
| 27,966 | <i>Coll1a1*</i> | 2.63E-24 | 3.71 | Variants involved in CP ^{58,59} |
| 32,060 | <i>Cryab*</i> | 5.97E-10 | 3.44 | EMT regulator ^{60,61} Apoptosis inhibitor ^{62,63} |
| 54,555 | <i>Adam12*</i> | 2.26E-19 | 2.9 | TGFβ-induced EMT regulator ⁶⁴ TGFβ signaling modulator ⁶⁵ Epithelial cell proliferation and apoptosis regulator ⁶⁶ |
| 22,483 | <i>Col2a1*</i> | 3.50E-06 | 2.68 | Mutations involved in cleft palate ^{58,59} |
| 31,391 | <i>L1cam*</i> | 1.59E-04 | 2.38 | Apoptosis resistance ⁶⁷ EMT regulator ⁶⁸ Cell adhesion ⁶⁹ |
| 26,253 | <i>Chrng*</i> | 1.40E-02 | 2.21 | Mutations associated with cleft palate ⁷⁰ |
| 5,148 | <i>Klf5</i> | 1.17E-03 | 2.19 | Epithelial proliferation promotor ⁷¹ Apoptosis inhibition ⁷² |
| 22,037 | <i>Clu*</i> | 4.54E-05 | 2.18 | TGFβ signaling modulator ⁷³ Apoptosis inhibitor ⁷⁴ TGFβ-induced EMT regulator ⁷⁵ |
| 24,529 | <i>Lox*</i> | 3.14E-06 | 2.15 | EMT regulator ^{76,77} TGFβ signaling target ⁷⁸ |
| 20,598 | <i>Nr-cam*</i> | 8.95E-03 | 2.15 | Cell proliferation and motility stimulator ⁷⁹ |
| 24,778 | <i>Fas*</i> | 4.34E-03 | 2.1 | EMT inducer ⁸⁰ Apoptosis mediator ^{81,82} |
| 28,763 | <i>Hspg2*</i> | 2.78E-04 | 2.03 | TGFβ signaling target ⁸³ |
| 22,512 | <i>Cldn1*</i> | 1.32E-03 | 2.01 | EMT promotor ^{84,85} Apoptotic regulator ^{86,87} |
| 20,758 | <i>Itgb4</i> | 4.08E-03 | 2 | EMT promotor ⁸⁸ Epithelial Cells mobility enhancer ^{89,90} Apoptosis inhibitor ⁹¹ |
| 35,799 | <i>Twist1*</i> | 2.85E-13 | -2.17 | CLP candidate gene ^{92,93} EMT marker ⁹⁴⁻⁹⁶ TGFβ signaling regulator ^{96,97} |
| 28,019 | <i>Pdgfc</i> | 1.74E-09 | -2.21 | CP candidate gene ^{1,98,99} TGFβ signaling target ¹⁰⁰ |
| 21,994 | <i>Wnt5a*</i> | 4.44E-04 | -2.48 | TGFβ signaling crosstalk ¹⁰¹⁻¹⁰³ EMT mediator ^{104,105} |
| 2,930 | <i>Ppp1r17</i> | 2.98E-02 | -2.52 | Embryonic palate development regulator ²⁵ |
| 48,450 | <i>Msx1</i> | 6.95E-10 | -2.66 | CP candidate gene ¹⁰⁶⁻¹⁰⁸ EMT mediator ¹⁰⁹ TGF-β superfamily pathways regulator ¹⁰⁶ |
| 40,310 | <i>Alx4</i> | 2.16E-13 | -2.67 | CP candidate gene ¹¹⁰ EMT mediator ^{111,112} |
| 59,022 | <i>Kcp</i> | 7.83E-10 | -2.67 | TGF-β superfamily pathways regulator ¹¹³ |
| 37,034 | <i>Pax1</i> | 3.22E-03 | -2.77 | Embryogenesis regulator ¹¹⁴ Apoptosis and differentiation inducer ¹¹⁵⁻¹¹⁷ |
| 18,486 | <i>Wnt9b</i> | 3.59E-05 | -4.88 | CLP candidate gene ^{118,119} Facial outgrowth and fusion promoter ¹²⁰ Cell proliferation, differentiation and cell polarity regulator ^{121,122} |
| ENSEMBLE gene ID | Gene name | Adjusted <i>p</i> value | Fold change (WT E16.5/WT E14.5) | Function |
| WT specific | | | | |
| 21,638 | <i>Ocln*</i> | 5.78E-10 | 2.84 | Tight junction component ¹²³ Adhesion, apoptosis, differentiation and homeostasis regulator in keratinocytes ¹²⁴ |
| 27,858 | <i>Tspan2*</i> | 2.86E-07 | 2.49 | Cell adhesion, proliferation, differentiation and migration regulator ²⁶ TGF-β signaling downstream effector ²⁶ |
| 21,678 | <i>F2rl1*</i> | 1.73E-03 | 2.26 | TGF-β signaling target ^{125,126} Cell proliferation enhancer ^{127,128} |
| 42,228 | <i>Lyn*</i> | 1.37E-07 | 2.17 | EMT mediator ¹²⁹ Apoptosis inhibitor ¹³⁰ TGF-β signaling target ¹³¹ |
| 303 | <i>Cdh1*</i> | 5.24E-03 | 2.09 | CLP candidate gene ^{5,132,133} |
| 63,727 | <i>Tnfrsf11b</i> | 3.11E-03 | 2.04 | TGF-β signaling pathways mediator ^{134,135} |
| 21,614 | <i>Vcan*</i> | 4.85E-30 | -2.66 | Cell proliferation, adhesion and apoptosis regulator ^{136,137} Apoptosis inhibitor ^{138,139} |
| 28,487 | <i>Bnc2</i> | 3.05E-05 | -2.76 | Embryonic craniofacial mesenchymal cell multiplication regulator ¹⁴⁰ |
| 30,498 | <i>Gas2</i> | 8.17E-05 | -3.06 | Cell cycle and apoptosis regulator ¹⁴¹ TGF-β signaling pathways target ¹⁴² |
| Continued | | | | |

| ENSEMBLE gene ID | Gene name | Adjusted <i>p</i> value | Fold change (WT E16.5/WT E14.5) | Function |
|------------------|----------------|-------------------------|---------------------------------|---|
| 41,911 | <i>Dlx1*</i> | 7.45E-09 | -3.25 | Craniofacial patterning controller ¹⁴³ TGF- β signaling pathways inhibitor ^{144,145} |
| 33,487 | <i>Fndc3c1</i> | 2.04E-21 | -4.58 | Ectoderm differentiation gene ¹⁴⁶ |

Table 2. Selected genes that are uniquely significantly differentially expressed (adjusted *p* value < 0.05) in the HM or WT groups. CP, Cleft Palate; CL, Cleft Lip; CLP, Cleft Lip and Palate; identified in mouse or in human (*).

as *Gsc* is downregulated in WT palates. Similarly, uniquely in the HM, *Tgf- β* signaling downregulates a separate homeobox gene, *Tlx2*, a transcription factor (Fig. 3B). The *Tlx2* gene encodes transcription factors essential in the development of neural-crest-derived cells suggesting a physiological role in the transcription-factor cascade underlying the differentiation of neuronal lineages during embryogenesis⁴². We did not find any direct relationship between TGF- β and *Tlx2* but based on the fact that *Tlx2* is crucial for neural-crest-derived cell development, it is therefore likely that palatal epithelia (which is an ectoderm derived cell) has limited or no role for *Tlx2* and therefore it is downregulated.

Our upstream regulator analysis identified targets of *Tgf- β* in the WT specific gene list. Exploring the involvement of these genes in the EMT and cell death mechanisms revealed limited involvement (*Cdh1* and *Lef1*) of the EMT pathway (Supplementary Fig. 4a) whereas nine apoptotic genes (*Ace*, *Cdh1*, *Dcn*, *Dlx2*, *Krt18*, *Pparg*, *Rasgrp*, *Sema7a*, and *Tnfrsf11b*) were directly regulated by TGF- β (Supplementary Fig. 5a). We also identified mechanistic networks, which are interaction networks of upstream regulators that explain the changes observed in the WT (or HM) specific genes. When we explored the functional characteristics of these interconnected upstream regulators, none of the genes were involved in the EMT (Supplementary Fig. 4b); and 12 out of 14 genes were involved in “cell death” (Supplementary Fig. 5b).

