Single cell transcriptomics and developmental trajectories of murine cranial neural crest cell fate determination and cell cycle progression

Yu Ji[1,2,3], Shuwen Zhang[1,2], Kurt Reynolds[1,2,3], Ran Gu[1,2], Moira McMahon[2], Mohammad Islam[2], Yue Liu[2], Taylor Imai[2], Rebecca Donham[2], Huan Zhao[2], Ying Xu[4], Diana Burkart-Waco[5], Chengji J. Zhou[1,2,3*]

[1]Department of Biochemistry and Molecular Medicine, University of California at Davis, School of Medicine, Sacramento, California 95817, USA

[2]Institute for Pediatric Regenerative Medicine, Shriners Hospitals for Children and UC Davis School of Medicine, Sacramento, California 95817, USA

[3]Biochemistry, Molecular, Cellular, and Developmental Biology (BMCDB) graduate program, University of California, Davis, California 95616, USA

[4]Cambridge-SU Genomic Resource Center, Medical College of Soochow University, Suzhou 215006, China

[5]DNA Technologies and Expression Analysis Core, Genome Center, University of California, Davis, California 95616, USA

[*]**Corresponding author: cjzhou@ucdavis.edu**

Abstract

Cranial neural crest (NC) cells migrate long distances to populate the future craniofacial regions and give rise to various tissues, including facial cartilage, bones, connective tissues, and cranial nerves. However, the mechanism that drives the fate determination of cranial NC cells remains unclear. Using single-cell RNA sequencing combined genetic fate mapping, we reconstructed developmental trajectories of cranial NC cells, and traced their differentiation in mouse embryos. We identified four major cranial NC cell lineages at different status: pre-epithelial-mesenchymal transition, early migration, NC-derived mesenchymal cells, and neural lineage cells from embryonic days 9.5 to 12.5. During migration, the first cell fate determination separates cranial sensory ganglia, the second generates mesenchymal progenitors, and the third separates other neural lineage cells. We then focused on the early facial prominences that appear to be built by undifferentiated, fast-dividing NC cells that possess similar transcriptomic landscapes, which could be the drive for the facial developmental robustness. The post-migratory cranial NC cells exit the cell cycle around embryonic day 11.5 after facial shaping is completed and initiates further fate determination and differentiation processes. Our results demonstrate the transcriptomic landscapes during dynamic cell fate determination and cell cycle progression of cranial NC lineage cells and also suggest that the transcriptomic regulation of the balance between proliferation and differentiation of the post-migratory cranial NC cells can be a key for building up unique facial structures in vertebrates.

Introduction

The craniofacial complex is one of the most diversified anatomical parts among vertebrates. Many studies have been focused on external factors like feeding behaviors that drive the selection of species-specific patterns of the craniofacial complex (Langenbach and van Eijden,

2001). Mutations that alter the molecular and cellular mechanisms of facial developmental processes could potentially provide the variation for selection, yet these mechanisms are still not well understood. The craniofacial complex comprise a cover of epithelial tissues, which are important signaling centers for craniofacial patterning (Hu and Marcucio, 2009), and mesenchymal tissues that build the substance of the complex and unique facial structures. Unlike the trunk mesenchyme that is derived from mesoderm, the facial mesenchyme mainly arises from cranial neural crest (NC) cells during development and evolution (Chai and Maxson, 2006, Dash and Trainor, 2020).

The NC cells are transiently existing multipotent stem cells and originated from the dorsal edge of the neural fold or neural tube during embryonic development (Ji et al., 2019, Crane and Trainor, 2006). After induction, the NC cells delaminate from the neuroectoderm and migrate long distances to generate numerous cell types (Shyamala et al., 2015). NC cells induced from the mid-diencephalon and hindbrain regions are named cranial or cephalic NC cells (Lumb et al., 2017). They migrate into the frontonasal prominence (FNP) and the first pharyngeal arch (Lumsden et al., 1991) to give rise to the paired lateral nasal prominences, the medial nasal prominences, the maxillary prominences, and the mandibular prominences, and eventually generate cartilage, bone, connective tissue, and other derivatives at the anterior region of the craniofacial complex and upper/lower jaws (Cordero et al., 2011, Chai and Maxson, 2006, La Noce et al., 2014).

When cranial NC cells arrive at the facial regions, the craniofacial complex is still very small. As the embryo develops, the NC cells undergo rapid proliferation to confer the correct shape and size of the craniofacial complex. Little is known about how the cranial NC cells generate different structures of the face. Some evidence shows that the fate of pre-migratory trunk NC cells can be predicted by the time they delaminate from the dorsal neural tube, as well as their ventral-dorsal position within the neural tube, suggesting there is an intrinsically

programmed pre-patterning that regulates the early fate restrictions of NC cells (Nitzan et al., 2013, Krispin et al., 2010). However, unlike the trunk NC cells, which undergo EMT individually, the cranial NC cells delaminate from the neural tube and migrate as a coherent sheet of cells (Ji et al., 2019, Alfandari et al., 2010). Therefore, it is not known if defined NC-derived parts within the craniofacial complex arise from pre-determined regions of the dorsal neural tube. On the other hand, NC cells always interact with the mesoderm and ectoderm in the first pharyngeal arch to generate a fully functioning jaw (Baker and Bronner-Fraser, 2001). Signals, such as Shh, Wnt, Fgf and Bmp, from the epithelial signaling centers could guide the outgrowth of the frontonasal structures (Jheon and Schneider, 2009, Foppiano et al., 2007, Kasberg et al., 2013, Szabo-Rogers et al., 2008, Creuzet et al., 2004), indicating that the environment within the FNP may be critical for establishing the fate restrictions of NC cells. Therefore, whether the intrinsic or extrinsic cues that regulate cranial NC cells to commit to a restricted cell fate still needs further investigations.

In this study, we combined single-cell RNA sequencing (scRNA-seq) and lineage tracing to investigate the cell fate decisions involved in cranial NC lineage cells. We identified four major cranial NC cell lineages at different status and found that the cell fate determination that separates neural and mesenchymal cell lineages happens twice. We then focused on the NC-derived early mesenchymal cells right before the formation of the bones, cartilage, and other tissues in the face. We demonstrate that the differentiation of NC-derived mesenchymal cells initiates at a relatively late developmental stage, after the outgrowth of the craniofacial complex. Our results indicate that the initiation of the differentiation of NC-derived mesenchymal cells is coupled to cell cycle exit, indicating that instead of pre-patterned before migration, cell cycle regulation may play a more critical role in cranial NC cell fate determination and differentiation.

Results

A single-cell graph of NC developmental process in the murine embryos

We combined lineage tracing and scRNA-seq to address the spatiotemporal dynamics and investigate the cell fate decisions involved in NC cell differentiation after they migrate to their target regions (Fig. 1A). To genetically label NC cells and their progenies, we crossed Wnt1-Cre mice with the Rosa26-tdTomato/eGFP (Rosa26-mT/mG) reporter mouse line. The Wnt1-Cre-mediated recombination has been shown to occur in cranial and cardiac neural crest cells, midbrain, and developing neural tube (Tavares and Clouthier, 2015). Therefore, it can label these tissues and their progeny cells with green fluorescence. Since the cranial NC cells migrate to their target region around E8.5 and the generation of the facial prominences is completed by E12.5 (Ji et al., 2020), we collected eGFP positive cells from E9.5, E10.5, E11.5, and E12.5 embryos using fluorescence-activated cell sorting (Fig. 1B, C). Therefore, our data covered most of the circuital stages for forming, fusion, and merging craniofacial prominences (Ji et al., 2019). Then we performed 10X single-cell RNA sequencing using live eGFP positive cells. The sequencing detected roughly 3,382, 3,380, 3,012, and 2,465 genes per cell from E9.5, E10.5, E11.5, and E12.5 embryos, respectively (Fig. S1).

In addition to the dorsal spinal cord, where the premigratory NC cells are located, the Wnt1-Cre is also expressed at the midlines of the midbrain, the caudal diencephalon, and the midbrain-hindbrain junction after the closure of the dorsal neural tube (Debbache et al., 2018). However, cells from the central nervous system can be distinguished from the NC-derived cells based on gene expression. We sorted out the non-NC cells based on their marker genes *in silico*. For example, the neural tube population expresses neural plate border specifier genes such as *Zic1*, *Zic3*, and *Pax3* (Soldatov et al., 2019), while cells from the developing midbrain, hindbrain, and cerebellum express *Wnt7b*, *Wnt8b*, *Hes5*, *Lhx1*, and *Lhx5* (Garda et al., 2002, Zhao et al., 2007). Moreover, epithelial cells (specifically expressing *Epcam*, *Krt8*, *Krt14*), endothelial cells (specifically expressing *Cdh5*), and red blood cells (specifically expressing

Hbb-y, *Hba-x*, *Hba-a2*), which are small clusters containing only eGFP negative cells, were also removed from further analysis.

After removing the non-NC cells from our dataset, a total of 19,640 single-cell transcriptomes passed quality control for further analysis, including 3,620, 4,284, 4,302, and 7,434 cells from E9.5, E10.5, E11.5, and E12.5 embryos, respectively. Using Uniform Manifold Approximation and Projection (UMAP), we analyzed the transcriptional heterogeneity of NC-derived cells. When the data from all four time points were analyzed together, we found that the NC cells clustered into four major cell populations.

Identification of the major NC-derived cell types

We identified the major NC-derived cell types based on the expression of known marker genes. The major NC-derived cell populations were identified as pre-EMT, which express *Wnt1* and *Wnt3a* (Ikeya et al., 1997), early migration, which express *Sox10* and *Foxd3* (Kirby and Hutson, 2010), NC-derived mesenchymal cells, which express *Twist1* and *Prrx2* (Soldatov et al., 2019), and the neural linage cells, which express *Tubb3* and *Elavl3* (Soldatov et al., 2019, Delile et al., 2019) (Fig. 2A, B, and Fig. S2, Table S1). Notably, a small cluster representing the cardiac neural crest cells that migrate to the outflow tract were identified based on their expression of *Tbx20* and *Acta2* (Singh et al., 2005). These results argue that our data captured a majority of known NC-derived cell types.

Previous studies showed that both pre-EMT and early migratory NC cells maintained multipotency in mice (Baggiolini et al., 2015). One important characteristic of multipotent stem cells is their ability to self-renew. Therefore, we analyzed the cell cycle states across the major cell populations using the Seurat scRNA-seq analysis pipeline (Fig. 2C, Fig. S3). Indeed, 64.2% of the pre-EMT and 86.7% of the early migratory neural crest cells are in the S or G2/M phase

of the cell cycle, suggesting they are fast-dividing cells consistent with the hypothesis that they are multipotent progenitor cells. In contrast, only 9.3% of neural linage cells are at the S or G2/M phase, indicating that the neural linage is no longer actively dividing. Interestingly, our results showed that 80.5% of the mesenchymal cells are at the S or G2/M phase, indicating that most mesenchymal cells are also proliferating, the frequency of which was dynamic across the sampled time points (Fig. 2D). The frequency of pre-EMT and early migratory NC cells declined over time, while mesenchymal and neural lineage cells increased dramatically after E11.5. These results suggest that our data represent the developmental process that NC progenitor cells lose multipotency and differentiate into various fates.

Reconstruction of developmental trajectories of NC cells

One of the benefits of scRNA-seq is that it allows mapping dynamic differentiation process by densely sampling cells at different stages, in our case E9.5, E10.5, E11.5, and E12.5. Therefore, these NC-derived cells sampled at different time points can be used to create a continually NC developmental trajectory. To reconstruct the progression that cells undergo during their differentiation from pre-EMT NC progenitor cells to fate determined neural or mesenchymal cells, we performed pseudotime analysis on all of the eGFP positive cells (Fig. 3A, B). The resulting lineage tree demonstrates the transcriptional changes associated with cell fate splits. After emigrating from the dorsal neural tube, the fate determination that separate neural and mesenchymal cell lineages happen twice (Fig. 3C). The first differentiation separates cranial sensory ganglia (expressing *Tlx2*, *Tlx3*, and *Neurog1*) (Logan et al., 1998) from progenitors of mesenchymal and the other neural linage, such as Schwann cell precursors (expressing *Pou3f1* and *Pu3f2*). The second fate determination generates the mesenchymal branch (Fig. 3D). Soldatov *et al.* also found that NC cell fate determination occurred through a progression of binary decisions, similar to our results (Soldatov et al., 2019).

Characterization of the mesenchymal cells at facial prominences

Once committed to a mesenchymal fate, cranial NC cells generate most of the bone and cartilage in the craniofacial regions. However, it remains unclear how post-migratory NC-derived mesenchymal cells are induced to differentiate into the specific craniofacial skeletal structures with the correct size and shape. To reveal the mechanisms that drive the fate determination of craniofacial mesenchymal cells, we re-clustered the 9,875 NC-derived mesenchymal cells collected from E9.5, E10.5, E11.5, and E12.5 embryos (Fig. 4A, B, Table S2). Using cluster-specific marker genes and wholemount *in situ* hybridization (WISH), we identified fifteen clusters associated with NC cells that migrate into the frontonasal prominences and the maxillary prominence of the first pharyngeal arch (Fig. 4C-E). For example, the aristaless-like homeobox *Alx* genes, *Alx1*, *Alx3,* and *Alx4*, are widely expressed in the midfacial complex at E10.5, and cells that highly express these genes are enriched in cluster m1 (mesenchymal 1), suggesting these clusters are from the midfacial prominences. By E11.5, *Alx1* and *Alx4* continue to be widely expressed in all three paired midfacial primordia, the lateral nasal prominence (LNP), the medial nasal prominence (MNP), and the maxillary prominence (MxP), suggesting that clusters m1, 3, 7, and 11 represent the mesenchymal cells at midfacial prominences. The expression of *Alx3* is limited to MNP, indicating that clusters m3 and 11 are the MNP. In contrast, *Barx1* is widely expressed in the mandibular prominence of the first pharyngeal arch, the second, third, fourth, and sixth pharyngeal arches at E10.5, suggesting clusters m2, 4, 5, 6, 8, 9, 10 are from the pharyngeal regions. At E11.5, a small part of ventral MxP also begins to express *Barx1*, and those cells are found in cluster m3. Together, using cluster-specific markers, we identified m1, 3, 7, and 11 as the mesenchymal cell populations of the midfacial prominences. Interestingly, cells from early developmental stages (E9.5 and E10.5) are enriched in cluster m1, while cells from late developmental stages (E11.5 and E12.5) are scattered in all four clusters, suggesting

that the transcriptional heterogeneity for cells from E11.5 and E12.5 is more complicated compared to E9.5 and E10.5 (Fig. 4B).

The differentiation of NC-derived craniofacial mesenchymal cells starts at a relatively late developmental stage

We next performed a pseudotime analysis of cells from clusters m1, 3, 7, and 11 using Monocle 2 to reconstruct the development processes (Trapnell et al., 2014). Monocle 2 performed an unsupervised analysis to order the cells and reconstructed a tree-like trajectory, beginning with E9.5 and E10.5 cells and ending with E11.5 and E12.5 cells (Fig. 5A, B). Notably, cells in the "root" branch are mostly from E9.5 and E10.5 embryos. By contrast, cells from E11.5 and E12.5 embryos are spread out in the other four states, indicating that the differentiation of craniofacial mesenchymal cells is likely to initiate at a later developmental stage (Fig. 5C).

The transcriptional complexity of NC lineage cells decreases with developmental progression (Jean-Baptiste et al., 2019, Saunders et al., 2019, Gulati et al., 2020). Dorrity et al. developed a mathematical model to calculate the developmental progression of a single cell based on the number of genes that were detected per cell (Dorrity, 2020). To test whether the differentiation of craniofacial mesenchymal cells starts after E11.5, we measured the developmental progression of craniofacial mesenchymal cells (Fig. 5F, G). The result shows that the transcriptional complexity decreases in E11.5 and E12.5 embryos, leading to a higher developmental progression score, while E9.5 and E10.5 embryos remain at early developmental stages (Fig. 5F, G). In addition, we identified that more than 10,000 genes that exhibit temporal expression patterns, which fall into two distinct gene groups (Fig. 5D, Table S3). One group of genes are highly expressed in the early development stages (E9.5 and E10.5). The other group of genes is highly expressed at the late development stages (E11.5 and E12.5) and includes

genes like *Msx1* and *Col1a1* that are critical for craniofacial development (Fig. 5E), suggesting the ossification of craniofacial mesenchymal cells initiates at E11.5.

