UC Irvine
UC Irvine Electronic Theses and Dissertations

Title

Transcriptomic and Genetic Investigations into Lineage Specification of the Neural Crest and Tailbud in Vertebrate Embryo Development

Permalink

https://escholarship.org/uc/item/9xr1r5qj

Author

Tatarakis, David

Publication Date

2021

Copyright Information

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA,
IRVINE

Transcriptomic and Genetic Investigations into Lineage Specification of the Neural Crest
and Tailbud in Vertebrate Embryo Development

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Biological Sciences

by

David Tatarakis

Dissertation Committee:
Professor Thomas Schilling, Chair
Professor Ken Cho
Professor Karina Cramer
Associate Professor Ali Mortazavi
Assistant Professor Zeba Wunderlich

2021

TABLE OF CONTENTS

LIST OF FIGURES

Page

ACKNOWLEDGEMENTS

Curriculum Vitae
David Tatarakis

Education

University of California Irvine, Irvine, CA   (Sep 2014-Mar 2021)
Ph.D. Developmental and Cell Biology

University of California Irvine, Irvine, CA   (Sep 2014-Mar 2018)
M.S. Biological Sciences

San Jose State University, San Jose, CA   (Aug 2009-May 2014)
B.S. in Biology (Concentration in Systems Physiology)
Minor: Chemistry
Graduated *Magna Cum Laude*

Publications

Varshney A*, Benedetti K*, Watters K, Shankar R, **Tatarakis D**, Villa DC, et al. "The receptor protein tyrosine phosphatase CLR-1 is required for synaptic partner recognition". *PLOS Genetics*. 2018 May 9;14(5):e1007312.

Tatarakis D, Tuttle Adam, Schilling T. "Lmo7a coordinates neural crest migration and lineage specification by regulating cell adhesion dynamics." *–in Revision at Elife*-

Tatarakis D, Cang Zixuan, Nie Qing, Schilling T. "Single cell transcriptomics elucidates emergence of lineages in migrating cranial neural crest cells and novel regulators" - *manuscript in preparation*-

---

Research Experience

Identification of novel genes coordinating neural crest migration and lineage specification through bulk and single cell RNA-sequencing.
- Uncovered emergence of differentiating lineages using single-cell RNA-seq and whole mount in situ HCR
- Identified putative ligand-receptor interactions driving differentiation by integrating bulk and single-cell RNA-seq data

Generating a single-cell RNA-seq atlas of vertebrate tailbud development
- Identified all cell types in the developing tailbud using single-cell RNA-seq datasets
- Discovered novel cell type transitions using pseudotemporal analysis and gene signature scoring

Characterization of a novel Lim domain protein driving neural crest migration in zebrafish.
- Generated a stable mutant line using CRISPR-Cas9 genome editing
- Analyzed intercellular dynamics of Lmo7a by designing, cloning, and synthesizing fusion RNA constructs
- Examined cell adhesion dynamics in vivo using immunohistochemistry and in situ HCR.

Fellowships and Grants

- Provost Ph.D. Fellowship (University of California Irvine)
- Graduate Dean Recruitment Fellowship (University of California Irvine)
- NIH Graduate Assistance for Areas of Academic Need (GAANN) Fellow (University of California Irvine) (2015-2018)
- NSF-Simons Center for Multiscale Cell Fate (CMCF) Research Fellow (2018-present)
- NSF-Simons Center for Multiscale Cell Fate (CMCF) Independent Research Grant (2019)
- UC Irvine School of Biological Sciences Susan V Bryant Graduate Fellowship Award (2020)

Presentations

The Novel Wnt Regulator Lmo7a Coordinates Neural Crest Cell Migration and Lineage Specification
- Southern California Zebrafish Annual Meeting 2019 (Talk)

Single Cell Transcriptomics Elucidates Emergence of Lineages in Migrating Neural Crest Cells
- Society for Craniofacial Genetics and Developmental Biology 2020 (Talk)

Single Cell Profiling and Genetic Analyses of Fate Determination in the Cranial Neural Crest
David Tatarakis, Adam Tuttle, Qing Nie, Thomas Schilling
- Society for Developmental Biology (SDB) Annual Meeting 2019 (Poster)
- International Zebrafish Conference 2018 (Poster)
- Society for Developmental Biology (SDB) West 2019 (Poster)

Abstract of the Dissertation

Transcriptomic and Genetic Investigations into Lineage Specification of the Neural Crest
and Tailbud in Vertebrate Embryo Development

by
David Tatarakis
Doctor of Philosophy in Biological Sciences
University of California, Irvine, 2021
Professor Thomas F. Schilling, Chair

In embryonic development, tissues arise from multipotent progenitor populations and
undergo lineage specification and terminal differentiation. Two examples of this are the
neural crest (NC) and the multiple progenitor cells of the tailbud. The NC is a
multipotent highly migratory cell type that gives rise to numerous tissues and cell types
in the adult animal, including pigmentation, cartilage and bone, neurons and glia, and
many others. Much is known about the mechanisms driving NC induction and migration,
but how lineage specification and migration are coordinated is still not well-understood.
To address outstanding questions in NC development, we took a multifaceted
approach, employing both traditional functional characterization of a novel Wnt regulator
and data-driven single-cell and bulk transcriptomics.

Here, I show that Lmo7a regulates both migration and specification of glial and pigment
lineages through interactions with focal adhesions (FAs) and canonical Wnt signaling.
Lmo7a is expressed in premigratory NC cell, and its loss results in large midline
aggregates of NC cells that go on to express markers of pigment and glial lineages.
These cells also show abnormally high levels of canonical Wnt signaling. I show that

Lmo7a is required for proper formation of FAs, suggesting a novel connection between cell adhesion and lineage specification in the NC through canonical Wnt signaling.

I also build a single-cell RNA-seq timeline of cranial NC development to investigate lineage decisions. I identify the precise timing of pigment, skeletal, and neural/glial lineage specification and identify several novel markers of each. Further, using combinatorial analysis of bulk and single-cell RNA-seq data, I computationally model the Wnt signaling dynamics that correspond to lineage specification. I identify a putative Wnt regulatory gene *atp6ap2* which encodes the (Pro)renin receptor ((P)RR) protein as a potential regulator of pigment development. Using CRISPR-Cas9 genome editing, I verify a role for ((P)RR) in pigment maturation potentially through canonical Wnt signaling.

I also investigate lineage relationships in the tailbud by constructing a comprehensive single-cell transcriptome map of all cells in the tailbud. I uncover evidence for a novel connection between neuromesodermal progenitors and both vascular endothelium and hematopoiesis. I also show that cell cycle arrest is indicative of a notochord fate in midline progenitor cells. In this thesis, I provide novel insights into lineage specification in the NC and tailbud.

Chapter I
Introduction and Background

A brief history of lineage tracing in developmental biology

The central question at the heart of all developmental biology is that of how a fully formed adult animal grows from a single-celled embryo. The observation over 100 years ago that cells do not arise spontaneously but instead come from the division of pre-existing cells inevitably led to the study of how these daughter cells acquire distinct identities. The earliest of these studies involved careful observation of invertebrates and revealed the then novel idea that the first cell divisions gave rise to cells that had a predetermined fate (Conklin, 1905). This process of tracking the terminal identities (also called "lineages") of progenitor cells became known as "lineage tracing." The apparent pre-determined nature of development informed early models, which described cell fate decisions as irreversible and progressively more restrictive (Waddington, 1957).

In the next decades, researchers would apply several lineage tracing approaches and techniques to characterize the development of many organisms. For example, the relatively simple nature of the *C. elegans* embryo allowed for a complete fate map to be constructed using time lapse imaging with phase contrast microscopy (Deppe et al., 1978; Sulston et al., 1983), but more complex organisms required additional methodology. Classical experiments with tissue transplantation revealed the embryonic origins of entire body segments and organs, exemplified by the discovery of Spemann's organizer through the engraftment of tissue between frog embryos (Spemann & Mangold, 1924) and the mapping of several organ-forming regions in gastrulating avian

embryos through duck/chick allografts (Waddington, 1932) and chick/quail chimeras (Le Douarin & Teillet, 1973). While these techniques were powerful for observing tissue-level lineage relationships, they were limited in their ability to provide cell fate information at the level of single cells.

The ability to track the fates of individual cells was aided by developments in cell labeling technologies. The injection of so-called "vital dyes," which label cells without impacting their viability (Axelrod, 1979; Vogt, 1929) including fluorescent dextrans (Gimlich & Braun, 1985; Weisblat & Shankland, 1985; Wetts & Fraser, 1988) or enzymes like horseradish peroxidase (HRP) (David A. Weisblat et al., 1978) enabled transient labeling of live cells and their progeny. Using this paradigm, it became possible to perform cell lineage tracing in a wide variety of organisms including many vertebrates. Researchers gained valuable insights into the embryonic origins of adult tissues and structures in *Xenopus laevis* (Eagleson et al., 1995; Eagleson & Harris, 1990), chicken (Garcia-Lopez et al., 2009; Serbedzija et al., 1989), mouse (Balakier & Pedersen, 1982; Johnson & Ziomek, 1981; Lawson et al., 1986; Rosenquist, 1971), and many teleost fish species (Ballard, 1973, 1982, 1986) including zebrafish (Kimmel et al., 1990). Despite their broad usefulness in tracking the fates of cells in many organisms, these cell labeling techniques had numerous limitations. Dye dilution reduced the specificity and longevity of labels, making it difficult to draw definitive conclusions about clonal relationships between cells over long periods of developmental time. Further, these techniques provided no information about molecular mechanisms or signaling events that drove the lineage relationships they uncovered.

The advent of genetic engineering in the late 20th century brought new, powerful techniques for lineage tracing in the form of cell labeling by genetic recombination. The first of these methods, heat shock-induced FLP-FRT, was employed to map clonal relationships between cells in *Drosophila* embryos by driving LacZ expression (Harrison & Perrimon, 1993). In mouse, the Cre-Lox system was developed first to drive expression of LacZ (Soriano, 1999) and later EGFP (Mao et al., 2001) in specific cell types/tissues. Later, Cre-Lox labeling was expanded to enable analysis of clonal relationships in mouse neurons and glia through the development of Brainbow (Livet et al., 2007), which was then adapted to other cell types and organisms (Pan et al., 2013; Hampel et al., 2011). These systems allowed for permanent genetic labeling of cells to track their progeny. These techniques, although orthogonal to earlier cell labeling approaches in their methodologies, achieved largely the same goals: observation of lineage relationships across developmental time. They still did not in and of themselves offer the ability to characterize the mechanisms or molecular pathways driving these cell fate decisions.

The 21st century gave rise to the "Omics revolution," where next generation sequencing (NGS) technologies offered the ability to measure gene expression at an unprecedented scale. RNA sequencing (RNA-seq) enabled researchers to measure the relative abundances of all mRNAs in a sample, allowing for investigations into gene expression signatures with orders of magnitude more information than with traditional quantitative PCR (qPCR) and microarray techniques (Lister et al., 2008; Mortazavi et al., 2008;

Nagalakshmi et al., 2008). In developmental biology, lineage tracing capabilities were furthered by advances in NGS technologies that could profile the transcriptomes of individual cells (Gierahn et al., 2017; Jaitin et al., 2014; Macosko et al., 2015; Picelli et al., 2014). These technologies were quickly deployed to explore the transcriptional diversity among cells in many tissues and developmental contexts (Haber et al., 2017; Klein et al., 2015; Zeisel et al., 2015). The enormous influx of transcriptomic data enabled researchers to investigate cell fate decisions in a completely new way by employing increasingly sophisticated computational techniques that inferred lineage relationships from single-cell RNA-seq data (Cannoodt et al., 2016; La Manno et al., 2018; Wolf et al., 2019). This allowed for lineages to be tracked along with the changes in gene expression that accompanied fate determination and other cellular behaviors in development. Several initiatives were undertaken to map the entirety of cellular transcriptional heterogeneity in whole adult animals including the Human Cell Atlas (Rozenblatt-Rosen et al., 2017) and Tabula Muris (Schaum et al., 2018). The first whole embryo single-cell transcriptomic map of early vertebrate development was achieved in zebrafish, where thousands of cells were sequenced from multiple stages ranging from pre-gastrulation to early larval stages (Briggs et al., 2018; Farrell et al., 2018; Wagner et al., 2018). These large-scale transcriptomic maps provide an invaluable resource to guide further, more detailed investigation into the specification of cell fates in particular tissues and organ systems.

Perspectives on lineage specification are ever-evolving, and many questions remain. The classical view of development as a series of progressively more restrictive lineage bifurcations is not sufficient to explain the complex dynamics at play. Some cell fate decisions may indeed be stable, while others may be reversible, demonstrated perhaps most dramatically by the discovery that adult cells can be reprogrammed to a pluripotent state by the expression of just a handful of factors The stability or reversibility of cell fate decisions (Takahashi & Yamanaka, 2006). Further, while some lineage tracing studies in invertebrates have achieved a complete hierarchical map of cell fate decisions, such efforts are orders of magnitude more challenging in vertebrates. This is because unlike invertebrate body plans which are largely epithelial, vertebrates are comprised of a huge number of mesenchymal cells that exhibit dynamic cellular behaviors and migrate through complex environments throughout development. Thus, to achieve a complete understanding of cell fate decisions in vertebrates, we will need to examine questions such as: How many cell types/states exist within developing tissues? How plastic are individual lineage decisions? What intercellular signals are involved in such decisions? How are migratory behaviors coordinated with lineage specification? In this thesis, I will present my work using traditional reverse genetic and molecular biology approaches along with scRNA-seq to address some of these difficult questions and advance our understanding of cell fate determination in neural crest and vertebrate tailbud development.

Neural crest induction and migration

The neural crest (NC) is a multipotent population of cells that is specified at the border of the neural plate and epithelium. These cells delaminate from the neuroepithelium and migrate throughout the embryo, eventually giving rise to numerous cell types and structures. (Mayor & Theveneau, 2012). The NC, first identified in 1868 by Wilhelm His in chick embryos (His, 1868) is present in all vertebrates, and the emergence of this cell type played a major role in the evolution of the head (Gans & Northcutt, 1983; Northcutt & Gans, 1983). NC cells are broadly divided into two main categories: cranial and trunk. Cranial NC cells form the craniofacial skeleton, the facial nerves, and pigmentation of the skin and eyes. Trunk NC cells form the dorsal root ganglia, pigmentation in the rest of the body, and several other structures including the enteric nervous system (Vega-Lopez et al., 2017).

NC induction is a complex process involving many gene interactions and inductive signals from surrounding tissues as the neural and nonneural ectoderm are segregated during neural tube formation. Among these are fibroblast growth factor (FGF), Wnt, and bone morphogenic protein (BMP) signaling molecules secreted from nonneural ectoderm and underlying axial mesoderm (Garnett et al., 2012; Meulemans & Bronner-Fraser, 2004). The combined signaling from these pathways leads to the expression of neural crest specifier genes at the folds of the neural plate including *Foxd3*, *Snail1/2*, *AP-2α*, *Sox10*, and *Twist* (Hopwood et al., 1989; LaBonne & Bronner-Fraser, 1998; Luo et al., 2003; Sasai et al., 2001; Wang et al., 2011). These factors and others interact in

numerous ways, promoting and maintaining each other's expression. These interactions form a complex gene regulatory network that is not yet fully understood.

Following induction, NC cells must delaminate from the neuroepithelium and migrate to their target destinations. This process involves an epithelial-to-mesenchymal transition (EMT) characterized by shifts in cadherin expression. Premigratory NC cells first downregulate E-cadherin and upregulate N-Cadherin and Cadherin6b. This first shift is then followed by a downregulation of N-cadherin and Cadherin6b and an upregulation of Cadherin7 and Cadherin11 (Nakagawa & Takeichi, 1995, 1998). These changes are driven by cooperation of the NC specifiers downstream of Wnt and BMP signaling (Burstyn-Cohen et al., 2004; Heuberger & Birchmeier, 2010; Stuhlmiller & García-Castro, 2012). Canonical Wnt signaling is particularly important as it drives expression of *Snail1*, *Snail2*, and *Twist1*, which then regulate the changing cadherin profiles of delaminating NC cells (Chalpe et al., 2010; Lander et al., 2013).

After delamination, the NC begins to migrate away from the dorsal neural tube, quickly diverging into discrete streams along different defined pathways. In the cranial NC, streams form corresponding to the pharyngeal arches (PAs) whereas in the trunk NC, streams are defined by their migratory path medial or lateral to the somites. Directed migration of these streams is the result of both guidance cues from surrounding tissues and cell-cell communication between the NC cells themselves (Szabó & Mayor, 2018). Ephrins (Eph) secreted from surrounding mesodermal tissues play a crucial role in constraining and guiding the migratory streams in both the trunk and cranial NC

(Santiago & Erickson, 2002; Smith et al., 1997; Wang & Anderson, 1997). Vascular endothelial growth factor (VEGF) and FGF serve as secreted signals to guide the migration of the NC migratory streams, the loss of either of which disrupts their directed migration (McLennan et al., 2010; Trokovic et al., 2005). Cxcl12/Sdf1 is expressed in the pharyngeal mesenchyme and signals to cranial NC cells through Cxcr4 to facilitate their migration, however it remains unclear if this signaling is merely a permissive signal for migration or a true directional chemotactic signal. The relative importance of this interaction for robust NC migration is also not clear (Olesnicky Killian et al., 2009; Theveneau et al., 2010).

In addition to these external signals, communication between NC cells also drives their directed migration. NC cells exhibit contact inhibition of locomotion (CIL) whereby cells that come into contact with each other repolarize to move away from one another. This behavior may be at least partially responsible for the migration of NC cells away from the NC-dense dorsal midline following EMT (Carmona-Fontaine et al., 2008; Teddy & Kulesa, 2004). Interestingly, this behavior appears to depend on the noncanonical Wnt/planar cell polarity (PCP) pathway, and involves the expression of Wnt11 in migrating NC cells (Theveneau et al., 2010). NC cells may also communicate through the complement pathway with NC cells expressing both the complement fragment C3a and its receptor C3aR leading to mutual attraction (Carmona-Fontaine et al., 2011). These types of cell-cell interactions inform the theory of leader-follower dynamics in NC collective migration, which holds that a small number of NC cells at the migratory front are specialized to guide the migration of the entire stream (McLennan et al., 2012,

2015; Morrison et al., 2017). However the nature of this leader cell population is not fully understood (Richardson et al., 2016; Theveneau & Linker, 2017).

Models of Lineage Specification in the Neural Crest

The unique combination of migratory behavior and multipotency in the NC has fascinated developmental biologist for many years and led to numerous efforts at mapping the contributions of NC cells to adult tissues through traditional lineage tracing techniques (Le Douarin & Teillet, 1973; LeDouarin, 1980; Serbedzija et al., 1989, 1990, 1992; Weston, 1970). A long-standing area of interest in NC biology concerns the timing and mechanism of lineage decisions. The NC as a whole has a remarkable ability to give rise to a huge array of cell types, but the ability of individual NC cells to contribute to these lineages is less clear (Dupin et al., 2018).

Some evidence supports the notion that NC cells are entirely unbiased towards any lineage until they reach their destination in the embryo. Experiments using quail NC cells cultured in vitro and then injected into chick embryos allowed for clonal lineage tracing that revealed multiple lineages emerging from a single NC progenitor cell (Bronner-Fraser et al., 1980). Other lineage tracing studies in avian embryos using vital dye and fluorescent protein labeling also indicated that some individual NC cells gave rise to multiple cell types in vivo (Marianne Bronner-Fraser & Fraser, 1988; McKinney et al., 2013). Other researchers have employed genetic recombination lineage tracing to
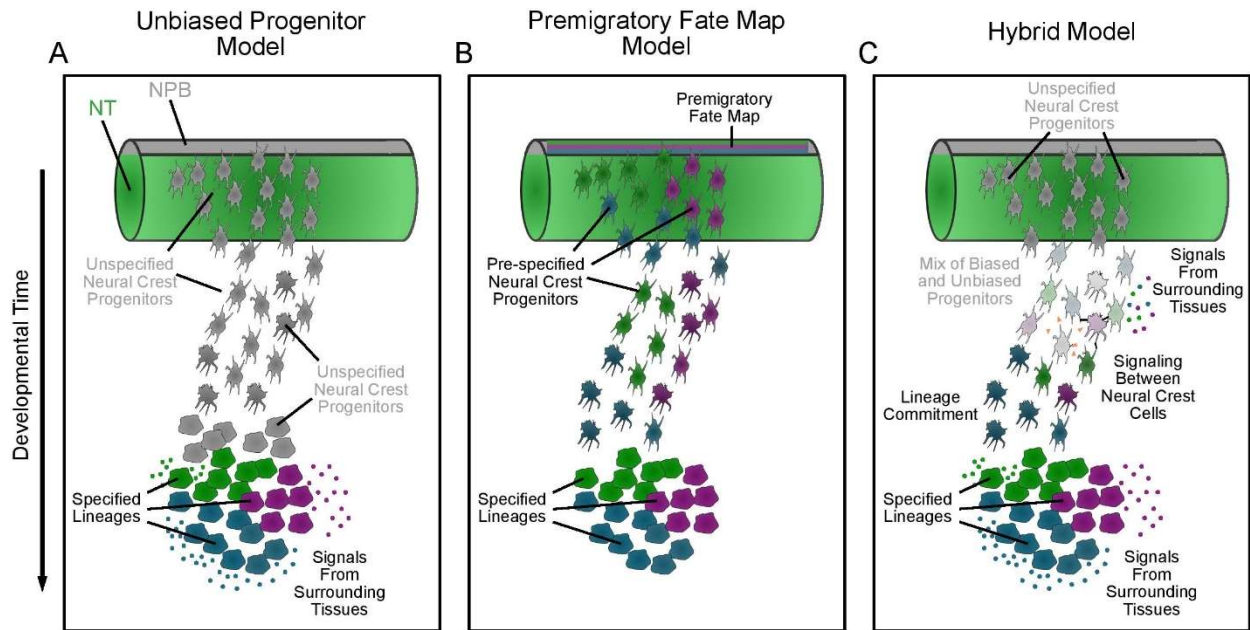
Figure 1.1: Simplified models of lineage specification in the NC. A) A model for lineage specification where NC cells migrate away from the neural tube (NT) as a homogenous population of unspecified progenitors. NC cells arrive at their target destination and then receive signals from the surrounding tissues to acquire the appropriate lineage identity. **B)** Another model where a fate map arises not long after NC induction. NC cells have predefined identities based on their position in the premigratory population. They begin already spatially segregated and migrate as committed progenitors and arrive at their target destinations to undergo terminal differentiation. **C)** A hybrid model incorporating aspects from both **A** and **B**. In this model, NC cells begin unspecified as they leave the neural ectoderm, but quickly become biased towards one or more lineages influenced by both extrinsic signals from the surrounding tissues and communication between NC cells themselves. The biased progenitors become spatially segregated and arrive at their target destinations to undergo lineage commitment and terminal differentiation owing partially to signals from the surrounding tissues.

examine this question in mouse. One such study found no apparent lineage restrictions in NC cells when recombined either during migration or prior to the onset of migration. Individual NC cells were recombined at multiple stages before and during migration, and in all cases only very few cells gave rise to lineage-restricted derivatives (Baggiolini et al., 2015). These findings support a model for NC lineage specification where premigratory NC cells are entirely multipotent and acquire their later identities through

signals from their environments after migration (**Fig 1.1 A**). Other evidence points toward a much different model of lineage specification, where a fate map exists in the early NC prior to migration (**Fig 1.1 B**). Lineage tracing studies in zebrafish and chick using vital dye labels showed that clones derived from individually-labeled premigratory NC cells were restricted to particular lineages based on their location (Krispin et al., 2010; Raible & Eisen, 1994; Schilling & Kimmel, 1994). Some in vitro studies also revealed clonal restrictions through serial cultures of single mammalian NC cells (Stemple & Anderson, 1992).