We identified genes regulated by *Tgf- β 3* in WT specific genes and showed that among these *Tgf- β 3* targets, *Cdh1* is upregulated and *Lef1* is downregulated (Supplementary Fig. 6). Loss of *Cdh1* is a key marker of EMT⁴³; and the expression of *Cdh1* can be repressed by the transcription factor *Lef1* in palatal EMT^{22,44}. Additionally, anti-EMT cell–cell adhesion genes (*Cdh1*, *Ocln*, *Tspan2*) are upregulated, which may suggest a potential relationship between *Tgf- β 3* signaling and suppression of palatal EMT. These findings may indicate two possible outcomes: (a) since palatal seam disintegration is complete at E16.5, the EMT markers are no longer expressed, or (b) EMT may not be a mechanism of palatal seam disintegration. Our findings, based on mechanistic network and upstream regulator analysis (Supplementary Figs. 4a, 4b, 4a, 4b, and 4), suggest that the latter is more probable as the number of apoptotic regulatory network genes are significantly more than the number of EMT genes.

Conclusion

Identifying transcripts that play key roles in regulating palatal development in critical stages has been a powerful approach to understanding how *Tgf- β 3* controls normal palatogenesis and how the lack of signaling (and its downstream signaling partners) is associated in induction of cleft palate. This study identifies potential CP-related genes based on differential expression between genotypes and gestational ages. The data presented in this work provide a strengthened understanding of the complex genetic mechanism of *Tgf- β 3*-regulated palatogenesis. In addition, we discussed the variations in gene expression in the absence of *Tgf- β 3* in HM implicated in cleft palate. Our results represent a comprehensive analysis of the gene profile in murine cleft palate due to the absence of *Tgf- β 3*. Further elucidation of the significantly up/downregulated genes will enhance our understanding of the mechanisms controlling palate development, thereby paving the way for prevention of cleft palate during development.

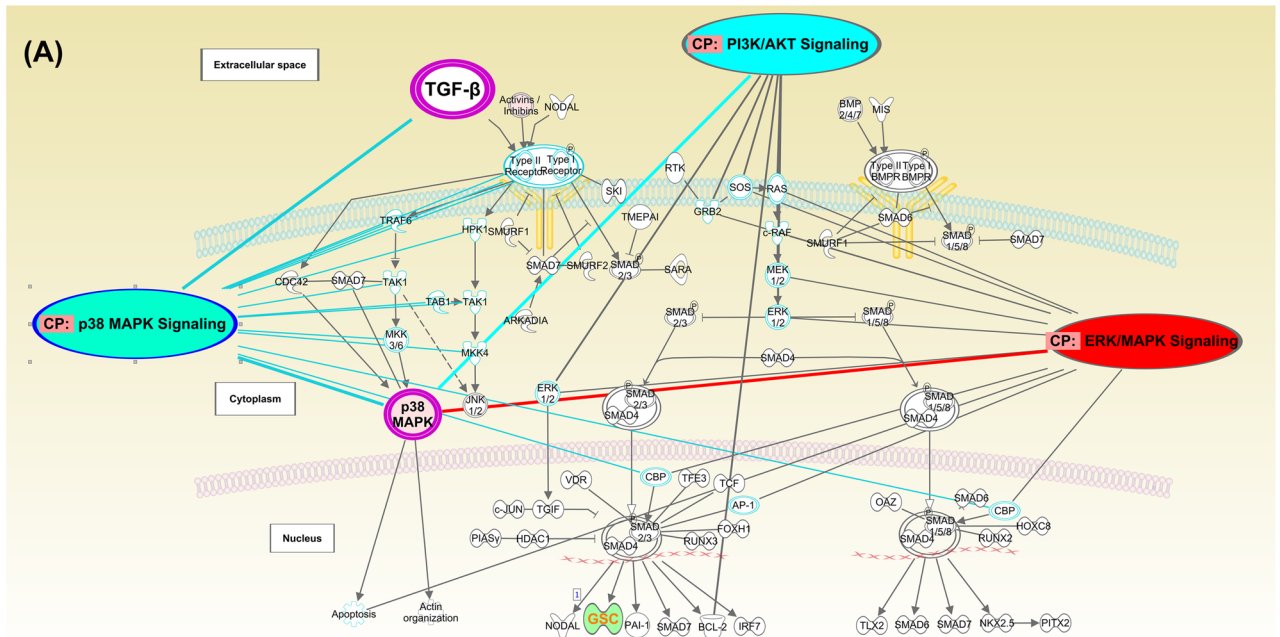
Methods

Animal selection and breeding. *Tgf- β 3* heterozygous (+/–) C57BL/6J breeder mice were obtained from Tom Doetschman (BIO5 Institute, University of Arizona, AZ). The reproduction and genotyping of *Tgf- β 3* –/– mice was conducted as previously described⁹. Mice were accommodated and subject to procedures at the University of Nebraska Medical Center (UNMC) College of Dentistry Animal Facilities under the approval of the UNMC Institution Animal Care and Use Committee (IACUC # 06–064). Null mutant embryos were generated by intercrossing *Tgf- β 3* heterozygous male and female mice in a Mendelian fashion.

Genomic DNA purification and genotyping. Palatal tissues were dissected under the NIKON SMZ1000 stereo microscope system (NIKON, Tokyo, Japan) from embryos collected on embryonic day (E) 14.5 and E16.5 following the identification of vaginal plugs, which are considered to be E0.5. Palatal samples were stored in RNAlater Stabilization Reagent (QIAGEN, Hilgen, Germany) to preserve the gene expression profile and individually labeled and matched with the corresponding tail tissue used for genotyping as detailed in our previous study¹².

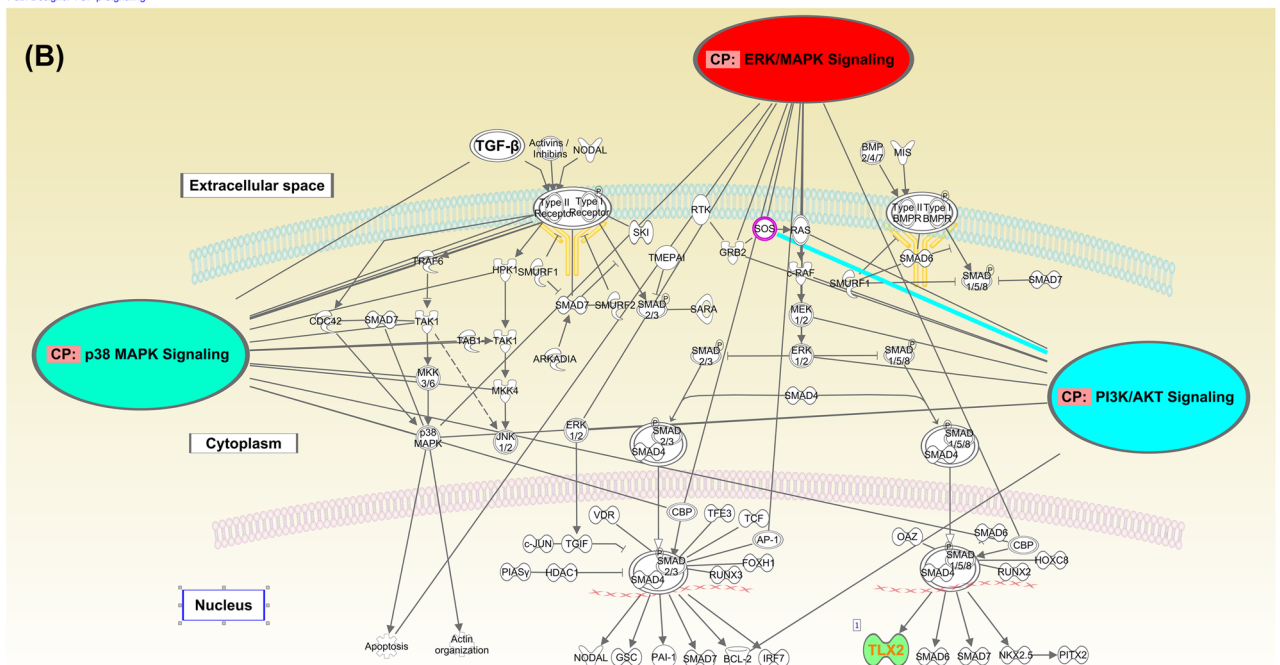
Extraction of RNA, construction of small RNA libraries, and RNA-Seq. Two biological replicates from each genotype and gestational stage were designed to ensure reproducibility and rule out the possibility of differences caused by technical procedures. Palatal shelves were harvested in pairs from eight fetuses out of four litters. Each sample consisted of two pairs of palatal shelves dissected from fetuses of the same genotype from the same litter. The total RNA was purified using Arcturus PicoPure RNA Isolation Kit (THERMOFISHER

Path Designer TGF-β Signaling



© 2000-2017 QIAGEN. All rights reserved.