To test whether the ossification of NC-derived craniofacial mesenchymal cells starts at E11.5, we examined the expression of genes regulating bone formation in cells at different time points along the developmental trajectory. The ossification of undifferentiated mesenchymal cells into bone cells begins with the formation of osteoprogenitors. This step is regulated by master transcription factors such as *Sox9*, *Runx2,* and *Msx1* (Javed et al., 2010). The expression of these genes increased after E11.5 along the trajectory (Fig. 5E). During the next stage of osteoblast development, the cells start to express genes like collagen and fibronectin genes (Rutkovskiy et al., 2016), which are known to be critical for the arrest of cell motility during the osteoblast-to-osteocyte transition (Shiflett et al., 2019). In our data, the expression of *Col1a1*, *Col2a1*, and *Col3a1* is increased after the first branch point along the trajectory (Fig. 5E), suggesting the ossification of craniofacial mesenchymal cells initiates after E11.5. Therefore, although NC-derived mesenchymal cells populate at the craniofacial complex and give rise to different prominences as early as E9.5, they remain at an undifferentiated, homogeny stage. The differentiation of NC-derived mesenchymal cells into different bones in the facial region does not initiate until E11.5, suggesting the fate of cranial NC cells might not be intrinsically programmed but is acquired from the environment.

The initiation of the ossification is coupled to cell cycle progression

To reveal the mechanism that triggers the differentiation of cranial NC-derived mesenchymal cells, we performed branch-dependent expression analysis at the first branch point (Fig. 6A, Table S4). The results showed that one group (Group 1 in Fig. 6A) of genes is highly expressed in the root branch, and the expression of these genes significantly decreased in both branches,

238 especially in cells at fate 1. Gene ontology (GO) enrichment analysis suggests that these genes
239 are involved in cell cycle regulation, chromosome segregation, and microtubule cytoskeleton
240 organization. In addition, many genes that inhibit the differentiation process are also
241 downregulated in both branches (Fig. 6B). For example, the inhibitor of DNA binding and cell
242 differentiation protein (Id1) has been shown to be expressed in embryonic and somatic stem
243 cells and sustains the stemness of these cells through inhibition of differentiation (Jankovic et
244 al., 2007, Nakashima et al., 2001, Nam and Benezra, 2009, Ying et al., 2003). *Id1* expression
245 was downregulated in both cell fates. *Stub1*, encoding an E3 ligase that negatively regulates
246 ossification by inducing the degradation of Runx2, is also down-regulated in both branches (Li et
247 al., 2008). These results suggest that at E11.5, the NC-derived craniofacial mesenchymal cells
248 adopt a ready-to-differentiate state.

249 Another group of genes (Group 3 in Fig. 6A) has been shown to be highly expressed in
250 cells at one branch (fate 1 in Fig. 5A and Fig. 6A). Functional analysis of genes in this group
251 suggests that they are involved in skeletal development, extracellular matrix organization, and
252 ossification (Fig. 6A). For example, bone-specific genes, such as collagens, *Sox9*, and *Msx1*,
253 are specifically up-regulated in fate 1. On the other hand, cell cycle genes, including *Top2a*,
254 *Cdk1*, *Cdc20*, and *Rrs1*, are upregulated in fate 2 (Fig. 6B). Moreover, 98.5% and 95% of cells
255 from E9.5 and E10.5 embryos are dividing cells, while only 62.5% and 62.7% of cells from
256 E11.5 and E12.5 embryos are dividing cells (Fig. 6C, D), indicating that a group of cells has
257 exited the cell cycle by E11.5. Given that the first branch point of the trajectory separates cells
258 before or after E11.5, our results indicate that this branch point might be a decision-making
259 point for a population of cells to exit the cell cycle and start to differentiate.

260 To address the reproducibility of our analysis, we isolated post-migratory NC cells
261 labeled by Sox10-Cre;Rosa26-mT/mG reporter from E11.5 embryos, and performed scRNA-seq
262 analysis (Hari et al., 2012). Indeed, the facial mesenchymal cells from the Sox10-Cre dataset

generated the same four clusters with a similar percentage of cells in each cluster as the Wnt1-Cre dataset. Moreover, the expression level of all genes, the patterns of marker genes were also similar between Wnt1-Cre and Sox10-Cre datasets (Fig. 7). Therefore, our analysis revealed highly reproducible NC-derived cell populations associated with facial prominences, which allows us to combine Wnt1-Cre and Sox10-Cre datasets for further analysis.

The facial mesenchymal cells from E11.5 embryos were populated into four clusters. Based on marker genes and published literature (Li et al., 2019), we found that m0, m1, and m2 contain cells from MNP and LNP (expressing *Alx1*, *Alx3*, *Alx4*, and *Pax7*), and m3 contains cells from MxP (expressing *Barx1* and *Asb4*) (Figs. 7F, 8A, B, Table S5). Cell cycle analysis indicated that 83.2% of cells in m2 and 67.7% in m3 are actively dividing cells while only 46.1% m0 and 44.9% m1 cells are undergoing mitosis (Fig. 8C, D). Further analysis revealed that cells in m0 and m1 exhibit more transcriptional complexity than cells in m2 and m3, indicating that m2 and m3 might be at an earlier developmental stage than m0 and m1 (Fig. 8E). During mouse facial development, the morphogenesis of the LNP- and MNP-derived structures progresses from E10.5 to E12.5. However, MxP cells continue to grow until E15.5 to give rise to the secondary palate (Ji et al., 2020). This is consistent with our results that, at E11.5, LNP and MNP cells (m0, m1, and m2) were grouped into three clusters at different differential stages while most of the MxP cells (m3) are still at an early developmental stage.

To further reveal the possible mechanisms that drive the differentiation of LNP and MNP cells, we analyzed the differentially expressed gene expression between m2 and m0 (Table S6) and m2 and m1 (Table S7). As a result, both *Twist1* and *Id2*, known to be required for cell proliferation during the early osteoblast differentiation stage (Javed et al., 2010, Sakata-Goto et al., 2012), were expressed at a higher level in m2. In contrast, many genes involved in cartilage formation, such as *Sox9*, *Col2a1*, *Itm2a*, *Igfbp5*, and *Col9a1*, are expressed at a higher level in m0 than in m2, suggesting cells in m0 are chondroprogenitors (Fig. 8F). Although cells in m1 do

not have many highly expressed genes compared to m2, *Igfbp5* and *Itm2a* were found to also highly express in m1. *Igfbp5* has been shown to stimulate bone cell growth (Miyakoshi et al., 2001). Studies in mice also showed that *Itm2a was* involved in osteogenic differentiation (Tuckermann et al., 2000). These results indicate that m1 might represent a transition stage between mesenchymal stem cells and chondroprogenitors. Moreover, we also found that *Wnt5a*, *Lef1*, and *Crabp1* were highly expressed in m2. In contrast, a WNT antagonist gene *Dkk2* was highly expressed in m0, suggesting Wnt and retinoic acid signaling might be essential for maintaining the self-renewal of mesenchymal stem cells.

Discussion

The facial region is mainly comprised of NC-derived cells. However, how cranial NC cells develop into the facial structures is still not entirely clear. We addressed this question by tracing the lineage of NC cells in mouse embryos. Using the single-cell transcriptomic data, we described a spatiotemporal molecular specification tree of post-migratory cranial NC cells, showing the fate determination process of NC-derived mesenchymal cells at the facial prominences. Our study indicates that the differentiation of NC-derived craniofacial mesenchymal cells initiates as late as E11.5, and the differentiation is coupled with the exit of the cell cycle.

In mice, the cranial NC cells migrate to the anterior of the embryos as early as E8.5 and build the frontonasal prominence (FNP) at E9.25. By E10.25, the FNP gives rise to the paired MNPs and LNPs. The MNPs continue to fuse with the MxP at E11.25, generating the upper jaw (Everson et al., 2018). Eventually, the cranial NC cells generate most of the bone and cartilage in the craniofacial region. However, how the fate of each NC cell is determined is unclear. One hypothesis was that NC cells contain an intrinsically programmed molecular facial patterning

"blueprint" when they delaminate from the neural tube. However, our single-cell transcriptome data shows that at E9.5 and E10.5, NC-derived mesenchymal cells in different facial prominences are similar to each other at the transcriptomic level, regardless of where they come from. At a later developmental stage, E11.5, the cells exhibit more transcriptional heterogeneity, suggesting that the fates of different cell populations are beginning to diverge. Reconstruction of a lineage trajectory of cranial NC lineage cells also indicates the differentiation of NC cells initiates at E11.5. Therefore, we propose that cranial NC cells maintain their differentiation potential until the morphology of the face is shaped. In support of our new hypothesis, Kaucka *et al.* showed that the shape of the face is mainly formed by local cellular divisions (Kaucka et al., 2016). The authors also found that the proliferation activity of NC cells resulted in cellular mixing in the facial tissue. They proposed that the fact that progenies of several NC cells locally mixed might guarantee the developmental robustness of facial complex. Our results support this hypothesis because we found that fast-dividing NC cells maintain their stemness. Therefore, hypothetically, mutations that happen in an NC cell can be counterbalanced by unaffected neighbor undifferentiated NC cells.

Terminal differentiation of many multipotent cells such as neural stem cells is associated with cell cycle exiting (Soufi and Dalton, 2016, Hardwick and Philpott, 2014). Our results showed that the NC-derived mesenchymal stem cell differentiation is also cell cycle-dependent. The molecular mechanisms that underline this phenomenon still need further study. One possible mechanism is the epigenetic landscape at the developmental genes that might change in the G1 phase. In pluripotent stem cells, most developmental genes are H3K4 and H3K27 trimethylated near their transcription start sites to be silenced by a polycomb-dependent mechanism (Bernstein et al., 2006). However, transcriptional leakiness of developmental genes could happen at the G1 phase, making the G1 phase an opportunity for differentiation (Singh et al., 2013). Consistent with this hypothesis, ablation of the plycomb complex-associated

methyltransferase gene *Ezh2* in NC cells causes various facial deformities (Kim et al., 2018, Schwarz et al., 2014). Also, BMP signaling has been shown to be required for guiding the outgrowth of facial prominences (Graf et al., 2016). G1-specific cyclin-dependent protein kinases (CDKs) have been shown to target transcription factors like Smad2/3, leading to the expression of their target genes in the G1 phase (Kim et al., 2018, Pauklin and Vallier, 2013). Cell cycle regulation has been known to play an important role in many development processes of NC cells. For example, trunk NC cells delaminate from the dorsal neural tube only in the S phase (Burstyn-Cohen and Kalcheim, 2002). *In vivo* studies also revealed the dividing activity of cranial NC cells is increased as they migrated into the branchial arches (Ridenour et al., 2014). These cells continue dividing to form the facial complex with the correct shape and size (Kaucka et al., 2016). Our study shows that at a relatively late developmental stage (E11.5), after building the morphology of the craniofacial structures, the NC cells exit the cell cycle and start to differentiate. This may indicate that the variety of craniofacial shapes and functions in different species might be regulated by the rates of proliferation and the time exiting the cell cycle.

The fact that cranial NC-derived mesenchymal cells keep rapidly dividing until E11.5 suggests some signals promoting fast NC cell proliferation. Comparing the gene expression of fast and slow dividing cells from MNPs and LNPs reveals that *Wnt5a* is highly expressed in the fast-dividing cells, which has been shown to orient the direction of cell division and outgrowth of facial structures (Kaucka et al., 2016). Our results show that NC-derived mesenchymal proliferation in the facial primordia is, at least partly, regulated by *Wnt5a*. This conclusion is supported by the fact that knocking out and overexpressing *Wnt5a* causes facial outgrowth deficiency (van Amerongen et al., 2012, Bakker et al., 2012, Ho et al., 2012), suggesting the gradient of *Wnt5a* needs to be precisely regulated within the facial primordia for developing a face with correct shape and size.

Our data highlight the similarity between cranial NC cells at early developmental stages at a single cell transcriptomic level, which could be the reason for the facial developmental robustness. Additionally, our data also reveal that NC cell differentiation is associated with the exiting of the cell cycle, and the regulation of cell proliferation might be a key step in the evolution of various craniofacial shapes and functions in vertebrates.

Materials and methods

Mouse strains and genetic fate mapping

All animal work was approved and permitted by the UC Davis Animal Care and Use Committee and conducted according to the NIH guidelines. NC cell-specific genetic tracing mouse lines, Wnt1-Cre and Sox10-Cre lines were previously described (Lewis et al., 2013, Chen et al., 2017, Hari et al., 2012). Both Wnt1-Cre and Sox10-Cre strains were crossed with the Rosa26-tdTomato/eGFP (Rosa26-mT/mG) reporter line to label NC cells with eGFP (Muzumdar et al., 2007). All the mouse strains were ordered from the Jackson Laboratory (stock numbers 022137, 025807, and 007676). Pregnant, timed-mated mice were euthanized with overdosed isoflurane (SAS, PIR001325-EA) prior to cesarean section. The day of conception was designated embryonic day 0.5 (E0.5). For genetic fate mapping, Wnt1-Cre;Rosa26-mT/mG and Sox10-Cre;Rosa26-mT/mG embryos were sampled at E10.5. The wholemount imaging was performed with Nikon A1 confocal laser microscope. Basic image processing and analysis were performed using NIS-Elements C software.

Wholemount *in situ* hybridization

E10.5 and E11.5 mouse embryo cDNA libraries were used to clone fragments of the coding sequence of mouse *Alx1*, *Alx3*, *Alx4*, and *Barx1*. The following primers were used: *Alx1*: forward primer 5′-GCGAGAAGTTTGCCCTGA-3′, reverse primer 5′-AAATGCGTGTCCGTTGGT-3′, *Alx3*: forward primer 5′-CTGTCTCATGTCTCCAGAGGG-3′, reverse primer 5′-TGTAGACTAGCACAGGGCAGAA-3′, *Alx4*: forward primer 5′-CCATCCTGGATTGGCAAC-3′, reverse primer 5′-GGGGGCCTGACTTTGACT-3′, *Barx1*: forward primer 5′-AGACAATTAAGGGCCAGACAAG-3′, reverse primer 5′-GTCCCCCACTGTGTCATAAAAT-3′.

The embryos were collected at E10.5 and E11.5 and fixed in 4% PFA overnight at 4 ˚C. As previously described, wholemount *in situ* hybridization was performed with digoxigenin-labeled antisense RNA probes (Volker et al., 2012). Briefly, embryos were digested with proteinase K (1:1000) for 6 min (for E10.5 embryos) or 20 min (for E11.5 embryos) and refixed in 4% PFA/0.25% glutaraldehyde for 20 min at room temperature. After 60 min of prehybridization at 67 ˚C, the embryos were hybridized with probes overnight at 67 ˚C. Embryos were incubated with 1:4,000 diluted alkaline phosphatase-conjugated anti-digoxigenin antibody (Roche, 11093274910) overnight at 4 ˚C. Alkaline phosphatase activity was detected using NBT/BCIP (Sigma, N6876 and 10760994001).

Single-cell RNA-sequencing analyses

Mouse embryos with Rosa26-mT/mG reporter were collected at E9.5, E10.5, E11.5, and E12.5 in ice-cold PBS and then dissociated using a cold-active protease (CAP) protocol (Adam et al., 2017). Briefly, embryos were incubated to 125 µl of cold protease solution [1.25 mg/ml *Bacillus Licheniformis* protease (Creative Enzymes, NATE0633) and 125 U/ml DNAseI (ThermoFisher, EN0521) in Dulbecco's phosphate-buffered saline (DPBS) with calcium and magnesium] in 4 ˚C

with trituration using a 1 ml pipet (10 s every 3 min). After 9, 12, 21, and 30 min incubation for E9.5, E10.5, E11.5, and E12.5 embryos, respectively, 1 ml ice-cold PBS with 15% fetal bovine serum (PBS/FBS) was added to the single-cell suspension. Cells were passed through a 35 μM cell strainer (Falcon, 352235). Cells were pelleted by 1200 g centrifuge for 5 min at 4 ˚C and re-suspended in 1 ml PBS with 1% FBS. The PBS/FBS wash was repeated one more time. 1000nM DAPI was added to the cell suspension to label dead cells. Fluorescent-activated cell sorting was performed to collect eGFP positive and DAPI negative cells. The cell concentration was adjusted to approximately 500 cells/ μl for 10x Genomics' single-cell RNA-seq.