These competing models of lineage specification have been difficult to reconcile, and it seems likely that both perspectives have validity. The true picture of NC lineage specification may involve a hybrid model (**Fig 1.1 C**), in which there is some degree of lineage bias in premigratory and/or early migrating NC cells with terminal lineage commitment being achieved through communication between NC cells and their environment as well as between NC cells. Such a model would be difficult to examine comprehensively using traditional techniques, but the advent of single-cell -omics approaches presents an exciting new paradigm to examine this issue.

Single-cell transcriptomic studies examining lineage specification in the NC
Recently, several groups have begun to apply these technologies to the questions of lineage specification and migration in the NC. One group profiled NC cells in different segments of the cranial migratory streams at multiple stages (HH stages 11, 13, and 15) in chick embryos using plate-based single-cell RNA-seq (scRNA-seq) (Morrison et al.,

2017). In this study, the researchers identified a putative population of "trailblazer" NC cells at the migratory front, marked by a unique transcriptional signature, which they had previously identified in bulk sequencing experiments (McLennan et al., 2015). Apart from this population, however, very little heterogeneity was observed, and no signs of lineage specification were evident.

Another study examined the premigratory cranial NC for signs of lineage specification in zebrafish (Lukoseviciute et al., 2018). Cells were FACS sorted from a *foxd3* genetrap transgenic, resulting in 94 single-cell libraries built using Smart-seq2. The data revealed a remarkably homogenous population of cells, with broad expression of NC specifier genes like *sox10*, and *tfap2a*. None of these cells expressed any markers that would indicate specification of later lineages. Interestingly, a very small group of cells captured displayed an entirely different expression profile, characterized by absence of NC specifier genes and high expression of pluripotency factors including *vox*, *vent*, and *cx43.4*. These cells notably also had low expression of *foxd3*. These cells were similar in their expression to cells marked by the *foxd3* transgenic marker at pre-gastrula stages, leading the authors to the conclusion that they represent a pluripotent stem-like population of NC cells.

An scRNA-seq analysis of premigratory and early migratory NC cells in chick utilized both plate-based Smart-seq2 sequencing and 10X Chromium droplet sequencing to examine transcriptional heterogeneity (Williams et al., 2019). FACS sorted NC cells, marked by an electroporated plasmid containing a Foxd3 enhancer-driven citrine

fluorescent marker were sequenced from 6SS and 7SS staged chick embryos. Their analysis of the Smart-seq2 data revealed 3 populations of NC cells. One which corresponded to an "early" NC population marked by NC specifier genes, and another resembled later NC cells, marked by supposed "mesenchymal"-like genes including *Lmo4*, C*xcr4*, and *Col9a4*. These early and late populations were largely segregated between the 6SS and 7SS stages, respectively. However, a third population was identified with an expression profile that shared some similarities with the early NC population, but also contained neural markers like *Sox3* and P*ax6*. This group was dubbed a "neural-NC" population, representing an early bias toward neural lineages. This population was also observed in their 10X data, along with a bona fide NC population and several other cell types. Interestingly, these other cell types, which included neuroepithelium and a *Pitx2*+ "mesenchymal" population were not identified in the Smart-seq2 data, despite identical FACS isolation. From these results, the authors conclude that cranial NC cells are comprised of two populations: an unspecified NC population and a pool of neural lineage-biased NC progenitors.

Most recently, a series of scRNA-seq experiments were performed in mouse, using 10X Chromium in a Wnt1-Cre transgenic line marking trunk and cranial NC cells (Soldatov et al., 2019). 1107 cells were FACS sorted and sequenced from E9.5 transgenic embryos, revealing several populations. Neural tube (NT) cells were identified along with pre-EMT NC cells. These populations had largely similar expression profiles, as would be expected, but they differed in that NT cells uniquely expressed *Pax6* while early NC cells expressed *Cdh11*. A bona fide NC population showed signs of NC specification,

with expression of genes including *Sox9* and *Foxd3*. RNA velocity indicated that these cells represented a smooth transition from neural tissue to migratory NC. Two major lineage branches were identified differentiating from the early migratory NC population: "mesenchymal" and autonomic nervous system. Another data set of cranial NC cells was obtained from the same transgenic line at E8.5 and E9.5. These data were made up largely of cephalic NC (cranial NC cells that migrate around the eyes and brain anterior to the pharyngeal arches [PAs]) along with smaller PA NC populations and a neuroepithelial population. Marker gene analysis revealed gene expression signatures of "mesenchymal" precursors at the later (E9.5) stage. These cells were marked by upregulation of *Twist1* and *Prrx1*. Interestingly, in both the trunk and cranial NC, no signatures of pigment progenitors were identified. This led to the to the conclusion that migrating NC cells first diverge between an autonomic and "mesenchymal" lineage branch, with pigment lineages being specified later.

These scRNA-seq studies each sought to examine the heterogeneity present in premigratory and/or migrating NC cells, but arrived at very different conclusions. Premigratory NC cells in zebrafish appeared almost entirely homogenous apart from a population of putative stem-like cells in the study by Lukoseviciute et al (2018), however this population was not identified in any of the other single-cell studies. Similarly, the studies in chick each identified unique populations among an otherwise largely homogenous population of NC cells. Morrison et al (2017) recovered a group of NC cells at the migratory front that expressed many genes previously described as part of a "trailblazer" identity but found no signs of lineage specification at any stage. Also in

chick, Williams et al (2019) reported a putative neural-biased group of NC cells present at early stages, but no indication of a trailblazer population. Further, none of these populations were identified in mouse by Soldatov et al (2019), who instead found signs of "mesenchymal" and neural/glial lineage specification arising during migration.

These discrepancies could, in part, be explained by species-specific differences, however they are more easily explained by methodology and differences of interpretation of the data. The earlier studies by Morrison et al (2017) and Lukoseviciute et al (2018) recovered a relatively small number of cells as they preceded droplet-based sequencing technologies, and more nuanced methods for dimensionality reduction beyond PCA were not yet available or widely used in single-cell analysis. Notably, while all of the studies utilized FACS to obtain pure populations of NC cells, the later droplet-based experiments clearly indicate that many non-NC cell types are captured, as the authors themselves indicate (Soldatov et al., 2019). This is a very common feature of sequencing experiments that rely on FACS. FACS gating is inadequate to obtain truly pure populations, necessitating careful subclustering in downstream analyses. For this reason, it is very likely that some of the populations identified in earlier studies may in fact be non-NC cell types mis-attributed to unique NC subpopulations. A reanalysis of their data and a recapitulation of the experiments using current methodologies would likely uncover information that may have been missed in these studies.

Another limitation of previous single-cell NC studies is the lack of temporal information. Global changes in gene expression as NC cells delaminate and migrate are difficult to

deconvolve from expression signatures that indicate lineage bifurcation when looking at a single snapshot in time. Further, different lineage decisions likely span a range of developmental time, making it unlikely to capture them all without a time course. Finally, computational methods for trajectory inference (pseudotime, RNA velocity), while powerful, are best used as hypothesis-generation tools. When used in a vacuum, they can sometimes imply lineage connections and differentiation that does not comport with experimentation. Validation and the addition of real developmental temporal information would be very helpful in evaluating the lineage bifurcations occurring throughout NC development.

Lineage relationships in the vertebrate tailbud

One aspect of development that is shared across all vertebrates is the presence of an embryonic tail. The tail is made up primarily of three tissues: the posterior neural tube, the notochord, and paraxial mesoderm (Beck, 2015). Also present in the tail are tissues just dorsal and ventral to the notochord, respectively: the floor plate and the hypochord, the latter of which exists exclusively in anamniotes (Löfberg & Collazo, 1997). These tissues grow and extend as the embryo elongates, eventually giving rise to many structures in the adult animal. The growth of the tail depends on cells residing in the caudal-most region of the tail, termed the "tailbud." Ordinarily, different embryonic tissues have distinct origins in the three germ layers, but the elongating tail is unique in that its tissues appear to derive from a common pool of cells in the tailbud (Noemi

Cambray & Wilson, 2007; Noemí Cambray & Wilson, 2002; McGrew et al., 2008). In vivo clonal lineage tracing revealed that the neural and mesodermal tissues in the tail indeed derive from a population of bipotent cells called neuromesodermal progenitors (NMPs) (Martin & Kimelman, 2012; Tzouanacou et al., 2009). The decision between neural and mesodermal lineages is controlled by differential levels of canonical Wnt signaling (**Fig 1.2**) (Bouldin et al., 2015; Garriock et al., 2015; Gouti et al., 2014; Takemoto et al., 2011). Low Wnt signaling pushes NMPs to a neural fate, while high
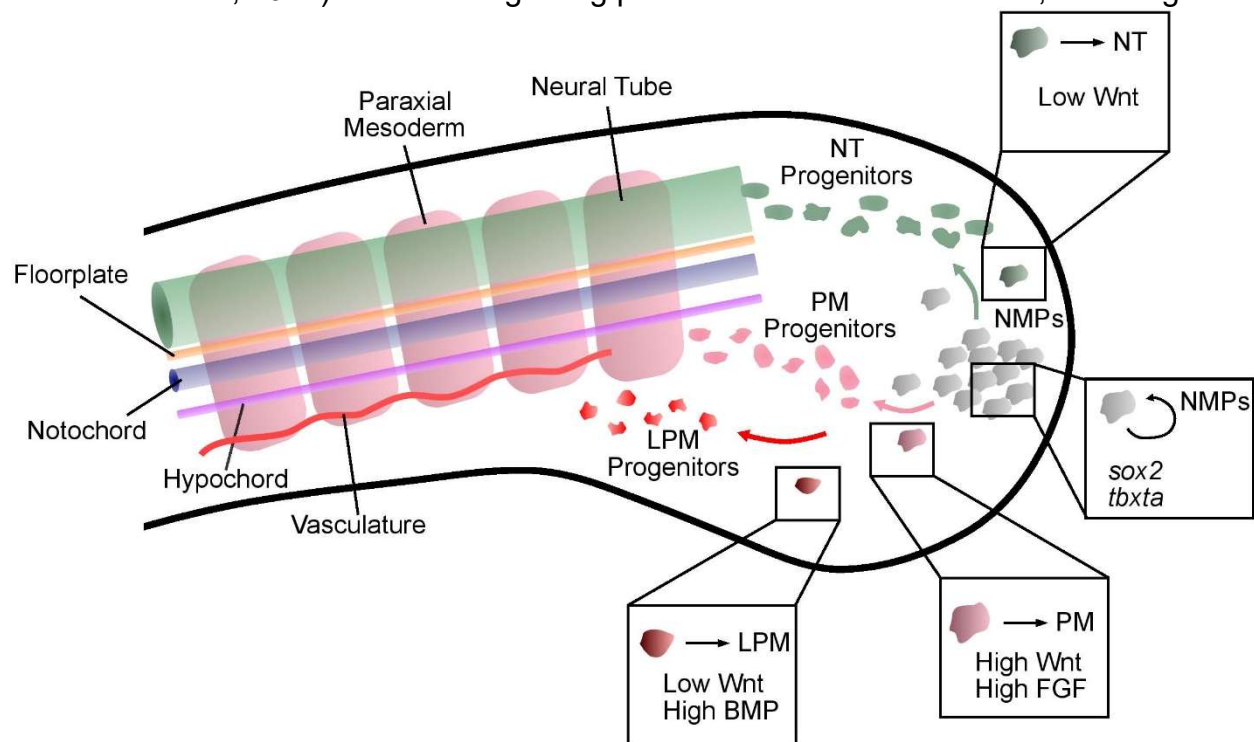
Figure 1.2: Lineage specification of neuromesodermal progenitor derivatives in the vertebrate tailbud. Diagram showing the primary tissues of the tail and the 3 major lineages canonically specified by neuromesodermal progenitors (NMPs) in the tailbud**.** NMPs exist as a mass of self-renewing multipotent cells in the caudal-most tip of the tailbud, sustained by expression of *sox2* and *tbxta*. NMPs give rise to both neural tube (NT) and mesodermal derivatives. This first bifurcation is controlled by canonical Wnt signaling. NMPs that receive high Wnt become mesodermal, while cells receiving low Wnt become NT. The mesodermal cells further bifurcate into paraxial mesoderm (PM) and lateral plate mesoderm (LPM). Mesodermal precursors that experience high Wnt and High FGF signaling become PM, while those that receive low Wnt signaling and high BMP signaling are driven to an LPM fate.

Wnt results in a mesodermal fate. The mesodermal descendants are further bifurcated into paraxial mesoderm, which generates the axial skeleton, musculature and dermis of the skin, intermediate mesoderm, which generates the kidneys and associated tubules, and lateral plate mesoderm, which generates blood and vasculature. Paraxial mesoderm is specified by sustained Wnt signaling and increased FGF signaling, and lateral plate mesoderm is induced by a loss of Wnt and an increase in BMP signaling (Row et al., 2018).

In addition to neural and mesodermal lineages, there is evidence that the NMPs may be responsible for the development of vascular endothelial tissues. Primary vasculature is derived from the lateral plate mesoderm specified during gastrulation, but in zebrafish, it has been proposed that a secondary round of vascularization occurs in the posterior region of the embryo (Martin & Kimelman, 2012). Genetic analysis of the vascular mutant *cloche* revealed that despite a loss of vasculature lineage markers in the trunk, a population of apparent vasculature progenitors remains in the tail (Liao et al., 1997; Thompson et al., 1998). In the limb bud, somitic mesoderm-derived cells have been shown to contribute to vasculature (Kardon et al., 2002), raising the possibility that the tailbud vasculature progenitors may derive from the NMPs. This lineage relationship has since been demonstrated in vivo (Row et al., 2018), but whether or not this population also gives rise to blood remains an open question.

The relationship between notochord, hypochord, and floor plate is also an active area of investigation. The notochord and floor plate are transient structures that serve as both a

structural scaffold and a source of signals to pattern the neural tube and surrounding

mesodermal tissues (Placzek & Briscoe, 2005; Stemple, 2005). The hypochord secretes

VEGF and other signals to pattern the developing dorsal aorta in fish and amphibians

(Löfberg & Collazo, 1997). These structures derive from a population of midline

progenitor cells (MPCs) which are specified during gastrulation (Catala et al., 1995,

1996; Melby et al., 1996; Selleck & Stern, 1991). Lineage tracing studies have revealed

species-specific differences in the potency of MPCs. In chick, the notochord and floor

plate arise from the same pool of cells, whereas in mouse, they appear to come from

two groups of MPCs that diverge during gastrulation (Jeong & Epstein, 2003; Selleck &

Stern, 1991). The origins of the hypochord are less understood, but it is generally

thought of as having an endodermal origin (Eriksson & Löfberg, 2000; Löfberg &

Collazo, 1997). However, recent work in the zebrafish has offered evidence to suggest

the existence of bipotential hypochord/notochord progenitors as well as notochord/floor

plate progenitors both residing in the tailbud (Row et al., 2016). Much remains to be

learned about the complex lineage relationships in the tailbud, presenting an exciting

opportunity to leverage the power of modern -omics technologies in discovering new

and difficult to identify cell populations in development. Two studies using targeted

single-cell RNA-seq have examining a few of the tailbud progenitor cell types (Chestnut

et al., 2020; Gouti et al., 2017). However there has yet to be an effort to profile the

transcriptomes of all cells present in the developing tailbud comprehensively.

In this thesis, I will present my efforts to uncover lineage relationships, characterize the

spatial and temporal components of cell fate decisions, and identify novel regulatory

genes driving and coordinating such decisions through multifaceted investigations using the zebrafish as a model organism. In chapter II, I characterize the role a novel regulator of cranial NC development, Lim domain only 7a (Lmo7a), which both facilitates migration of groups of NC cells and influences their lineage decisions through canonical Wnt signaling. In chapter III, I use a novel approach to explore lineage decisions in the cranial NC using single-cell RNA-seq while maintaining both spatial and temporal information. I then bolster my findings using in vivo spatial transcriptomic imaging and CRISPR-Cas9 genome editing, uncovering a novel regulator of pigment cell development. Finally, in chapter IV, I use single-cell RNA-seq analysis to build a complete transcriptomic atlas of the developing tailbud and potentially uncover a long-hypothesized connection between NMPs and hematopoietic lineages.

Chapter II

The role of Lmo7a in coordinating migration and lineage specification in cranial NC cells

Introduction

During embryonic development, cell migration and lineage specification must be tightly coordinated. Many of the mechanisms driving progenitor cell migration also regulate differentiation toward specific lineages (He et al., 2018; McBeath et al., 2004). In vertebrates, this is particularly true for mesenchymal cell populations such as migratory neural crest (NC) cells, which emerge from the neural ectoderm and disperse to generate an extraordinary variety of cell types throughout the body including neurons, glia, pigment cells, cartilage, and bone. Despite extensive studies of both intrinsic and extrinsic factors that specify these different fates, how the acquisition of distinct NC lineages relate to their migratory behaviors remains largely unclear (Kalcheim & Kumar, 2017).

NC cells are induced at the neural plate border during neural tube closure through a combination of Wnt, BMP, FGF, and Notch signaling (Stuhlmiller & García-Castro, 2012). They subsequently undergo epithelial-mesenchymal transition (EMT) and migrate away from the dorsal midline along distinct trajectories throughout the embryo (Kerosuo & Bronner-Fraser, 2012; Mayor & Theveneau, 2012). Among these migratory paths, are cranial NC cell streams that populate the pharyngeal arches (PAs) to form the facial skeleton, as well as distinct lateral and medial pathways in the trunk in which

NC cells form pigment cells in the skin or sensory neurons and glia in peripheral nerves, respectively. Some in vivo lineage tracing and in vitro clonal analyses have suggested that these fates depend entirely on the migratory environments and final destinations of NC cells (Baggiolini et al., 2015; M. Bronner-Fraser et al., 1991; Dupin et al., 2010). Other experiments have provided evidence for early lineage specification in premigratory NC and a link between initial position, migratory path, and cell fate (Krispin et al., 2010; Schilling & Kimmel, 1994; Stemple & Anderson, 1992). In recent years, the advent of single cell transcriptomics has given us a more detailed picture of the degree to which NC cell fates are both dynamic and heterogeneous during migration (Lukoseviciute et al., 2018; Morrison et al., 2017; Soldatov et al., 2019).

Canonical Wnt signaling plays important roles in inducing NC cells, promoting their migration, and driving lineage decisions at later stages. Tight regulation of signaling levels is vital for proper initiation of migration (Ahsan et al., 2019; Hutchins & Bronner, 2018; Maj et al., 2016) and also biases cells toward pigment versus glial cell fates through regulation of genes such as *Sox10/Foxd3* and *Pax3/7,* respectively (Curran et al., 2010; Dorsky et al., 1998; Minchin & Hughes, 2008). We previously demonstrated novel roles for Ovol1a and Rbc3a/Dmxl2 in promoting NC migration, due at least in part to changes in responses to Wnt signaling (Piloto & Schilling, 2010; Tuttle et al., 2014). Both are specifically expressed in premigratory NC cells, and loss-of-function of either gene disrupts the migration of subsets of NC cells, which form aggregates in the dorsal midline and acquire pigment cell fates. Ovol1a is a direct Wnt target (Li et al., 2002), while Rbc3a/Dmxl2 controls Wnt receptor trafficking, and both alter the localization of

cadherins (such as Cdh2) important for migration. These results establish a link between adhesive mechanisms that govern NC migration, Wnt signaling, and the decisions that lead toward specific cell fates.

In a microarray screen to identify genes downregulated in *tfap2a/g*-deficient zebrafish embryos, we discovered the Lim-domain-only 7 gene *lmo7a* (Hoffman et al., 2007; Tuttle et al., 2014). Lim-domain proteins vary in structure and cellular function and include the Lim-domain-only (Lmo) subclass. Lmo1-4 contain no annotated functional domains apart from multiple Lim domains and function as nuclear transcriptional co-regulators important in cancer progression (Matthews et al., 2013; Sang et al., 2014). Lmo4 promotes EMT of NC and neuroblastoma cells through direct binding to Snail and Slug transcription factors (Ferronha et al., 2013; Ochoa et al., 2012). A protein containing four and one-half Lim domains (FHL2) interacts with β-catenin (β-cat) to either increase or decrease levels of TCF/LEF dependent transcription, depending on the cellular context (Hamidouche et al., 2008; B. Martin et al., 2002). Other Lim-domain family members contain multiple functional domains and have more divergent roles.

Lmo7, despite its name, contains a calponin homology (CH) domain, a PDZ domain, and a single Lim domain. In mammals, multiple functions have been described for Lmo7, including promoting expression of myogenic transcription factors in skeletal muscle cells (Dedeic et al., 2011; Holaska et al., 2006) and regulating afadin-nectin-E-cadherin junctions in epithelial cells (Ooshio et al., 2004) and in the cuticular plate of the cochlea (Du et al., 2019). Interestingly, its function in myogenesis requires entry into the

nucleus and transcriptional regulation, while its epithelial functions involve interactions with the actin cytoskeleton and membrane-associated proteins. Lmo7 also influences cancer cell metastasis, such as the expression of myocardin-related transcription factors through regulation of Rho-dependent actin dynamics at the cell membranes of breast cancer cells (Hu et al., 2011; Nakamura et al., 2005; Teixeira et al., 2014). A paralog of Lmo7, LIMCH1, regulates cell migration through roles in focal adhesion (FA) formation and actomyosin dynamics (Lin et al., 2017). Lmo7 can localize to FAs and act as a shuttling protein to mediate integrin signaling in HeLa cells and mouse embryonic fibroblasts (Holaska et al., 2006; Wozniak et al., 2013). Both LIMCH1 and Lmo7 are associated with poor prognosis in human lung cancer (Karlsson et al., 2018), and expression of Lmo7 (also called PCD1) is associated with increased metastasis in numerous human cancers (Furuya et al., 2002; Kang et al., 2000; Sasaki et al., 2003).

Here, we show that zebrafish Lmo7a promotes NC migration and modulates lineage decisions through interactions with canonical Wnt signaling, similar to Ovol1a and Rbc3a/Dmxl2. Lmo7a is expressed in premigratory NC cells where it localizes to cell membranes, and loss-of-function leads to aggregation of subsets of NC cells at the dorsal midline. These cells show elevated nuclear β-cat as well as altered paxillin (Pxn) localization and reduced phosphorylated focal adhesion kinase (pFAK). Furthermore, analysis of gene expression in these NC cell aggregates reveals that the cells adopt identities of pigment and glial progenitors, but not other NC lineages. Our results suggest that Lmo7a has a dual role in promoting migration of NC cells and regulating

lineage decisions through modulation of canonical Wnt signaling and cell adhesion

dynamics.
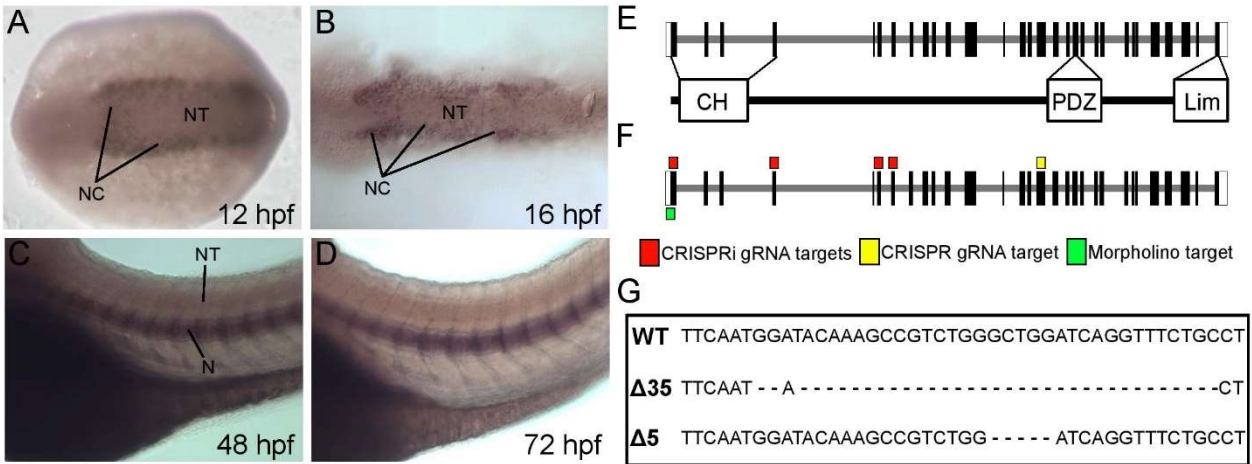
Figure 2.1: *lmo7a* **expression and domain structure. (A-B)** Whole mount in situ hybridization (ISH) for *lmo7a* mRNA (dorsal views, anterior to the left). Expression in premigratory NC cells at 12 hpf (**A**) and migratory NC cells at 16 hpf (**B**). (**C-D**) Whole mount ISH at 48 hpf (**C**) and 72 hpf (**D**) (lateral views), shows expression in the notochord and somite boundaries. **(E)** The *lmo7a* genomic locus consists of 33 exons. The full-length protein contains calponin homology (CH), PDZ, and Lim domains. **(F)** An antisense morpholino targeted the first exon, while CRISPR and CRISPRi gRNAs targeted exons 1,4,6, and 7. **(G)** Sequences for two CRISPR deletion alleles generated in Exon 16 compared to WT sequence.