Path Designer TGF-β Signaling



© 2000-2017 QIAGEN. All rights reserved.

Figure 3. (A) IPA TGF-β signaling canonical pathway analysis overlaying the WT specific gene list: genes significantly differentially expressed (adjusted $p < 0.05$) with a $|FC| > 2.0$ uniquely in WT, E16.5 versus E14.5. Pink implies upregulation and green implies downregulation. Genes that are also involved in major non-Smad pathways (ERK/MAPK, p38MAPK and PI3-AKT) in palatogenesis are indicated with links to the corresponding pathways. (B) IPA TGF-β signaling canonical pathway analysis for the HM specific gene list: genes significantly differentially expressed (adjusted $p < 0.05$) with a $|FC| > 2.0$ uniquely in HM, E16.5 versus E14.5. Green implies downregulation. Genes that are also involved in major non-Smad pathways (ERK/MAPK, p38MAPK and PI3-AKT) in palatogenesis are indicated with links to the corresponding pathways.

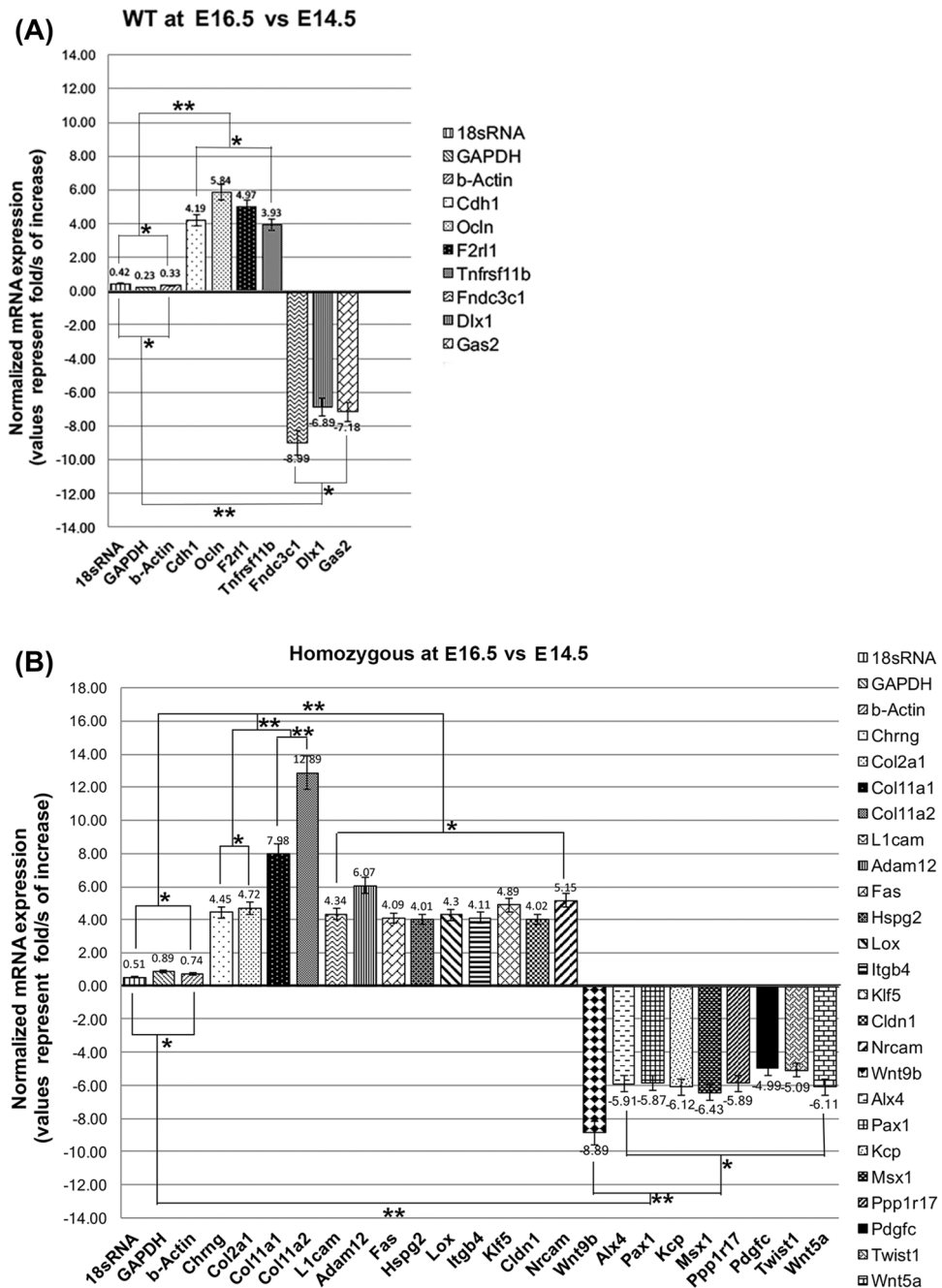


Figure 4. (A) Illustration of the level of mRNA expression showing fold change (log₂) expression of the WT gene mRNA comparative to reference control genes (18rSRNA, GAPDH, and β-actin). The bar heights demonstrate mean expression of the genes in the WT samples. And the error bars suggest 95% confidence interval estimates of the mean expressions. One asterisk designates statistically significant difference between the means of a sample set in comparison to the mean of the control sample set (*p* value < 0.05); two asterisks indicate statistically significant difference (*p* value < 0.01). (B) Illustration of the level of mRNA expression showing fold change (log₂) expression of HM gene mRNA comparative to reference control genes (18rSRNA, GAPDH, and β-actin). The bar heights demonstrate mean expression of the genes in HM samples. And the error bars suggest 95% confidence interval estimates of the mean expressions. One asterisk designates statistically significant difference between the means of a sample set in comparison to the mean of the control sample set (*p* value < 0.05); two asterisks indicate statistically significant difference (*p* value < 0.01).

| ENSEMBLE gene ID | Gene name | Adjusted <i>p</i> value | Fold change (HM E16.5/WT E16.5) | Function |
|------------------|------------------|-------------------------|---------------------------------|--|
| 26,908 | <i>Eif2s3y</i> * | 4.39E-76 | 16.75 | Mutations associated with CLP ¹⁴⁷ |
| 20,592 | <i>Kdm5d</i> * | 2.35E-95 | 14.13 | Mutations associated with palate deformities ¹⁴⁸ |
| 22,290 | <i>Uty</i> * | 4.81E-45 | 8.69 | Mutations associated with palate deformities ¹⁴⁹ |
| 26,900 | <i>Ddx3y</i> * | 1.77E-25 | 5.78 | Mutations associated with cleft palate ¹⁵⁰ TGF- β signaling pathways target ¹⁵¹ |
| 20,248 | <i>Serpinb3a</i> | 0.000523 | 2.15 | Epithelial hyperplasia and hyperkeratosis ^{152,153} |
| 245,026 | <i>Col6a6</i> * | 0.001059 | 2.13 | Epithelial hyperkeratosis ¹⁵⁴ |
| 16,682 | <i>Krt4</i> | 0.000106 | 2.09 | Epithelial integrity gene ¹⁵⁵ Epithelial cell differentiation modulator ¹⁵⁶ |
| 333,564 | <i>Fndc3e1</i> | 0.000915 | 2.02 | Ectoderm differentiation gene ¹⁴⁶ |
| 57,294 | <i>Rps27</i> * | 0.000126 | -2.75 | Mutations associated with cleft palate ¹⁵⁷ |

Table 3. Epithelial-specific genes that are significantly differentially expressed (adjusted *p* value < 0.05, |FC| > 2.0) at E16.5 in HM versus WT. CLP, Cleft Lip and Palate; identified in mouse or in human (*).

SCIENTIFIC, San Francisco, CA) to reliably extract high-quality RNA from a few cells by following the manufacturer's protocols. Purity and concentration were measured by ultraviolet spectroscopy (NANODROP, Wilmington, DE). The RNA integrity evaluation, construction of libraries, and validation was performed as described in our previous study¹².