The Cell Ranger Single Cell software (http://10xgenomics.com/) was used to align reads and generate feature-barcode matrices. Cell clusters and marker genes were identified using Seurat_ 3.2.0 (Butler et al., 2018, Stuart et al., 2019). Initial cell filtering selected cells that expressed >2000 reads and contained <10% mitochondrial genes. Normalization was performed by the "NormalizeData" function in Seurat. The "FindVariableFeatures" function was used to calculate a subset of highly variable features (10000 genes) for future analysis. We used the "CellCycleScoring" function in Seurat to score the cell cycle phase of every cell (Nestorowa et al., 2016). The cell that highly express G2/M- or S-phase markers were annotated as G2/M- or S-phase, respectively. Other cells were annotated as G1 phase cells. Clustering was performed with the "RunUMAP" function in Seurat using significant principal components determined by the JackStraw plot. For each cluster, Marker genes were determined with Seurat's "FindAllMarkers" function using genes detected in at least 25% of cells and a fold change threshold of 1.8. Sub-clustering of the mesenchymal cluster was performed as above. The developmental processes were calculated as Dorrity *et al.* descripted (Dorrity, 2020).

Single-cell trajectory reconstruction

The single-cell trajectories were reconstructed using Monocle_ 2.10.1 (Chen et al., 2019, Trapnell et al., 2014, Qiu et al., 2017). A nonlinear reconstruction algorithm, Discriminative Dimensionality Reduction with Trees (DDRTree), was used to reconstruct the single-cell trajectories with genes differentially expressed across four different time points. The state contains cells from E9.5 embryos were set as time zero, and other cells were ordered across the trajectory. Differently expressed genes across pseudotime were selected with a q value less than 0.01. Differential expression analysis between states at branch 1 was performed using the "Beam" function in Monocle. Differently expressed genes were clustered by pseudotime expression patterns to draw the heatmaps.

Data availability

The single cell RNA-seq datasets have been deposited in the Single-Cell Portal of the Broad Institute under accession number SCP1367.

Acknowledgments

We are grateful to Drs. Bruce Draper and Lesilee Rose (UC Davis BMCDB program) for their supports, Bridget Mclaughlin and Jonathan Van Dyke (UC Davis Flow Cytometry Shared Resource) for conducting fluorescence-activated cell sorting.

Additional information

Funding

Funder	Grant reference number	Author
NIH/NIDCR	R01DE026737	Chengji Zhou

NIH/NINDS	R01NS102261	Chengji Zhou
Shriners Hospitals for Children	85105	Chengji Zhou

The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

452

453 Author contributions

454 YJ, Conception and design, Acquisition of data, Analysis and interpretation of data, Drafting or
455 revising the article; SZ, KR, Acquisition of data, Analysis and interpretation of data, Drafting or
456 revising the article; RG, MM, MI, YL, TI, RD, HZ, DB, Acquisition of data; YX, Analysis and
457 interpretation of data, Drafting or revising the article; CZ, Conception and design, Acquisition of
458 data, Analysis and interpretation of data, Drafting or revising the article.

459

References

ADAM, M., POTTER, A. S. & POTTER, S. S. 2017. Psychrophilic proteases dramatically reduce single-cell RNA-seq artifacts: a molecular atlas of kidney development. *Development,* 144**,** 3625-3632.

ALFANDARI, D., COUSIN, H. & MARSDEN, M. 2010. Mechanism of Xenopus cranial neural crest cell migration. *Cell Adh Migr,* 4**,** 553-60.

BAGGIOLINI, A., VARUM, S., MATEOS, J. M., BETTOSINI, D., JOHN, N., BONALLI, M., ZIEGLER, U., DIMOU, L., CLEVERS, H., FURRER, R. & SOMMER, L. 2015. Premigratory and migratory neural crest cells are multipotent in vivo. *Cell Stem Cell,* 16**,** 314-22.

BAKER, C. V. & BRONNER-FRASER, M. 2001. Vertebrate cranial placodes I. Embryonic induction. *Dev Biol,* 232**,** 1-61.

BAKKER, E. R., RAGHOEBIR, L., FRANKEN, P. F., HELVENSTEIJN, W., VAN GURP, L., MEIJLINK, F., VAN DER VALK, M. A., ROTTIER, R. J., KUIPERS, E. J., VAN VEELEN, W. & SMITS, R. 2012. Induced Wnt5a expression perturbs embryonic outgrowth and intestinal elongation, but is well-tolerated in adult mice. *Dev Biol,* 369**,** 91-100.

BERNSTEIN, B. E., MIKKELSEN, T. S., XIE, X., KAMAL, M., HUEBERT, D. J., CUFF, J., FRY, B., MEISSNER, A., WERNIG, M., PLATH, K., JAENISCH, R., WAGSCHAL, A., FEIL, R., SCHREIBER, S. L. & LANDER, E. S. 2006. A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell,* 125**,** 315-26.

BURSTYN-COHEN, T. & KALCHEIM, C. 2002. Association between the cell cycle and neural crest delamination through specific regulation of G1/S transition. *Dev Cell,* 3**,** 383-95.

BUTLER, A., HOFFMAN, P., SMIBERT, P., PAPALEXI, E. & SATIJA, R. 2018. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol,* 36**,** 411-420.

CHAI, Y. & MAXSON, R. E., JR. 2006. Recent advances in craniofacial morphogenesis. *Dev Dyn,* 235**,** 2353-75.

CHEN, G., ISHAN, M., YANG, J., KISHIGAMI, S., FUKUDA, T., SCOTT, G., RAY, M. K., SUN, C., CHEN, S. Y., KOMATSU, Y., MISHINA, Y. & LIU, H. X. 2017. Specific and spatial labeling of P0-Cre versus Wnt1-Cre in cranial neural crest in early mouse embryos. *Genesis,* 55.

CHEN, G., NING, B. & SHI, T. 2019. Single-Cell RNA-Seq Technologies and Related Computational Data Analysis. *Front Genet,* 10**,** 317.

CORDERO, D. R., BRUGMANN, S., CHU, Y., BAJPAI, R., JAME, M. & HELMS, J. A. 2011. Cranial neural crest cells on the move: their roles in craniofacial development. *Am J Med Genet A,* 155A**,** 270-9.

CRANE, J. F. & TRAINOR, P. A. 2006. Neural crest stem and progenitor cells. *Annu Rev Cell Dev Biol,* 22**,** 267-86.

CREUZET, S., SCHULER, B., COULY, G. & LE DOUARIN, N. M. 2004. Reciprocal relationships between Fgf8 and neural crest cells in facial and forebrain development. *Proc Natl Acad Sci U S A,* 101**,** 4843-7.

DASH, S. & TRAINOR, P. A. 2020. The development, patterning and evolution of neural crest cell differentiation into cartilage and bone. *Bone,* 137**,** 115409.

DEBBACHE, J., PARFEJEVS, V. & SOMMER, L. 2018. Cre-driver lines used for genetic fate mapping of neural crest cells in the mouse: An overview. *Genesis,* 56**,** e23105.

DELILE, J., RAYON, T., MELCHIONDA, M., EDWARDS, A., BRISCOE, J. & SAGNER, A. 2019. Single cell transcriptomics reveals spatial and temporal dynamics of gene expression in the developing mouse spinal cord. *Development,* 146.

DORRITY, M. Temperature stress introduces variability in embryogenesis via cell type-specific effects on developmental rate SDB 79th Annual Meeting, 2020.

EVERSON, J. L., FINK, D. M., CHUNG, H. M., SUN, M. R. & LIPINSKI, R. J. 2018. Identification of sonic hedgehog-regulated genes and biological processes in the cranial neural crest mesenchyme by comparative transcriptomics. *BMC Genomics,* 19**,** 497.

FOPPIANO, S., HU, D. & MARCUCIO, R. S. 2007. Signaling by bone morphogenetic proteins directs formation of an ectodermal signaling center that regulates craniofacial development. *Dev Biol,* 312**,** 103-14.

GARDA, A. L., PUELLES, L., RUBENSTEIN, J. L. & MEDINA, L. 2002. Expression patterns of Wnt8b and Wnt7b in the chicken embryonic brain suggest a correlation with forebrain patterning centers and morphogenesis. *Neuroscience,* 113**,** 689-98.

GRAF, D., MALIK, Z., HAYANO, S. & MISHINA, Y. 2016. Common mechanisms in development and disease: BMP signaling in craniofacial development. *Cytokine Growth Factor Rev,* 27**,** 129-39.

GULATI, G. S., SIKANDAR, S. S., WESCHE, D. J., MANJUNATH, A., BHARADWAJ, A., BERGER, M. J., ILAGAN, F., KUO, A. H., HSIEH, R. W., CAI, S., ZABALA, M., SCHEEREN, F. A., LOBO, N. A., QIAN, D., YU, F. B., DIRBAS, F. M., CLARKE, M. F. & NEWMAN, A. M. 2020. Single-cell transcriptional diversity is a hallmark of developmental potential. *Science,* 367**,** 405-411.

HARDWICK, L. J. & PHILPOTT, A. 2014. Nervous decision-making: to divide or differentiate. *Trends Genet,* 30**,** 254-61.

HARI, L., MIESCHER, I., SHAKHOVA, O., SUTER, U., CHIN, L., TAKETO, M., RICHARDSON, W. D., KESSARIS, N. & SOMMER, L. 2012. Temporal control of neural crest lineage generation by Wnt/beta-catenin signaling. *Development,* 139**,** 2107-17.

HO, H. Y., SUSMAN, M. W., BIKOFF, J. B., RYU, Y. K., JONAS, A. M., HU, L., KURUVILLA, R. & GREENBERG, M. E. 2012. Wnt5a-Ror-Dishevelled signaling constitutes a core developmental pathway that controls tissue morphogenesis. *Proc Natl Acad Sci U S A,* 109**,** 4044-51.

HU, D. & MARCUCIO, R. S. 2009. A SHH-responsive signaling center in the forebrain regulates craniofacial morphogenesis via the facial ectoderm. *Development,* 136**,** 107-16.

IKEYA, M., LEE, S. M., JOHNSON, J. E., MCMAHON, A. P. & TAKADA, S. 1997. Wnt signalling required for expansion of neural crest and CNS progenitors. *Nature,* 389**,** 966-70.

JANKOVIC, V., CIARROCCHI, A., BOCCUNI, P., DEBLASIO, T., BENEZRA, R. & NIMER, S. D. 2007. Id1 restrains myeloid commitment, maintaining the self-renewal capacity of hematopoietic stem cells. *Proc Natl Acad Sci U S A,* 104**,** 1260-5.

JAVED, A., CHEN, H. & GHORI, F. Y. 2010. Genetic and transcriptional control of bone formation. *Oral Maxillofac Surg Clin North Am,* 22**,** 283-93, v.

JEAN-BAPTISTE, K., MCFALINE-FIGUEROA, J. L., ALEXANDRE, C. M., DORRITY, M. W., SAUNDERS, L., BUBB, K. L., TRAPNELL, C., FIELDS, S., QUEITSCH, C. & CUPERUS, J. T. 2019. Dynamics of Gene Expression in Single Root Cells of Arabidopsis thaliana. *Plant Cell,* 31**,** 993-1011.

JHEON, A. H. & SCHNEIDER, R. A. 2009. The cells that fill the bill: neural crest and the evolution of craniofacial development. *J Dent Res,* 88**,** 12-21.

JI, Y., GARLAND, M. A., SUN, B., ZHANG, S., REYNOLDS, K., MCMAHON, M., RAJAKUMAR, R., ISLAM, M. S., LIU, Y., CHEN, Y. & ZHOU, C. J. 2020. Cellular and developmental basis of orofacial clefts. *Birth Defects Res*.

JI, Y., HAO, H., REYNOLDS, K., MCMAHON, M. & ZHOU, C. J. 2019. Wnt Signaling in Neural Crest Ontogenesis and Oncogenesis. *Cells,* 8.

KASBERG, A. D., BRUNSKILL, E. W. & STEVEN POTTER, S. 2013. SP8 regulates signaling centers during craniofacial development. *Dev Biol,* 381**,** 312-23.

KAUCKA, M., IVASHKIN, E., GYLLBORG, D., ZIKMUND, T., TESAROVA, M., KAISER, J., XIE, M., PETERSEN, J., PACHNIS, V., NICOLIS, S. K., YU, T., SHARPE, P., ARENAS, E., BRISMAR, H., BLOM, H., CLEVERS, H., SUTER, U., CHAGIN, A. S., FRIED, K., HELLANDER, A. & ADAMEYKO, I. 2016. Analysis of neural crest-derived clones reveals novel aspects of facial development. *Sci Adv,* 2**,** e1600060.

KIM, H., LANGOHR, I. M., FAISAL, M., MCNULTY, M., THORN, C. & KIM, J. 2018. Ablation of Ezh2 in neural crest cells leads to aberrant enteric nervous system development in mice. *PLoS One,* 13**,** e0203391.

KIRBY, M. L. & HUTSON, M. R. 2010. Factors controlling cardiac neural crest cell migration. *Cell Adh Migr,* 4**,** 609-21.

KRISPIN, S., NITZAN, E., KASSEM, Y. & KALCHEIM, C. 2010. Evidence for a dynamic spatiotemporal fate map and early fate restrictions of premigratory avian neural crest. *Development,* 137**,** 585-95.

LA NOCE, M., MELE, L., TIRINO, V., PAINO, F., DE ROSA, A., NADDEO, P., PAPAGERAKIS, P., PAPACCIO, G. & DESIDERIO, V. 2014. Neural crest stem cell population in craniomaxillofacial development and tissue repair. *Eur Cell Mater,* 28**,** 348-57.

LANGENBACH, G. E. & VAN EIJDEN, T. M. 2001. Mammalian feeding motor patterns. *American Zoologist,* 41**,** 1338-1351.

LEWIS, A. E., VASUDEVAN, H. N., O'NEILL, A. K., SORIANO, P. & BUSH, J. O. 2013. The widely used Wnt1-Cre transgene causes developmental phenotypes by ectopic activation of Wnt signaling. *Dev Biol,* 379**,** 229-34.

LI, H., JONES, K. L., HOOPER, J. E. & WILLIAMS, T. 2019. The molecular anatomy of mammalian upper lip and primary palate fusion at single cell resolution. *Development,* 146.

LI, X., HUANG, M., ZHENG, H., WANG, Y., REN, F., SHANG, Y., ZHAI, Y., IRWIN, D. M., SHI, Y., CHEN, D. & CHANG, Z. 2008. CHIP promotes Runx2 degradation and negatively regulates osteoblast differentiation. *J Cell Biol,* 181**,** 959-72.

LOGAN, C., WINGATE, R. J., MCKAY, I. J. & LUMSDEN, A. 1998. Tlx-1 and Tlx-3 homeobox gene expression in cranial sensory ganglia and hindbrain of the chick embryo: markers of patterned connectivity. *J Neurosci,* 18**,** 5389-402.

LUMB, R., BUCKBERRY, S., SECKER, G., LAWRENCE, D. & SCHWARZ, Q. 2017. Transcriptome profiling reveals expression signatures of cranial neural crest cells arising from different axial levels. *BMC Dev Biol,* 17**,** 5.

LUMSDEN, A., SPRAWSON, N. & GRAHAM, A. 1991. Segmental origin and migration of neural crest cells in the hindbrain region of the chick embryo. *Development,* 113**,** 1281-91.

MIYAKOSHI, N., RICHMAN, C., KASUKAWA, Y., LINKHART, T. A., BAYLINK, D. J. & MOHAN, S. 2001. Evidence that IGF-binding protein-5 functions as a growth factor. *J Clin Invest,* 107**,** 73-81.

MUZUMDAR, M. D., TASIC, B., MIYAMICHI, K., LI, L. & LUO, L. 2007. A global double-fluorescent Cre reporter mouse. *Genesis,* 45**,** 593-605.

NAKASHIMA, K., TAKIZAWA, T., OCHIAI, W., YANAGISAWA, M., HISATSUNE, T., NAKAFUKU, M., MIYAZONO, K., KISHIMOTO, T., KAGEYAMA, R. & TAGA, T. 2001. BMP2-mediated alteration in the developmental pathway of fetal mouse brain cells from neurogenesis to astrocytogenesis. *Proc Natl Acad Sci U S A,* 98**,** 5868-73.

NAM, H. S. & BENEZRA, R. 2009. High levels of Id1 expression define B1 type adult neural stem cells. *Cell Stem Cell,* 5**,** 515-26.

NESTOROWA, S., HAMEY, F. K., PIJUAN SALA, B., DIAMANTI, E., SHEPHERD, M., LAURENTI, E., WILSON, N. K., KENT, D. G. & GOTTGENS, B. 2016. A single-cell resolution map of mouse hematopoietic stem and progenitor cell differentiation. *Blood,* 128**,** e20-31.