Results

Lmo7a is expressed in and required for migration of subsets of NC cells

Zebrafish *lmo7a* was identified in a microarray screen of *tfap2a/g*-deficient embryos,

which lack NC cells (Hoffman et al., 2007; Tuttle et al., 2014). Whole mount in situ

hybridization (ISH) first detected *lmo7a* expression in cranial NC cells at 12 hours post-

fertilization (hpf), just prior to the onset of NC cell migration (**Fig 2.1 A**). Expression

persisted at 16 hpf in migrating NC cells in the PAs and between the eyes (**Fig 2.1 B**),

but was no longer detected at 24 hpf. At later embryonic stages (48-72 hpf), expression

was restricted to the notochord and somite boundaries (**Fig 2.1 C, D**).

Lmo7a contains CH, PDZ and Lim domains (**Fig 2.1 E**). Two deletions in *lmo7a*, -5 bp

and -35 bp, were generated by CRISPR-Cas9 gene editing using a guide RNA targeting

exon 16 just upstream of the PDZ domain (**Fig 2.1 F, G**). Trans-heterozygous mutants

carrying the *Tg(-4.9sox10:nEOS)* transgene (hereafter referred to as sox10:nEOS,

which labels the nuclei of pre-migratory and migrating NC cells) appeared largely

normal but NC cells formed small aggregates (5-10 cells/aggregate; ~10-20

cells/embryo) at the dorsal midline of the neural tube extending along the anterior-

posterior (A-P) axis from the midbrain-hindbrain boundary to the anterior spinal cord

(**Fig 2.2 A,B**). Such aggregates closely resemble the phenotypes of *rbc3a$^{-/-}$* and *ovol1a*-

deficient embryos (Piloto & Schilling, 2010; Tuttle et al., 2014).

Similar to *lmo7a* mutants, knockdown of *lmo7a* using an antisense morpholino targeting the translation start site (*lmo7a*-MO) (**Fig 2.1 H**) in sox10:nEOS fish resulted in NC aggregates (5-30 cells/aggregate; ~50-100 cells/embryo) at the dorsal midline in ~93% of injected embryos compared to sibling embryos injected with a control MO (**Fig 2.2 D-E,G-H**). Aggregates became distinct by 18 hpf while other surrounding NC cells migrated away from them and ventrally into the PAs at approximately the same rate as
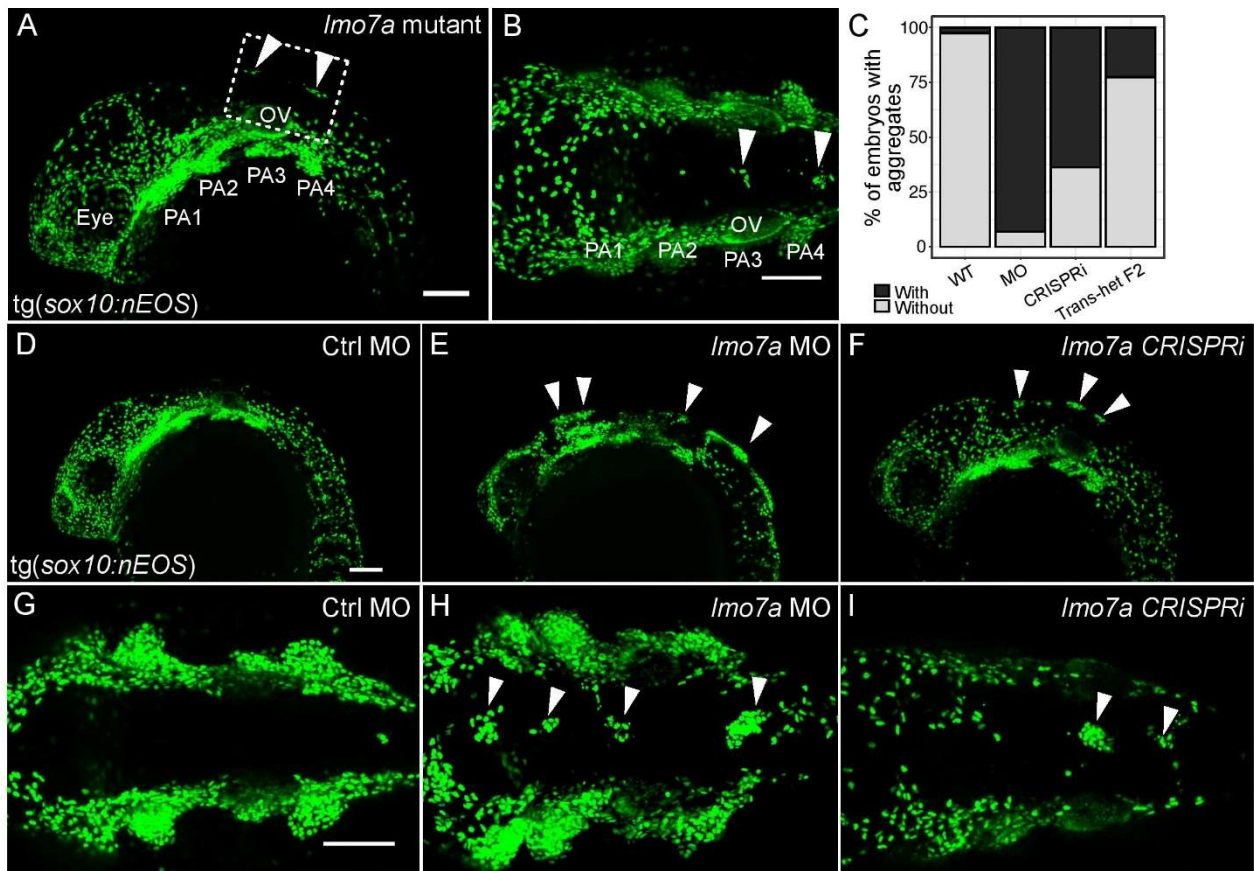
Figure 2.2: *lmo7a* knockdown disrupts migration of subsets of NC cells. (A-B, D-I) Whole mount live confocal images of 24 hpf *sox10:nEOS* embryos. **(A-B)** NC aggregates form along the dorsal midline in *lmo7a* trans-heterozygous mutants, in contrast to WT siblings. **(C)** Percentages of embryos displaying >10 NC cells at the dorsal midline in various *lmo7a* gene perturbations. **(D,G)** Embryos injected with 4 ng of control morpholino (MO), **(E,H)** Embryos injected with 4 ng of antisense MO targeting *lmo7a* **(F,I)** Embryos injected with *dCas9* mRNA and 4 gRNAs targeting the coding region of *lmo7a*. NC cell aggregates at the dorsal midline (arrowheads). **(A, D-F)** Lateral views. **(B, G-I)** Dorsal views. Scale bars = 100 µm **(A,B,D,G)**. PA=Pharyngeal Arch, OV=Otic Vesicle

WT cells. CRISPR inhibition (CRISPRi) produced similar NC cell aggregates in ~64% of injected embryos (**Fig 2.2F,I**). With CRISPRi, expression of *lmo7a* at 12 hpf was nearly undetectable (**Fig 2.3A**). These results provide independent confirmation that Lmo7a function is required for subsets of NC cells to migrate.
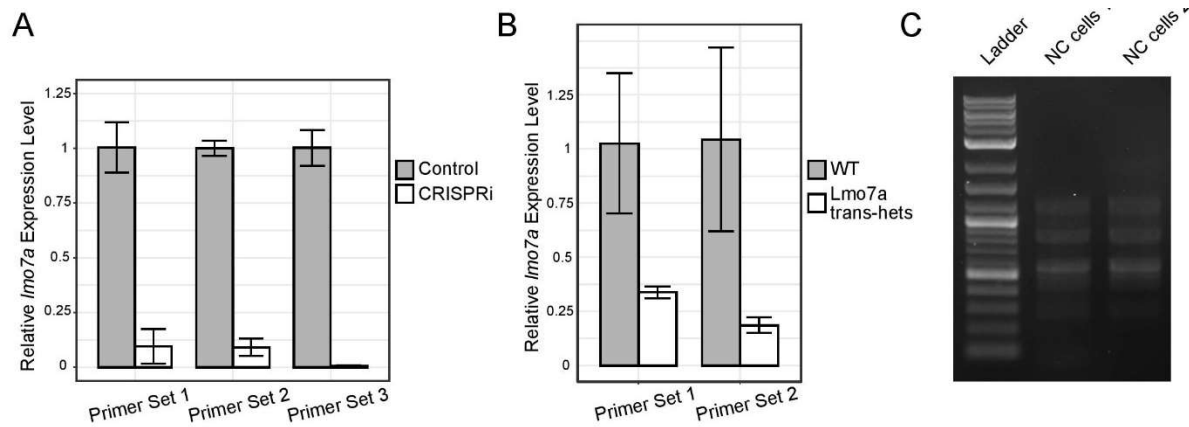
Figure 2.3: Relative expression of *lmo7a* in CRISPRi and mutant embryos. A) Bar plots showing expression levels of *lmo7a* in WT vs CRISPRi embryos at 12 hpf based on qPCR using 3 different primer sets targeting different segments of the coding region. **B)** Bar plots showing expression levels of *lmo7a* in WT vs transhet mutant embryos at 12 hpf based on qPCR using 2 primer sets. **C)** PCR of *lmo7a* in FACS sorted NC cells. Multiple transcripts are detected in pure NC populations. 2 NC cell preps were analyzed independently. For **A-B)** Bars indicate mean fold change of 3 biological replicates as compared to mean WT expression. Fold changes calculated by taking $2^{\wedge}\Delta\Delta$ CT for each replicate. Error bars indicate mean ±SD.

Lmo7a localizes to NC cell membranes and its function requires the calponin homology domain

Lmo7 was previously shown to function at the membrane in epithelial cells where it interacts with adherens junctions and/or focal adhesions (FAs), and in the nucleus in muscle cells, where it binds the transcription factor Emerin (Holaska et al., 2006; Ooshio et al., 2004; Wozniak et al., 2013). To determine the subcellular localization of Lmo7a in

A tg(*sox10:lyn-TdTom*)

B sfGFP-lmo7a

C Merge

D *lmo7a* MO

E MO + full mRNA

tg(*sox10:nEOS*)

F MO + no-Lim

G MO + no-CH

H

I

Full Transcript — CH —— PDZ — Lim

No Lim — CH —— PDZ —

No CH —— PDZ — Lim
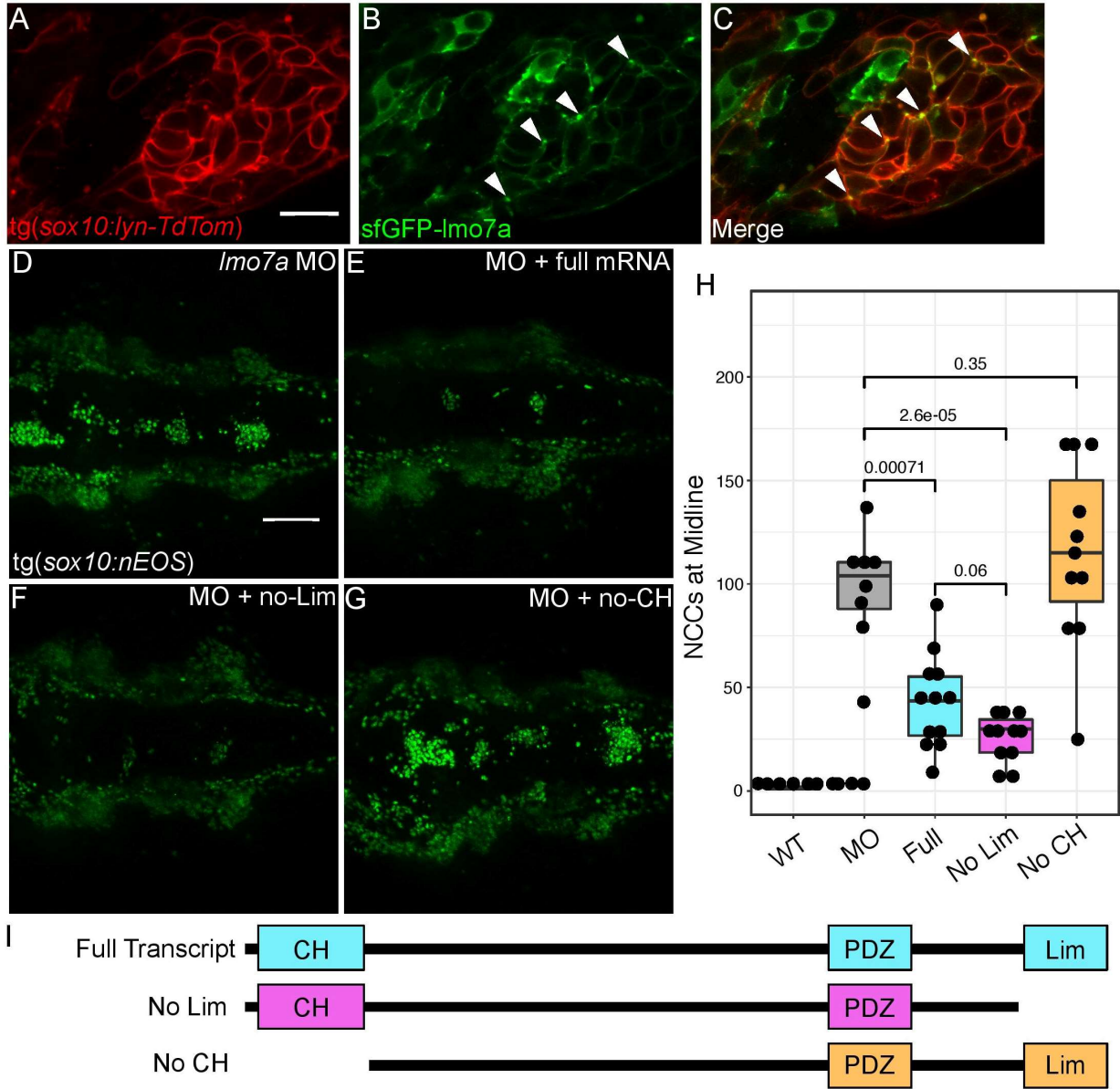
NC cells, we generated a fusion construct encoding superfolder GFP (sfGFP) fused to

the N-terminus of Lmo7a, *sfGFP-lmo7a*. This mRNA was injected at the 1-cell stage into

Tg(-4.9sox10:lyn-tdTomato) embryos (hereafter referred to as *sox10:lyn-tdTom*) to mark

NC cell membranes. At 18 hpf, the *sfGFP-lmo7a* fusion protein was restricted to bright

puncta that co-localized with *sox10:lyn-tdTom*, with little to no expression detected in

cell nuclei (**Fig 2.4 A-C**), indicating a potential role at the membrane.

Figure 2.4: Lmo7a localizes to the plasma membrane and requires the CH domain in NC. (A-C) Whole mount live confocal images of WT *sox10:lyn-TdTomato* embryos injected with *sfGFP-lmo7a* fusion mRNA. sfGFP-Lmo7a puncta at the plasma membrane of NC cells (arrowheads).(lateral views, anterior to the left). **(D-G)** Whole mount live confocal images of 24 hpf *sox10:nEOS* embryos (dorsal views, anterior to the left). **(D)** *lmo7a* MO-injected embryos display large NC aggregates. Both full length mRNA **(E)** and mRNA lacking the Lim domain **(F)** reduce the number of NC cells in aggregates, while mRNA lacking the CH domain **(G)** does not. **(H)** Quantification of phenotype severity based on median number of NC cells at midline/embryo (n = 8 embryos per condition). **(I)** Schematics of rescue transcripts. *Tg(sox10:lyn-tdTomato)* For **(H)**: Medians: MO = 104 cells/embryo, MO+full mRNA = 43.5 cells/embryo, MO+noLim mRNA = 30 cells/embryo, MO+noCH mRNA = 115 cells/embryo. Line indicates median, boxes indicate IQR, whiskers indicate IQR*1.5, point indicates outlier. Kruskal-Wallis ANOVA: p-value <0.0001. Posthoc Wilcoxon tests (BH corrected) p-values indicated on plot. Scale bars = 20 μm **(A)** and 100 μm **(D)**.

Lmo7a contains CH, PDZ and Lim domains, all implicated in protein-protein interactions (**Fig 2.1 G**). To test requirements for these domains in NC cell migration, rescue experiments were performed in an *lmo7a*-MO background. Co-injection of *lmo7a*-MO with mRNA encoding full-length *lmo7a* significantly reduced the severity of the migration phenotype as quantified by the number of NC cells that failed to migrate away from the midline by 24 hpf (MO median = 104 cells/embryo, MO+*lmo7a*-full median = 43.5 cells/embryo; p<0.001) (**Fig 2.4 D, E, H**). Next, mRNAs encoding full-length *lmo7a* lacking either the Lim or the CH domain were co-injected with *lmo7a*-MO. Interestingly, removal of the Lim domain caused no significant change in the ability of injected *lmo7a* mRNA to rescue the *lmo7a*-MO phenotype (MO+*lmo7a-noLim* median = 30 cells/embryo; p<0.0001) while *lmo7a* lacking the CH domain failed to rescue (MO+*lmo7a*-full median = 115 cells/embryo; p=0.35) (**Fig 2.4 F, G, H**). The CH domain mediates interactions with the actin cytoskeleton. These data suggest that the CH domain but not the Lim domain is required for *lmo7a's* function in early NC cell

migration, further supporting a role for Lmo7a at the membrane in NC cells and potential interactions with the cytoskeleton.

NC cells in *lmo7a*-deficient embryos have aberrant accumulation of focal adhesion components

Based on its membrane localization in NC cells and known roles in FAs (Holaska et al., 2006; Wozniak et al., 2013), we hypothesized that Lmo7a likely plays a role in FAs essential for proper filopodial dynamics and migration. To test this idea we next examined filopodial extension and FA formation. We utilized a double transgenic line expressing both a fluorescent paxillin (Pxn) fusion protein under the control of a ubiquitous promoter, *Tg(β-actin:Pxn-EGFP)* as well as *sox10:lynTdTomato* to facilitate live, time-lapse imaging of Pxn-based adhesion complexes in NC cells. Embryos were imaged from 16-19 hpf, when cranial NC cells migrate into the PAs. In WT embryos, NC cells dynamically change morphology as they migrate, extending and retracting filopodia. Transient accumulations of Pxn-EGFP were visible along the membranes of these NC cells during migration (**Fig 2.5A-C**). In contrast, in *lmo7a*-deficient embryos, NC cells in aggregates that remained at the midline maintained a rounded morphology with few to no filopodial projections and contained large aggregates of Pxn-EGFP that appeared cytoplasmic (**Fig 2.5D-F**). To investigate if Lmo7a colocalizes with Pxn, we injected *mCherry-Lmo7a* mRNA into WT *tg(B-actin:Pxn-EGFP)* embryos. Overlap of mCherry-Lmo7a puncta and Pxn-EGFP puncta was observed at the membranes of migratory WT NC cells (**Fig 2.5G-I**). These observations suggest that in *lmo7a*-deficient

embryos, Pxn-FA complexes fail to form properly at the membrane and that NC cells consequently fail to extend directional projections that facilitate migration.
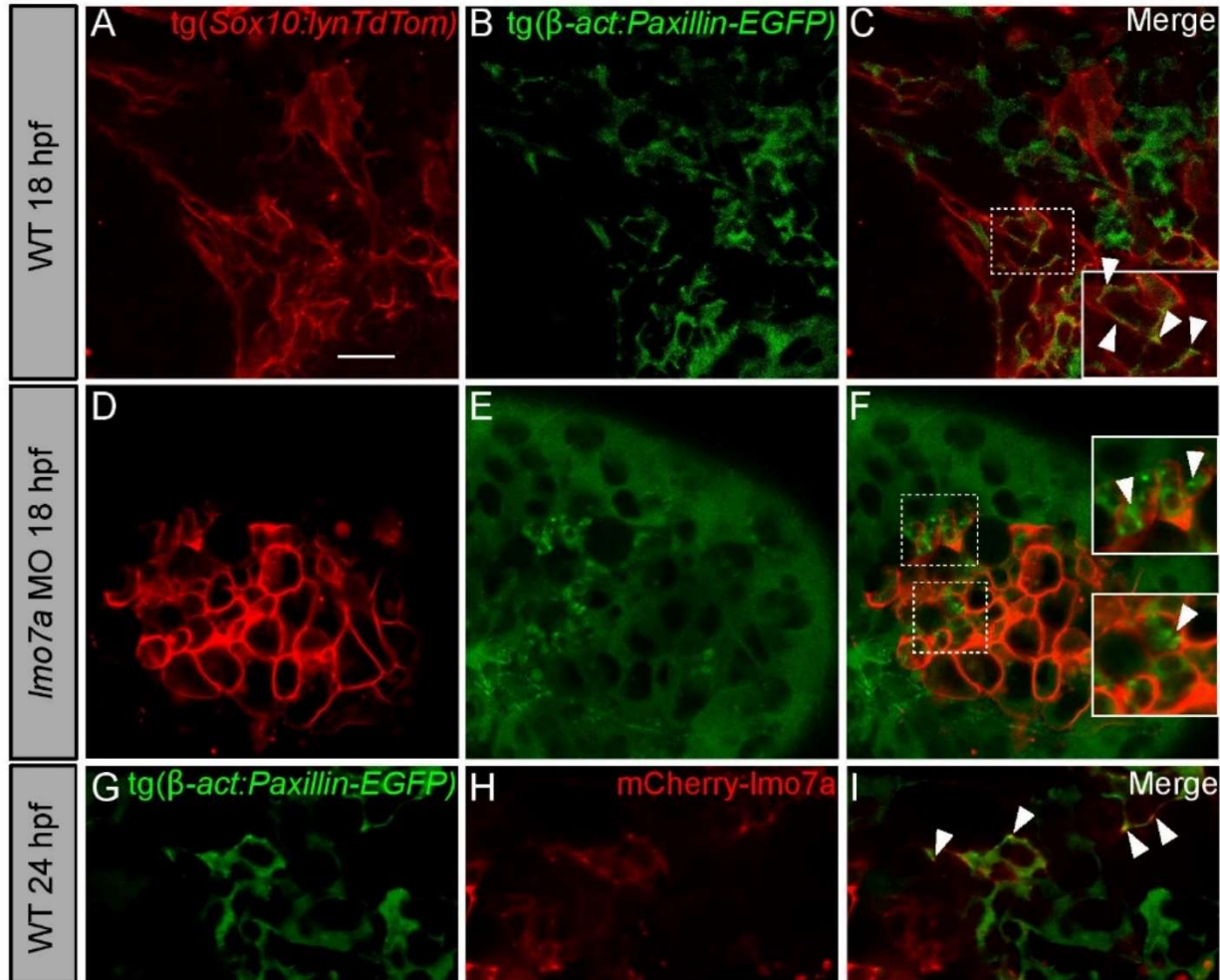
Figure 2.5: *lmo7a* **knockdown results in abnormal NC cell morphology and aggregation of paxillin complexes. (A-F)** Whole mount live confocal images of *Tg(sox10:lyn-tdTomato; β-actin:paxillin-EGFP)* double transgenic embryos. **(A-C)** NC cells at 18 hpf display long cytoplasmic protrusions (filopodia) and localization of Pxn-EGFP along the plasma membrane (arrowheads) in WT. **(D-F)** NC cells in dorsal midline aggregates in embryos injected with *lmo7a*-MO. Cells are rounded and accumulate Pxn-EGFP in the cytoplasm. **(G-I)** Whole mount live confocal images of transgenic *β-actin:paxillin-EGFP* embryos injected with *mCherry-lmo7a* mRNA. mCherry-Lmo7a puncta colocalize with Paxillin-EGFP in WT NC cells at 24 hpf **(I)** (arrowheads). Scale bar = 15 µm