RNA-seq analysis. RNA-seq data was obtained for four groups: two genotypes, *Tgf- β 3* $-/-$ HM and *Tgf- β 3* $+/+$ WT samples, profiled at two time points, embryonic days E14.5 and E16.5. Each group was represented by two biological replicates, resulting in eight samples. Each sample was run in two lanes on the ILLUMINA HISEQ2000 next generation sequencer using the 2×101 bp paired-end mode. Due to the high correlation coefficient between them ($r > 0.995$) and in order to increase the coverage per biological sample and reduce the lane effect⁴⁵, lane data for each sample were pooled at the read level.

Raw reads were analyzed with FASTQC (V. 0.11.5) for quality control⁴⁶. Overrepresented (e.g., adapter and similar technical) sequences remaining in the raw reads were assessed and subsequently removed using TRIMMOMATIC (V. 0.36) in the palindrome mode based on default alignment detection and scoring parameters⁴⁷. Trimmomatic was also used for low quality base filtering. Maximum information quality filtering was employed with a minimum average read quality threshold of 25. Following technical sequence and low-quality base removal, reads that were shorter than 36 bp were filtered out.

Transcript quantification was done based on the GRCm38.p5 reference genome using Salmon (v. 0.8.2) with default parameters⁴⁸. Salmon uses sample-specific models, such as correction for GC-content bias, that improve the accuracy of transcription abundance estimates. We used transcripts per million (TPM) in Salmon's output as the normalized relative abundance measure employed in our downstream analysis. Differential gene expression analysis was done using DESeq2 (Love, Huber et al. 2014), which has been shown to perform well in experimental designs with few replicates⁴⁹. The RNA-seq data used in this paper were deposited under NCBI's Gene Expression Omnibus (GEO) database (Accession No.: GSE109838).

Clustering of samples and/or genes was done using the Unweighted Pair Group Method with Arithmetic Mean (UPGMA) method (also known as hierarchical clustering) with Pearson's correlation as the distance measure⁵⁰. The expression data matrix was row-normalized prior to the application of average linkage clustering. The Database for Annotation, Visualization and Integrated Discovery (DAVID) v6.7⁵¹ was used for functional analysis of the gene lists, interrogating the Biological Process (BP), Molecular Function (MF), and Cellular Component (CC) Gene Ontology (GO) categories⁵², and the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways⁵³. Biologically relevant categories that are overrepresented in the gene set and, therefore, may be of further interest were assessed using the Expression Analysis Systematic Explorer (EASE) score in the DAVID tool. The EASE score is the upper bound of the distribution of jackknife iterative resampling of Fisher's exact probabilities with Bonferroni multiple testing correction. Categories containing low numbers of genes are underweighted so that the EASE score is more robust than the Fisher exact test. The EASE score is a significance level, with smaller EASE scores indicating increasing confidence in overrepresentation. We selected GO categories that have EASE scores of 0.05 or lower as significantly overrepresented.

We further analyzed the differentially expressed gene lists using the Ingenuity Pathway Analysis (IPA; QIAGEN Inc., <https://www.qiagenbioinformatics.com/products/ingenuity-pathway-analysis>) software. IPA is based on the manual curation of scientific literature to identify pathways, networks, and functional categories that are significantly represented in the input gene list⁵⁴. The computational analysis methods used in IPA are based on enrichment approaches^{55,56} where pathways or functional groups in which the input gene lists are overrepresented are identified. By the same token, IPA identifies upstream regulators (e.g., transcription factors, microRNAs, kinases, compounds, or drugs) and generates interaction networks (based on known interactions identified in the literature) that best explain the transcriptional changes observed in the input gene list.

Confirmation of differentially expressed genes with qRT-PCR. To verify differentially expressed genes in *Tgf-β3* WT and HM samples, qRT-PCR was undertaken as previously described^{19,57}. Embryonic palates were extracted from at E14.5 and E16.5, and RNA extraction was conducted as previously described^{19,57} using Arcturus PicoPure RNA Isolation Kit (THERMOFISHER SCIENTIFIC, San Francisco, CA) to reliably extract high-quality RNA from a few cells. RNA (500 ng) was converted to cDNA using Invitrogen Superscript IV VIL0 Master Mix (THERMOFISHER SCIENTIFIC, San Francisco, CA) that generated a significant cDNA yield at high temperatures in less time. An additional preamplification step was performed using TaqMan PreAmp Master Mix (THERMOFISHER SCIENTIFIC, San Francisco, CA) with a custom preamplification pool of genes of interest to amplify small amounts of cDNA without introducing amplification bias. Samples were preamplified for 14 cycles with thermal cycling conditions of 95 °C for 15 s and 60 °C for 4 min followed by immediate placement on ice. Finally, samples were diluted with TE buffer (pH 8.0) to 1:20 and were placed on 96-well custom array plates in technical triplicate and qRT-PCR executed with TaqMan Fast Advanced Master Mix (THERMOFISHER SCIENTIFIC, San Francisco, CA) reagents. Polymerase chain reaction conditions were run in 40 cycles at 95 °C for 15 s, and 60 °C for 60 s. Following subtraction of technical repeat's gene-specific Δ^{Ct} value from the housekeeping gene's Δ^{Ct} value, the data were, then, analyzed using analysis of variance (ANOVA) on the repeat-normalized $\Delta\Delta^{Ct}$ values, including the control group; and translate effects were estimated from the ANOVA onto the multiplicative scale. The values of 29 genes tested were normalized by adjusting for the concentration of established housekeeping gene, like 18S rRNA, glyceraldehyde 3-phosphate dehydrogenase (GAPDH) and β -actin and the Δ^{Ct} values of a naive/vehicle group.

Ethics approval and consent to participate. *Tgf-β3* heterozygous (+/−) C57BL/6 J breeder mice were obtained from Tom Doetschman (BIO5 Institute, University of Arizona, AZ). The reproduction and genotyping of *Tgf-β3* −/− mice was conducted as previously described⁹. The live mice experiments were performed in accordance with the guidelines and regulations of the University of Nebraska Medical Center (UNMC) College of Dentistry Animal Facilities under the approval of the UNMC Institution Animal Care and Use Committee (IACUC # 18-088).

Data availability

The RNA-seq data used in this paper were deposited under NCBI's Gene Expression Omnibus (GEO) database (Accession No.: GSE109838).

Received: 6 May 2020; Accepted: 17 August 2020

Published online: 10 September 2020

References

- Jugessur, A., Farlie, P. & Kilpatrick, N. The genetics of isolated orofacial clefts: from genotypes to subphenotypes. *Oral Dis.* **15**, 437–453 (2009).
- Schutte, B. C. & Murray, J. C. The many faces and factors of orofacial clefts. *Hum. Mol. Genet.* **8**, 1853–1859 (1999).
- Bush, J. O. & Jiang, R. Palatogenesis: morphogenetic and molecular mechanisms of secondary palate development. *Development* **139**, 231–243. <https://doi.org/10.1242/dev.067082> (2012).
- Nawshad, A. Palatal seam disintegration: to die or not to die? That is no longer the question. *Dev. Dyn.* **237**, 2643–2656. <https://doi.org/10.1002/dvdy.21599> (2008).
- Meng, L., Bian, Z., Torensma, R. & Von den Hoff, J. W. Biological mechanisms in palatogenesis and cleft palate. *J. Dent. Res.* **88**, 22–33. <https://doi.org/10.1177/0022034508327868> (2009).
- Nawshad, A., Lagamba, D., Polad, A. & Hay, E. D. Transforming growth factor- β signaling during epithelial-mesenchymal transformation: implications for embryogenesis and tumor metastasis. *Cells Tissues Organs* **179**, 11–23. <https://doi.org/10.1159/000084505> (2005).
- Sun, D., Vanderburg, C. R., Odierna, G. S. & Hay, E. D. TGF β 3 promotes transformation of chicken palate medial edge epithelium to mesenchyme in vitro. *Development* **125**, 95–105 (1998).
- Lidral, A. C. *et al.* Association of MSX1 and TGF β 3 with nonsyndromic clefting in humans. *Am. J. Hum. Genet.* **63**, 557–568 (1998).
- Proetzel, G. *et al.* Transforming growth factor- β 3 is required for secondary palate fusion. *Nat. Genet.* **11**, 409–414 (1995).
- Hu, L. *et al.* TGF β 3 regulates periderm removal through Δ Np63 in the developing palate. *J. Cell. Physiol.* **230**, 1212–1225 (2015).
- Richardson, R. *et al.* p63 exerts spatio-temporal control of palatal epithelial cell fate to prevent cleft palate. *PLoS Genet.* **13**, e1006828 (2017).
- Ozturk, F., Li, Y., Zhu, X., Guda, C. & Nawshad, A. Systematic analysis of palatal transcriptome to identify cleft palate genes within TGF β 3-knockout mice alleles: RNA-Seq analysis of TGF β 3 Mice. *BMC Genom.* **14**, 113. <https://doi.org/10.1186/1471-2164-14-113> (2013).
- Potter, A. S. & Potter, S. S. Molecular anatomy of palate development. *PLoS ONE* **10**, e0132662 (2015).
- Marazita, M. L. The evolution of human genetic studies of cleft lip and cleft palate. *Ann. Rev. Genom. Hum. Genet.* **13**, 263–283 (2012).
- Brown, N. L. *et al.* Microarray analysis of murine palatogenesis: temporal expression of genes during normal palate development. *Dev. Growth Differ.* **45**, 153–165 (2003).
- Nawshad, A., LaGamba, D. & Hay, E. D. Transforming growth factor β (TGF β) signalling in palatal growth, apoptosis and epithelial mesenchymal transformation (EMT). *Arch. Oral Biol.* **49**, 675–689. <https://doi.org/10.1016/j.archoralbio.2004.05.007> (2004).
- Xu, X., Bringas, P. Jr., Soriano, P. & Chai, Y. PDGFR- α signaling is critical for tooth cusp and palate morphogenesis. *Dev. Dyn.* **232**, 75–84. <https://doi.org/10.1002/dvdy.20197> (2005).
- Lan, Y., Xu, J. & Jiang, R. Cellular and molecular mechanisms of palatogenesis. *Curr. Top. Dev. Biol.* **115**, 59–84. <https://doi.org/10.1016/bs.ctdb.2015.07.002> (2015).
- LaGamba, D., Nawshad, A. & Hay, E. D. Microarray analysis of gene expression during epithelial-mesenchymal transformation. *Dev. Dyn.* **234**, 132–142. <https://doi.org/10.1002/dvdy.20489> (2005).
- Zhu, X., Ozturk, F., Pandey, S., Guda, C. B. & Nawshad, A. Implications of TGF β on transcriptome and cellular biofunctions of palatal mesenchyme. *Front. Physiol.* **3**, 85. <https://doi.org/10.3389/fphys.2012.00085> (2012).