NITZAN, E., KRISPIN, S., PFALTZGRAFF, E. R., KLAR, A., LABOSKY, P. A. & KALCHEIM, C. 2013. A dynamic code of dorsal neural tube genes regulates the segregation between neurogenic and melanogenic neural crest cells. *Development,* 140**,** 2269-79.

PAUKLIN, S. & VALLIER, L. 2013. The cell-cycle state of stem cells determines cell fate propensity. *Cell,* 155**,** 135-47.

QIU, X., MAO, Q., TANG, Y., WANG, L., CHAWLA, R., PLINER, H. A. & TRAPNELL, C. 2017. Reversed graph embedding resolves complex single-cell trajectories. *Nat Methods,* 14**,** 979-982.

RIDENOUR, D. A., MCLENNAN, R., TEDDY, J. M., SEMERAD, C. L., HAUG, J. S. & KULESA, P. M. 2014. The neural crest cell cycle is related to phases of migration in the head. *Development,* 141**,** 1095-103.

RUTKOVSKIY, A., STENSLOKKEN, K. O. & VAAGE, I. J. 2016. Osteoblast Differentiation at a Glance. *Med Sci Monit Basic Res,* 22**,** 95-106.

SAKATA-GOTO, T., TAKAHASHI, K., KISO, H., HUANG, B., TSUKAMOTO, H., TAKEMOTO, M., HAYASHI, T., SUGAI, M., NAKAMURA, T., YOKOTA, Y., SHIMIZU, A., SLAVKIN, H. & BESSHO, K. 2012. Id2 controls chondrogenesis acting downstream of BMP signaling during maxillary morphogenesis. *Bone,* 50**,** 69-78.

SAUNDERS, L. M., MISHRA, A. K., AMAN, A. J., LEWIS, V. M., TOOMEY, M. B., PACKER, J. S., QIU, X., MCFALINE-FIGUEROA, J. L., CORBO, J. C., TRAPNELL, C. & PARICHY, D. M. 2019. Thyroid hormone regulates distinct paths to maturation in pigment cell lineages. *Elife,* 8.

SCHWARZ, D., VARUM, S., ZEMKE, M., SCHOLER, A., BAGGIOLINI, A., DRAGANOVA, K., KOSEKI, H., SCHUBELER, D. & SOMMER, L. 2014. Ezh2 is required for neural crest-derived cartilage and bone formation. *Development,* 141**,** 867-77.

SHIFLETT, L. A., TIEDE-LEWIS, L. M., XIE, Y., LU, Y., RAY, E. C. & DALLAS, S. L. 2019. Collagen Dynamics During the Process of Osteocyte Embedding and Mineralization. *Front Cell Dev Biol,* 7**,** 178.

SHYAMALA, K., YANDURI, S., GIRISH, H. C. & MURGOD, S. 2015. Neural crest: The fourth germ layer. *J Oral Maxillofac Pathol,* 19**,** 221-9.

SINGH, A. M., CHAPPELL, J., TROST, R., LIN, L., WANG, T., TANG, J., MATLOCK, B. K., WELLER, K. P., WU, H., ZHAO, S., JIN, P. & DALTON, S. 2013. Cell-cycle control of developmentally regulated transcription factors accounts for heterogeneity in human pluripotent cells. *Stem Cell Reports,* 1**,** 532-44.

SINGH, M. K., CHRISTOFFELS, V. M., DIAS, J. M., TROWE, M. O., PETRY, M., SCHUSTER-GOSSLER, K., BURGER, A., ERICSON, J. & KISPERT, A. 2005. Tbx20 is essential for cardiac chamber differentiation and repression of Tbx2. *Development,* 132**,** 2697-707.

SOLDATOV, R., KAUCKA, M., KASTRITI, M. E., PETERSEN, J., CHONTOROTZEA, T., ENGLMAIER, L., AKKURATOVA, N., YANG, Y., HARING, M., DYACHUK, V., BOCK, C., FARLIK, M., PIACENTINO, M. L., BOISMOREAU, F., HILSCHER, M. M., YOKOTA, C., QIAN, X., NILSSON, M., BRONNER, M. E., CROCI, L., HSIAO, W. Y., GUERTIN, D. A., BRUNET, J. F., CONSALEZ, G. G., ERNFORS, P., FRIED, K., KHARCHENKO, P. V. & ADAMEYKO, I. 2019. Spatiotemporal structure of cell fate decisions in murine neural crest. *Science,* 364.

SOUFI, A. & DALTON, S. 2016. Cycling through developmental decisions: how cell cycle dynamics control pluripotency, differentiation and reprogramming. *Development,* 143**,** 4301-4311.

STUART, T., BUTLER, A., HOFFMAN, P., HAFEMEISTER, C., PAPALEXI, E., MAUCK, W. M., 3RD, HAO, Y., STOECKIUS, M., SMIBERT, P. & SATIJA, R. 2019. Comprehensive Integration of Single-Cell Data. *Cell,* 177**,** 1888-1902 e21.

SZABO-ROGERS, H. L., GEETHA-LOGANATHAN, P., NIMMAGADDA, S., FU, K. K. & RICHMAN, J. M. 2008. FGF signals from the nasal pit are necessary for normal facial morphogenesis. *Dev Biol,* 318**,** 289-302.

TAVARES, A. L. & CLOUTHIER, D. E. 2015. Cre recombinase-regulated Endothelin1 transgenic mouse lines: novel tools for analysis of embryonic and adult disorders. *Dev Biol,* 400**,** 191-201.

TRAPNELL, C., CACCHIARELLI, D., GRIMSBY, J., POKHAREL, P., LI, S., MORSE, M., LENNON, N. J., LIVAK, K. J., MIKKELSEN, T. S. & RINN, J. L. 2014. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat Biotechnol,* 32**,** 381-386.

TUCKERMANN, J. P., PITTOIS, K., PARTRIDGE, N. C., MERREGAERT, J. & ANGEL, P. 2000. Collagenase-3 (MMP-13) and integral membrane protein 2a (Itm2a) are marker genes of chondrogenic/osteoblastic cells in bone formation: sequential temporal, and spatial expression

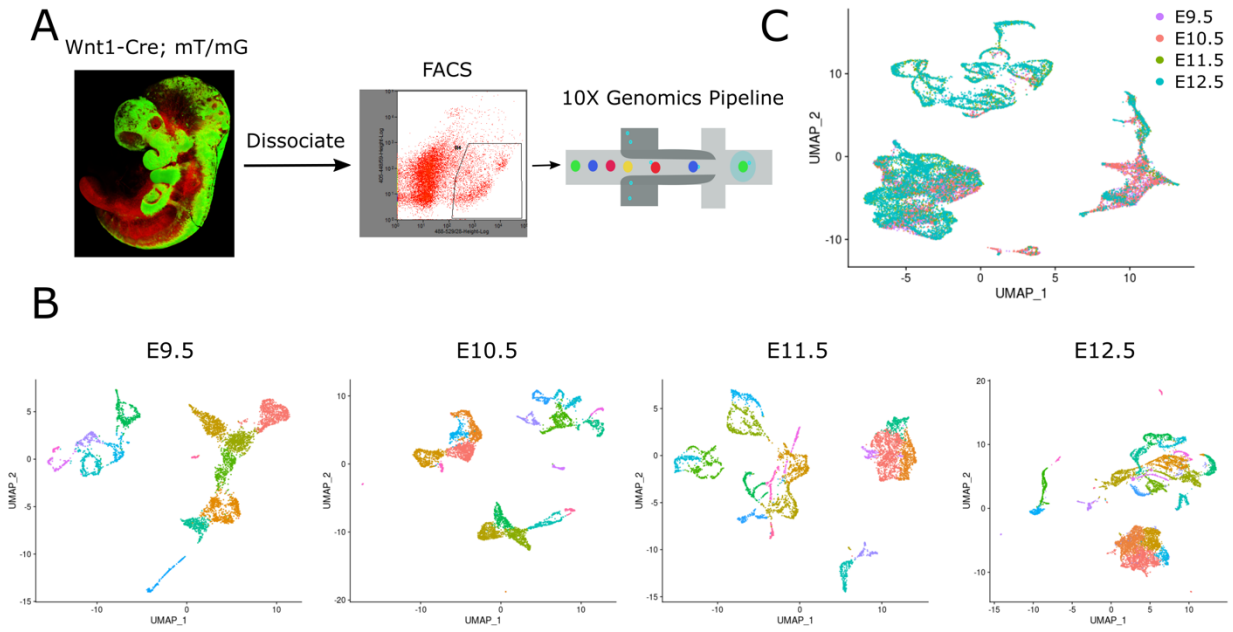
649 of Itm2a, alkaline phosphatase, MMP-13, and osteocalcin in the mouse. *J Bone Miner Res,* 15**,**
650 1257-65.
651 VAN AMERONGEN, R., FUERER, C., MIZUTANI, M. & NUSSE, R. 2012. Wnt5a can both activate and
652 repress Wnt/beta-catenin signaling during mouse embryonic development. *Dev Biol,* 369**,** 101-
653 14.
654 VOLKER, L. A., PETRY, M., ABDELSABOUR-KHALAF, M., SCHWEIZER, H., YUSUF, F., BUSCH, T., SCHERMER,
655 B., BENZING, T., BRAND-SABERI, B., KRETZ, O., HOHNE, M. & KISPERT, A. 2012. Comparative
656 analysis of Neph gene expression in mouse and chicken development. *Histochem Cell Biol,* 137**,**
657 355-66.
658 YING, Q. L., NICHOLS, J., CHAMBERS, I. & SMITH, A. 2003. BMP induction of Id proteins suppresses
659 differentiation and sustains embryonic stem cell self-renewal in collaboration with STAT3. *Cell,*
660 115**,** 281-92.
661 ZHAO, Y., KWAN, K. M., MAILLOUX, C. M., LEE, W. K., GRINBERG, A., WURST, W., BEHRINGER, R. R. &
662 WESTPHAL, H. 2007. LIM-homeodomain proteins Lhx1 and Lhx5, and their cofactor Ldb1,
663 control Purkinje cell differentiation in the developing cerebellum. *Proc Natl Acad Sci U S A,* 104**,**
664 13182-6.

665

666

667 **Figures and legends**

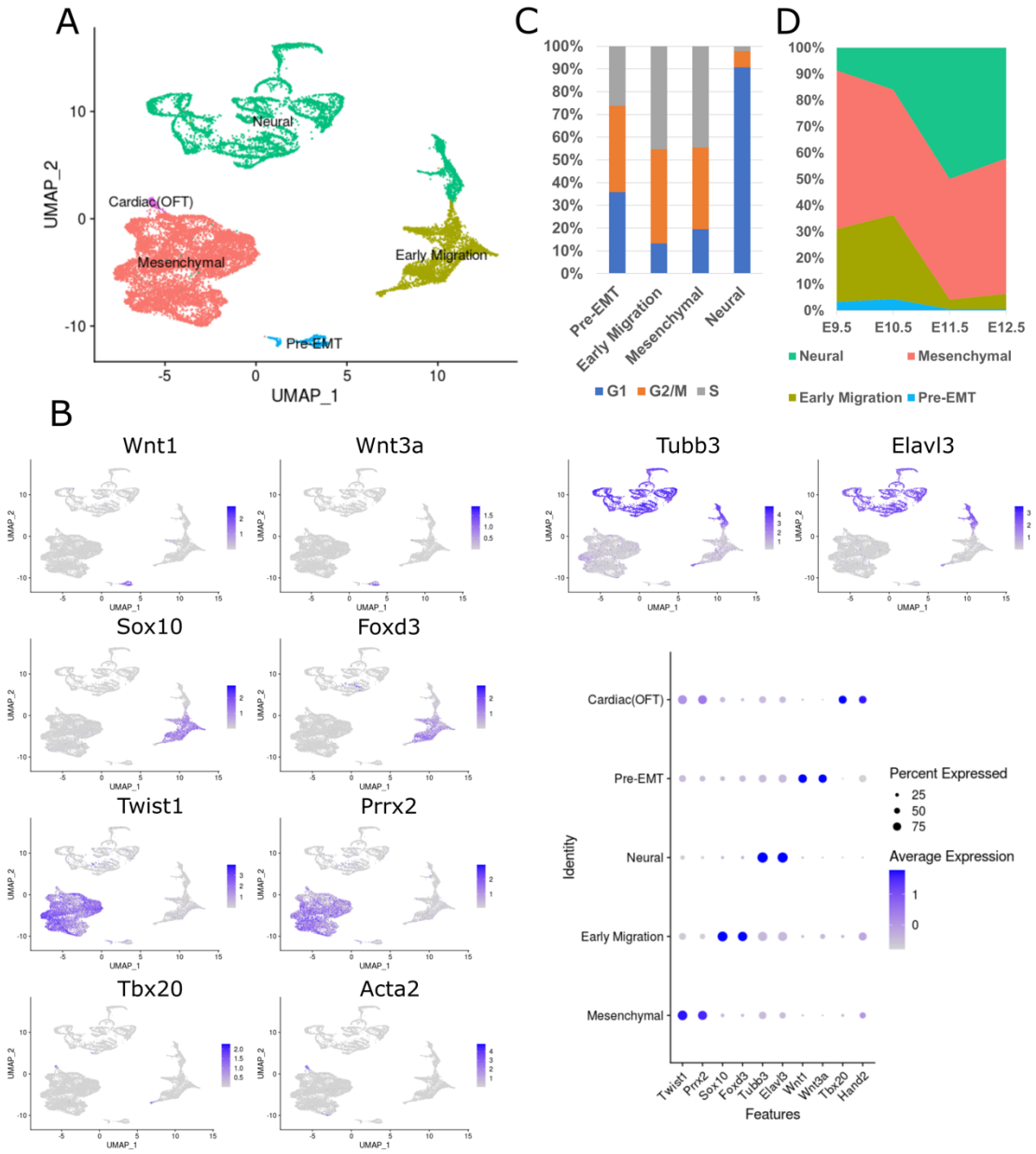
668 **Figure 1.** Single-cell transcriptional atlas of mouse NC cells. (A) Experimental workflow. NC
669 cells were genetically labeled by Wnt1-Cre;Rosa26-mT/mG reporter. EGFP positive cells from
670 E9.5, E10.5, E11.5, and E12.5 embryos were collected using fluorescent-activated cell sorting
671 and performed 10X single-cell RNA sequencing. (B) UMAP plot for each time point. Non-NCCs
672 were excluded from the following analysis based on expressed marker genes. (C) UMAP plot of
673 19,640 eGFP positive cells from four-time points. Cells are colored by time point.



674

675

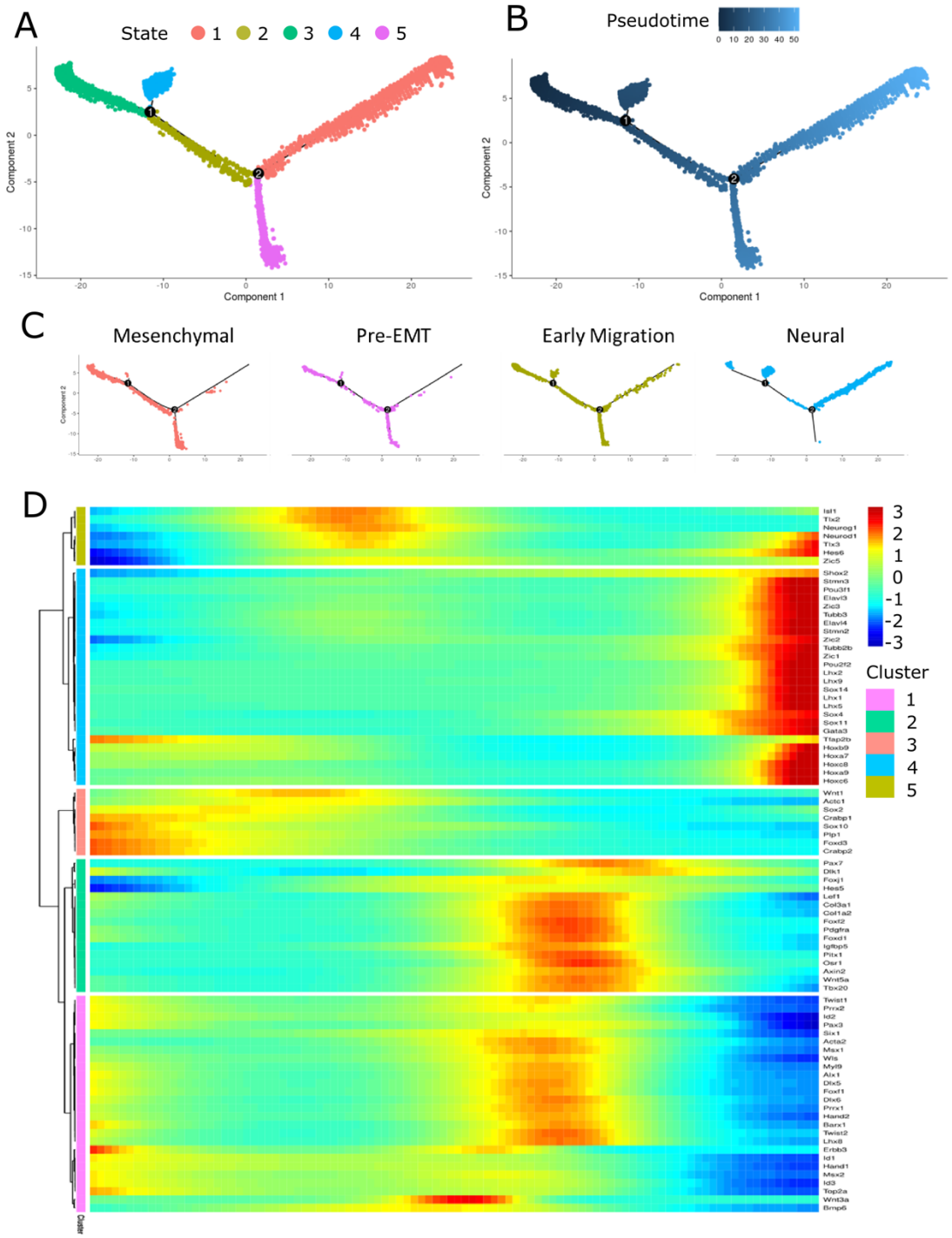
676 **Figure 2.** Single-cell RNA-seq identifies major NC-derived cell types. (A) UMAP plot of NC cells
677 from four-time points. Cells are colored by major NC-derived cell types inferred from expressed
678 marker genes. (B) Feature plots and dot plots of marker genes for each major cell population.
679 (C) Histogram showing the fraction of different cell types in G1 (Blue), G2M (Orange), or S
680 (Grey) phase. (D) The fraction of cell type at each time point, showing a decrease of pre-EMT
681 and early migration NC cells after E11.5.