To investigate if this Pxn accumulation is indicative of changes in FA dynamics, we examined localization of phosphorylated Focal Adhesion Kinase (pFAK) using a polyclonal anti-pFAK (pY576) antibody. At 18 hpf, many bright puncta can be seen along the membranes of migrating NC cells in embryos injected with a control MO (Ctrl MO) (**Fig 2.6 A-D**). While some puncta can be seen in the membranes of midline aggregate NCs in *lmo7a*-MO-injected embryos (**Fig 2.6 E-H**), the number of puncta per cell was significantly lower than controls (Ctrl MO mean = 13.1/NC cell/ n = 7 embryos,
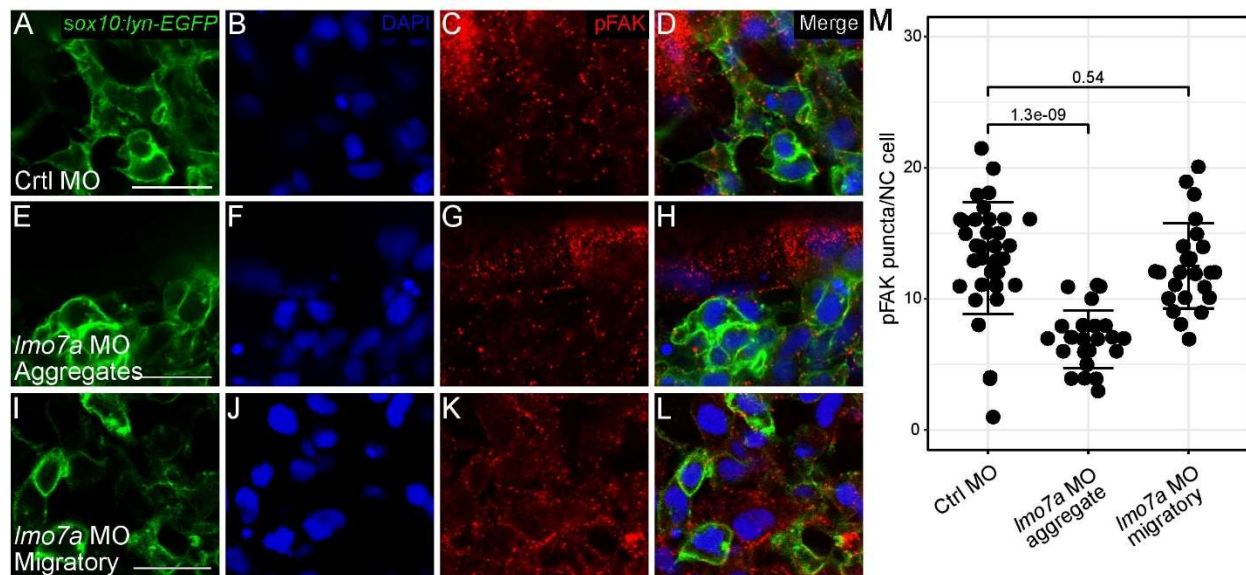
Figure 2.6: lmo7a-deficient NC aggregates have decreased levels of phosphorylated focal adhesion kinase. (A-H) Whole mount confocal images of immunostaining with anti-pFAK antibody (red) and DAPI (blue) in sox10:lynEGFP transgenic embryos (green). **(A-D)** NC cells in Ctrl MO-injected embryos show many bright pFAK puncta along their membranes. **(E-H)** Aggregate NC cells in lmo7a MO-injected embryos show fewer pFAK puncta in their membranes. **(I-L)** Migratory NC cells in lmo7a MO-injected embryos show unchanged pFAK localization compared to Ctrl embryos. **(M)** Quantification of pFAK puncta per NC cell in Ctrl MO and lmo7a MOinjected embryos. Dots represent the number of puncta in individual cells (Ctrl MO n=7 embryos, mean=13.1 puncta/cell; lmo7a MO n=5 embryos, mean=6.2 puncta/aggregate NCcell, mean = 12.5 puncta/migratory NC cell). Ctrl vs lmo7a MO aggregate T-test p-value = 1.3e-9, Ctrl vs migratory lmo7a MO NC cells = 0.54. Line indicates mean. Error bars indicate ±SD. Scale bars = 20 μm.

lmo7a MO non migrating mean = 6.92/NC cell/ n = 5 embryos, p=1.3e-9) (**Fig 2.6 M**). Notably,

in *lmo7a*-deficient embryos, non-NC accumulation of pFAK appeared unaltered as did pFAK

localization in migrating NC cells (mean = 12.5/NC cell, p=0.54) (**Figure 2.6 I-M**).

**NC cell transplantations reveal non-cell-autonomous effects of *lmo7a*-deficiency
on cell migration independent of EMT**

NC cells in *lmo7a*-deficient embryos form large aggregates, with single NC cells rarely

remaining at the midline. This could suggest effects on some non-cell-autonomous factors such

as EMT or leader-follower dynamics. We first investigated influences on EMT through
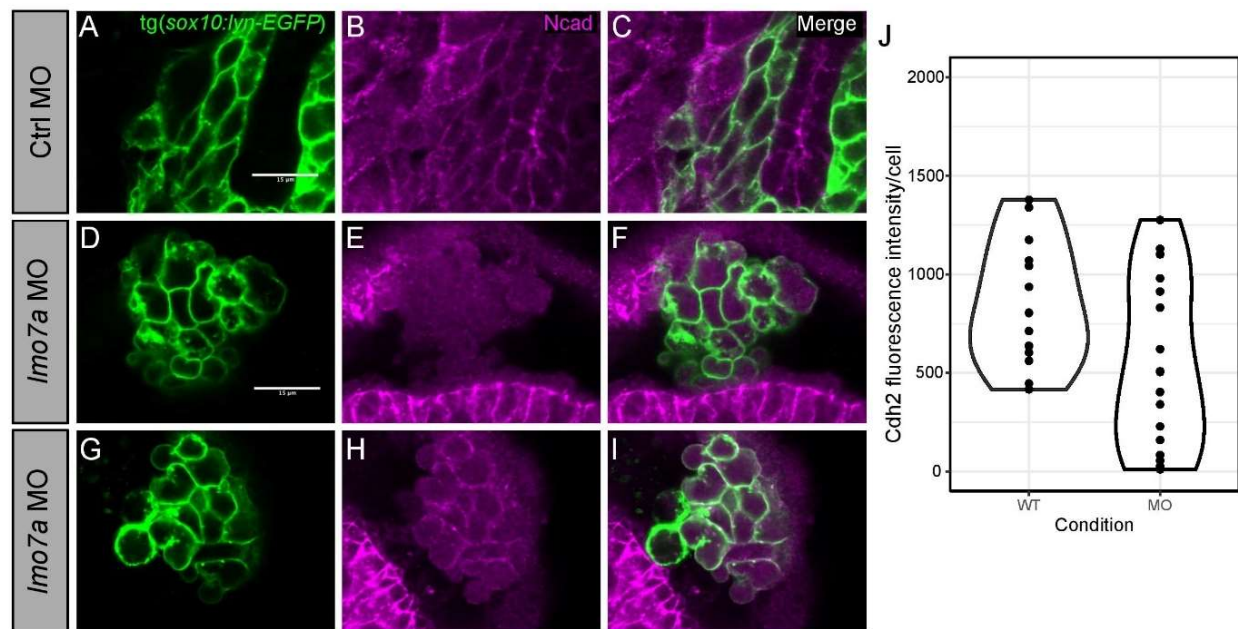
Figure 2.7: N-cadherin expression is not reduced consistent with loss of migration. (A-I) Whole mount confocal images of immunostaining with anti-N-cad antibody (magenta) in sox10:lynEGFP transgenic embryos (green). (A-C) NC cells in Ctrl MO injected embryos show N-cad staining along their membranes. (D-F) Some NC cell aggregates in lmo7a MO-injected embryos show reduced N-cad staining, while (G-I) other aggregates have normal N-cad membrane staining. (I) There is a bimodal distribution of NC aggregate cells that have normal levels and those that have reduced levels. Dots represent the N-cad fluorescence intensity in individual cells calculated as membrane intensity - (membrane area x background intensity). (Ctrl MO n=4 embryos, 4 cells/embryo, lmo7a MO n=4 embryos, 4 cells/embryo). Scale bars = 15 μm.

examination of N-cadherin expression. In *lmo7a*-MO-injected embryos, some aggregates showed reduced N-cad staining at the plasma membrane as compared to migrating NC cells in WT embryos (**Fig 2.7 A-F**). However, other aggregates displayed normal N-cad localization (**Fig 2.7 G-I**). The lack of any correlation between N-cad localization and the formation of aggregates suggests the defects seen in *lmo7a*-deficient embryos are not the result of a change in EMT (**Fig 2.7 J**).

Next, we investigated the possibility that the loss of migration in a small subset of NC cells could have non-cell-autonomous effects on surrounding NC cells. We utilized cell transplantation to examine how WT NC cells behave in an *lmo7a*-deficient embryo. At the embryonic shield stage, cells at the position known to give rise to NC were transplanted from WT *Tg(Sox10:nEOS)* embryos into stage-matched control and MO-injected *Tg(Sox10:lynTdTomato)* embryos. By 24 hpf, a large number of WT donor (nEOS+) cells could be seen in the pharyngeal arches of control embryos (**Fig 2.8 A-C, G**) with only ~2.4% of donor cells and ~0.2% of host cells remaining at the midline (**Fig 2.8 G**). In MO-injected embryos, both WT donor cells and *lmo7a*-deficient host cells remained at the dorsal midline in aggregates (**Fig 2.8 D-G**) (Donor cells = ~22%, p = 0.0032; Host cells = ~40.5%, p = 0.0014).
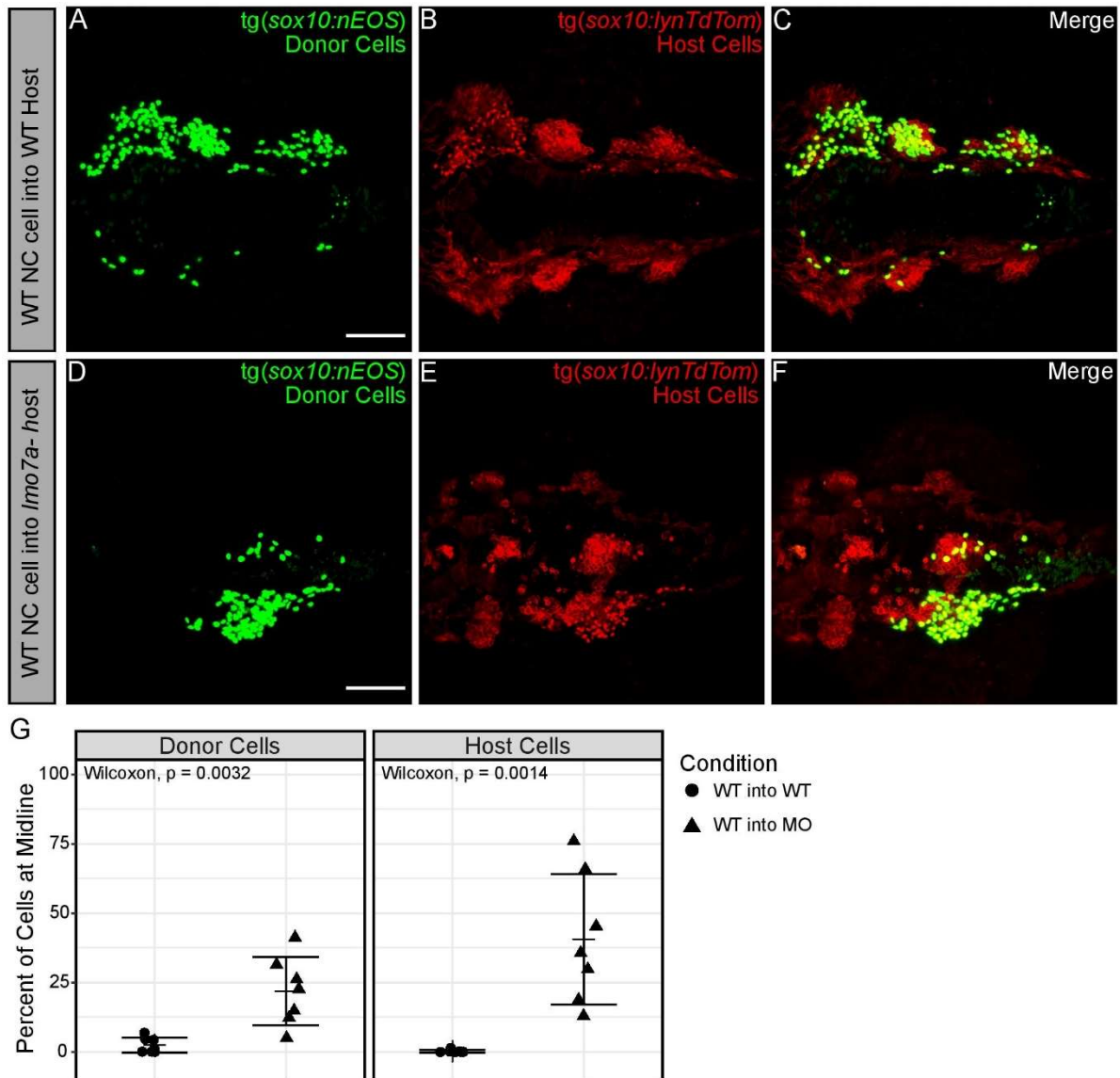
Figure 2.8: Transplantation of WT NC cells into Lmo7a-deficient embryos reveals non-cell-autonomous effects of Lmo7a loss. A-C) Micrographs showing transplantation of WT NC cells into WT host embryos. Donor embryos express *sox10:nEOS* and host embryos express *sox10:lyntdTomato*. **D-F)** Micrographs showing transplantation of WT NC cells into *lmo7a*-MO host embryos. Donor embryos express *sox10:nEOS* and host embryos express *sox10:lyntdTomato*. **G)** Dotplots showing percentage of cells remaining at the midline in transplant experiments. Dots represent WT cells into WT embryos, triangles represent WT cells into MO embryos. Left plot shows percentagages of donor cells at the midline in each experiment, right plot shows percentages of host cells at the midline. (WT into WT n=7 embryos, Donor mean=, Host mean= ; WT into MO n=7 embryos, Donor mean= , Host mean=). Donor cells wilcox p-value=0.0032, Host cells wilcox p-value=0.0014. Lines indicates means. Error bars indicate mean ±SD. Scale bars = 50 µm.
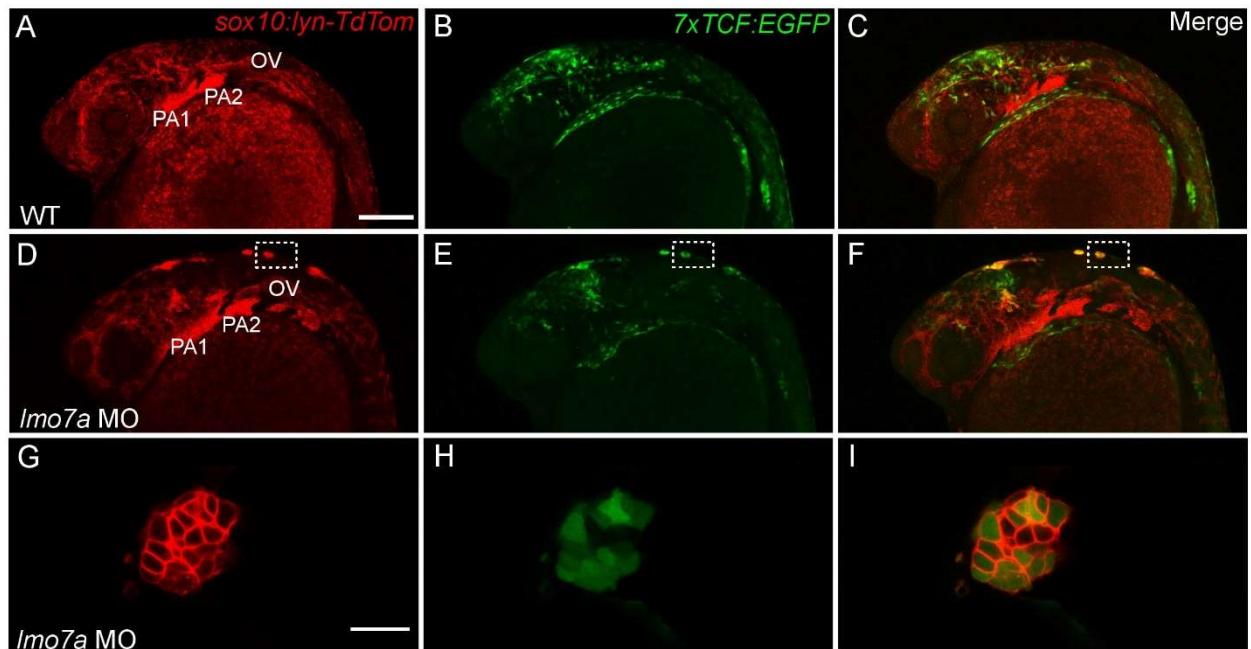
Figure 2.9: *lmo7a* knockdown increases canonical Wnt signaling in NC midline aggregates. Whole mount live confocal images of double transgenic *Tg(sox10:lyn-tdTomato; 7XTCF:EGFP)* embryos at 24 hpf. **(A-C)** In WT embryos, EGFP is detected in migratory NC cells around the midbrain, anterior hindbrain and PAs 1 and 2. **(D-F)** In *lmo7a* MO injected embryos EGFP is detected in NC aggregates at the dorsal midline regardless of anterior-posterior position. **(G-I)** At higher magnification, most cells within each aggregate are positive for EGFP. **(A-F)** Lateral views. **(G-I)** Dorsal view. Scale bars = 150 μm **(A-C)** 25 μm **(G-I)** and 15 μm **(J-N)** . PA=Pharyngeal Arch, OV=Otic Vesicle

Lmo7a deficiency elevates canonical Wnt signaling in NC cell aggregates

We have previously shown that Ovol1a and Rbc3a/Dmxl2 regulate NC cell migration, at least in part, by regulating canonical Wnt signaling (Piloto & Schilling, 2010; Tuttle et al., 2014). To investigate if the NC migratory defects observed in *lmo7a*-deficient embryos also alter Wnt signaling, we examined a canonical Wnt reporter line, *Tg(7xTCF:EGFP)*.
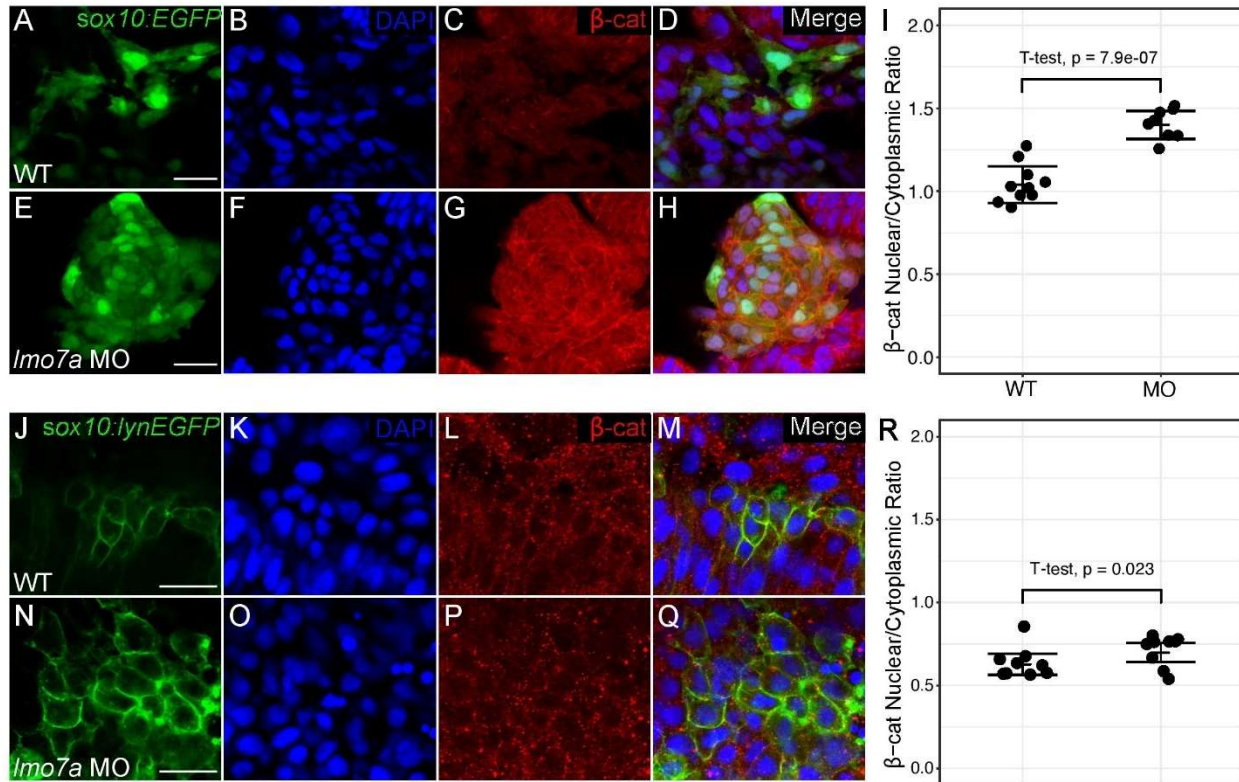
Figure 2.10: *lmo7a*-deficient NC aggregates have increased nuclear localization of β-catenin. (A-H) Whole mount confocal images of immunostaining with anti-β-cat antibody (red) and DAPI (blue) in 24 hpf *sox10:EGFP* transgenic embryos (green). **(A-D)** In migratory NC cells in WT embryos, β-cat staining is both in the nucleus and cytoplasm. **(E-H)** Aggregate NC cells in *lmo7a* MO-injected embryos display increased nuclear β-cat relative to cytoplasmic staining. **(I)** Nuclear β-cat levels for each cell were quantified as the mean fluorescence intensity in the nucleus divided by the mean fluorescence intensity in the cytoplasm in 10 individual cells per embryo. (WT n=10 embryos, mean=1.04; *lmo7a* MO n=8 embryos, mean=1.40). T-test p-value=7.913e-07. Line indicates mean. Error bars indicate ±SD. **(J-Q)** Whole mount confocal images of immunostaining with anti-β-cat antibody (red) and DAPI (blue) in 12 hpf *sox10:lynEGFP* transgenic embryos (green). Premigratory NC cells in WT embryos **(J-M)** and in *lmo7a* Mo-injected embryos **(N-Q)** both display low levels of β-cat staining in the nucleus. **(R).** (WT n=9 embryos, mean=0.63; *lmo7a* MO n=9 embryos, mean=0.70). T-test p-value=0.023. Line indicates mean. Error bars indicate ±SD. Scale bars = 20 μm.