21. Liu, X. Q. *et al.* Transcriptional gene expression profile of human nasopharynx. *Int. J. Mol. Med.* **14**, 409–420 (2004).
22. Nawshad, A. & Hay, E. D. TGF β 3 signaling activates transcription of the LEF1 gene to induce epithelial mesenchymal transformation during mouse palate development. *J. Cell Biol.* **163**, 1291–1301 (2003).
23. Lin, M. F. *et al.* FOXC2, regulated by miR-1271-5p, promotes cell growth and indicates unfavorable prognosis in hepatocellular carcinoma. *Int. J. Biochem. Cell Biol.* **88**, 155–161. <https://doi.org/10.1016/j.biocel.2017.05.019> (2017).
24. Endo, S. *et al.* Molecular identification of human G-substrate, a possible downstream component of the cGMP-dependent protein kinase cascade in cerebellar Purkinje cells. *Proc. Natl. Acad. Sci.* **96**, 2467–2472 (1999).
25. Weston, W. M. *et al.* Phosphatase regulation of gene expression during development of the palate. *Life Sci.* **71**, 1849–1862 (2002).
26. Zhao, J. *et al.* Selective expression of TSPAN2 in vascular smooth muscle is independently regulated by TGF- β 1/SMAD and myocardin/serum response factor. *FASEB J.* **31**, 2576–2591 (2017).
27. Smane, L. & Pilmene, M. in *SHS Web of conferences* (EDP Sciences).
28. Martinez-Alvarez, C. *et al.* Snail family members and cell survival in physiological and pathological cleft palates. *Dev. Biol.* **265**, 207–218 (2004).
29. Rivarola de Gutierrez, E., Innocenti, A. C., Cipitelli, M. J., Salomon, S. & Vargas-Roig, L. M. Determination of cytokeratins 1, 13 and 14 in oral lichen planus. *Med. Oral Patol. Oral Cir. Bucal.* **19**, e359–e365. <https://doi.org/10.4317/medoral.19289> (2014).
30. Kasai, Y. *et al.* Brush biopsy of human oral mucosal epithelial cells as a quality control of the cell source for fabrication of transplantable epithelial cell sheets for regenerative medicine. *Regen. Ther.* **4**, 71–77. <https://doi.org/10.1016/j.reth.2016.02.008> (2016).
31. Katou, F. *et al.* Differential expression of cornified cell envelope precursors in normal skin, intraorally transplanted skin and normal oral mucosa. *Br. J. Dermatol.* **148**, 898–905. <https://doi.org/10.1046/j.1365-2133.2003.05288.x> (2003).
32. Kitase, Y. & Shuler, C. F. Palatal adhesion is dependent on Src family kinases and p38MAPK. *Int. J. Dev. Biol.* **58**, 335–341 (2014).
33. Lane, J. *et al.* Tak1, Smad4 and Trim33 redundantly mediate TGF- β 3 signaling during palate development. *Dev. Biol.* <https://doi.org/10.1016/j.ydbio.2014.12.006> (2014).
34. Xu, X. *et al.* Ectodermal Smad4 and p38 MAPK are functionally redundant in mediating TGF- β /BMP signaling during tooth and palate development. *Dev. Cell* **15**, 322–329. <https://doi.org/10.1016/j.devcel.2008.06.004> (2008).
35. Cui, X. M. *et al.* TGF- β 3-dependent SMAD2 phosphorylation and inhibition of MEE proliferation during palatal fusion. *Dev. Dyn.* **227**, 387–394 (2003).
36. Cui, X. M. *et al.* Overexpression of Smad2 in Tgf- β 3-null mutant mice rescues cleft palate. *Dev. Biol.* **278**, 193–202 (2005).
37. Dudas, M., Nagy, A., Laping, N. J., Moustakas, A. & Kaartinen, V. Tgf- β 3-induced palatal fusion is mediated by Alk-5/Smad pathway. *Dev. Biol.* **266**, 96–108 (2004).
38. Shiomi, N., Cui, X. M., Yamamoto, T., Saito, T. & Shuler, C. F. Inhibition of SMAD2 expression prevents murine palatal fusion. *Dev. Dyn.* **235**, 1785–1793 (2006).
39. Yamada, G. *et al.* Nasal and pharyngeal abnormalities caused by the mouse goosecoid gene mutation. *Biochem. Biophys. Res. Commun.* **233**, 161–165 (1997).
40. Hartwell, K. A. *et al.* The Spemann organizer gene, Goosecoid, promotes tumor metastasis. *Proc. Natl. Acad. Sci. USA* **103**, 18969–18974. <https://doi.org/10.1073/pnas.0608636103> (2006).
41. Watabe, T. *et al.* Molecular mechanisms of Spemann's organizer formation: conserved growth factor synergy between Xenopus and mouse. *Genes Dev.* **9**, 3038–3050 (1995).
42. Borghini, S. *et al.* The TLX2 homeobox gene is a transcriptional target of PHOX2B in neural-crest-derived cells. *Biochem. J.* **395**, 355–361. <https://doi.org/10.1042/BJ20051386> (2006).
43. Aclouque, H., Thiery, J. P. & Nieto, M. A. The physiology and pathology of the EMT: meeting on the epithelial–mesenchymal transition. *EMBO Rep.* **9**, 322–326 (2008).
44. Nawshad, A., Medici, D., Liu, C. C. & Hay, E. D. TGF β 3 inhibits E-cadherin gene expression in palate medial-edge epithelial cells through a Smad2-Smad4-LEF1 transcription complex. *J. Cell Sci.* **120**, 1646–1653. <https://doi.org/10.1242/jcs.003129> (2007).
45. Auer, P. L. & Doerge, R. W. Statistical design and analysis of RNA sequencing data. *Genetics* **185**, 405–416. <https://doi.org/10.1534/genetics.110.114983> (2010).
46. Andrews, S. FastQC: a quality control tool for high throughput sequence data. <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/> (2010).
47. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for illumina sequence data. *Bioinformatics* **30**, 2114–2120. <https://doi.org/10.1093/bioinformatics/btu170> (2014).
48. Patro, R., Duggal, G., Love, M. I., Irizarry, R. A. & Kingsford, C. Salmon provides fast and bias-aware quantification of transcript expression. *Nat. Methods* **14**, 417–419. <https://doi.org/10.1038/nmeth.4197> (2017).
49. Schurch, N. J. *et al.* How many biological replicates are needed in an RNA-seq experiment and which differential expression tool should you use?. *RNA* **22**, 839–851 (2016).
50. Sneath, P. *Numerical Taxonomy: The Principles and Practice Of Numerical Classification* (W. H. Freeman, New York, 1973).
51. da Huang, W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4**, 44–57. <https://doi.org/10.1038/nprot.2008.211> (2009).
52. Ashburner, M. *et al.* Gene ontology: tool for the unification of biology: the gene ontology consortium. *Nat. Genet.* **25**, 25–29. <https://doi.org/10.1038/75556> (2000).
53. Kanehisa, M. & Goto, S. KEGG: kyoto encyclopedia of genes and genomes. *Nucl. Acids Res.* **28**, 27–30 (2000).
54. Kramer, A., Green, J., Pollard, J. Jr. & Tugendreich, S. Causal analysis approaches in ingenuity pathway analysis. *Bioinformatics* **30**, 523–530. <https://doi.org/10.1093/bioinformatics/btt703> (2014).
55. Huang, D. W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4**, 44 (2009).
56. Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. USA* **102**, 15545–15550. <https://doi.org/10.1073/pnas.0506580102> (2005).
57. Zhu, X., Ozturk, F., Liu, C., Oakley, G. G. & Nawshad, A. Transforming growth factor- β activates c-Myc to promote palatal growth. *J. Cell. Biochem.* **113**, 3069–3085 (2012).
58. Nikopensius, T. *et al.* Genetic variants in COL2A1, COL11A2, and IRF6 contribute risk to nonsyndromic cleft palate. *Birth Defects Res. A Clin. Mol. Teratol.* **88**, 748–756 (2010).
59. Melkonian, M. *et al.* Collagen XI sequence variations in nonsyndromic cleft palate, Robin sequence and micrognathia. *Eur. J. Hum. Genet.* **11**, 265 (2003).
60. Malin, D., Petrovic, V., Strekalova, E., Sharma, B. & Cryns, V. L. α B-crystallin: Portrait of a malignant chaperone as a cancer therapeutic target. *Pharmacol. Ther.* **160**, 1–10 (2016).
61. Ishikawa, K. *et al.* α B-Crystallin regulates subretinal fibrosis by modulation of epithelial–mesenchymal transition. *Am. J. Pathol.* **186**, 859–873 (2016).
62. Kamradt, M. C., Chen, F., Sam, S. & Cryns, V. L. The small heat shock protein α B-crystallin negatively regulates apoptosis during myogenic differentiation by inhibiting caspase-3 activation. *J. Biol. Chem.* **277**, 38731–38736 (2002).
63. Nahomi, R. B. *et al.* Chaperone peptides of α -crystallin inhibit epithelial cell apoptosis, protein insolubilization, and opacification in experimental cataracts. *J. Biol. Chem.* **288**, 13022–13035 (2013).