682

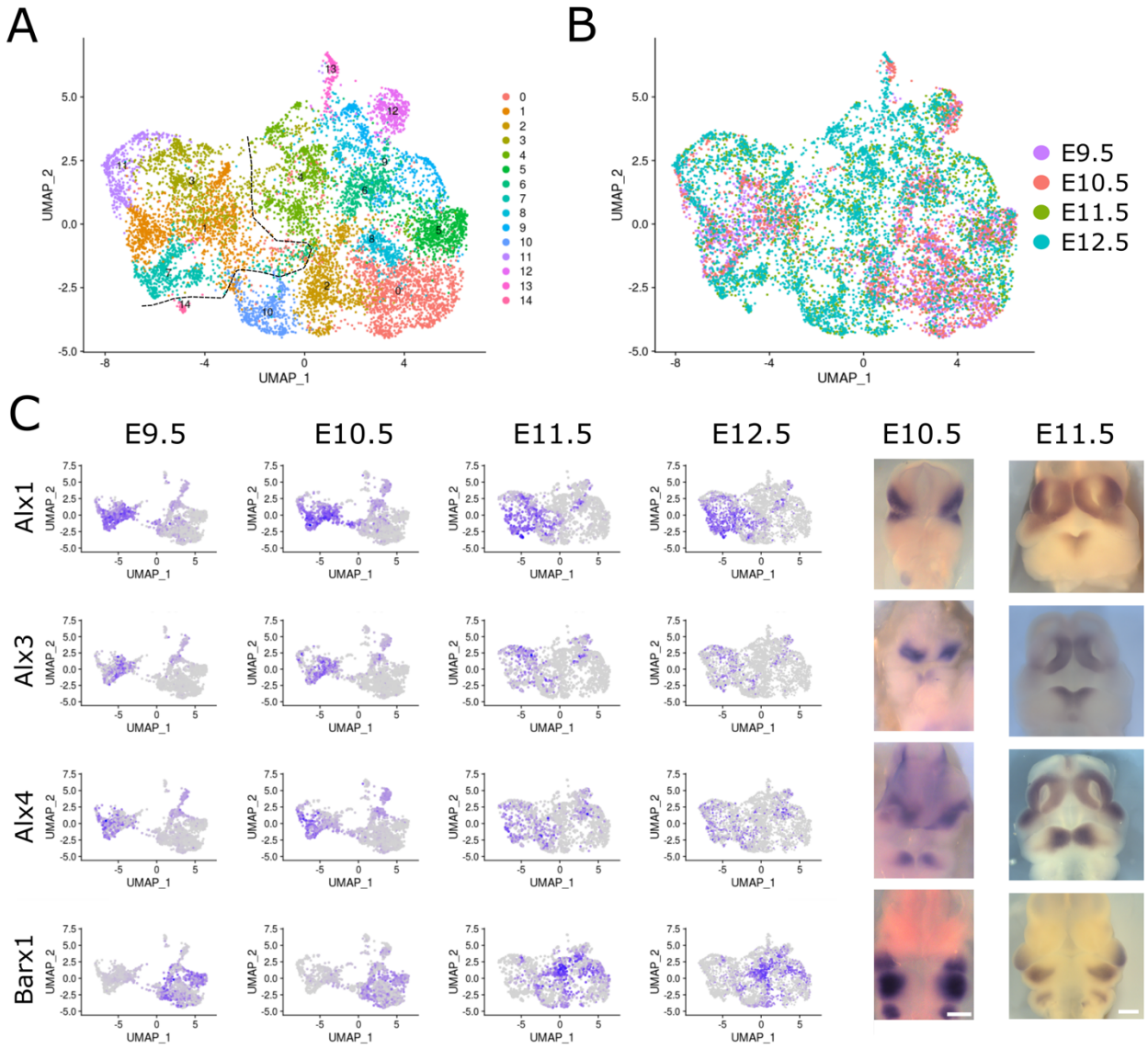
683

684 **Figure 3.** Developmental trajectories of NC cells. (A, B) Monocle pseudotime trajectory of NC
685 cells from four-time points. Cells are colored by state (A) or pseudotime (B). (C) Monocle
686 pseudotime trajectory showing the progression of pre-EMT, early migrating NC cells,
687 mesenchymal, and neural lineages derived from NC cells. (D) Expression heatmap showing
688 gene markers that link NC clusters to developmental states.



689

690 **Figure 4.** Single-cell RNA-seq identifies major NC-derived mesenchymal cells at facial
691 primordia. (A, B) UMAP plot of 9,875 NC-derived mesenchymal cells showing 15 different cell
692 populations. Cells are colored by mesenchymal clusters (A) or time point (B). (C) Identifying the
693 mesenchymal clusters using wholemount *in situ* hybridization. Left panels, feature plots for
694 marker genes. Right panels, wholemount *in situ* hybridization results for indicated genes at
695 E10.5 and E11.5. Scale bar: 0.5mm



696

697

Figure 5. The differentiation of NC-derived craniofacial mesenchymal cells starts at late developmental stages. (A-C) Monocle pseudotime trajectory of NC-derived craniofacial mesenchymal cells from four-time points. Cells are colored by state (A), pseudotime (B), or time point (C). (D) Global analysis of gene expression along the trajectory identified over 10,000 genes exhibits temporal expression patterns. (E) Pseudotime kinetics show the expression of known gene markers of the cell cycle and differentiation process. (F, G) Feature plot (F) and histograms (G) showing the developmental progression increased after E11.5 using a developmental progression score for each cell. Wilcoxon rank sum test was used to compare samples from each time point with E9.5 sample. *** = $p< e^{-14}$

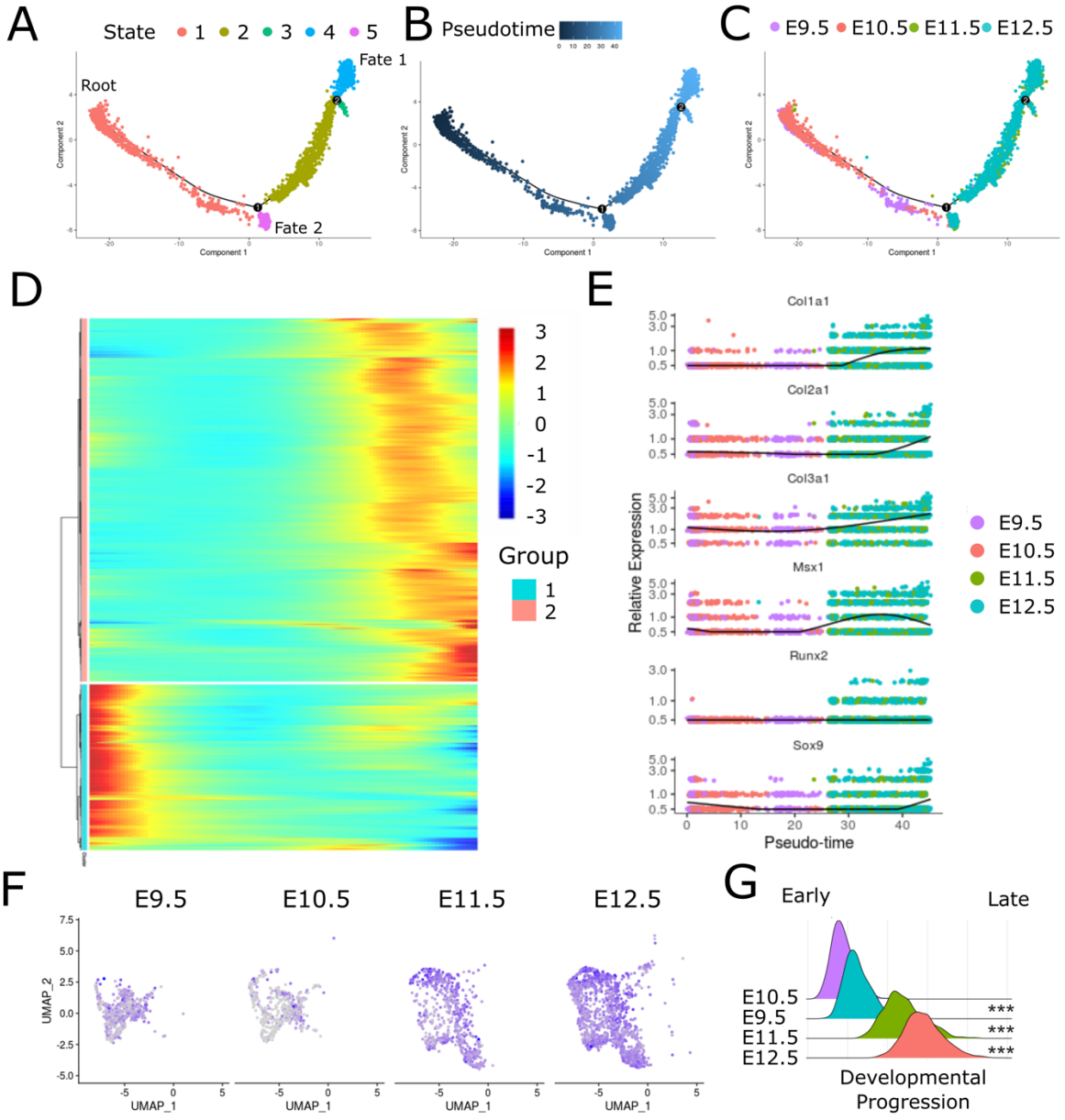


Figure 6. Branch 1 is a decision-making point governing whether a cell exits the cell cycle (A) Expression heatmap showing genes expressed in a branch 1 -dependent manner. Pre-branch refers to the cells before the branch. Gene ontology (GO) enrichment analysis for genes from each cell fate is shown (Right panel). (B) Pseudotime kinetics show the expression of known gene markers of the cell cycle and differentiation process from the root of the trajectory to fate 1 (solid line) or fate 2 (dotted line). (C) Pseudotime trajectory of craniofacial mesenchymal cells from four-time points. Cells are colored by the cell cycle phase. (D) Histogram showing the fraction of cells in G1 (Blue), G2M (Orange), or S (Grey) phase at each time point.

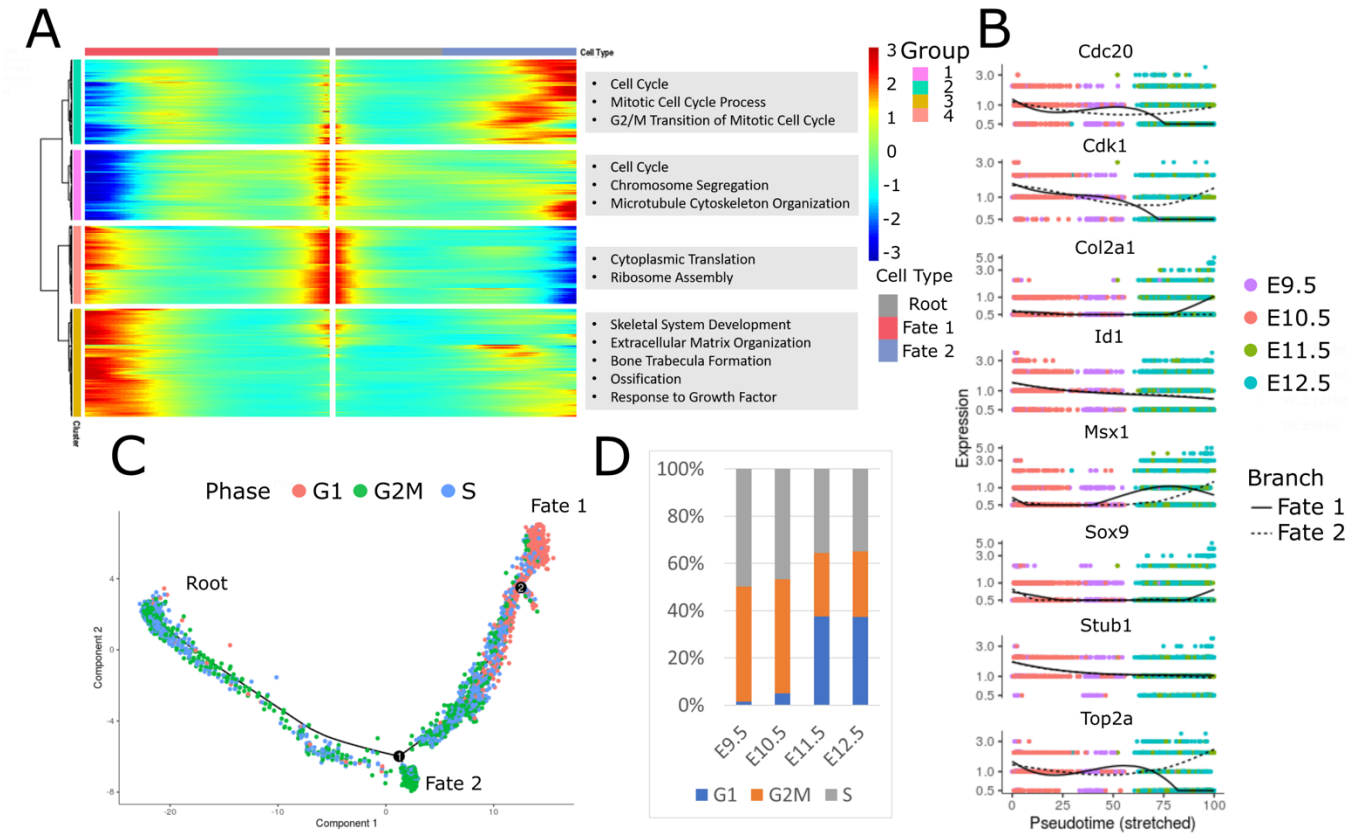


Figure 7. High reproducibility between Wnt1-Cre and Sox10-Cre scRNA-seq dataset at E11.5. (A) Confocal imaging shows Sox10-Cre marks migratory NC cells and their progeny. (B) UMAP plot shows the transcriptional dynamics are similar between Wnt1-Cre and Sox10-Cre scRNA-seq dataset. (C) Distributions of the number of UMIs and genes detected in the two datasets is similar. (D) Scatter plot shows the average expression of each gene in each cell is similar between the two datasets. (E) Pie charts show that the percentages of each cell type are similar. (F) Feature plots of marker genes show the expression patterns of these genes are similar between the two datasets.