In WT embryos, cranial NC cells in PA1 and PA2 as well as over the midbrain expressed the Wnt reporter at 24 hpf (**Fig 2.9 A-C**). In contrast, most of the NC cells within aggregates in *lmo7a*-deficient embryos were TCF-GFP-positive, regardless of their

A-P position (**Fig 2.9 D-F**). Levels of Wnt responses varied dramatically between cells within an aggregate (**Fig 2.9 G-I**).

To confirm the apparent increase in canonical Wnt signaling, we analyzed the subcellular localization of β-cat in WT and *lmo7a*-deficient NC cells at 24 hpf (**Fig 2.10 A-H**). Similar to our previous results for Ovol1a and Rbc3a, in *lmo7a*-deficient embryos, NC cells in the midline aggregates displayed increased levels of β-cat in the nucleus as compared to WT NC cells (**Fig 2.10 I**).

lmo7a-deficient NC cell aggregates co-express pigment and glia markers

Increased canonical Wnt signaling can drive NC cells toward pigment cell fates (Curran et al., 2010). To determine if this is the case in NC aggregates in *lmo7a*-deficient embryos, we performed in situ Hybridization Chain Reaction (HCR) for genes that mark different NC lineages at 24 hpf (**Fig 2.11**). In *lmo7a*-MO injected embryos, expression of the skeletogenic marker *dlx2a* was restricted to the PAs, similar to WT, and was excluded from midline NC aggregates (**Fig 2.11 B, F**). In contrast, all of these aggregates expressed *mitfa*, which labels melanocyte progenitors (**Fig 2.11 C, G**). These results suggest that elevated canonical Wnt signaling in NC cells that aggregate at the midline in *lmo7a*-deficient embryos promotes their differentiation as pigment cells, similar to our previous results for Ovol1a and Rbc3a/ Dmxl2.
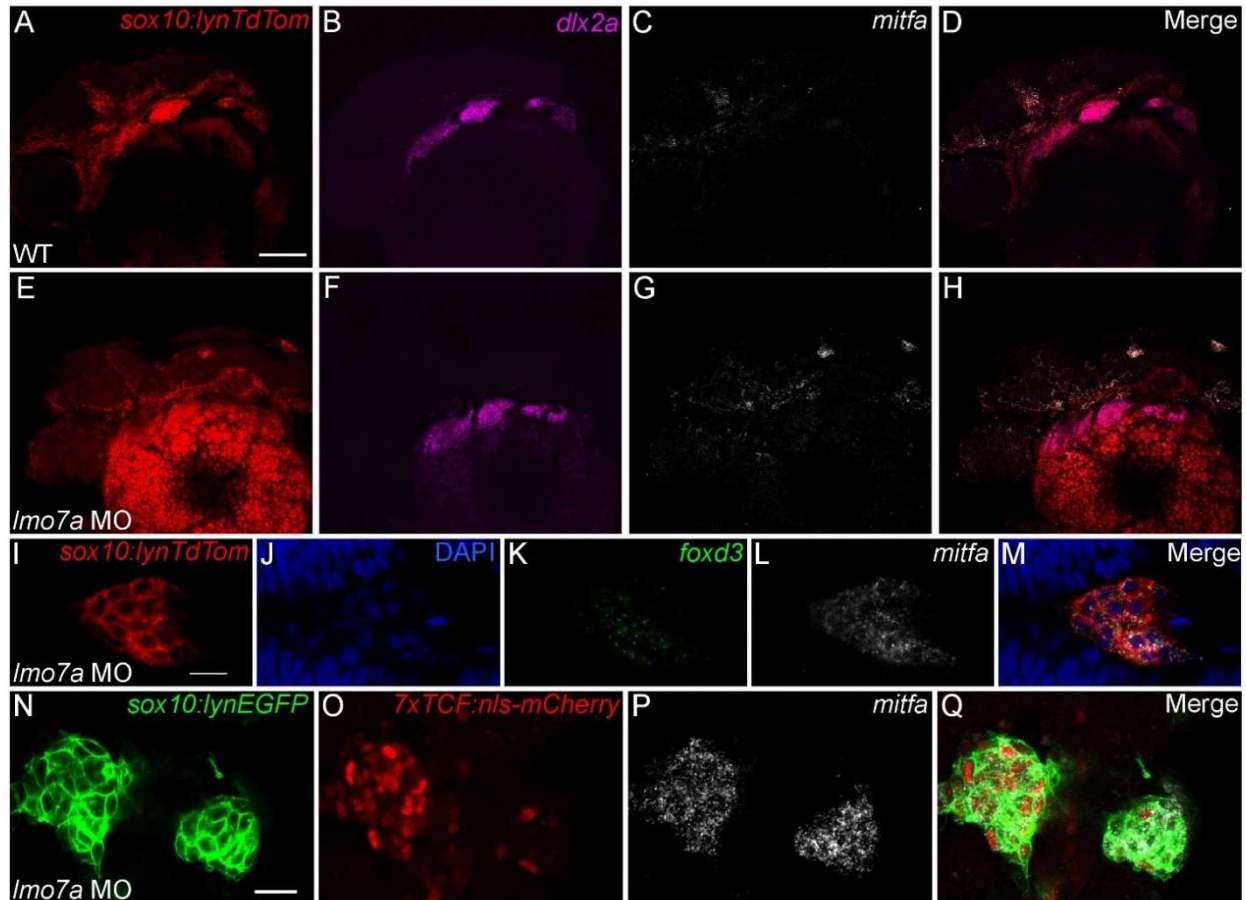
Figure 2.11: NC aggregates adopt a bipotential pigment/glial cell fate. (A-H) Whole mount confocal images of in situ hybridization chain reaction (HCR) for NC lineage markers in transgenic *Tg(sox10:lyn-tdTomato)* embryos. **(A-D)** In WT embryos, *dlx2a* is expressed in NC cells within the PAs and *mitfa* expression is observed in many NC cells outside the PAs. **(E-H)** In *lmo7a* MO-injected embryos, *dlx2a* expression is unaffected, but *mitfa* expression marks NC aggregates at the dorsal midline. **(I-M)** Whole mount confocal images of HCR for NC lineage markers. **(I)** NC aggregates in *lmo7a* MO-injected embryos express both the glial progenitor marker *foxd3* **(J)** and the melanocyte progenitor marker *mitfa* **(L),** and some cells co-express both markers (M). **(N-Q)** Whole mount confocal images of HCR for NC lineage markers in *lmo7a* MO-injected double transgenic *Tg(sox10:lyn-EGFP; 7XTCF:nls-mCherry)*. Both mCherry+ and mCherry- NC cells in midline aggregates express *mitfa*. Scale bars = 100 μm **(A)** 15 μm **(I)** and 20 μm **(N)**.

40

However, surprisingly, NC aggregates in *lmo7a*-deficient embryos were also positive for *foxd3* mRNA, which at this stage in WT marks Schwann cell precursors (**Fig 2.11 I-M**). Furthermore, many NC cells within an aggregate clearly co-expressed both *mitfa* and *foxd3*. Wnt signaling promotes Mitf and represses Foxd3 in the context of lineage specification (Curran et al., 2010). Therefore, to assess the Wnt responses occurring in these apparent bipotential *mitfa/foxd3* double-positive NC cells we examined their expression of *7xTCF:nls-mCherry.* Although expression is mosaic in the NC aggregates of *lmo7a*-deficient embryos (**Fig 2.11 N-O**), all mCherry- cells also expressed *mitfa* (**Fig 2.11 P**).

Bulk RNA-seq of NC cell aggregates reveals unique expression signature distinct from that of *ovol1a*-deficient aggregates.

L*mo7a*-depleted NC cells fail to migrate, increase their Wnt signaling, and express both pigment and glial lineage markers, but how are these phenotypes connected? And how do these aggregates differ in their gene expression profiles from those found in *ovol1a*-deficient embryos? To address these questions, we analyzed their transcriptomic profiles using RNA-seq. We took advantage of the photoconvertible nature of the EOS protein in the *sox10:nEOS* transgenic line in order to isolate select groups of NC cells for transcriptomic analysis. Briefly, NC aggregates in *lmo7a*-MO and *ovol1a*-MO injected embryos were photoconverted from green to red with a 405nm laser at 24 hpf, converted and unconverted NC cells were subsequently isolated via fluorescence activated cell sorting (FACS), and total RNA was extracted and used to construct cDNA libraries for RNA-sequencing (RNA-seq) (**Fig 2.12 A-B**).

Differential gene expression analyses were performed comparing both the entire NC

population and isolated NC aggregates of *lmo7a*- and *ovol1a*-deficient embryos with NC

cells from stage matched WT embryos. In *lmo7a* morphants, >1000 genes differed

significantly from WT and in *ovol1a* morphants >500 genes differed from WT (FDR

<0.05). The top 500 genes in each condition were clustered based on their expression

profiles (**Fig 2.12 B**). Several genes involved in NC migration had decreased expression

in the NC cell aggregates including *ednrab*, *cxcr4b*, and the Wnt response modulator

draxin. Aggregates also showed increased expression of multiple pigment specifiers

including *mitfa, gch2*, and *pax7a* as well as the glial markers *foxd3* and *sox10*, further

supporting the notion that these cells adopt a glial/pigment identity. In addition to these

known NC regulators, several novel genes had altered expression. Expression of the

matrix metalloproteinase *mmp9* was severely reduced while *mmp2* was upregulated.

While neither of these genes have been implicated in NC development, *mmp17b* plays
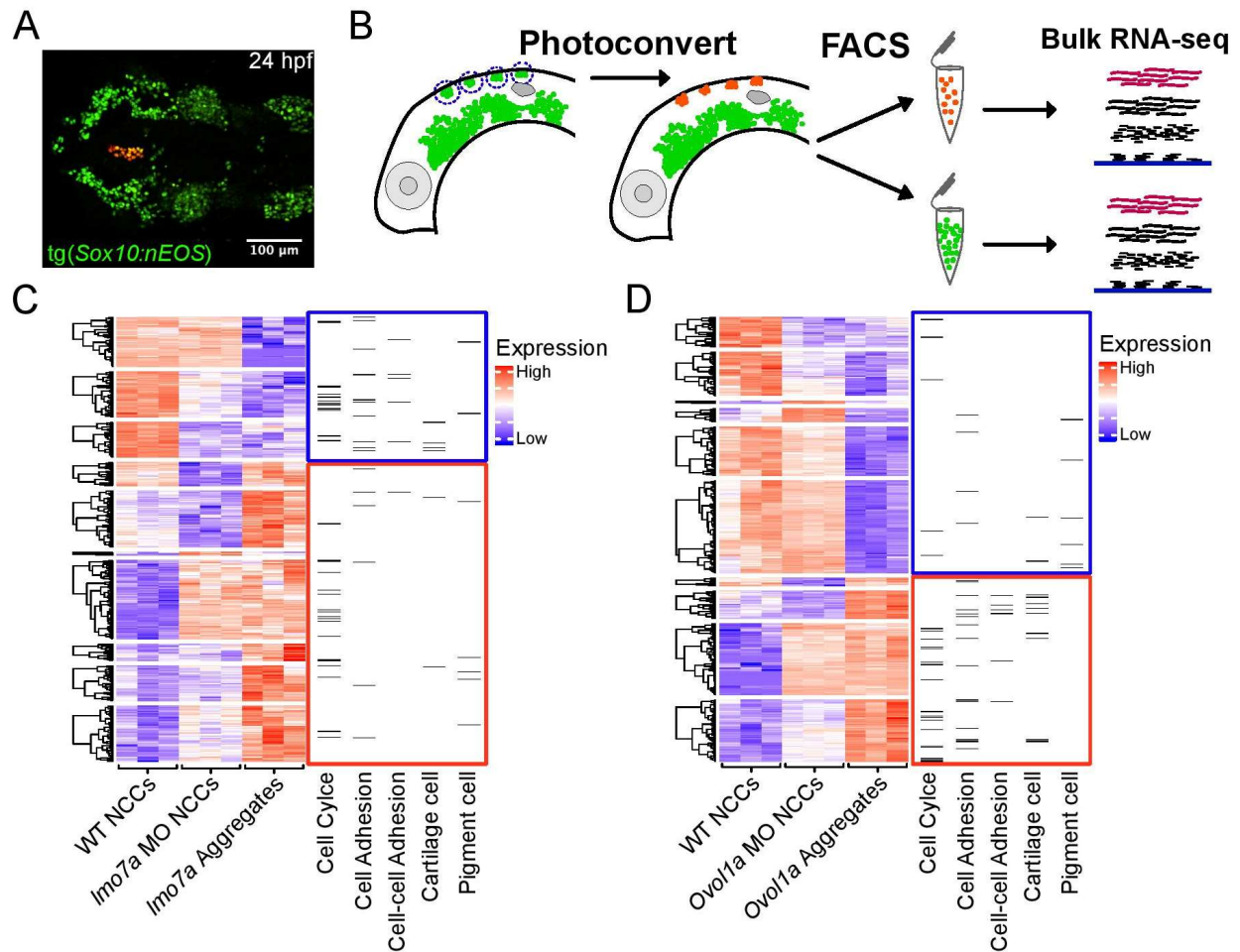
Figure 2.12: Bulk RNA-seq of lmo7a and ovol1a morphant aggregates reveals unique transcriptional signatures. A) Live fluorescent image showing photoconverted NC cell aggregate in morphant embryo. B) Diagram showing experimental design. Photoconverted aggregates and non-converted NC cells in both lmo7a and ovol1a morphant embryos were isolated via FACS and then each population was combined for bulk RNA-sequencing. C-D) Heatmaps showing the top 500 DE genes in each morphant condition compared to WT NC cells. Side annotation bars show which genes are associated with GO terms for cell cycle, cell adhesion, cell-cell adhesion, cartilage development, and pigment development. lmo7a MO aggregates show downregulation of many cell adhesion genes and cartilage development genes, whereas ovol1a MO aggregates show upregulation of many cell adhesion genes and cartilage development genes and downregulation of pigment development genes.

an important role in promoting NC migration (Leigh et al, 2013). Aggregates also misexpress a number of NC cadherins including decreased *cdh11* and increased *cdh1* and *pcdh10a.* Finally, *lmo7a*-deficient aggregates express higher levels of the Rho-

GTPase *rhoub*, which is known to have NC expression (Thisse et al, 2004), but has no previously described role in NC development. Interestingly, *ovol1a*-deficient aggregates showed increased expression of cartilage development and cell adhesion genes (**Fig 2.12 C**). This strikingly different expression profile is surprising, considering the migratory defect so closely resembles that seen in *lmo7a*-deficient embryos.

Discussion

NC cells integrate many signals as they migrate to arrive at their final positions and generate the correct cell types. The gene regulatory network for NC specification and differentiation has been well studied, but the relationship between NC migration and fate is still poorly understood (Kalcheim & Kumar, 2017). Here we show novel roles for Lmo7a in regulating both NC migration and lineage specification through its functions in modulating adhesion and canonical Wnt signaling. This is the first evidence for a role for Lmo7 in NC development, and given the complex and modular nature of its function in other contexts, we can hypothesize many potential mechanisms. In our model (**Fig 2.13**), Lmo7a serves to direct formation of FAs in a similar manner to its previously described function in cultured cells (Holaska et al., 2006; Wozniak et al., 2013). Loss of Lmo7a would then interfere with formation of the Pxn-FA complex at the membrane, thereby disrupting the ability of NC cells to migrate in a coordinated manner. This is consistent with its highly restricted expression to migrating NC cells as well as its subcellular localization to the membrane where it regulates Pxn localization. We hypothesize that this loss of proper Pxn-FA formation may disrupt a complex similar to the recently described NC Cadherin11-FA complex, which includes Pxn and β-cat (Langhe et al., 2016), thereby increasing the amount of free β-cat in the cell and consequently elevating canonical Wnt signaling (**Fig 2.13**). This would be consistent

with previously described roles for Lmo7 in regulating FA formation (Wozniak et al 2013) and in linking cadherins to afadin-nectin junctions (Ooshio et al., 2004).

Lmo7a regulates NC cell migration and FA localization

Lmo7 has been implicated in muscle and epithelial development as well as cancer
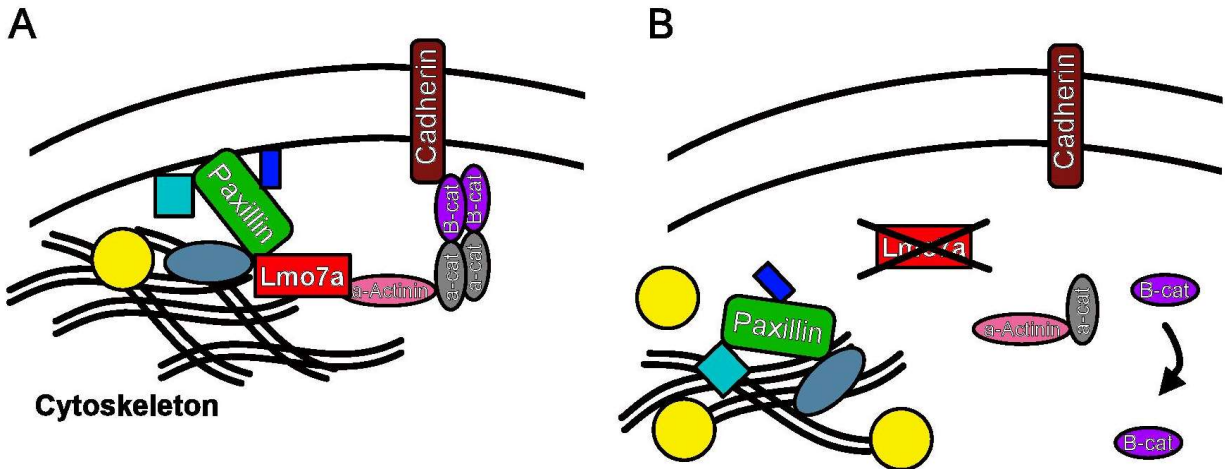
Figure 2.13. Hypothetical model for Lmo7a functions in NC cells (A) Lmo7a promotes proper assembly of the Pxn-FA complex in coordination with a Cadherin-β-Cat complex similar to a classical adherens junction. Both associate with elements of the actin cytoskeleton. (B) In the absence of Lmo7a, the Pxn-FA complex fails to localize correctly and accumulates in the cytoplasm. This also disrupts the Cadherin-β-Cat complex, resulting in loss of directional migration and increased free β-Cat levels.

metastasis. The mechanism of its function and its subcellular localization differ in each of these contexts. In muscle, Lmo7 interacts with transcriptional regulators in the nucleus, while in normal epithelia and breast cancer cells it interacts with transmembrane signaling molecules and the actin cytoskeleton (Dedeic et al., 2011; Holaska et al., 2006; Hu et al., 2011; Nakamura et al., 2005; Ooshio et al., 2004; Teixeira et al., 2014). A large number of alternatively spliced Lmo7 transcripts have been identified, which further increases the range of this multifaceted protein's potential functions.

We found that an sfGFP-Lmo7a fusion protein localized to the plasma membrane of zebrafish NC cells, with no detectable expression in the nucleus. In addition, injection of mRNA encoding an Lmo7a protein lacking the Lim domain rescued NC migration at similar levels to full length mRNA, while mRNA lacking the CH domain did not. CH domains are well known to mediate interactions with the cytoskeleton through binding with F-actin (Korenbaum & Rivero, 2002). Taken together these results strongly suggest a role for Lmo7a at the membrane, perhaps similar to its role in connecting the cytoskeleton to adherens junctions in epithelial cells (Ooshio et al., 2004).

In breast cancer cells, Lmo7 functions to modify the cytoskeleton and promote transcription of metastatic genes. Consistent with a similar structural role for Lmo7a in NC cells, we found mislocalization and aggregation of Pxn, an essential scaffolding protein for FA formation. Directional FAs have been observed in NC cells cultured in vitro (Toro-Tapia et al., 2018), and proper FA-Integrin (Itg) signaling is crucial for directed NC migration and cardiac outflow tract formation in vivo (Dai et al., 2013). LIMCH1, a related protein with a very similar domain structure as Lmo7, controls cell migration by regulating FA formation and actomyosin dynamics (Lin et al., 2017). Lmo7 itself has been shown to localize to FAs and act as a shuttling protein to mediate Itg signaling in HeLa cells and mouse embryonic fibroblasts (Wozniak et al., 2013). The essential nature of its CH domain in NC migration supports this notion, as CH domains have been found to bind Pxn (Sjöblom et al., 2008). Our results suggest a role for Lmo7a in promoting NC cell migration through regulation of FA dynamics. This

requirement appears to be specific to subsets of NC cells. This finding is similar to loss of Fscn1-dependant filopodia in premigratory NC cells, which leads to a loss of ventrolateral migration of subsets of NC cells (Boer et al., 2015). The selective loss of NC migration in *lmo7a*-deficient embryos reinforces the notion that requirements for specific migratory regulators are heterogeneous in the NC and possibly lineage-specific. This is an exciting avenue for research into the in vivo relationship between FAs and NC migration and lineage decisions.

NC aggregates elevate Wnt signaling secondary to a loss of migratory capacity

After NC migration at 24 hpf, unmigrated NC aggregates in *lmo7a*-deficient embryos display high levels of Wnt signaling as measured using both the transgenic Wnt reporter, *Tg(7xTCF:EGFP)* and nuclear β-cat protein. This may have important consequences both for the migratory behaviors as well as the fates of the cells. By this stage most wild-type NC cells as well as migrated NC cells in *lmo7a*-deficient embryos have significantly downregulated Wnt signaling. In addition to its roles in NC induction and migration, Wnt signaling also regulates the fate decision between pigment and glial lineages. Bipotential pigment/glial progenitors downregulate *foxd3* and upregulate *mitfa* in response to elevated Wnt, biasing them toward a pigment cell fate (Curran et al., 2010). Surprisingly, despite high levels of Wnt signaling, the NC aggregates in *lmo7a*-deficient embryos express both *foxd3* and *mitfa*, often in the same cells. This is in contrast to previous findings showing that when either *rbc3a/dmxl2* or *ovol1a* are eliminated, NC cells aggregate at the dorsal midline, upregulate responses to Wnt, and

all become pigment cells. Comparative analysis with other known Wnt-regulatory

mutants may help to disentangle the effects on migration and lineage specification.

How these functions of Lmo7a in adhesion lead to specific effects on distinct NC

derivatives is less clear. Either Lmo7a functions to regulate the pigment/glial fate

decision directly, or it may affect migration of a subset of NC cells that is distinct from

the populations affected by *rbc3a/dmxl2* and *ovol1a*. NC aggregates in lmo7a-deficient

embryos appear strikingly similar to loss of *rbc3a* or *ovol1*, yet distinct in that they

contain both pigment and glial progenitors (Piloto & Schilling, 2010; Tuttle et al., 2014).

In the case of *ovol1a*, the migration defect observed in NC cells showed a clearly

correlates with elevated Wnt signaling through downstream Wnt effector genes and

adoption of pigment cell fate. Similarly, in the case of *rbc3a/* dmxl2, effects on trafficking

of Frizzled-7 receptors correlates with elevated nuclear β-cat and pigment cell identity.

However, our results with lmo7a call into question a direct connection between NC

aggregation, Wnt signaling, and lineage specification. It is possible that this reflects

either differential premigratory positioning of cells affected in each context or some yet

unknown secondary transcriptional roles for one or more of these genes. Comparative

transcriptomic analyses may help to resolve these differences, and further elucidate the

distinct roles of these novel regulators in NC development. Preliminary bulk RNA-seq

comparisons indicate profiles for NC aggregates that are distinct between *lmo7a* and

ovol1a-deficient embryos. However, This data alone is not sufficient to disentangle the

differences between them (partially due to complications arising from infiltrating cell

types in FACS sorting). Single-cell comparative RNA-seq between these gene

perturbations at the stages leading up to and following aggregation would help to uncover the expression dynamics that accompany an aggregation phenotype in individual cells, perhaps revealing why these effects are partially non-cell-autonomous. Furthermore, this information would allow us to probe how responses to Wnt signaling differ between NC aggregate cells that adopt pigment vs glial identities. Our results continue to underscore the complex nature of the manifold roles for Wnt signaling in NC development.