64. Ruff, M. *et al.* The disintegrin and metalloprotease ADAM12 is associated with TGF- β -induced epithelial to mesenchymal transition. *PLoS ONE* **10**, e0139179 (2015).
65. Atfi, A. *et al.* The disintegrin and metalloproteinase ADAM12 contributes to TGF- β signaling through interaction with the type II receptor. *J. Cell Biol.* **178**, 201–208 (2007).
66. Rocks, N. *et al.* The metalloproteinase ADAM-12 regulates bronchial epithelial cell proliferation and apoptosis. *Cell Prolif.* **41**, 988–1001 (2008).
67. Schäfer, H. *et al.* TGF- β 1-dependent L1CAM expression has an essential role in macrophage-induced apoptosis resistance and cell migration of human intestinal epithelial cells. *Oncogene* **32**, 180 (2013).
68. Kiefel, H. *et al.* EMT-associated up-regulation of L1CAM provides insights into L1CAM-mediated integrin signalling and NF- κ B activation. *Carcinogenesis* **33**, 1919–1929 (2012).
69. Samatov, T. R., Wicklein, D. & Tonevitsky, A. G. L1CAM: cell adhesion and more. *Prog. Histochem. Cytochem.* **51**, 25–32 (2016).
70. Vogt, J. *et al.* CHRNG genotype-phenotype correlations in the multiple pterygium syndromes. *J. Med. Genet.* **49**, 21–26. <https://doi.org/10.1136/jmedgenet-2011-100378> (2012).
71. Li, M. *et al.* Krüppel-like factor 5 promotes epithelial proliferation and DNA damage repair in the intestine of irradiated mice. *Int. J. Biol. Sci.* **11**, 1458 (2015).
72. Farrugia, M. *et al.* Regulation of anti-apoptotic signaling by Kruppel-like factors 4 and 5 mediates lapatinib resistance in breast cancer. *Cell Death Dis.* **6**, e1699 (2016).
73. Lee, K.-B. *et al.* Clusterin, a novel modulator of TGF- β signaling, is involved in Smad2/3 stability. *Biochem. Biophys. Res. Commun.* **366**, 905–909 (2008).
74. Zhang, H. *et al.* Clusterin inhibits apoptosis by interacting with activated Bax. *Nat. Cell Biol.* **7**, 909 (2005).
75. Shiota, M. *et al.* Clusterin mediates TGF- β -induced epithelial–mesenchymal transition and metastasis via Twist1 in prostate cancer cells. *Cancer Res.* **72**, 5261–5272 (2012).
76. Thomas, C. & Karnoub, A. E. Lysyl oxidase at the crossroads of mesenchymal stem cells and epithelial–mesenchymal transition. *Oncotarget* **4**, 376 (2013).
77. Kasashima, H. *et al.* Lysyl oxidase is associated with the epithelial–mesenchymal transition of gastric cancer cells in hypoxia. *Gastric Cancer* **19**, 431–442 (2016).
78. Hong, H.-H., Uzel, M. I., Duan, C., Sheff, M. C. & Trackman, P. C. Regulation of lysyl oxidase, collagen, and connective tissue growth factor by TGF- β 1 and detection in human gingiva. *Lab. Invest. J. Tech. Methods Pathol.* **79**, 1655–1667 (1999).
79. Conacci-Sorrell, M. *et al.* The shed ectodomain of Nr-CAM stimulates cell proliferation and motility, and confers cell transformation. *Cancer Res.* **65**, 11605–11612 (2005).
80. Zheng, H. *et al.* Fas signaling promotes motility and metastasis through epithelial–mesenchymal transition in gastrointestinal cancer. *Oncogene* **32**, 1183 (2013).
81. Zhang, B. *et al.* Functional polymorphisms in FAS and FASL contribute to increased apoptosis of tumor infiltration lymphocytes and risk of breast cancer. *Carcinogenesis* **28**, 1067–1073 (2006).
82. Tanaka, A. *et al.* Hyperoxia-induced LC3B interacts with the Fas apoptotic pathway in epithelial cell death. *Am. J. Respir. Cell Mol. Biol.* **46**, 507–514 (2012).
83. Ichimaru, Y., Krimmer, D. I., Burgess, J. K., Black, J. L. & Oliver, B. G. TGF- β enhances deposition of perlecan from COPD airway smooth muscle. *Am. J. Physiol. Lung Cell. Mol. Physiol.* **302**, L325–L333 (2011).
84. Bhat, A. A., Ahmad, R., Uppada, S. B., Singh, A. B. & Dhawan, P. Claudin-1 promotes TNF- α -induced epithelial–mesenchymal transition and migration in colorectal adenocarcinoma cells. *Exp. Cell Res.* **349**, 119–127 (2016).
85. Suh, Y. *et al.* Claudin-1 induces epithelial–mesenchymal transition through activation of the c-Abl-ERK signaling pathway in human liver cells. *Oncogene* **32**, 4873 (2013).
86. Akasaka, H. *et al.* Anti-apoptotic effect of claudin-1 in tamoxifen-treated human breast cancer MCF-7 cells. *BMC Cancer* **10**, 548 (2010).
87. Hoewel, T., Macek, R., Swisshelm, K. & Kubbies, M. Reexpression of the TJ protein CLDN1 induces apoptosis in breast tumor spheroids. *Int. J. Cancer* **108**, 374–383 (2004).
88. Masugi, Y. *et al.* Upregulation of integrin β 4 promotes epithelial–mesenchymal transition and is a novel prognostic marker in pancreatic ductal adenocarcinoma. *Lab. Invest.* **95**, 308 (2015).
89. Colburn, Z. T. & Jones, J. C. α 6 β 4 integrin regulates the collective migration of epithelial cells. *Am. J. Respir. Cell Mol. Biol.* **56**, 443–452 (2017).
90. Wang, H., Jin, H. & Rapraeger, A. C. Syndecan-1 and Syndecan-4 capture epidermal growth factor receptor family members and the α 3 β 1 integrin via binding sites in their ectodomains novel statins prevent kinase capture and inhibit α 6 β 4-integrin-dependent epithelial cell motility. *J. Biol. Chem.* **290**, 26103–26113 (2015).
91. Weaver, V. M. *et al.* β 4 integrin-dependent formation of polarized three-dimensional architecture confers resistance to apoptosis in normal and malignant mammary epithelium. *Cancer Cell* **2**, 205–216 (2002).
92. Bacon, W. *et al.* Genetic origin of non-syndromic cleft lip and palate. TWIST, a candidate gene? Research protocol. *L'Orthodontie française* **78**, 249–255 (2007).
93. Nuckolls, G. H., Shum, L. & Slavkin, H. C. Progress toward understanding craniofacial malformations. *Cleft Palate Craniofacial J.* **36**, 12–26 (1999).
94. Hintsala, H. R., Haapasaari, K. M., Soini, Y. & Karihtala, P. An immunohistochemical study of NFE2L2, KEAP1 and 8-hydroxy-2'-deoxyguanosine and the EMT markers SNAI2, ZEB1 and TWIST1 in metastatic melanoma. *Histol. Histopathol.* **32**, 129–136 (2017).
95. Yoon, N. A. *et al.* Tristetraprolin suppresses the EMT through the down-regulation of Twist1 and Snail1 in cancer cells. *Oncotarget* **7**, 8931 (2016).
96. Fan, Q. *et al.* Twist induces epithelial–mesenchymal transition in cervical carcinogenesis by regulating the TGF- β /Smad3 signaling pathway. *Oncol. Rep.* **34**, 1787–1794 (2015).
97. Menon, D. R. *et al.* TGF- β 1 and TNF- α differentially regulate Twist1 mediated resistance towards BRAF/MEK inhibition in melanoma. *Pigment Cell Melanoma Res.* **26**, 912–916 (2013).
98. Ding, H. *et al.* A specific requirement for PDGF-C in palate formation and PDGFR- α signaling. *Nat. Genet.* **36**, 1111 (2004).
99. Han, J., Xiao, Y., Lin, J. & Li, Y. PDGF-C controls proliferation and is down-regulated by retinoic acid in mouse embryonic palatal mesenchymal cells. *Birth Defects Res. B Dev. Reprod. Toxicol.* **77**, 438–444 (2006).
100. Zwerner, J. P. & May, W. A. PDGF-C is an EWS/FLI induced transforming growth factor in Ewing family tumors. *Oncogene* **20**, 626 (2001).
101. Hasegawa, D. *et al.* Wnt5a induces collagen production by human periodontal ligament cells through TGF β 1-mediated upregulation of periostin expression. *J. Cell. Physiol.* **230**, 2647–2660 (2015).
102. Kumawat, K. *et al.* Noncanonical WNT-5A signaling regulates TGF- β -induced extracellular matrix production by airway smooth muscle cells. *FASEB J.* **27**, 1631–1643 (2013).
103. Jin, E.-J. *et al.* Wnt-5a is involved in TGF- β 3-stimulated chondrogenic differentiation of chick wing bud mesenchymal cells. *Int. J. Biochem. Cell Biol.* **38**, 183–195 (2006).
104. Taki, M. *et al.* Downregulation of Wnt4 and upregulation of Wnt5a expression by epithelial–mesenchymal transition in human squamous carcinoma cells. *Cancer Sci.* **94**, 593–597 (2003).