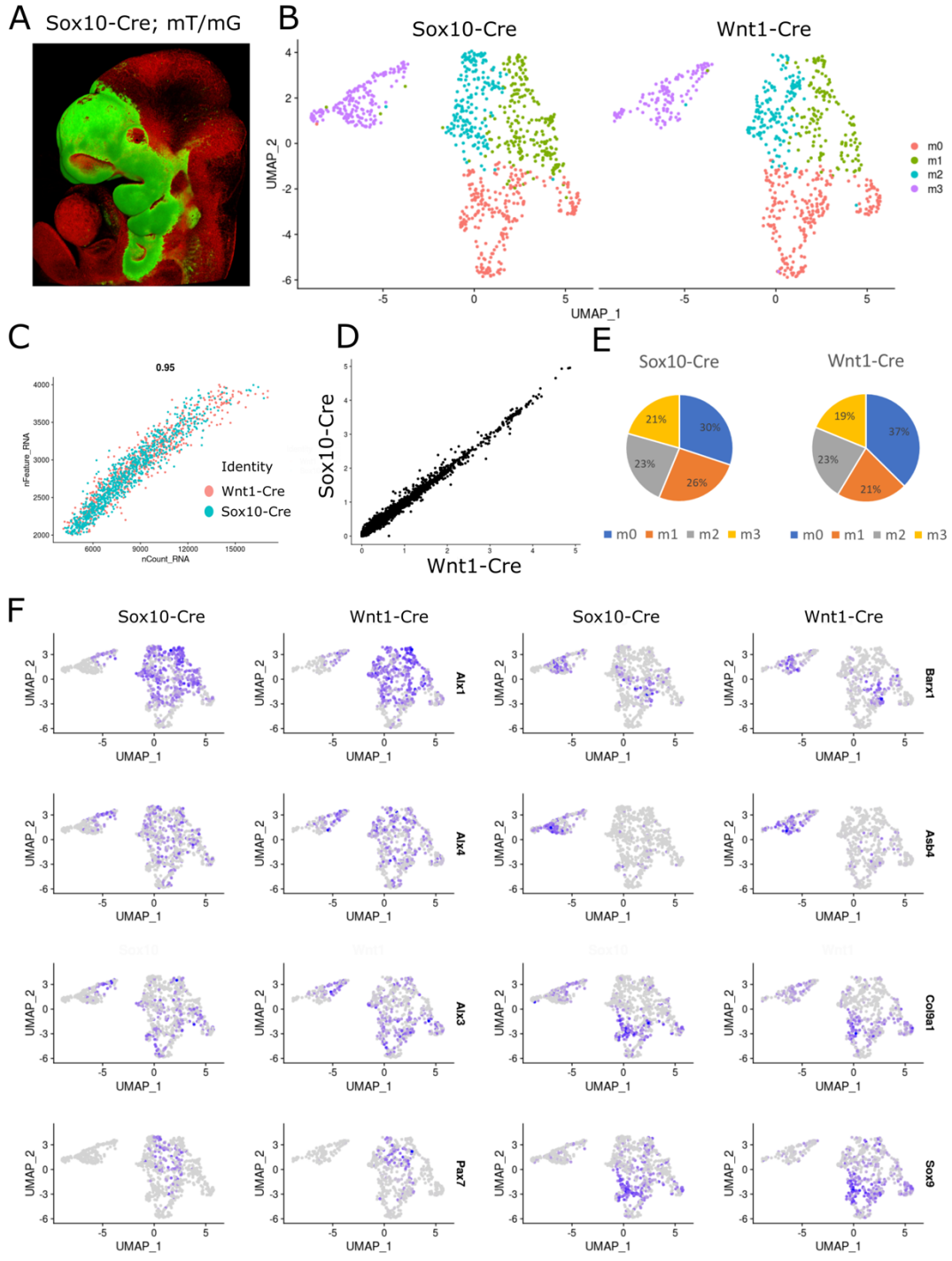
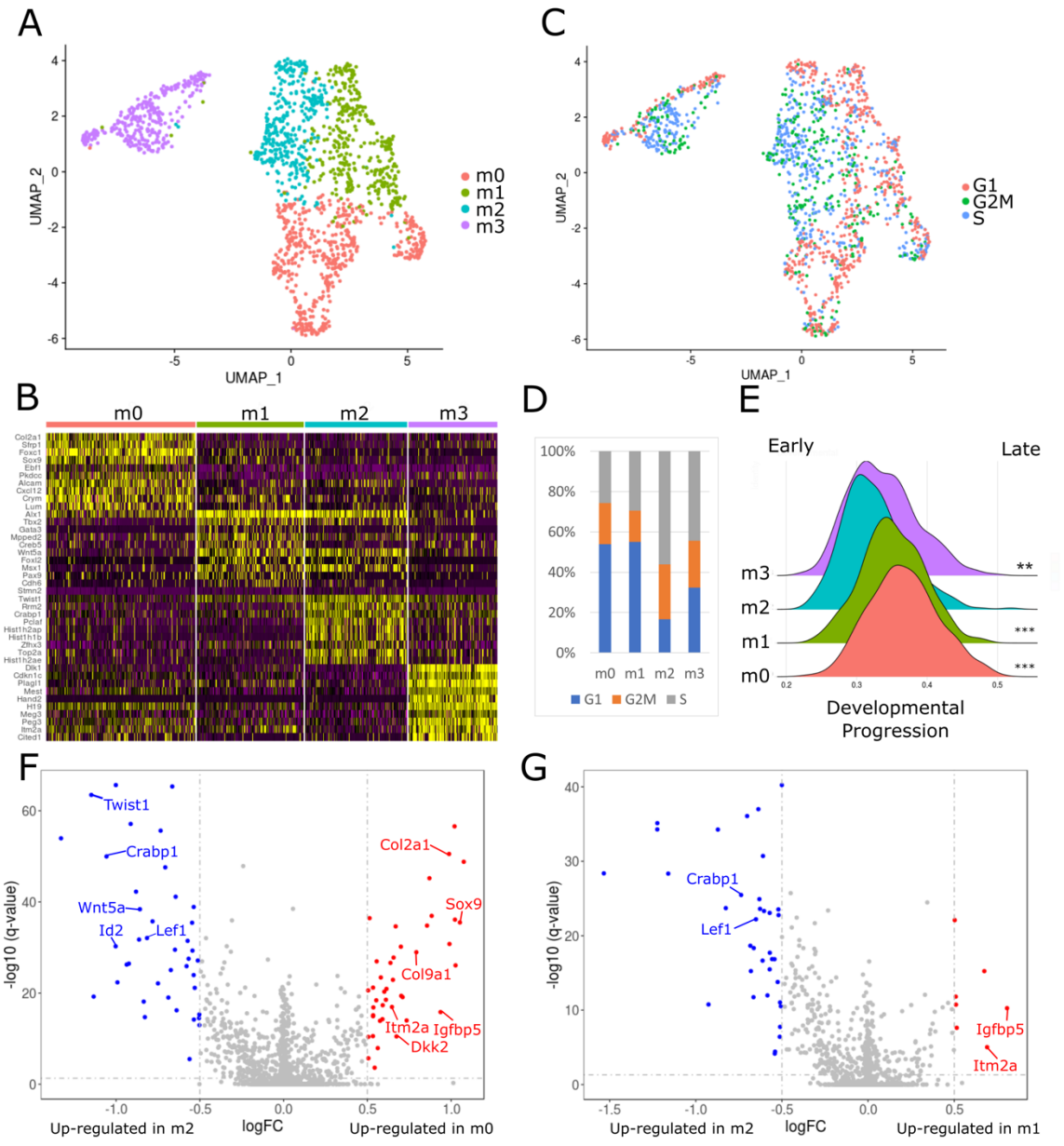
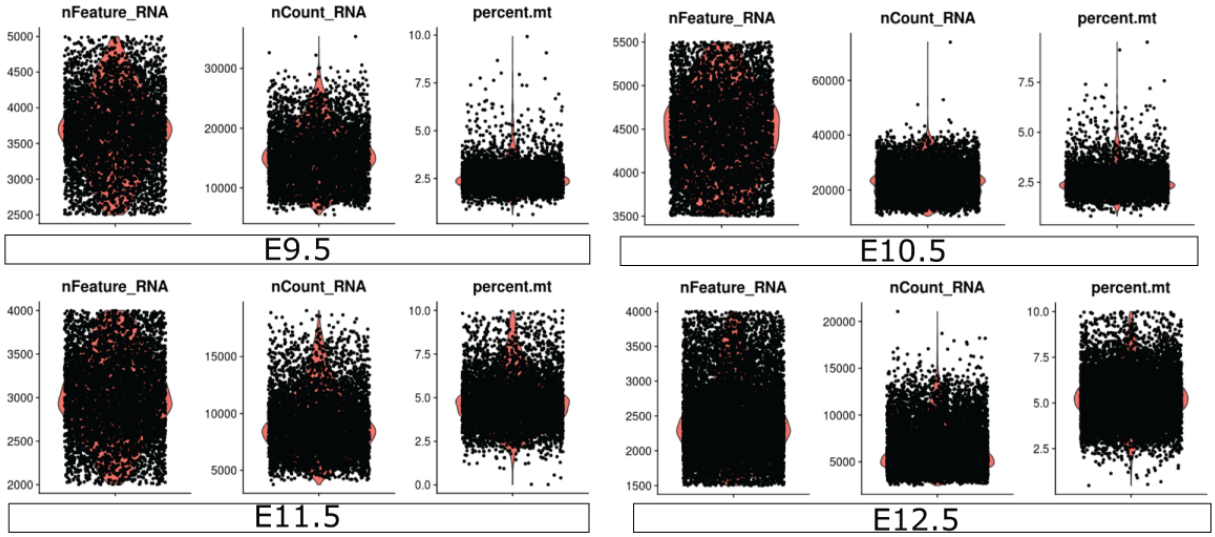


Figure 8. The development of craniofacial mesenchymal cells is associated with cell cycle progression at E11.5. (A) UMAP plot of NC-derived craniofacial mesenchymal cells from E11.5 embryos. (B) Gene expression heatmap of top 10 marker genes for each cluster. (C) Segregation of craniofacial mesenchymal cells by cell cycle phase. (D) Quantification of the proportion of different cell types in G1 (Blue), G2M (Orange), or S (Grey) phase. (E) Histograms showing the developmental progression score for each cell type. Wilcoxon rank sum test was used to compare cells from each cluster with m2. *** = p< e^{-14}, ** = p< e^{-3} (F, G) Volcano plots showing differentially expressed genes between m2 and m0 (F) or m2 and m1 (G). Highly expressed genes (avg_logFC > 0.5, P adj < 0.05) are colored in blue (in m2) or red (in m0 or m1) dots, respectively.



740 **Figure S1.** Violin plots for quality control metrics of single cell RNA-seq, including the number of
741 unique molecular identifiers (UMIs) (Left), number of genes (Middle), and percentage of
742 mitochondrial genes (Right) at each age of E9.5~E12.5.

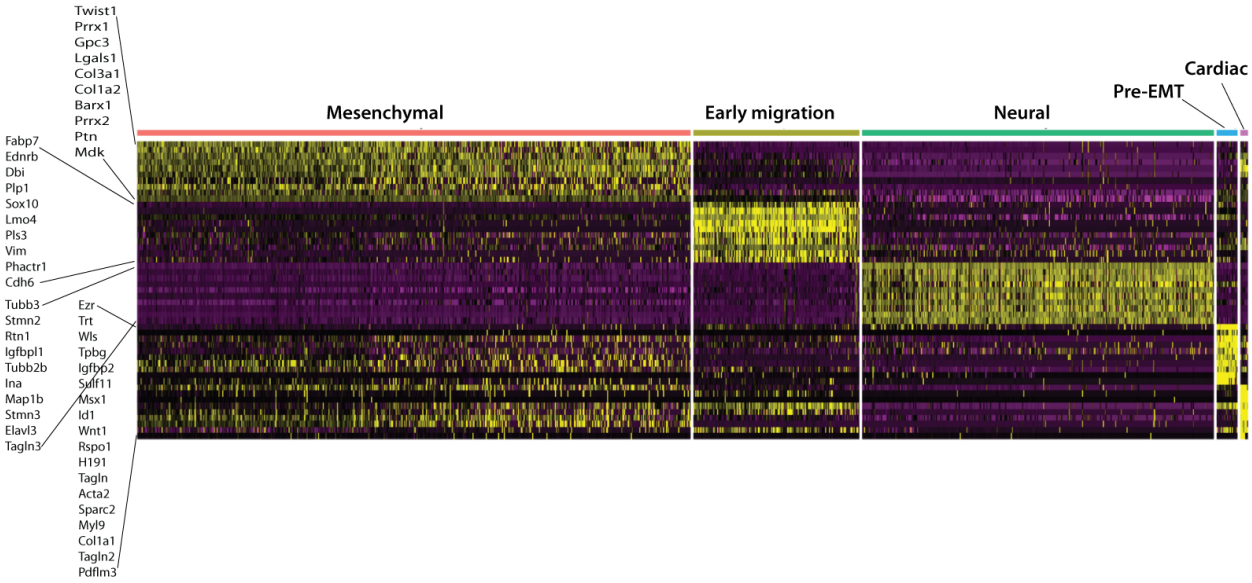


743

744

745 **Figure S2.** Heatmap showing the expression of the top 10 marker genes for each cell type. The

746 identity of each cluster is labeled on the top of each column.



747

748

Figure S3. Heatmap of genes used to define the cell cycle phase for each cell.

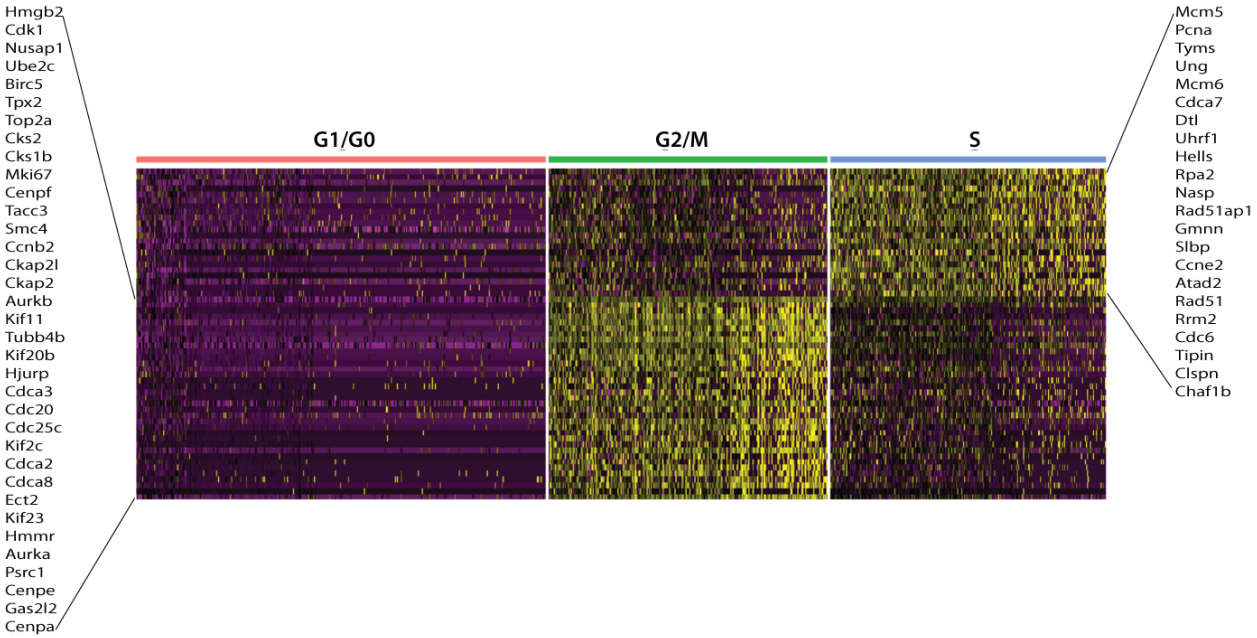


Table S1. Top marker genes for major NC-derived cell types.

Table S2. Top marker genes for NC-derived mesenchymal clusters.

"p_val" and "p_val_adj" represent the probability for a gene to mark the cluster. avg_logFC, log2 of the ratio of the average expression level of a gene for all cells within the cluster compare to all the other cells. "pct.1" and "pct.2" are the percent of cells in or out the cluster that express the gene.

Table S3. Top gene lists in groups 1 and 2 that exhibit temporal expression patterns, related to figure 5.

Table S4. Top gene lists in groups 1~4 from BEAM Analysis of Branch Point 1, related to figure 6.

Table S5. Top marker genes for E11.5 facial mesenchymal clusters m0~m3.

Table S6. Differentially expressed genes between m0, m2.