Conserved roles for Lmo7 in coordinating cell migration and fate

Cells that undergo EMT and migrate require a spatiotemporal balance of several interconnected cellular processes including signal transduction, actomyosin activity, and FA assembly/disassembly. Taken together, our results suggest that Lmo7a forms a novel link between these processes in migrating NC cells. While previous work has shown important roles for Lmo7 in regulating cell-cell junctions in epithelial cells (Du et al., 2019; Ooshio et al., 2004) we propose novel roles for Lmo7a in Wnt signaling in NC cells as well as in cell migration through regulation of Pxn-FAs (Figure 7). Similarly migration of mesenchymal stem cells (MSCs) and their commitment to form osteoblasts are both regulated by cytoskeletal reorganization through FAs and Itgs (Khang et al., 2012), under the control of canonical Wnt signaling (He et al., 2018).

Mammalian Lmo7 has been shown to localize to FAs and influence Itg signaling in vitro (Holaska et al., 2006; Wozniak et al., 2013) and its relative, LIMCH1, regulates FA formation and actomyosin dynamics (Lin et al., 2017). Both are associated with

increased cancer metastasis (Furuya et al., 2002; Hu et al., 2011; Kang et al., 2000; Nakamura et al., 2005; Sasaki et al., 2003; Teixeira et al., 2014). While mammalian Lmo4 has been implicated in NC and cancer EMT, by directly regulating Snail and Slug, it is both structurally and functionally quite distinct from Lmo7 (Ferronha et al., 2013; Ochoa et al., 2012). Lmo7 may have gained a specific role in NC cells in regulating the association of Pxn and FAs with the actin cytoskeleton during migration, as well as linking these to Cdh and β-cat membrane/nuclear ratios, thereby indirectly influencing Wnt signaling. Future studies should examine if LMO7 plays a role in cell migration in other embryonic cells in which it is expressed (e.g. axial mesoderm that forms the notochord) and as a potential therapeutic target in cancer metastasis.

Materials and Methods

Zebrafish husbandry and transgenic lines

Embryos were obtained from natural breeding and staged as described in Kimmel et al,
1995. All zebrafish lines were maintained according to standard protocols (Westerfield,
2000). Transgenic lines used in this study include *Tg(-4.9sox10:nEOS)*[w18] (Curran et al.,
2010), *Tg(-4.9sox10:lyn-tdTomato)*[ir1040](Schilling et al., 2010), *Tg(-4.9sox10:lyn-EGFP)*[ir866](Schilling et al., 2010), *Tg(-7.2sox10:EGFP)*[ir937](Schilling et al., 2010; Wada et
al., 2005) , *Tg(β-actin:Pxn-EGFP)*[mai1](Goody et al., 2010), *Tg(7XTCF:EGFP)*[ia4](Moro et
al., 2012), and *Tg(7XTCF:nls-mCherry)*[ia5](Moro et al., 2012). HCR, chemical inhibition,
and WNT-reporter analysis experiments were performed in double transgenic embryos
derived from crosses of *Tg(7XTCF:EGFP)*[ia4] to *Tg(-4.9sox10:lyn-tdTomato)*[ir1040]. For
Pxn-EGFP live imaging, *Tg(β-actin:Pxn-EGFP)*[mai1] was crossed to *Tg(-4.9sox10:lyn-tdTomato)*[ir1040] and double-transgenic progeny were analyzed. For live imaging,
embryos were mounted in 1% low-melt agarose dissolved in embryo medium containing
1% tricaine. Images were taken on either a Nikon C1 confocal or a Leica SP8 confocal.
Images were processed using ImageJ (NIH).

In situ hybridization and hybridization chain reaction

A 291 bp clone corresponding to a retained intron in the *lmo7a* genomic locus, was
obtained from OpenBiosystems (EST BI845812) and used to generate an antisense
DIG-labeled probe, which was then used for ISH as previously described (Thisse and
Thisse, 2008). *foxd3*, *mitfa*, and *dlx2a* HCR probes were ordered from Molecular

Technologies (Los Angeles, CA) using the accession numbers NM_131290.2,

NM_130923.2, and NM_131311.2 respectively. Whole mount HCR was carried out as

described (Choi et al., 2014).

Immunohistochemistry

For β-cat staining, embryos were fixed in 4% paraformaldehyde (PFA) for 1.5 hours at

room temperature, permeabilized in PBS/1%Triton-x/1%DMSO for 1 hour at room

temperature (RT), washed with PBS/0.1%Triton-x/1%DMSO, and blocked with 5%

Donkey Serum for 1 hour at RT. Embryos were then incubated overnight at 4°C with

1:200 β-cat antibody (GeneTex) in block. For pFAK stains, embryos embryos were fixed

in 4% PFA for 1 hour at room temperature, permeabilized in acetone at -20°C for 5

minutes and then PBS/1%Triton-x/1%DMSO for 30 minutes at RT, washed with

PBS/0.1%Triton-x/1%DMSO, and blocked with 5% Donkey Serum/2% BSA for 1 hour at

RT. Embryos were then incubated overnight with 1:200 pFAK antibody (Invitrogen) in

block.

Molecular cloning

Coding sequences of *lmo7a* were amplified from cDNA isolated from 16 hpf zebrafish

embryos. cDNA was generated using Protoscript II First Strand Synthesis Kit (New

England Biolabs). Lmo7a rescue constructs and lmo7a:sfGFP fusion constructs were

generated by Gibson Assembly (Gibson et al, 2009) into the pCS2+ vector.

Morpholino and mRNA microinjections

Antisense morpholino oligonucleotide (MO) targeting lmo7a (5'-

TCGCCACTCCATCACCGGTCAACGT-3') and Control MO (Gene Tools) were

dissolved in nanopure water prior to injection. For all mRNA injection experiments,

mRNAs were transcribed in vitro using the mMessage mMachine kit (Ambion). All

injections were performed at the 1-cell stage, and a volume of 1 nl was injected in all

cases. For XAV939 treatment experiments, *lmo7a* MO was injected at 3 ng/embryo

because of DMSO and XAV939 toxicity. For all other experiments, it was injected at 4

ng/embryo. Control MO was injected at 4 ng/embryo. In all cases, 400 pg of mRNA was

injected.

CRISPR and CRISPRi gRNA

For CRISPR-Cas9 injections, templates were generated using a different scaffold

chimeric primer design (Varshney et al, 2015) and multiple target-specific primers as

described (Wu et al, 2018). gRNAs were then synthesized using T7 MegaShortScript kit

(Ambion). gRNAs were incubated with Cas9 protein (IDT) at 37°C for 5 minutes and

injected into embryos at the 1-cell stage. 150 pg gRNA and 800 pg Cas9 protein were

injected. For CRISPRi injections, templates for gRNAs were generated by PCR using

the scaffold chimeric primer design (Larson et al, 2013) and a target-specific primer

containing a T7 promoter sequence. *dCas9* mRNA was synthesized from the pT3T-nls-

dCas9-nls plasmid (Rossi et al, 2015) using mMessage mMachine kit (Ambion). *dCas9*

mRNA was coinjected with gRNAs targeting multiple sites along the *lmo7a* coding

region. 75 pg gRNAs and 450 pg *dcas9* mRNA were injected.

Genotyping

The CRISPR-Cas9 induced Δ5 and Δ35 deletions were identified by PCR of gDNA

isolated from individual juvenile fish using primers targeting Exon16 of the *lmo7a*

genomic locus. Heterozygous F1 fish were incrossed to generate transheterozygous F2

fish.

Quantitative RT-PCR

For qPCR, total RNA was extracted from biological triplicates of de-yolked CRISPRi-

injected and *dCas9*-injected control embryos using trizol (Invitrogen) and purified using

Direct-zol RNA Miniprep kit (Zymo). cDNA was generated using Protoscript II First

Strand cDNA Synthesis kit (New England Biolabs). cDNAs were diluted 1:20 and then

qPCR was carried out in triplicate using the Luna Universal qPCR Master Mix (New

England Biolabs). 3 primer sets targeted different exon-exon junctions in the *lmo7a*

coding region and primers targeting *rps13a* as the control housekeeping gene were

used for normalization. qPCR was performed on a LightCycler 480 (Roche). Fold

changes were determined by calculating ΔΔCT values for each biological replicate.

Statistical analyses

For mRNA rescues, medians were compared using Kruskal-Wallis ANOVA and posthoc Wilcoxon tests. For β-cat and pFAK stains, normal distribution assumptions were tested using Shapiro-Wilk tests and means were compared using Welch's two sample T-tests. Statistical tests were carried out using R, and plots were generated in R using the ggplot2 package. Other R packages used include: plyr, dplyr, reshape2, ggpubr, ggsignif, FSA, plotly, plotrix, and gridExtra.

Bulk RNA-seq library construction and sequencing

RNA was extracted from cell lysates using the RNEasy Micro Kit (Qiagen). cDNA libraries were generated following the Smart Seq II protocol (Picelli et al, 2014). Libraries were then sequenced on a NovaSeq6000 (Illumina) at a depth of ~20M reads per sample.

Chapter III

A single-cell transcriptomic timeline of cranial NC development reveals timing of lineage decisions and novel regulators of NC development

Introduction

A fundamental question in developmental biology is how embryonic cells acquire their fates as part of tissues undergoing rapid growth and cell rearrangements. A dramatic example of this in vertebrates is the neural crest (NC), a highly migratory and multipotent transient cell population. NC cells are specified at the border of the neuroepithelium and epidermis, undergo an epithelial-mesenchymal transition (EMT) and migrate extensively throughout the embryo to generate a wide variety of cell types including cartilage, bone, neurons, glia, pigment, and many others (LeDouarin, 1980; Weston, 1970). Despite many studies addressing the origins and emergence of these cell types during NC migration the degree to which individual NC cells are multipotent or biased toward particular lineages, as well as the timing of their lineage decisions, remains unclear (Dupin et al., 2018; Kalcheim & Kumar, 2017; Prasad et al., 2019).

Lineage tracing experiments in vivo have produced evidence both for NC multipotency as well as early lineage restrictions prior to migration. Cell labeling experiments in the chick have shown that single premigratory NC cells can give rise to multiple fates (Bronner-Fraser & Fraser, 1988; Bronner-Fraser et al., 1980; McKinney et al., 2013), while in zebrafish many single labeled cranial NC cells generate cell-type restricted clones depending on their initial premigratory locations (Schilling & Kimmel, 1994). Cell labeling of trunk NC cells in chick and zebrafish embryos has similarly provided evidence for a

spatiotemporal map of NC cell fate independent of migratory environment (Krispin et al., 2010; Raible & Eisen, 1994), while lineage tracing in mice using genetically encoded fluorescent tags suggests that early NC cells show no apparent lineage restrictions (Baggiolini et al., 2015). A variety of in vitro studies, for example serially isolating single NC cells and assaying the array of cell types they can generate in culture have also shown restricted potency (Stemple & Anderson, 1992). These seemingly inconsistent findings highlight the need for further investigation into the timing and mechanisms of lineage specification in the NC.

Recent studies have revisited this issue with new single-cell transcriptomic and epigenomic approaches. Single cell RNA sequencing (scRNA-seq) and ATAC-seq data from FAC-sorted premigratory and migrating NC cells in mice suggest that glial/ectomesenchymal lineages arise during migration (Soldatov et al., 2019) while similar studies of chick NC cells suggest an earlier split between neural and non-neural lineages (Williams et al., 2019). Interestingly, some of the single cell evidence points to a population of stem-like premigratory NC cells in zebrafish(Lukoseviciute et al., 2018) and this is supported by spatial genomic analyses of gene expression in the chick dorsal neural tube (Lignell et al., 2017). However, another independent single cell NC study in chick shows no evidence for similar early stem-like subpopulations (Morrison et al., 2017). Thus, many questions remain as to precisely when and where distinct cell lineages emerge in the NC, as well as the mechanisms that regulate their fate decisions.

To address these questions, we have generated scRNA-seq datasets for cranial NC cells of the first pharyngeal (mandibular) arch in zebrafish. Unlike previous single cell studies, which lack spatial information and rely on pseudotime to infer temporal changes in gene expression, our study focuses on the mandibular NC subpopulation over a closely spaced series of stages (every few hours). Our scRNA-seq data suggest that premigratory NC cells are relatively homogenous at the transcriptional level. Gene expression signatures of pigment and skeletal lineages emerge mid-migration and a neural/glial signature becomes apparent slightly later during migration. We identify several novel markers of these lineages and confirm their expression and spatial segregation within the migrating NC using in situ hybridization. Emergence of these lineage-specific gene expression signatures correlates with shifts in Wnt signaling between NC cells and surrounding tissues, which we show by computationally integrating bulk RNA-seq data from Wnt reporters with the scRNA-seq data. Our findings provide valuable insights into the specification of lineages in the cranial NC and uncover novel regulators of cell fate decisions.

Results

Single cell RNA-seq of cranial NC cells reveals heterogeneity in gene expression during migration

To profile heterogeneity in gene expression over time in a defined population of cranial NC cells, we constructed a single cell transcriptomic timeline for cells that contribute to the first pharyngeal arch (PA, mandibular) over the period that they migrate. We isolated cells of the first PA migratory stream using a transgenic photoconvertible marker of NC cells, *tg(sox10:nEOS)* (**Fig 3.1 A**). Photoconverted cells were isolated via FACS at 6 embryonic stages from the onset to essentially completion of NC migration in this stream: 12, 14, 18, 20, 24, and 30 hours post fertilization (hpf) (**Fig 3.1 B**). Single cell cDNA libraries were then constructed via the 10X Genomics platform and sequenced. Clustering and marker gene analysis using Seurat (Stuart et al., 2019) revealed a large NC population and a few smaller populations of likely non-NC cell types including endodermal, mesodermal/cardiac, neuroectodermal, and epithelial progenitors (**Fig 3.1 C**; **Fig 3.2 A-C**). These presumptive non-NC cells were removed from further analyses to obtain a pure population of NC cells. Following clustering and dimensionality reduction using Seurat, these cells largely clustered within their time point identities (**Fig 3.1 D**). Marker genes obtained through differential gene expression analysis using Seurat revealed general trends in temporal gene expression dynamics (**Fig 3.1 E**). For example, expression of early NC specifier genes including *sox10* was very strong in the 12 and 14 hpf populations, but considerably reduced later. Conversely, expression of many pigment-specifiers including *mitfa*, as well as markers of skeletal lineages including *col2a1b* progressively increased at mid-late migration stages (**Fig 3.1 F**).
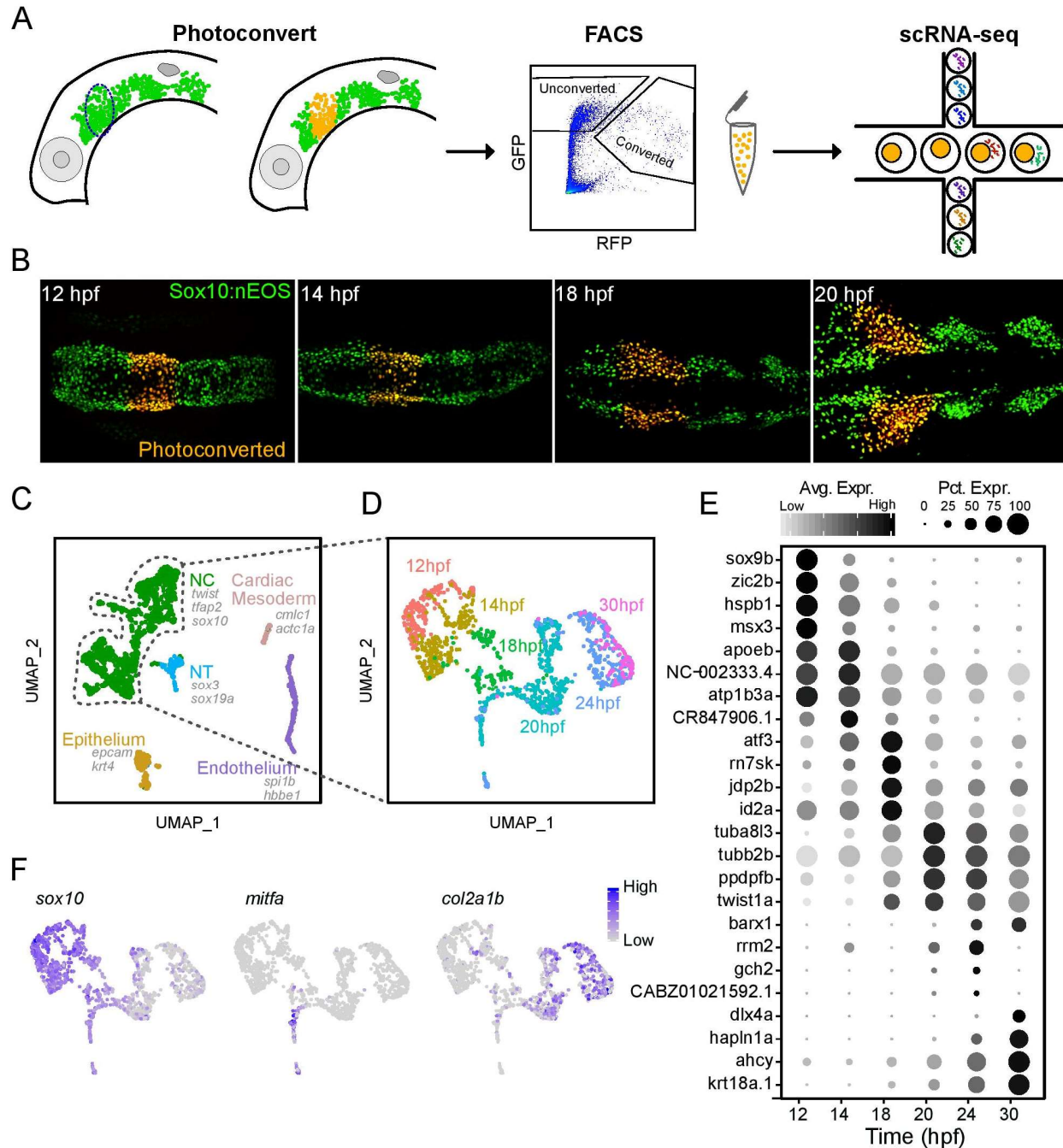
Differentiation of lineages is characterized by their expression of unique transcription factors (TFs) that in turn drive expression of downstream differentiation programs. For example, in the pigment lineage of NC, master regulatory TFs like Mitf precede expression of later enzymes involved in pigment production such as tyrosinase. Therefore, a temporal analysis of TFs can inform the specification of lineages prior to their

Figure 3.1: Time course scRNA-seq of isolated first pharyngeal arch NC cells reveals broad changes in gene expression across developmental time. A) Diagram showing experimental set up. NC cells expressing photoconvertible nEOS were photoconverted and isolated via FACS, and then sequenced using 10X Genomics Drop-seq. **B)** Live fluorescent images showing photoconverted first PA migratory stream at 4 of 6 timepoints. **C)** Unbiased clustering and UMAP embedding showing NC and other cell types isolated. All 6 timepoints were aggregated and analyzed together. Canonical markers for each cell type are displayed. **D)** Pure NC subpopulation was subclustered and reanalyzed. UMAP displays time point identities for each cell. **E)** Dotplot displaying markers for each time point. **F)** UMAPs displaying relative expression levels for presumptive early NC, pigment, and cartilage lineages marked by *sox10*, *mitfa*, and *col2a1b* respectively.

terminal differentiation. In order to investigate the mechanisms driving the emergence of the pigment and skeletal developmental branches, we used pseudotime analysis driven by differentially expressed TFs. Dimensionality reduction and clustering with a comprehensive list of TFs (adapted from Vaquerizas et al., 2009**; Fig 3.3 A**) resulted in smooth connections between clusters. Combined lineage trajectory and pseudotime analysis using Monocle3 (Cao et al., 2019; Trapnell et al., 2014) identified TFs with differential expression across developmental time (**Fig 3.3 B**). Louvain clustering uncovered 6 distinct modules of gene expression, in total consisting of 91 TFs (**Fig 3.3 C, D**). Three of these TF modules distinctly overlapped with the early NC, pigment, and skeletal populations. Using these TFs along with other, established canonical markers, we built lists of genes that mark early NC, pigment, skeletal, and neural/glial lineages (Table S1).