105. Zhang, Y. *et al.* EGF-reduced Wnt5a transcription induces epithelial-mesenchymal transition via Arf6-ERK signaling in gastric cancer cells. *Oncotarget* **6**, 7244 (2015).
106. Zhang, Z. *et al.* Rescue of cleft palate in Msx1-deficient mice by transgenic Bmp4 reveals a network of BMP and Shh signaling in the regulation of mammalian palatogenesis. *Development* **129**, 4135–4146 (2002).
107. Hilliard, S. A., Yu, L., Gu, S., Zhang, Z. & Chen, Y. P. Regional regulation of palatal growth and patterning along the anterior-posterior axis in mice. *J. Anat.* **207**, 655–667. <https://doi.org/10.1111/j.1469-7580.2005.00474.x> (2005).
108. Satokata, I. & Maas, R. Msx1 deficient mice exhibit cleft palate and abnormalities of craniofacial and tooth development. *Nat. Genet.* **6**, 348–356. <https://doi.org/10.1038/ng0494-348> (1994).
109. Chen, Y.-H., Ishii, M., Sucov, H. M. & Maxson, R. E. Msx1 and Msx2 are required for endothelial-mesenchymal transformation of the atrioventricular cushions and patterning of the atrioventricular myocardium. *BMC Dev. Biol.* **8**, 75 (2008).
110. Beverdam, A., Brouwer, A., Reijnen, M., Korving, J. & Meijlink, F. Severe nasal clefting and abnormal embryonic apoptosis in Alx3/Alx4 double mutant mice. *Development* **128**, 3975–3986 (2001).
111. Hudson, R., Taniguchi-Sidle, A., Boras, K., Wiggan, O. N. & Hamel, P. A. Alx-4, a transcriptional activator whose expression is restricted to sites of epithelial-mesenchymal interactions. *Dev. Dyn.* **213**, 159–169 (1998).
112. Yuan, H. *et al.* HOXB13 and ALX4 induce SLUG expression for the promotion of EMT and cell invasion in ovarian cancer cells. *Oncotarget* **6**, 13359 (2015).
113. Soofi, A. *et al.* The kielin/chordin-like protein (KCP) attenuates high-fat diet-induced obesity and metabolic syndrome in mice. *J. Biol. Chem.* **292**, 9051–9062 (2017).
114. Sonnesen, L., Nolting, D., Kjaer, K. W. & Kjaer, I. Association between the development of the body axis and the craniofacial skeleton studied by immunohistochemical analyses using collagen II, Pax9, Pax1, and Noggin antibodies. *Spine* **33**, 1622–1626 (2008).
115. Su, H. Y. *et al.* An epigenetic marker panel for screening and prognostic prediction of ovarian cancer. *Int. J. Cancer* **124**, 387–393 (2009).
116. Cheng, S.-J. *et al.* Hypermethylated ZNF582 and PAX1 are effective biomarkers for detection of oral dysplasia and oral cancer. *Oral Oncol.* **62**, 34–43 (2016).
117. Hassan, Z. K., Hafez, M. M., Kamel, M. M., Rahman, A. & Zekri, N. Human papillomavirus genotypes and methylation of CADM1, PAX1, MAL and ADCYAP1 genes in epithelial ovarian cancer patients. *Asian Pac. J. Cancer Prev.* **18**, 169–176 (2017).
118. Juriloff, D. M., Harris, M. J., McMahon, A. P., Carroll, T. J. & Lidral, A. C. Wnt9b is the mutated gene involved in multifactorial nonsyndromic cleft lip with or without cleft palate in A/WySn mice, as confirmed by a genetic complementation test. *Birth Defects Res. A Clin. Mol. Teratol.* **76**, 574–579. <https://doi.org/10.1002/bdra.20302> (2006).
119. Juriloff, D. M. & Harris, M. J. Mouse genetic models of cleft lip with or without cleft palate. *Birth Defects Res. A Clin. Mol. Teratol.* **82**, 63–77. <https://doi.org/10.1002/bdra.20430> (2008).
120. Lan, Y. *et al.* Expression of Wnt9b and activation of canonical Wnt signaling during midfacial morphogenesis in mice. *Dev. Dyn.* **235**, 1448–1454. <https://doi.org/10.1002/dvdy.20723> (2006).
121. Cadigan, K. M. & Nusse, R. Wnt signaling: a common theme in animal development. *Genes Dev.* **11**, 3286–3305 (1997).
122. Wodarz, A. & Nusse, R. Mechanisms of Wnt signaling in development. *Ann. Rev. Cell Dev. Biol.* **14**, 59–88 (1998).
123. Mir, H. *et al.* Occludin deficiency promotes ethanol-induced disruption of colonic epithelial junctions, gut barrier dysfunction and liver damage in mice. *Biochim. Biophys. Acta (BBA) Gen Subj.* **1860**, 765–774 (2016).
124. Rachow, S. *et al.* Occludin is involved in adhesion, apoptosis, differentiation and Ca²⁺-homeostasis of human keratinocytes: implications for tumorigenesis. *PLoS ONE* **8**, e55116 (2013).
125. Tsai, S.-H. *et al.* TGF- β inhibits IL-1 β -activated PAR-2 expression through multiple pathways in human primary synovial cells. *J. Biomed. Sci.* **16**, 97 (2009).
126. Xiang, Y. *et al.* Expression of proteinase-activated receptors (PAR)-2 in articular chondrocytes is modulated by IL-1 β , TNF- α and TGF- β . *Osteoarthr. Cartil.* **14**, 1163–1173 (2006).
127. Xie, L. *et al.* Enhanced proliferation of human hepatoma cells by PAR-2 agonists via the ERK/AP-1 pathway. *Oncol. Rep.* **28**, 1665–1672 (2012).
128. Al-Eryani, K. *et al.* Protease-activated receptor 2 modulates proliferation and invasion of oral squamous cell carcinoma cells. *Hum. Pathol.* **46**, 991–999 (2015).
129. Choi, Y.-L. *et al.* LYN is a mediator of epithelial-mesenchymal transition and a target of dasatinib in breast cancer. *Cancer Res.* **70**, 2296–2306 (2010).
130. Wei, S. *et al.* Critical role of Lyn kinase in inhibition of neutrophil apoptosis by granulocyte-macrophage colony-stimulating factor. *J. Immunol.* **157**, 5155–5162 (1996).
131. Li, G. *et al.* Lyn mitigates mouse airway remodeling by downregulating the TGF- β 3 isoform in house dust mite models. *J. Immunol.* **191**, 5359–5370 (2013).
132. Letra, A., Menezes, R., Granjeiro, J. M. & Vieira, A. R. AXIN2 and CDH1 polymorphisms, tooth agenesis, and oral clefts. *Birth Defects Res. A Clin. Mol. Teratol.* **85**, 169–173 (2009).
133. Song, H. *et al.* Association of single-nucleotide polymorphisms of CDH1 with nonsyndromic cleft lip with or without cleft palate in a northern Chinese Han population. *Medicine* **96**, e5574 (2017).
134. Toffoli, B. *et al.* Osteoprotegerin promotes vascular fibrosis via a TGF- β 1 autocrine loop. *Atherosclerosis* **218**, 61–68 (2011).
135. Thirunavukkarasu, K. *et al.* Stimulation of osteoprotegerin (OPG) gene expression by transforming growth factor- β (TGF- β) mapping of the OPG promoter region that mediates TGF- β effects. *J. Biol. Chem.* **276**, 36241–36250 (2001).
136. Zimmermann, D. R., Dours-Zimmermann, M. T., Schubert, M. & Bruckner-Tuderman, L. Versican is expressed in the proliferating zone in the epidermis and in association with the elastic network of the dermis. *J. Cell Biol.* **124**, 817–825 (1994).
137. Wu, Y. *et al.* Versican V1 isoform induces neuronal differentiation and promotes neurite outgrowth. *Mol. Biol. Cell* **15**, 2093–2104 (2004).
138. Sheng, W. *et al.* The roles of versican V1 and V2 isoforms in cell proliferation and apoptosis. *Mol. Biol. Cell* **16**, 1330–1340 (2005).
139. LaPierre, D. P. *et al.* The ability of versican to simultaneously cause apoptotic resistance and sensitivity. *Cancer Res.* **67**, 4742–4750 (2007).
140. Vanhoutteghem, A. *et al.* Basonuclin 2 has a function in the multiplication of embryonic craniofacial mesenchymal cells and is orthologous to disco proteins. *Proc. Natl. Acad. Sci.* **106**, 14432–14437 (2009).
141. Sgorbissa, A., Benetti, R., Marzinotto, S., Schneider, C. & Brancolini, C. Caspase-3 and caspase-7 but not caspase-6 cleave Gas2 in vitro: implications for microfilament reorganization during apoptosis. *J. Cell Sci.* **112**, 4475–4482 (1999).
142. Manzow, S., Brancolini, C., Marks, F. & Richter, K. H. Expression of growth arrest-specific (Gas) genes in murine keratinocytes: Gas2 is specifically regulated. *Exp. Cell Res.* **224**, 200–203 (1996).
143. Qiu, M. *et al.* Role of the Dlx homeobox genes in proximodistal patterning of the branchial arches: mutations of Dlx-1, Dlx-2, and Dlx-1 and -2 alter morphogenesis of proximal skeletal and soft tissue structures derived from the first and second arches. *Dev. Biol.* **185**, 165–184 (1997).
144. Chiba, S. *et al.* Homeoprotein DLX-1 interacts with Smad4 and blocks a signaling pathway from activin A in hematopoietic cells. *Proc. Natl. Acad. Sci.* **100**, 15577–15582 (2003).
145. Starkova, J. *et al.* Up-regulation of homeodomain genes, DLX1 and DLX2, by FLT3 signaling. *Haematologica* **96**, 820–828 (2011).