Table S7. Differentially expressed genes between m1, m2.

Highly expressed genes (avg_logFC > 0.2, P adj < 0.05) in m0, m1, or m2 are shown. The percentage of cells that have detected level of gene expression in each cluster (pct.1 and pct.2) are shown.

Table S1	p_val	avg_logFC	pct.1	pct.2	p_val_adj	cluster
Twist1	0	1.837950537	0.923	0.433	0	Mesenchymal
Prrx1	0	1.718982678	0.895	0.452	0	Mesenchymal
Gpc3	0	1.443583933	0.945	0.456	0	Mesenchymal
Lgals1	0	1.336392147	0.986	0.599	0	Mesenchymal
Col3a1	0	1.297483355	0.936	0.521	0	Mesenchymal
Col1a2	0	1.196412671	0.962	0.531	0	Mesenchymal
Barx1	0	1.153533664	0.662	0.323	0	Mesenchymal
Prrx2	0	1.142081491	0.852	0.29	0	Mesenchymal
Ptn	0	1.123755488	0.954	0.651	0	Mesenchymal
Mdk	0	1.087947385	1	0.825	0	Mesenchymal
H19	0	1.036826715	0.895	0.533	0	Mesenchymal
Tpm1	0	1.035623992	0.934	0.643	0	Mesenchymal
Igfbp4	0	1.006166922	0.956	0.632	0	Mesenchymal
Msx1	0	0.973900235	0.741	0.393	0	Mesenchymal
Lpar1	0	0.94218044	0.871	0.443	0	Mesenchymal
Cdh11	0	0.941894922	0.886	0.516	0	Mesenchymal
Gas1	0	0.931569071	0.886	0.525	0	Mesenchymal
Peg3	0	0.873679186	0.87	0.645	0	Mesenchymal
Alx1	0	0.869351114	0.651	0.292	0	Mesenchymal
Fabp7	0	1.786473609	0.962	0.357	0	Early Migration
Ednrb	0	1.529837076	0.962	0.346	0	Early Migration
Dbi	0	1.309816675	0.998	0.897	0	Early Migration
Plp1	0	1.309460685	0.95	0.337	0	Early Migration
Sox10	0	1.067943647	0.949	0.24	0	Early Migration
Lmo4	0	1.051364362	0.969	0.745	0	Early Migration
Pls3	0	1.045389778	0.914	0.383	0	Early Migration
Vim	0	1.042342624	0.998	0.885	0	Early Migration
Phactr1	0	1.007878875	0.903	0.409	0	Early Migration
Cdh6	0	0.96435694	0.871	0.39	0	Early Migration
Postn	0	0.941535298	0.945	0.478	0	Early Migration
Erbb3	0	0.897573765	0.913	0.217	0	Early Migration
Foxd3	0	0.89232875	0.893	0.213	0	Early Migration
Sparc	0	0.882325379	0.976	0.686	0	Early Migration
Rhob	0	0.848181188	0.941	0.537	0	Early Migration
Rgcc	0	0.847376283	0.796	0.398	0	Early Migration
Metrn	0	0.843090468	0.954	0.459	0	Early Migration
Dnajc1	0	0.820848538	0.959	0.497	0	Early Migration
Tubb3	0	3.196231028	0.997	0.722	0	Neural
Stmn2	0	2.744425076	0.941	0.652	0	Neural
Rtn1	0	2.661743025	0.97	0.556	0	Neural
Igfbpl1	0	2.456986185	0.858	0.548	0	Neural
Tubb2b	0	2.280173965	1	0.912	0	Neural

Ina	0	2.223342976	0.838	0.529	0	Neural
Map1b	0	2.223260464	0.996	0.823	0	Neural
Stmn3	0	2.136427725	0.913	0.567	0	Neural
Elavl3	0	2.117870037	0.983	0.618	0	Neural
Tagln3	0	2.025686207	0.976	0.547	0	Neural
Tuba1a	0	1.887266269	1	0.993	0	Neural
Ckb	0	1.84327919	0.982	0.722	0	Neural
Nefm	0	1.838292157	0.763	0.605	0	Neural
Ebf1	0	1.813158281	0.904	0.679	0	Neural
Elavl4	0	1.79265133	0.941	0.504	0	Neural
Mllt11	0	1.757808348	0.948	0.554	0	Neural
Crmp1	0	1.749022743	0.943	0.54	0	Neural
Nhlh2	0	1.721136315	0.813	0.501	0	Neural
Nefl	0	1.704633963	0.728	0.609	0	Neural
Dcx	0	1.701401224	0.935	0.622	0	Neural
Clic6	0	0.651215592	0.967	0.067	0	Pre-EMT
Pifo	0	0.553802765	0.967	0.058	0	Pre-EMT
C1qtnf3	0	0.540598443	0.907	0.112	0	Pre-EMT
Deup1	0	0.520932528	0.929	0.063	0	Pre-EMT
1500015O10Rik	0	0.508643191	0.956	0.175	0	Pre-EMT
Gmnc	0	0.428575115	0.888	0.04	0	Pre-EMT
Crb3	0	0.343304214	0.907	0.085	0	Pre-EMT
1700012B09Rik	0	0.319688479	0.891	0.11	0	Pre-EMT
Ccdc113	0	0.272468863	0.951	0.105	0	Pre-EMT
Cfap206	0	0.262485757	0.945	0.099	0	Pre-EMT
Dnajb13	4.66E-302	0.270954702	0.921	0.247	9.31E-299	Pre-EMT
1700016K19Rik	2.20E-263	0.275156039	0.926	0.145	4.40E-260	Pre-EMT
Rbm47	8.01E-254	0.480341361	0.921	0.153	1.60E-250	Pre-EMT
Ccno	2.13E-246	0.489108294	0.915	0.234	4.26E-243	Pre-EMT
Foxj1	9.16E-242	0.858098323	0.956	0.15	1.83E-238	Pre-EMT
Sfn	8.79E-241	0.620060877	0.921	0.177	1.76E-237	Pre-EMT
Ezr	1.03E-238	2.273323909	0.995	0.462	2.06E-235	Pre-EMT
AC154683.1	7.23E-233	0.876625528	0.945	0.23	1.45E-229	Pre-EMT
Pthlh	2.31E-224	0.342670286	0.948	0.379	4.61E-221	Pre-EMT
Gata4	1.24E-131	0.395465864	0.648	0.064	2.47E-128	Cardiac(OFT)

Gata6	1.16E-92	1.082319789	0.841	0.159	2.32E-89	Cardiac(OFT)
Tbx20	1.33E-83	0.815028396	0.731	0.128	2.65E-80	Cardiac(OFT)
H191	3.17E-81	1.688768303	1	0.713	6.34E-78	Cardiac(OFT)
Tagln	1.30E-80	2.643572077	0.903	0.369	2.61E-77	Cardiac(OFT)
Acta2	2.52E-79	2.687111999	0.869	0.344	5.03E-76	Cardiac(OFT)
Sparc2	2.92E-76	1.420431613	1	0.728	5.84E-73	Cardiac(OFT)
S100a111	2.81E-75	1.381788161	0.972	0.601	5.62E-72	Cardiac(OFT)
Myl12a1	9.91E-75	0.987316952	1	0.817	1.98E-71	Cardiac(OFT)
Myl91	5.22E-73	1.754909249	0.938	0.583	1.04E-69	Cardiac(OFT)
Tpm11	1.40E-72	1.585025889	0.986	0.788	2.79E-69	Cardiac(OFT)
Fstl12	1.76E-71	1.232137331	0.993	0.713	3.51E-68	Cardiac(OFT)
Col1a11	2.19E-71	1.419467781	0.966	0.638	4.37E-68	Cardiac(OFT)
Anxa51	3.38E-70	1.361708054	0.959	0.602	6.76E-67	Cardiac(OFT)
Rhoc3	1.01E-69	1.11376957	0.972	0.602	2.02E-66	Cardiac(OFT)
Fn11	1.00E-68	1.189440549	0.979	0.662	2.01E-65	Cardiac(OFT)
Tpm21	2.16E-68	1.262410417	0.945	0.561	4.32E-65	Cardiac(OFT)
Lgals11	7.34E-64	1.331547992	0.993	0.792	1.47E-60	Cardiac(OFT)
Col3a11	1.74E-63	1.404975276	0.979	0.728	3.48E-60	Cardiac(OFT)

768

Table S2	p_val	avg_logFC	pct.1	pct.2	p_val_adj	cluster
Pou3f3	4.23E-303	0.467173	0.864	0.562	8.46E-300	0
Pantr1	1.05E-298	0.592121	0.898	0.602	2.10E-295	0
Pcp4	1.35E-298	0.263894	0.725	0.29	2.69E-295	0
Meis1	1.17E-182	0.357107	0.866	0.559	2.35E-179	0
Meis2	1.38E-166	0.457654	0.949	0.701	2.77E-163	0
Foxl2	9.55E-266	0.620202	0.769	0.511	1.91E-262	1
Alx1	7.92E-257	0.863108	0.885	0.615	1.58E-253	1
Igf1	1.58E-208	0.595412	0.867	0.679	3.16E-205	1
Irx3	2.05E-189	0.581111	0.848	0.614	4.11E-186	1
Irx5	4.74E-134	0.483192	0.822	0.633	9.47E-131	1
Lhx8	0	1.087502	0.965	0.549	0	2
Lhx6	5.24E-284	0.492885	0.85	0.519	1.05E-280	2
Pcdh10	8.80E-228	0.613855	0.863	0.579	1.76E-224	2
Prrx1	3.99E-226	0.687277	1	0.883	7.98E-223	2
Tfap2c	3.83E-222	0.314842	0.793	0.456	7.65E-219	2
Sox9	1.59E-195	0.879273	0.902	0.633	3.19E-192	3
Foxc1	3.29E-157	0.607397	0.769	0.418	6.57E-154	3
Crym1	1.93E-141	0.85107	0.748	0.469	3.86E-138	3
Col2a1	5.42E-140	0.704274	0.872	0.684	1.08E-136	3
Six2	4.63E-128	0.633662	0.87	0.721	9.26E-125	3
Sfrp2	6.00E-142	1.477298	0.802	0.734	1.20E-138	4
Tsc22d1	6.40E-87	0.495903	0.949	0.949	1.28E-83	4
Col2a11	1.23E-86	1.02935	0.761	0.695	2.46E-83	4
Barx11	4.65E-67	1.342572	0.686	0.659	9.30E-64	4
Mecom	9.73E-61	0.656984	0.642	0.581	1.95E-57	4
Dlx5	4.89E-264	0.821311	0.949	0.614	9.77E-261	5
Dlx6os11	8.95E-203	0.473625	0.846	0.51	1.79E-199	5
Dlx6	8.45E-178	0.556757	0.862	0.529	1.69E-174	5
Tac1	2.10E-148	0.381195	0.821	0.532	4.20E-145	5
Ccnd1	5.29E-148	0.511618	0.988	0.814	1.06E-144	5
Hand2	1.12E-190	0.829173	0.85	0.51	2.23E-187	6
Pitx11	2.63E-129	0.775259	0.833	0.587	5.27E-126	6
5033428I22Rik	1.81E-110	0.344276	0.711	0.47	3.61E-107	6
Satb2	9.24E-106	0.266098	0.728	0.483	1.85E-102	6
Dlk1	1.25E-87	0.783428	0.862	0.69	2.50E-84	6
Wnt5a	1.73E-212	0.909328	0.951	0.722	3.47E-209	7
Alx12	1.80E-195	0.991865	0.952	0.627	3.59E-192	7
Crabp1	6.15E-172	0.992566	0.984	0.869	1.23E-168	7
Sox11	1.59E-141	0.565017	1	0.979	3.17E-138	7
Pax7	1.76E-136	0.604201	0.746	0.485	3.53E-133	7
Meox2	4.60E-108	0.472637	0.747	0.427	9.20E-105	8
Igfbp3	2.59E-84	0.761039	0.717	0.438	5.18E-81	8

Efemp1	4.46E-84	0.433892	0.77	0.552	8.91E-81	8
Igfbp5	1.67E-77	0.768754	0.858	0.677	3.34E-74	8
Tcf15	1.10E-76	0.264764	0.671	0.369	2.21E-73	8
Meg32	2.88E-147	1.40046	0.986	0.927	5.75E-144	9
H193	1.73E-122	0.878608	0.974	0.893	3.47E-119	9
Dlk12	1.09E-108	1.343368	0.874	0.693	2.18E-105	9
Cdkn1c1	2.83E-90	1.054203	0.996	0.982	5.66E-87	9
Lpar1	5.67E-66	0.518055	0.907	0.866	1.13E-62	9
Lef11	1.33E-224	1.567127	0.953	0.679	2.65E-221	10
Twist1	1.72E-181	0.897208	1	0.918	3.43E-178	10
Cxcl141	1.78E-177	1.949817	0.9	0.641	3.55E-174	10
Trps11	3.55E-144	1.288034	0.878	0.704	7.11E-141	10
Zeb21	5.57E-130	1.07641	0.866	0.685	1.11E-126	10
Akap12	4.30E-192	1.604027	0.917	0.512	8.60E-189	11
Alcam	3.15E-190	1.840846	0.926	0.67	6.31E-187	11
Col1a11	3.33E-184	1.36234	0.986	0.801	6.67E-181	11
Sparc1	2.29E-168	1.178145	0.99	0.878	4.58E-165	11
Vim1	8.58E-168	0.987026	0.998	0.935	1.72E-164	11
Plagl11	3.47E-194	1.375955	0.976	0.638	6.93E-191	12
Peg34	1.78E-184	1.325317	0.997	0.867	3.57E-181	12
Dlk13	4.35E-177	1.504826	0.992	0.691	8.70E-174	12
Hand22	6.17E-176	1.163096	0.951	0.518	1.23E-172	12
Itm2a4	1.93E-171	1.325358	0.989	0.748	3.85E-168	12
Tmsb4x2	5.50E-87	1.689991	1	0.985	1.10E-83	13
Sparc2	3.82E-80	1.310559	1	0.881	7.64E-77	13
Gata6	8.63E-77	1.038987	0.794	0.206	1.73E-73	13
Anxa51	4.21E-72	1.373198	0.942	0.684	8.42E-69	13
Myl92	1.62E-71	1.630096	0.942	0.729	3.23E-68	13
Alx13	3.09E-28	1.467363	0.949	0.644	6.19E-25	14
Kcnmb2	6.15E-28	1.093351	0.808	0.457	1.23E-24	14
Plk23	4.04E-24	1.045568	0.923	0.712	8.08E-21	14
2010111I01Rik2	4.33E-21	0.915288	0.949	0.677	8.65E-18	14
Foxd14	1.69E-19	0.810664	0.91	0.612	3.37E-16	14

769

Table S3	Group		Group
Col9a1	1	Col2a1	2
Mapk14	1	Sox5	2
Zbtb16	1	Sox6	2
Srf	1	Rarg	2
Poc1a	1	Col11a1	2
Carm1	1	Pthlh	2
Zmpste24	1	Runx2	2
Kat2a	1	Gdf5	2
Mapk3	1	Bmp4	2
Dlx2	1	Rara	2
Snai1	1	Msx2	2
Matn3	1	Scx	2
Mycn	1	Bmp2	2
Twist1	1	Shox2	2
Cdx1	1	Rflna	2
Trp53	1	Rarb	2
Rhoa	1	Tgfbr2	2
Mcph1	1	Rflnb	2
Notch2	1	Wnt5a	2
Sbds	1	Wnt9a	2
Cited2	1	Trps1	2
Foxc1	1	Col11a2	2
Sp5	1	Thbs3	2
Insig1	1	Trip11	2
Nsd2	1	Lrp6	2
Tulp3	1	Thbs1	2
Rpl13	1	Msx1	2
Fli1	1	Hspg2	2
Cadm1	1	Mef2c	2
Ppib	1	Axin2	2
Prpsap2	1	Sulf2	2
Tfap2a	1	Sulf1	2
Bbx	1	Ift80	2
Tjp2	1	Por	2
Fzd4	1	Chsy1	2
Csnk1d	1	Smad1	2
Rspo1	1	Fgf18	2
Dixdc1	1	Bmp6	2
Stk11	1	Fgf9	2
Tcf7	1	Tgfb1	2
Ccdc88c	1	Nog	2
Myc	1	Snai2	2