Distinct gene expression signatures arise progressively during NC migration

To examine the temporal emergence of lineages in more detail, we utilized the partition-based graph abstraction (PAGA) method (Wolf et al., 2019), which determines topology of the data and computes a measure of cell connectivity. The PAGA graph was then used
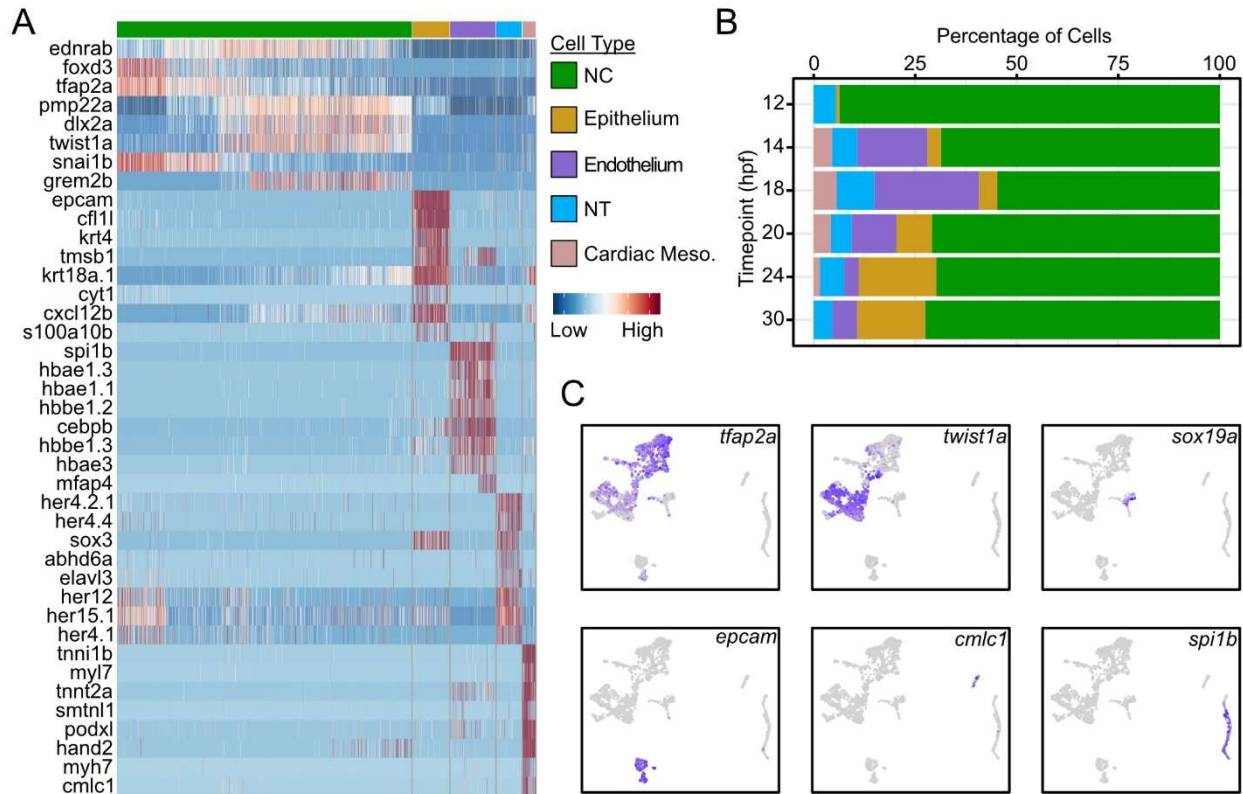
Figure 3.2: Expression profiles of non-NC cell types. A) Heatmap showing top marker genes in each cell type in Fig. 1C. **B)** Stacked barplots showing the number of each cell type recovered in each time point. **C)** Feature plots showing expression of strong marker genes for each cell type overlayed with UMAP.

to initialize force directed layout in ForceAtlas2 (Jacomy et al., 2014) for single-cell embedding that preserves global topology and connectivity (**Fig 3.4 A**). Module scoring revealed groups of cells enriched for markers of each lineage (**Fig 3.4 B**). We then applied probabilistic cell identification using CellAssign driven by these lineage markers (Zhang et al., 2019). This analysis revealed cells identified as early NC, pigment, skeletal, neural/glial, and another group that belonged to none of these lineages, which we labeled as "transitional" (**Fig 3.4 C**). Expression of most markers was restricted to single lineages, though a few were shared between lineages (**Fig 3.5 A**). Transitional cells showed a unique expression profile, distinct both from any specific lineage as well as earlier NC
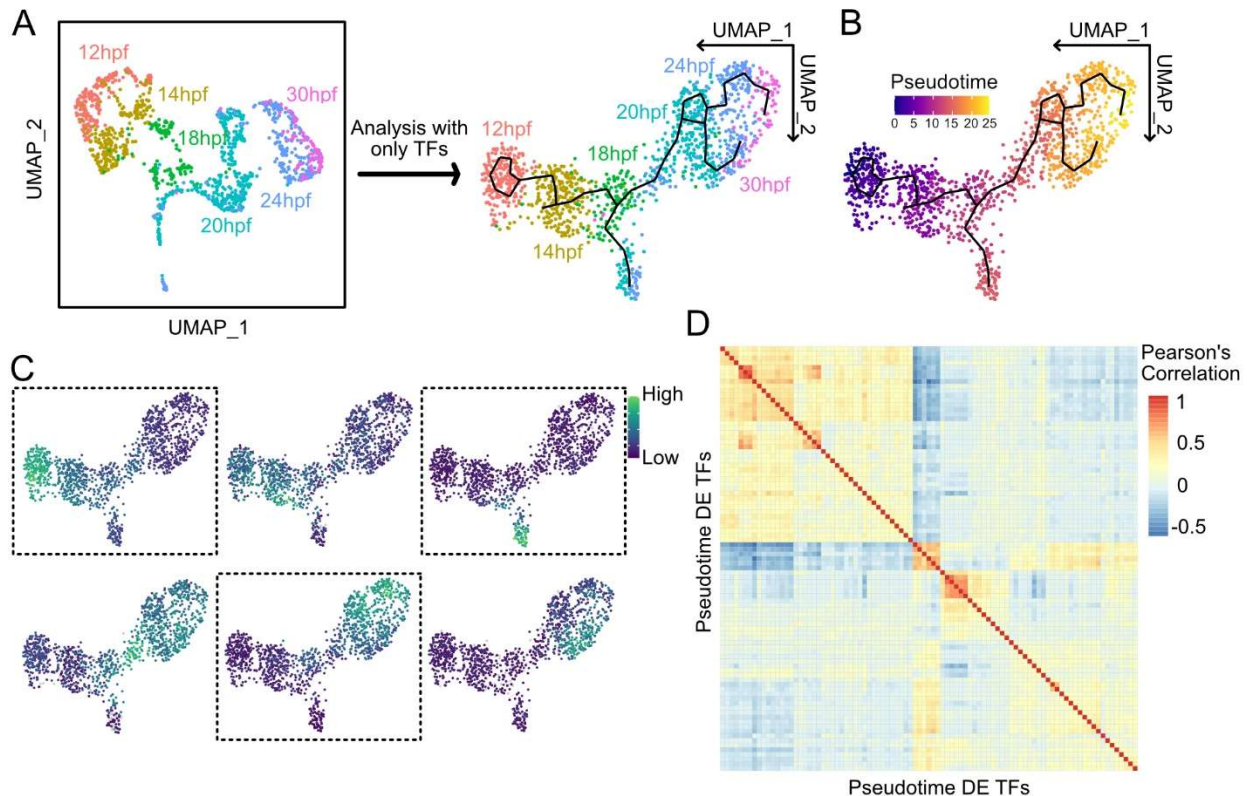
Figure 3.3: Temporal analysis of transcription factors in the NC. A) NC data set from Fig. 1D re-analyzed using only TFs. The new UMAP shows the same broad structure but smoothens connections between time points (black lines). **B)** The same UMAP overlayed with pseudotime values calculated using Monocle3. **C)** Six modules of TFs were identified with distinct expression profiles across pseudotime. Feature plots showing expression scores for each module overlayed with the UMAP. 3 modules correspond to early NC, skeletal, and pigment branches (dashed line boxes), totaling 91 TFs. **D)** Heatmap showing correlation between each of the 91 TFs making up the modules. Hierarchical clustering reveals groups of co-expressed TFs.

cells. PAGA embedding and connectivity were then used to calculate pseudotime and identify branch points. Importantly, when ordered by pseudotime, the cells analyzed in this manner arranged in a way that largely recapitulated real developmental time (**Fig 3.5 B**), suggesting that branch points correspond closely to the temporal emergence of each lineage (**Fig 3.4 D**). These analyses suggest that subpopulations with gene expression signatures of pigment and skeletal lineages appear first, while the neural/glial progenitor subpopulation arises later.
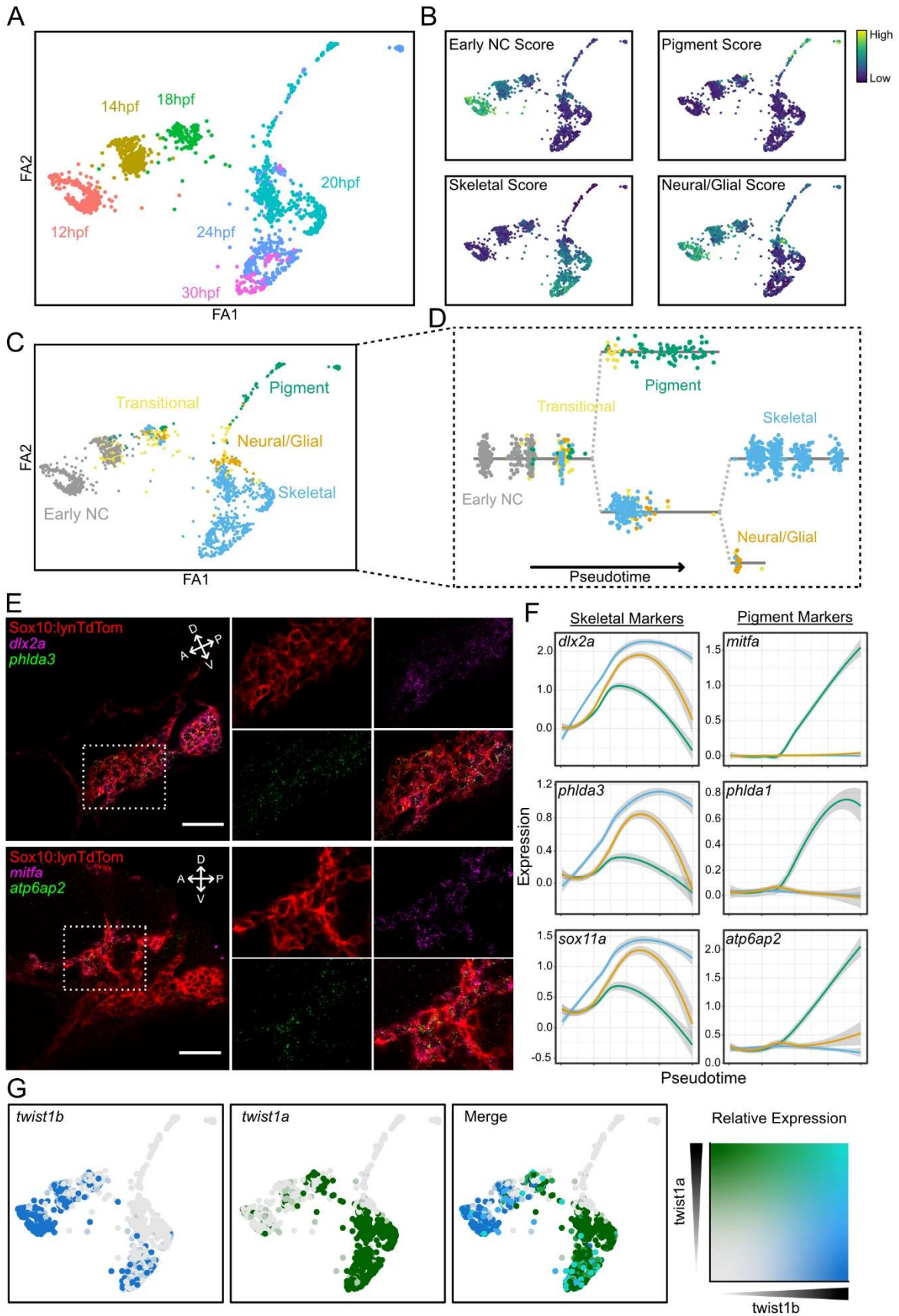
Figure 3.4: Pseudotemporal analysis and in situ Hybridization Chain Reaction (isHCR) reveal emergence of NC lineages during migration. A) ForceAtlas (FA) embedding of single cell timeline initialized using PAGA graph with time point identities of individual cells labeled and color-coded. **B)** Module scores, color code distinct from A, for early NC, pigment, skeletal, and neural/glial lineages overlaid on FA embedding. **C)** CellAssign probabilistic cell type labels computed using module score genes overlaid on FA embedding, color code distinct from A or B. **D)** Pseudotime was calculated using PAGA graph. Cells are arranged according to pseudotime values along the x-axis, with pseudotime scaled within each branch. Branch points are split on the y-axis and connected by dashed lines. Color code corresponding to C. **E)** Confocal micrographs of isHCR for novel markers of skeletal and pigment progenitors in PA1 at 24 hpf in *tg(sox10:lynTdTomato)* embryos where NC plasma membranes are marked by tdTomato (red). Expression of *phlda3* (green in upper panels) overlaps with *dlx2a* (magenta) in the ventral portion of the migratory stream, and *atp6ap2* expression (green in lower panels) overlaps with *mitfa* (magenta) in the dorsal portion. **F)** Local regression graphs for novel and canonical markers of early NC, pigment, and skeletal progenitors, showing expression levels (Y axis) across pseudotime (X axis). Lines indicate moving averages for each branch. Temporal expression profiles for *phlda3* and *sox11a* strongly resemble that of *dlx2a*, while profiles of *phlda1* and *atp6ap2* resemble that of *mitfa*. **G)** Feature plots showing expression of *twist1b* and *twist1a*. First two panels show individual expression patterns (blue and green respectively). Third panel shows overlapping expression in teal.

Skeletal and pigment lineage markers included many that were expected, such as *dlx2a* and *mitfa,* respectively. However, numerous genes with no previously-characterized role in NC were co-expressed in these lineages as well. Among these, *sox11a* and *phlda3* were expressed specifically in putative skeletal progenitors, while *atp6ap2* and *phlda1* were expressed in putative pigment progenitors (**Fig 3.5 C**). To confirm the specificity of their expression in vivo, we performed *in situ* Hybridization Chain Reaction (isHCR). Strikingly, as predicted, *phlda3* was co-expressed with *dlx2a* while *atp6ap2* expression overlapped with *mitfa* (**Fig 3.4 E**). *sox11a* and *phlda1* had similar expression patterns in vivo (**Fig 3.6 A, B**). Based on their expression in pseudotime, each of these genes has a temporal expression profile that closely matches canonical skeletal and pigment progenitor markers (**Fig 3.4 F**). Taken together with their spatial patterns of expression with isHCR these genes are strong candidates for novel factors involved in the
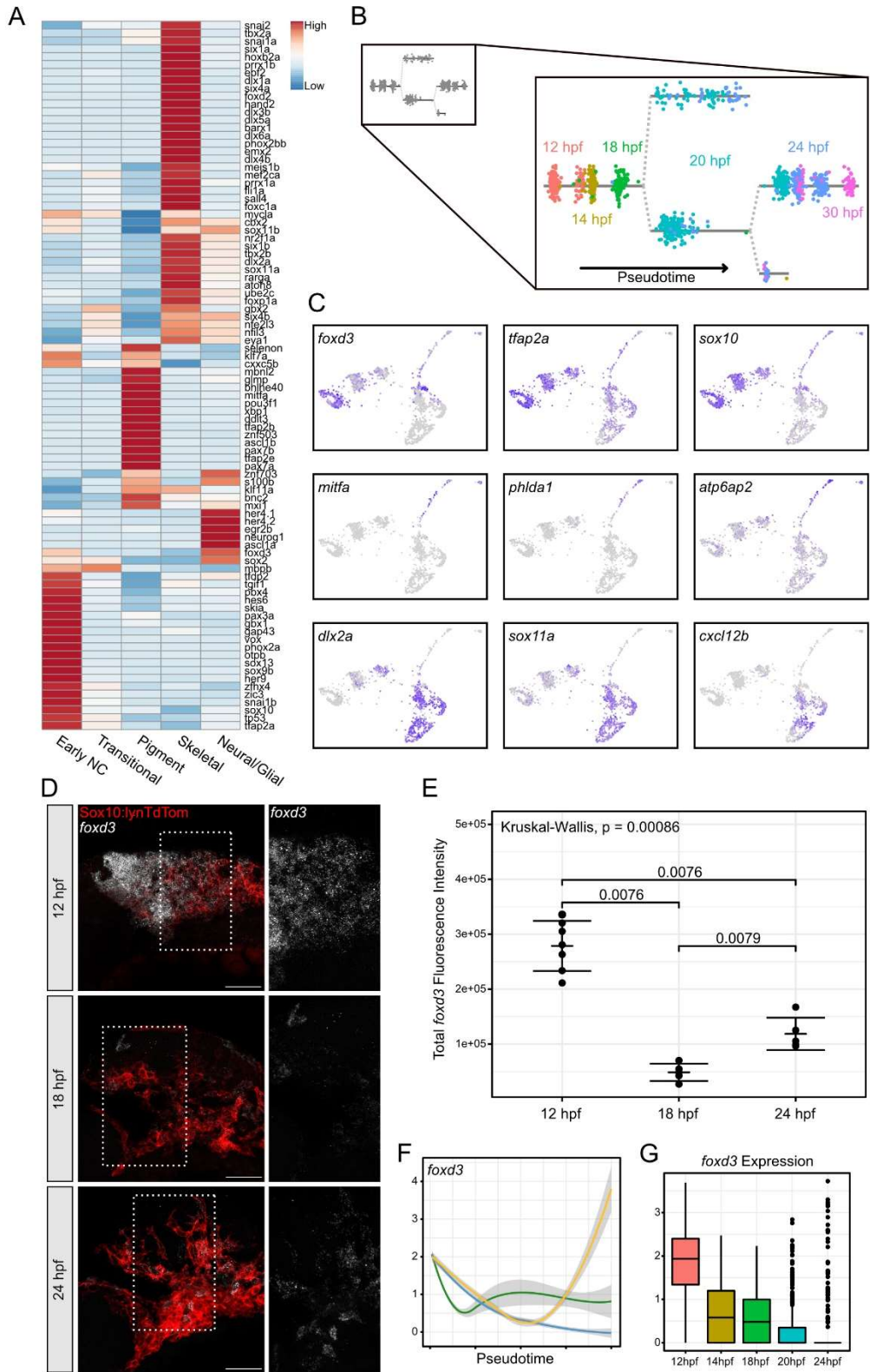
development of these NC lineages. The emergence of a skeletal progenitor signature also

clearly correlates with a shift in expression of two Twist1 orthologs; early NC cells express

Figure 3.5: Marker gene profiles of NC lineages and in vivo temporal expression patterns of *foxd3*. A) Heatmap showing average expression of top lineage markers in each NC subtype. **B)** Dendroplot with cells ordered according to pseudotime and colored according to real developmental time point identity. **C)** Feature plots showing relative expression (dark purple indicates high levels) of lineage markers overlayed with ForceAtlas (FA embedding. **D)** Confocal micrographs of isHCR for *foxd3* expression (white) at 12, 18, and 24 hpf in *tg(sox10:lynTdTomato)* embryos where NC plasma membranes are marked by tdTomato (red). Enlarged region is the PA1 migratory stream. **E)** Quantification of *foxd3* expression from isHCR images in PA1 using corrected total fluorescence intensity (Y axis) at 12 hpf (n = 7 embryos, mean = 278640), 18 hpf (n = 5 embryos, mean = 48478), and 24 hpf (n = 5 embryos, mean = 118584). Dots represent means in individual embryos; lines represent means within conditions. Error bars represent mean ±SD. **F)** Local regression graph for *foxd3* expression levels (Y axis) in pigment (green), glial (yellow) and skeletal (blue) progenitors across pseudotime (X axis). Lines indicate moving averages for each branch. **G)** Barplots showing expression levels of *foxd3* (Y axis) across developmental time points (X axis). Lines indicate means, boxes indicate IQR, whiskers indicate IQR*1.5, points indicate outliers. For micrographs, Scale bars = 50µm.

twist1b, while later putative skeletal progenitors express *twist1a*. Notably, neither pigment nor neural/glial branches express either Twist1 ortholog (**Fig 3.4 G**).

The TF *foxd3* plays numerous roles in NC development, including NC induction as well as specification of the neural/glial lineage (Dottori et al., 2001; Lister et al., 2006; Montero-Balaguer et al., 2006; Stewart et al., 2006). Consistent with this, we observed *foxd3* expression in early NC cells, reduced expression in transitional cells, and later upregulated expression in neural/glial cells in our scRNA-seq dataset (**Fig 3.5 A,C**). To confirm these temporal dynamics of *foxd3* expression in vivo, we performed isHCR and found that expression was high at 12 hpf, significantly lower by 18 hpf, and increased again by 24 hpf (**Fig 3.5 D-E**), consistent with our pseudotemporal analysis in scRNA-seq data (**Fig 3.5 F**). In addition, while changes in expression levels at 12 and 18 hpf both occurred broadly across all NC cells, the later increase at 24 hpf was restricted to a small subset of cells (**Fig 3.5 G**). These observations are consistent with previous work and the
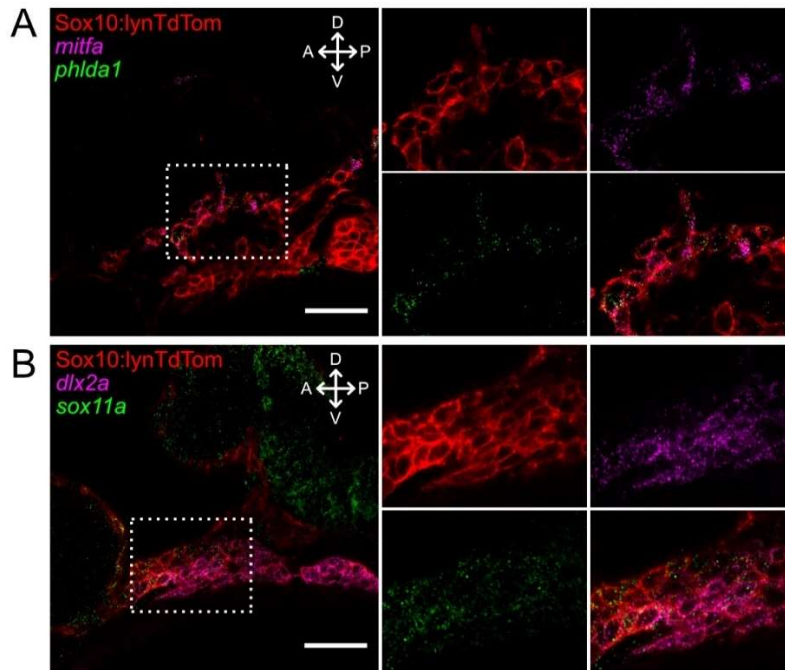
Figure 3.6: isHCR confirms additional novel markers of skeletal and pigment lineages. Confocal micrographs of isHCRs co-labeled for phlda1 (green) and mitfa (magenta) in putative pigment progenitors **(A)** and sox11a (green) and dlx2a (magenta) in putative skeletal progenitors **(B)** at 24 hpf in *tg(sox10:lynTdTomato)* embryos where NC plasma membranes are marked by tdTomato (red).

interpretation that *foxd3* shifts from a broad early NC marker to a specific neural/glial lineage specifier, lending strong support that other temporal and lineage-specific features of our scRNA-seq analyses are accurate.

A distinct gene expression signature defines a transitional population of NC cells mid-migration

The analysis of single-cell expression profiles near branchpoints in our pseudotime analyses pinpointed 18 hpf as the stage at which heterogeneity begins to increase in the cranial NC of PA1 in zebrafish (**Fig 3.7 A**). To gain deeper insights into the apparent rapid developmental changes taking place at 18 hpf, we generated a second, larger scRNA-seq data set at this stage, containing 420 NC cells. Dimensionality reduction and clustering in Seurat resulted in 5 clusters with distinct expression profiles **(Fig 3.7 B)**. One cluster showed a clear pigment progenitor expression signature, while two clusters had expression profiles indicative of the skeletal lineage. Another cluster appeared relatively undifferentiated, with low level expression of
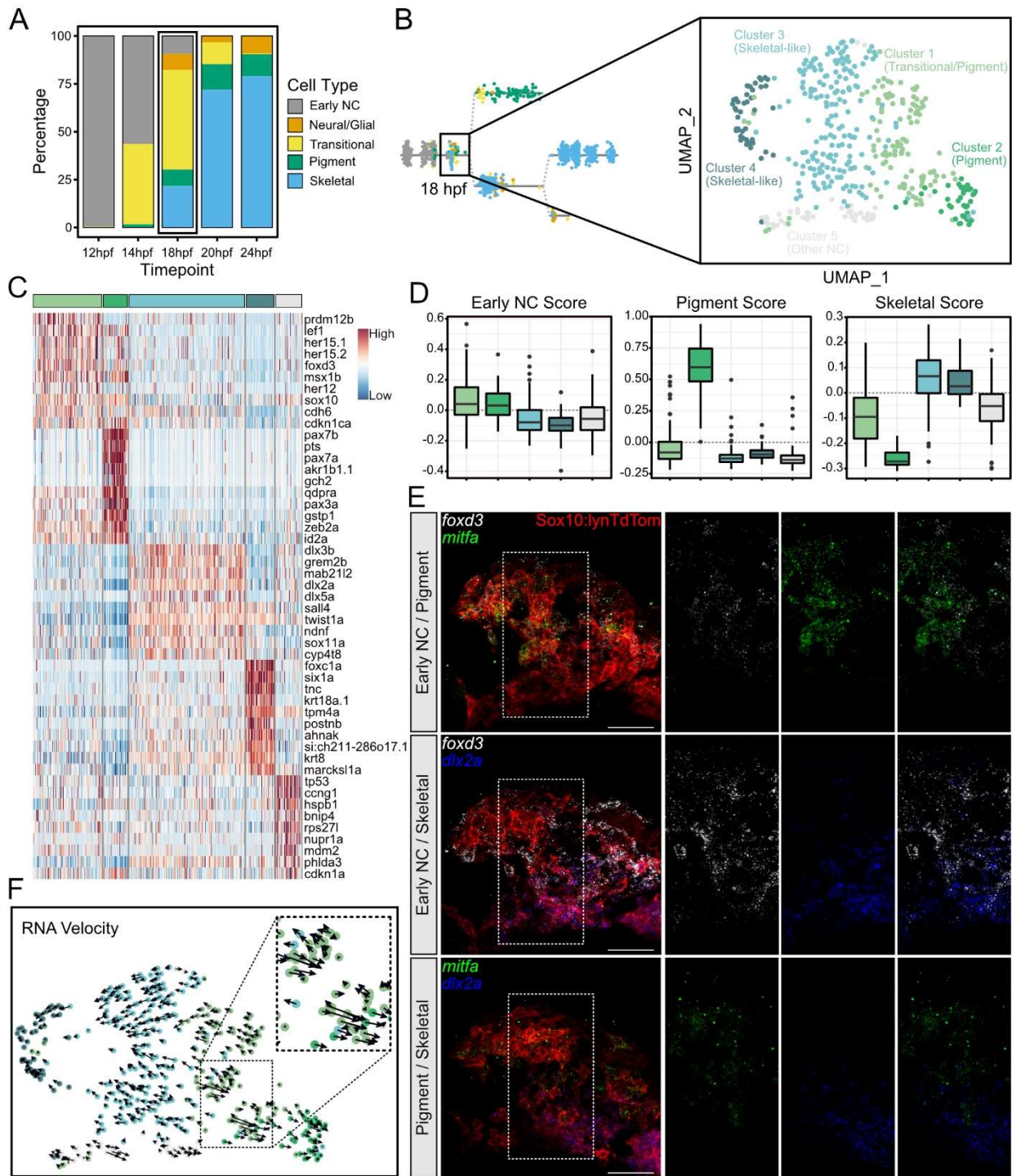
both pigment and skeletal lineage markers. The fifth cluster was even less distinct, marked by heat shock and cell cycle genes, and possibly stressed/unhealthy cells that were missed initially. However, notably these cells expressed some unique factors

Figure 3.7: Analysis of 18 hpf scRNA-seq dataset shows emergence of distinct lineages as well as transitional cell states. A) Stacked barplot showing percentages of cell types at each time point from 12-24 hpf, color code corresponding to Fig. 2C. **B)** Based on heterogeneity at 18 hpf a new larger scRNA-seq data set was generated at this stage. Dimensionality reduction and unbiased clustering revealed 5 distinct clusters: 2 skeletal-like clusters, a pigment cluster, a transitional cluster, and a cluster showing signs of cell stress, color code distinct from A. **C)** Heatmap showing top 10 marker genes for each cluster, colored bar across the top indicates color code corresponding to B. **D)** Boxplots showing early NC, pigment, and skeletal module scores for each cluster, color coding corresponding to B and C. **E)** Confocal micrographs of isHCR for known markers of early NC (*foxd3*, white), pigment (*mitfa*, green), and skeletal (*dlx2a*, blue) progenitors in *tg(sox10:lynTdTomato)* embryos where NC plasma membranes are marked by tdTomato (red). Expression of *foxd3* overlaps strongly with *mitfa* and to a lesser extent with *dlx2a*. However, *mitfa* and *dlx2a* do not overlap. **F)** RNA velocity showing differentiation trajectories (arrows) for all cells. The transitional population (inset) is split between progressing toward the skeletal (blue) and pigment (green) clusters.

including *mmp2* perhaps indicative of a unique phenotype within the migratory stream

(**Fig 3.7 C, D, Fig 3.8A**).