146. Pearton, D. J. *et al.* Elf5 counteracts precocious trophoblast differentiation by maintaining Sox2 and 3 and inhibiting Hand1 expression. *Dev. Biol.* **392**, 344–357 (2014).
147. Borck, G. *et al.* eIF2 γ mutation that disrupts eIF2 complex integrity links intellectual disability to impaired translation initiation. *Mol. Cell* **48**, 641–646 (2012).
148. Brookes, E. *et al.* Mutations in the intellectual disability gene KDM5C reduce protein stability and demethylase activity. *Hum. Mol. Genet.* **24**, 2861–2872 (2015).
149. Premi, S., Srivastava, J., Panneer, G. & Ali, S. Startling mosaicism of the Y-chromosome and tandem duplication of the SRY and DAZ genes in patients with turner syndrome. *PLoS ONE* **3**, e3796 (2008).
150. Blok, L. S. *et al.* Mutations in DDX3X are a common cause of unexplained intellectual disability with gender-specific effects on Wnt signaling. *Am. J. Hum. Genet.* **97**, 343–352 (2015).
151. Luo, Q., Nieves, E., Kzhyshkowska, J. & Angeletti, R. H. Endogenous transforming growth factor- β receptor-mediated Smad signaling complexes analyzed by mass spectrometry. *Mol. Cell. Proteom.* **5**, 1245–1260 (2006).
152. Sivaprasad, U. *et al.* A nonredundant role for mouse Serpinb3a in the induction of mucus production in asthma. *J. Allergy Clin. Immunol.* **127**, 254–261 (2011).
153. Sivaprasad, U. *et al.* SERPINB3/B4 contributes to early inflammation and barrier dysfunction in an experimental murine model of atopic dermatitis. *J. Investig. Dermatol.* **135**, 160–169 (2015).
154. Heo, W. I. *et al.* Identification of novel candidate variants including COL6A6 polymorphisms in early-onset atopic dermatitis using whole-exome sequencing. *BMC Med. Genet.* **18**, 8 (2017).
155. Ness, S. L. *et al.* Mouse keratin 4 is necessary for internal epithelial integrity. *J. Biol. Chem.* **273**, 23904–23911 (1998).
156. Opitz, O. G., Jenkins, T. D. & Rustgi, A. K. Transcriptional regulation of the differentiation-linked human K4 promoter is dependent upon esophageal-specific nuclear factors. *J. Biol. Chem.* **273**, 23912–23921 (1998).
157. Gazda, H. T. *et al.* Ribosomal protein L5 and L11 mutations are associated with cleft palate and abnormal thumbs in Diamond-Blackfan anemia patients. *Am. J. Hum. Genet.* **83**, 769–780 (2008).

Acknowledgements

This study was financially supported by a NIDCR, NIH grant (R01DE017986) to Dr. Ali Nawshad. Jingpeng Liu was supported by a China Scholar Council (CSC) scholarship (Beijing, China). The authors thank Ms. Marian J. Schmid of the College of Dentistry, University of Nebraska Medical Center (UNMC), Lincoln, Nebraska for assistance with animal handling, housing and maintenance in the facility.

Author contributions

The study design: A.N. and H.H.O.; study conduct: J.L., S.K.C., and K.M.W.; data collection: J.L., K.M.W., S.K.C., and M.A.; data analysis: J.L., S.K.C., K.M.W., and M.A.; data interpretation: J.L., S.K.C., K.M.W., M.A., A.N., and H.H.O.; drafting manuscript: J.L., S.K.C., K.M.W., A.N., and H.H.O.; revising manuscript content: J.L., A.N., and H.H.O.; approving the final version of manuscript: A.N. and H.H.O.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41598-020-71636-0>.

Correspondence and requests for materials should be addressed to A.N.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020