Sox2	1	Tgfb2	2
Gid8	1	Runx3	2
Emd	1	Gli2	2
Ctnnbip1	1	Prrx1	2
Cby1	1	Ddrgk1	2
Cdc73	1	Otor	2
Csnk2a1	1	Uncx	2
Pin1	1	Creb3l2	2
Ddx3x	1	Chst11	2
Hdac1	1	Bmpr1a	2
Rac1	1	Bmp1	2
Peg12	1	Snx19	2
Csnk1g2	1	Pth1r	2
Nxn	1	Six2	2
Depdc1b	1	Fgfr3	2
Ppm1a	1	Gli3	2
Otulin	1	Gnas	2
Sdhaf2	1	Ext1	2
Ccne1	1	Col1a1	2
Aspm	1	Stc1	2
Csnk2a2	1	Sik3	2
Pten	1	Ltbp3	2
Hmga2	1	Grem1	2
Ube2b	1	Ep300	2
Ruvbl1	1	Smad9	2
Nle1	1	Mef2d	2
Rnf220	1	Col27a1	2
Etv2	1	Scube2	2
Ruvbl2	1	Glg1	2
Mesd	1	Anxa6	2
Snx3	1	Smad5	2
Vps35	1	Osr2	2
Fermt2	1	Ctnnb1	2
Gpc5	1	Ror2	2
Wdr61	1	Rspo2	2
Lrrfip2	1	Pkd1	2
G3bp1	1	Smad3	2
Ctdnep1	1	Fgf2	2
Grk6	1	Thra	2
Sdc1	1	Fbxw4	2
Cela1	1	Dicer1	2
Arl6	1	Bmp7	2
Csnk2b	1	Fgfr1	2

Rnf138	1	Prkca	2
Mbd2	1	Satb2	2
Plk1	1	Bmpr2	2
Cenpe	1	Mia3	2
Aurka	1	Lnpk	2
Cdca5	1	Atp7a	2
Mad2l1	1	Ctsk	2
Fbxo5	1	Gdf6	2
Sgo2a	1	Tgfbr1	2
Cdt1	1	Nfib	2

770

Table S4	Group
Crabp1	1
Crabp2	1
AC160336.1	1
Prtg	1
Ldha	1
Grrp1	1
Srm	1
Lin28a	1
Dctpp1	1
Aldoa	1
Rgcc	2
Fabp7	2
Ttyh1	2
Car11	2
Nell2	2
Postn	2
Gpm6b	2
Sparc	2
Zbtb20	2
Hey2	2
Map1b	3
Hist3h2ba	3
Apoe	3
Nfia	3
Anxa2	3
Ckb	3
Rspo1	3
Scrn1	3
Crip1	3
Arl4a	3
Twist1	4
Gsta4	4
Meis2	4
Cd24a	4
Prrx2	4
Slc16a3	4
Sfrp2	4
Mest	4
Cited2	4
Aprt	4
Aard	4
Id1	4

Cth	4
Ptgis	4
Hacd1	4
Naf1	4

771

772

Table S5	p_val	avg_logFC	pct.1	pct.2	p_val_adj	cluster
Col2a1	3.77E-99	1.181788	0.856	0.467	7.53E-96	m0
Sfrp1	7.82E-89	0.915514	0.861	0.519	1.56E-85	m0
Foxc1	1.13E-79	0.86289	0.757	0.355	2.26E-76	m0
S100a11	5.87E-59	0.635084	0.96	0.867	1.17E-55	m0
Sox9	2.12E-58	0.997222	0.761	0.503	4.23E-55	m0
Ebf1	6.53E-55	0.884611	0.82	0.584	1.31E-51	m0
Pkdcc	1.20E-54	0.779568	0.825	0.58	2.41E-51	m0
Six2	1.97E-51	0.637623	0.789	0.514	3.95E-48	m0
Zic2	5.40E-51	0.319418	0.427	0.066	1.08E-47	m0
Calca	9.57E-47	0.732756	0.579	0.307	1.91E-43	m0
Alx1	1.34E-46	0.783765	0.908	0.64	2.68E-43	m1
Mdk	2.32E-39	0.323551	1	1	4.64E-36	m1
Mfap4	9.50E-25	0.386263	0.966	0.935	1.90E-21	m1
Id3	1.31E-22	0.421061	0.974	0.891	2.63E-19	m1
Gata3	3.28E-22	0.444604	0.479	0.255	6.56E-19	m1
Tbx2	6.13E-22	0.407679	0.65	0.429	1.23E-18	m1
Creb5	7.38E-21	0.476501	0.607	0.424	1.48E-17	m1
Mpped2	2.76E-19	0.462925	0.774	0.6	5.53E-16	m1
Wnt5a	2.52E-17	0.423862	0.762	0.62	5.03E-14	m1
Gata2	5.39E-17	0.366436	0.521	0.328	1.08E-13	m1
Stmn1	5.06E-68	0.538458	1	0.991	1.01E-64	m2
Fabp5	2.44E-61	0.62328	1	0.9	4.89E-58	m2
Lix1	3.86E-59	0.757052	0.804	0.369	7.72E-56	m2
Crabp1	3.07E-56	0.871867	0.943	0.687	6.14E-53	m2
Twist1	1.17E-54	0.780003	0.973	0.836	2.33E-51	m2
Lef1	2.80E-51	0.771265	0.714	0.394	5.61E-48	m2
Pax1	1.19E-50	0.6487	0.571	0.213	2.38E-47	m2
Stmn2	7.13E-49	0.78273	0.557	0.204	1.43E-45	m2
Rrm2	3.71E-44	0.756943	0.744	0.433	7.41E-41	m2
AI646519	7.29E-40	0.322885	0.473	0.13	1.46E-36	m2
Dlk1	1.80E-126	2.165033	0.954	0.587	3.59E-123	m3
Cdkn1c	8.01E-112	1.569123	1	0.989	1.60E-108	m3
Plagl1	6.13E-110	1.265317	0.937	0.531	1.23E-106	m3
Mest	4.34E-108	1.079631	0.997	0.761	8.67E-105	m3
Hand2	1.18E-101	1.087069	0.776	0.232	2.36E-98	m3
H19	2.63E-95	1.118661	0.987	0.847	5.26E-92	m3
Meg3	8.66E-89	1.192614	0.997	0.914	1.73E-85	m3
Peg3	4.64E-88	1.270328	0.967	0.789	9.27E-85	m3
Gpc3	4.45E-85	0.829088	0.997	0.83	8.90E-82	m3
Nrk	2.63E-70	0.903391	0.828	0.475	5.26E-67	m3

773

Table S6	p_val	avg_logFC	pct.1	pct.2	p_val_adj
Sfrp1	1.28E-60	1.020448	0.859	0.507	2.57E-57
Nr2f2	1.48E-54	0.98802	0.758	0.268	2.96E-51
Col2a1	7.75E-53	1.075514	0.859	0.562	1.55E-49
Foxc1	3.14E-49	0.870546	0.764	0.366	6.27E-46
Pkdcc	5.71E-41	0.883909	0.826	0.559	1.14E-37
Scx	1.83E-40	0.513053	0.697	0.34	3.66E-37
Ebf1	3.87E-40	1.022873	0.824	0.582	7.74E-37
Sox9	1.59E-39	1.052514	0.762	0.458	3.18E-36
Foxp2	7.85E-39	0.856256	0.719	0.329	1.57E-35
Islr	1.15E-38	0.669	0.752	0.464	2.31E-35
Cxcl12	8.49E-35	0.990234	0.766	0.539	1.70E-31
Sparc	3.39E-34	0.700054	0.966	0.885	6.77E-31
Col9a1	5.47E-33	0.792423	0.687	0.363	1.09E-29
H19	8.17E-32	0.656543	0.943	0.79	1.63E-28
S100a11	5.53E-31	0.555594	0.964	0.896	1.11E-27
Foxp1	1.20E-30	0.637863	0.873	0.602	2.39E-27
Crym	4.23E-30	1.026478	0.683	0.346	8.45E-27
Cdkn1c	1.88E-27	0.580529	0.992	0.983	3.76E-24
Thbs1	6.30E-27	0.652725	0.644	0.38	1.26E-23
Col9a2	3.22E-25	0.533942	0.65	0.389	6.43E-22
Cldn11	6.08E-25	0.614006	0.582	0.305	1.22E-21
Selenom	1.33E-24	0.504677	0.891	0.758	2.67E-21
Tmsb4x	2.49E-24	0.601038	0.994	0.994	4.99E-21
Tgfb2	2.03E-23	0.702159	0.75	0.553	4.05E-20
Calca	3.95E-23	0.710456	0.588	0.363	7.90E-20
Mest	1.51E-22	0.610764	0.855	0.715	3.03E-19
Tgfbi	1.96E-22	0.553116	0.723	0.493	3.93E-19
Nr2f1	2.33E-21	0.591666	0.754	0.524	4.65E-18
Itm2a	5.66E-21	0.646689	0.812	0.7	1.13E-17
Malat1	7.10E-21	0.535658	0.998	0.997	1.42E-17
Igfbp5	7.49E-20	0.936769	0.703	0.55	1.50E-16
Fbn2	3.76E-19	0.534481	0.913	0.853	7.52E-16
Aebp1	6.07E-19	0.534053	0.598	0.357	1.21E-15
Hist3h2ba	2.70E-18	0.590156	0.693	0.516	5.39E-15
Foxd1	5.24E-18	0.734427	0.588	0.343	1.05E-14
Maf	6.25E-18	0.575489	0.606	0.398	1.25E-14
Apoe	1.41E-14	0.53326	0.723	0.605	2.82E-11
Dkk2	1.66E-14	0.67385	0.576	0.349	3.31E-11
Tsc22d1	2.08E-14	0.505721	0.921	0.882	4.15E-11
Gm42418	6.01E-12	0.562094	1	1	1.20E-08
Col4a1	1.03E-09	0.506056	0.723	0.646	2.06E-06
Cthrc1	1.23E-07	0.543544	0.679	0.628	0.000247

Lix1	1.13E-69	-1.0017	0.226	0.793	2.26E-66
Stmn1	2.31E-69	-0.66493	0.988	1	4.63E-66
Twist1	1.63E-67	-1.14838	0.78	0.971	3.27E-64
Stmn2	3.81E-61	-0.9131	0.095	0.556	7.61E-58
Fabp5	1.15E-59	-0.73428	0.877	1	2.30E-56
Ptn	5.52E-58	-1.32899	0.887	0.971	1.10E-54
Crabp1	5.27E-54	-1.05751	0.606	0.922	1.05E-50
Sox11	1.28E-51	-0.70656	0.875	0.991	2.56E-48
Rrm2	2.80E-46	-0.88112	0.376	0.758	5.61E-43
Pax1	3.62E-45	-0.64457	0.147	0.565	7.24E-42
Nasp	6.43E-43	-0.53667	0.915	0.971	1.29E-39
Wnt5a	1.99E-42	-0.85797	0.531	0.81	3.97E-39
Id2	9.28E-40	-0.78257	0.618	0.87	1.86E-36
H2afz	1.85E-39	-0.54631	0.99	1	3.69E-36
Lef1	4.20E-36	-0.81557	0.455	0.706	8.39E-33
Pax3	9.03E-36	-0.86303	0.483	0.718	1.81E-32
Tyms	1.72E-35	-0.57306	0.57	0.847	3.45E-32
Zfhx3	2.78E-34	-1.00285	0.671	0.865	5.56E-31
Anxa2	1.55E-33	-0.64857	0.721	0.922	3.09E-30
Dnajc9	2.43E-33	-0.54418	0.552	0.833	4.85E-30
Gap43	1.47E-31	-0.56717	0.37	0.695	2.95E-28
Tnfaip6	3.66E-31	-0.51224	0.388	0.72	7.33E-28
Pclaf	1.74E-30	-0.92385	0.537	0.767	3.48E-27
Msx1	2.83E-30	-0.93684	0.374	0.654	5.65E-27
Prrx2	6.25E-30	-0.57944	0.695	0.885	1.25E-26
Alx1	4.53E-29	-0.67345	0.626	0.859	9.05E-26
Smc4	5.69E-28	-0.53691	0.469	0.767	1.14E-24
Hist1h1b	2.22E-26	-0.99212	0.307	0.597	4.45E-23
2010111I01Rik	3.81E-26	-0.75023	0.497	0.715	7.62E-23
Hist1h1e	3.66E-25	-0.53061	0.315	0.611	7.31E-22
Hist1h2ap	3.00E-23	-1.1341	0.505	0.689	6.01E-20
Hmgb2	4.87E-23	-0.68815	0.85	0.916	9.74E-20
Hist1h2ae	3.80E-22	-0.83527	0.394	0.631	7.59E-19
Zeb2	3.08E-20	-0.63898	0.495	0.718	6.16E-17
Crabp2	2.47E-19	-0.50323	0.695	0.867	4.95E-16
Top2a	9.68E-19	-0.82848	0.515	0.68	1.94E-15
Ifitm1	1.62E-18	-0.50506	0.269	0.527	3.23E-15
Mki67	3.28E-18	-0.53571	0.473	0.677	6.57E-15
Nbl1	5.49E-17	-0.50253	0.558	0.735	1.10E-13
Ube2c	1.54E-09	-0.56163	0.471	0.605	3.09E-06

774

775

Table S7	p_val	avg_logFC	pct.1	pct.2	p_val_adj
Mfap4	4.07E-26	0.502695	0.967	0.922	8.14E-23
Foxp2	2.86E-19	0.674452	0.612	0.329	5.73E-16
Epha4	7.66E-16	0.508901	0.749	0.553	1.53E-12
Hist3h2ba	9.11E-15	0.51013	0.678	0.516	1.82E-11
Igfbp5	2.71E-14	0.806362	0.689	0.55	5.41E-11
Creb5	1.19E-11	0.51449	0.592	0.444	2.39E-08
Itm2a	4.71E-09	0.690702	0.785	0.7	9.43E-06
Stmn1	2.92E-44	-0.50051	0.992	1	5.85E-41
Fabp5	5.01E-41	-0.63687	0.876	1	1.00E-37
Tyms	4.22E-40	-0.70201	0.504	0.847	8.43E-37
Pclaf	3.76E-39	-1.22321	0.391	0.767	7.52E-36
Hist1h1b	2.62E-38	-1.22349	0.149	0.597	5.23E-35
Rrm2	2.75E-38	-0.87167	0.38	0.758	5.49E-35
Dek	9.78E-35	-0.61002	0.796	0.948	1.96E-31
Hist1h2ap	2.10E-32	-1.53365	0.38	0.689	4.20E-29
Top2a	2.27E-32	-1.15987	0.328	0.68	4.55E-29
Crabp1	1.70E-29	-0.73685	0.683	0.922	3.40E-26
Hist1h1e	6.09E-29	-0.63085	0.24	0.611	1.22E-25
Hmgb2	9.85E-28	-0.82583	0.807	0.916	1.97E-24
Smc4	1.25E-27	-0.62717	0.471	0.767	2.50E-24
Clspn	1.47E-27	-0.52086	0.229	0.576	2.94E-24
Spc24	2.40E-27	-0.60391	0.369	0.72	4.81E-24
Gmnn	4.32E-27	-0.57106	0.355	0.706	8.63E-24
Rrm1	8.55E-27	-0.51978	0.38	0.744	1.71E-23
Lef1	3.13E-26	-0.65071	0.397	0.706	6.25E-23
Vim	1.12E-22	-0.68345	0.879	0.974	2.25E-19
Mki67	2.30E-22	-0.66355	0.394	0.677	4.60E-19
Smc2	9.85E-22	-0.57038	0.46	0.738	1.97E-18
Hist1h4d	6.97E-21	-0.55714	0.242	0.548	1.39E-17
Cdca8	7.20E-21	-0.54295	0.339	0.657	1.44E-17
Cdk1	1.12E-20	-0.61241	0.391	0.654	2.24E-17
Atad2	1.66E-19	-0.57118	0.377	0.643	3.32E-16
Ube2c	2.84E-19	-0.68019	0.3	0.605	5.67E-16
Tubb2b	8.30E-18	-0.52476	0.656	0.888	1.66E-14
Pax1	5.21E-16	-0.58407	0.344	0.565	1.04E-12
S100a6	9.09E-16	-0.66506	0.402	0.646	1.82E-12
Spc25	4.64E-15	-0.51107	0.339	0.571	9.29E-12
Hist1h2ae	8.66E-15	-0.92622	0.46	0.631	1.73E-11
Birc5	1.47E-14	-0.5065	0.419	0.663	2.94E-11
Stmn2	8.97E-12	-0.51079	0.35	0.556	1.79E-08
S100a10	1.91E-10	-0.51272	0.667	0.844	3.82E-07
Alcam	1.89E-08	-0.54002	0.534	0.7	3.77E-05

Enpp2	3.51E-08	-0.54154	0.295	0.484	7.02E-05

776