To confirm our scRNA-seq evidence for early NC lineage specification events at 18 hpf

in vivo, we again performed isHCR, for genes thought to mark either a pigment or skeletal

progenitor identity. *foxd3* expression was relatively broad across the migratory stream,

while both *mitfa* and *dlx2a* marked subsets of NC cells, consistent with the idea that some

transitional cells exist at this time point for both lineages. However, *mitfa* and *dlx2a* isHCR

labeling did not overlap in NC cells of the PA1 migratory stream, with *dlx2+* NC cells

largely occupying the ventral region of the stream and *mitfa+* NC cells being restricted to

the dorsal region (**Fig 3.7 E**). We also performed RNA velocity analysis (La Manno et al.,

2018), and this also confirmed that NC differentiation trajectories within the transitional

population largely progressed toward either the pigment or skeletal clusters (**Fig 4F**).

Thus, our in vivo isHCR and scRNA-seq data at 18 hpf both indicated the presence of NC
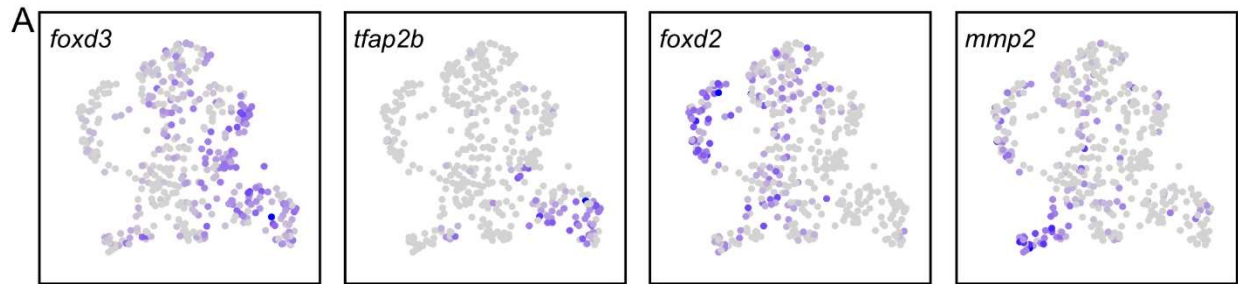
Figure 3.8: Additional marker genes in 18 hpf neural crest subpopulations. A) Feature plots showing expression of 4 marker genes in 18 hpf neural crest.

cells with pigment and skeletal progenitor signatures, with many cells still in transition. Transitional cells were identified by a unique expression profile characterized by overlapping expression of early NC markers like *sox10*, *foxd3*, and *pax3a* and lowlevel expression of later lineage markers including *dlx2a*, *mitfa*, and *sox11a*. Other lineage markers like *pax7b* and *dlx5a* were absent and likely only appear later as these subpopulations further differentiate.

Combined bulk and scRNA-seq reveals Wnt signaling dynamics in developing NC

Many different Wnt ligands and receptors facilitate communication between NC cells and their surrounding tissues during induction, migration, and lineage specification (De Calisto et al., 2005; Ji et al., 2019; Maj et al., 2016; J. Wu et al., 2003). To gain insights into the dynamics of Wnt ligand-receptor interactions during cranial NC migration, we generated transcriptomic profiles of Wnt-responsive NC cells using a transgenic line of zebrafish carrying a fluorescent Wnt reporter, tg(*7xTCF:mcherry*), thought to primarily respond to canonical Wnt signaling (Moro et al., 2012). By combining this with tg(*sox10:egfp*) to mark NC cells, we FACS sorted double-transgenic NC cells exhibiting high levels of Wnt responsiveness ("high-Wnt", mCherry+) from those with much lower Wnt responses
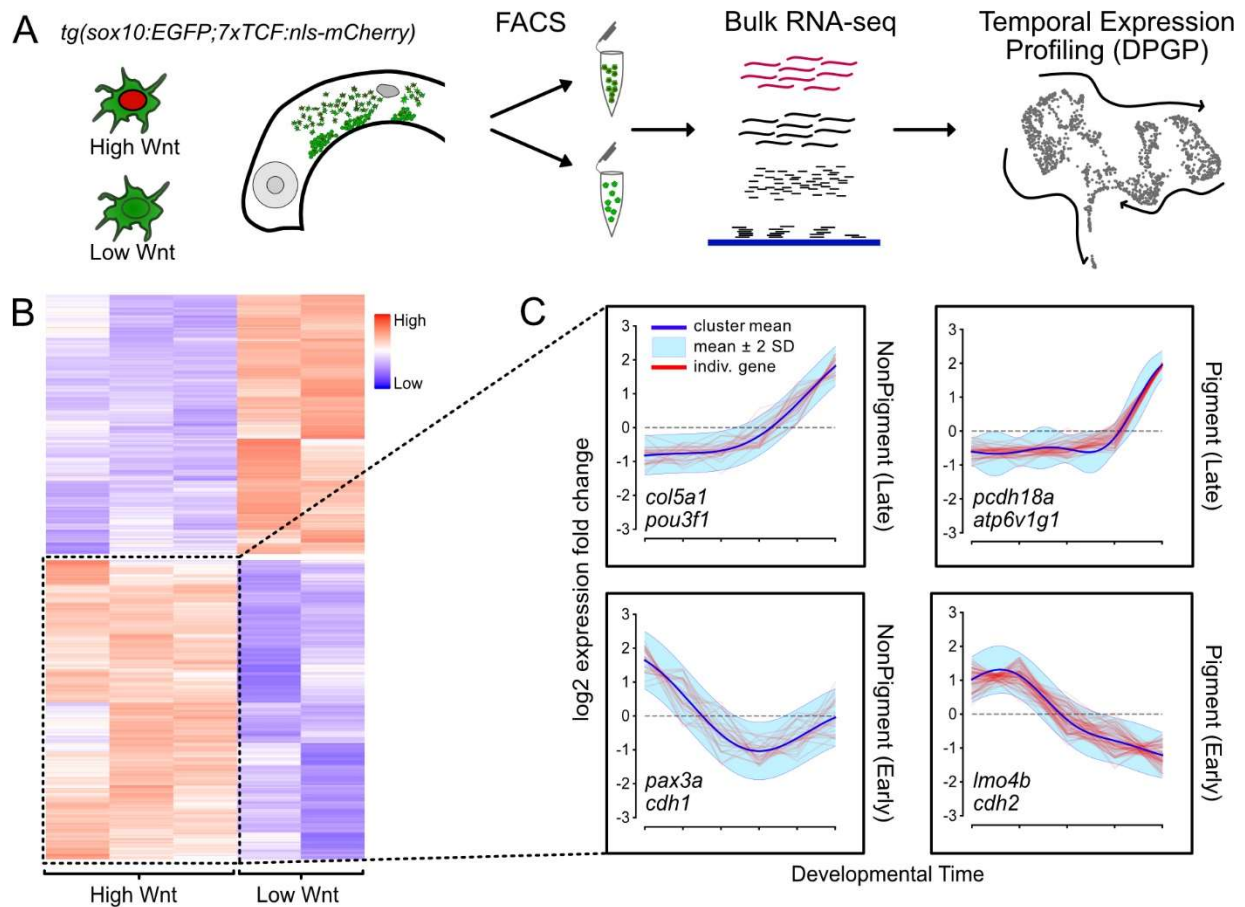
Figure 3.9: Bulk RNA-seq of a transgenic Wnt reporter combined with scRNA-seq analysis reveals temporal kinetics of Wnt signaling. A) Diagram showing experimental design. Lateral views of the head of a sox10:EGFP;7xTCF:nls-mCherry transgenic, showing migrating NC cells with either high (high Wnt) or low (low Wnt) mCherry levels were isolated via FACS and then combined for bulk RNA sequencing. Differentially expressed (DE) genes were then analyzed using our scRNA-seq timeline to uncover temporal patterns. **B)** Heatmap showing DE genes between high-Wnt and low-Wnt cells. **C)** Wnt-upregulated genes were clustered based on their temporal profiles in the scRNA-seq timeline using DPGP (Dirichlet Process Gaussian Process). The timeline (X axis) was split into pigment (left panels) and non-pigment (right panels) branches to separate out Wnt-regulated genes that specifically drive pigment specification. For each branch, two major clusters of genes were identified as having early or late expression signatures. Example genes from each temporal cluster are inset with the plots.

("low-Wnt", mCherry-). Bulk RNA libraries were then constructed and sequenced from the

high and low Wnt-responsive NC populations (**Fig 3.9 A**). 1917 genes were differentially

expressed (DE) between the two populations (**Fig 3.9 B**). From these DE genes, 1027
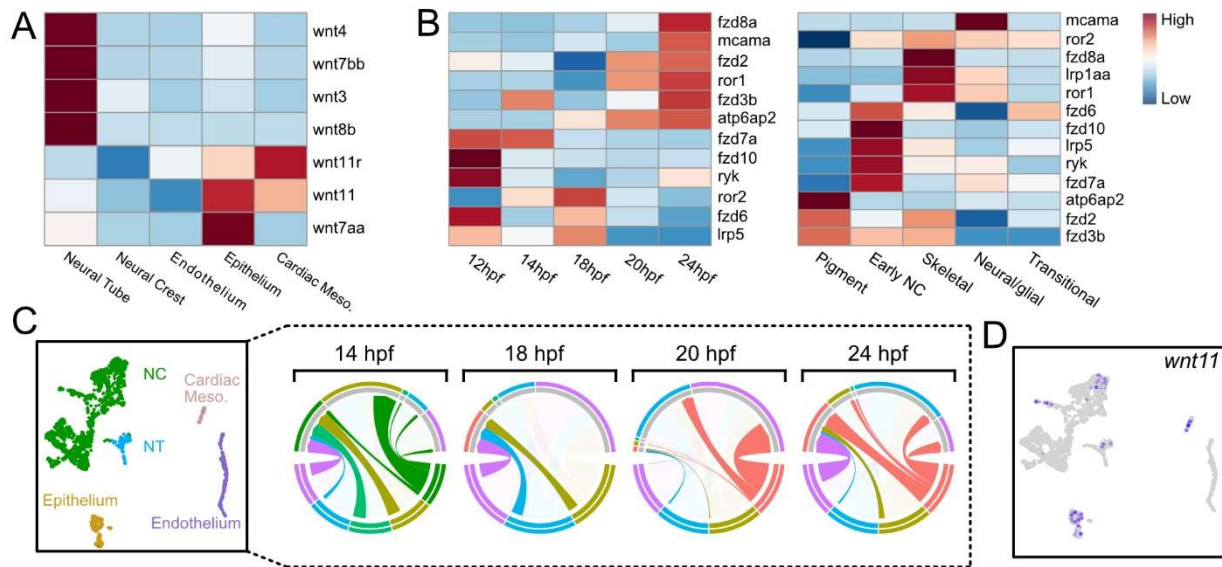
Figure 3.10: Cell-cell signaling analysis suggests changes in source of Wnt signaling across development in the NC. A) Heatmap showing average expression of Wnt ligands in each cell type in the single cell timeline. **B)** Heatmaps showing average expression of Wnt receptors in NC cells both at different timepoints and in different lineages. **C)** Signaling analysis for cell-cell communication between NC cells and other cell types through SoptSC, color code corresponding to Fig. 1C. The bottom half of each circos plot represents signal-sending clusters; the top half represents signal-receiving clusters. Arrows are colored according to signal-sending cell type. Width of arrows indicates probability of a signaling event between the two cell types. **D)** Feature plot showing expression of *wnt11* in all cell types except endothelial, with NC expression confined mostly to early NC cells and pigment cells.

genes were upregulated in the high-Wnt population. Because Wnt signaling has such varied functions across all stages of NC development, we attempted to disentangle the temporal role of each gene in the profile. We utilized our scRNA-seq timeline and the tool DPGP (Dirichlet Process Gaussian Process) (McDowell et al., 2018) to identify shared temporal profiles among groups of genes in the bulk transcriptomic data by gene clustering. Since canonical Wnt signaling promotes pigment cell fates (Curran et al., 2010), we analyzed the pigment and non-pigment branches separately. The combined analyses identified 358 Wnt-responsive genes that were organized into groups

corresponding to early and late development in both pigment and non-pigment trajectories (**Fig 3.9 C**).

Having analyzed Wnt responses in migrating cranial NC cells, we next investigated heterogeneity in the expression of Wnt receptors in NC cells as well as Wnt ligands across all cell types in our FACS sorted samples. Several Wnt ligands, *wnt4*, *wnt7bb*, *wnt3*, and *wnt8b*, were expressed strongly in cells with a neuroectodermal gene expression profile, while *wnt7aa*, *wnt11*, and *wnt11r* were expressed in cells with an epithelial profile. *wnt11* and *wnt11r* were also expressed in the mesodermal/cardiac subpopulation (**Fig 3.10 A**). Interestingly, there was a dramatic shift in expression of Wnt receptors in the NC over the time course of migration (**Fig 3.10 B**). Notably, at 18 hpf, expression of *fzd7a*, highly expressed in early NC cells, was reduced, while expression of other genes, such as *atp6ap2*, was strongly upregulated in cells showing a pigment progenitor gene expression signature. Expression of several other canonical and noncanonical Wnt receptors increased at 20-24 hpf, near the end of NC migration into the mandibular arch, including *ror1*, *fzd3b*, and *fzd2*. The Wnt receptor profile of each NC subtype was even more striking, particularly the pigment-specific expression of *atp6ap2,* skeletal-specific expression of *fzd8a*, and neural/glial-specific expression of *mcama* (**Fig 3.10 B**).

We also performed cell-cell communication network analyses on our NC scRNA-seq datasets using the SoptSC package (Wang et al., 2019), which predicts signaling between cells in scRNA-seq data through expression of pathway components. To facilitate this analysis, we generated an unbiased list of directional NC Wnt targets (up
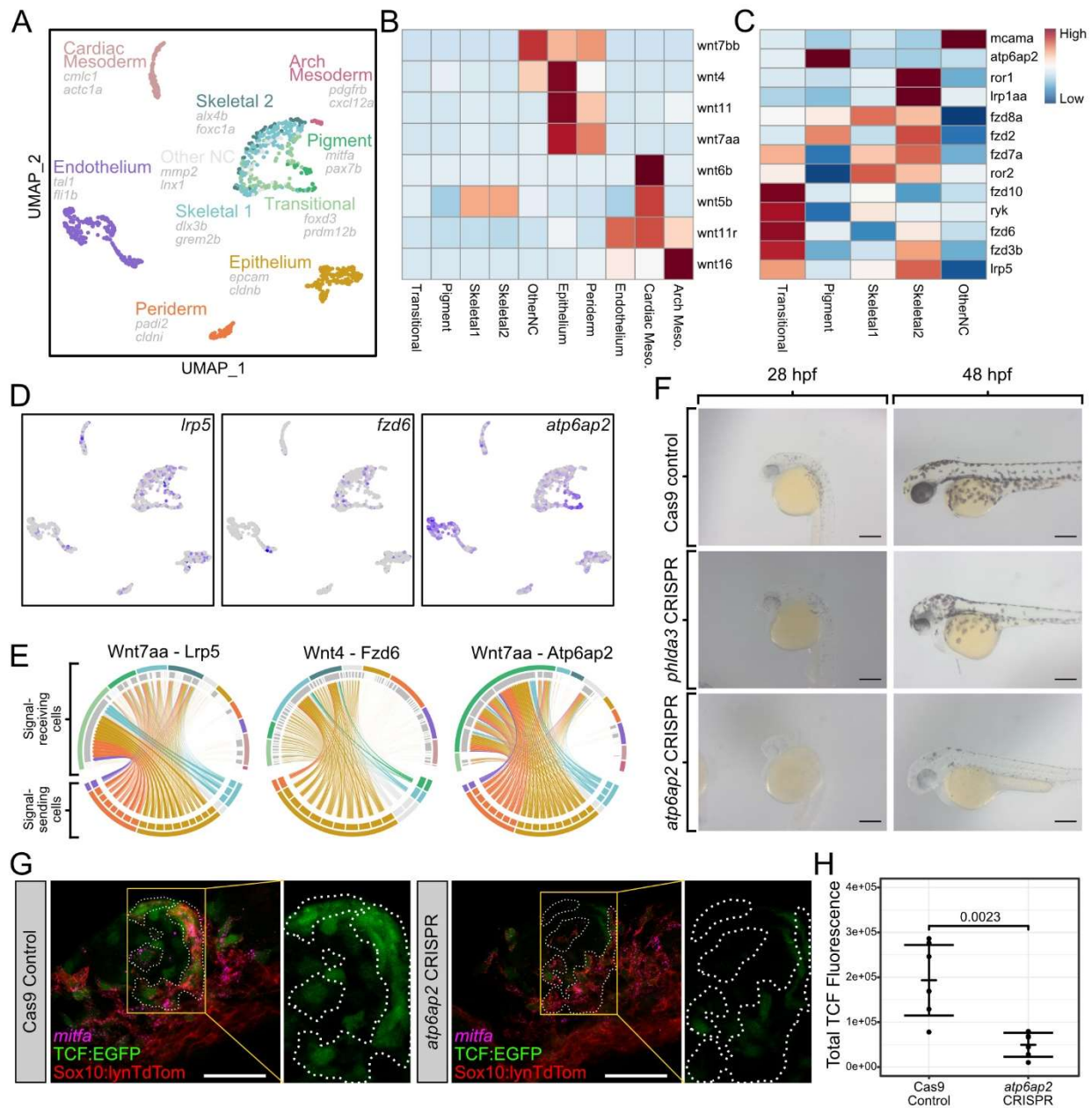
or downregulated) by filtering DE genes from our bulk Wnt-reporter RNA-seq data using the "Wnt signaling" GO term. With the full set of ligands, receptors, and Wnt-regulated genes expressed in our scRNA-seq datasets, we calculated signaling probabilities between each NC cell subtype and surrounding tissues at 14, 18, 20, and 24 hpf. The resulting signaling probabilities suggest changes in the sources of signals as well as

Figure 3.11: Emergence of pigment and skeletal lineages are marked by unique Wnt signals from epithelial tissues. A) UMAP showing all cell types isolated from photoconverted PA1 NC cells. Cell types were determined from top differentially expressed (DE) genes and combined with NC subtype identities from our NC-specific analysis. **B)** Heatmap showing average expression of Wnt ligands in each cell type. **C)** Heatmap showing average expression of Wnt receptors in each cell type. **D)** Feature plots showing expression of Wnt receptors *lrp5*, *fzd6*, and *atp6ap2*. **E)** Circos plots showing imputed signaling interactions. Top segments of each circle indicate signal-receiving cells. Bottom segments indicate signal-sending cells. Each line indicates a signaling event between two individual cells. Lines are colored according to the cell type identity of the signal-sending cell. Colors are consistent with UMAP. **F)** Live embryos injected with either Cas9 alone (top row) or Cas9 and 4 gRNAs targeting *phlda3* (middle row) or *atp6ap2* (bottom row) shown in lateral views at 28 hpf (left column) and 48 hpf (right column). Note reduced black melanocyte pigmentation with atp6ap2 CRISPR. **G)** Confocal micrographs showing Wnt-responsive TCF-EGFP (green, outlined by dashed white lines) and *mitfa* (magenta) expression in PA1 at 24 hpf in *tg(sox10:lynTdTomato;7xTCF:EGFP)* embryos where NC plasma membranes are marked by tdTomato (red). **H)** Quantification of 7xTCF:EGFP using corrected total fluorescence intensity between control embryos (n = 7 embryos, mean = 193407) and *atp6ap2* CRISPR embryos (n = 6 embryos, mean = 49567). Wilcox p-value = 0.0023. Dots represent mean in individual embryos. Lines represent mean within conditions. Error bars represent mean ±SD.

their strength over time (**Fig 3.10 C**). Surprisingly, we also found evidence for Wnt-dependent communication between NC cells, weak at 14 hpf but stronger at 24 hpf.

Because initial specification of at least some lineages appears to be taking place by 18 hpf, we investigated Wnt signaling dynamics at this stage in our larger 18 hpf data set. Analysis of all cells revealed several non-NC cell types expressing Wnt ligands including cardiac mesoderm, endothelium, and epithelium as well as a small arch mesodermal population and a periderm population (**Fig 3.11 A**). Wnt ligands had uneven distribution across cell types. *wnt6b*, *wnt11r,* and *wnt16* were mostly expressed in the mesodermal populations, while *wnt7bb*, *wnt4*, *wnt11*, and *wnt7aa* were enriched in epithelial and peridermal cells (**Fig 3.11 B**). Wnt receptors also had distinct

expression patterns across NC cell subtypes. Several receptors including *fz6* and *lrp5* were expressed strongly in the transitional population but also in pigment and skeletal cells. Pigment cells again had very strong enrichment of *atp6ap2* (**Fig 3.11 C, D**). SoptSC analyses at 18 hpf revealed three notable lineage-specific interactions: signaling from epithelial and peridermal cells to transitional NC cells via *wnt7aa* and *lrp5*, signaling from epithelial cells to skeletal NC lineages via *wnt4* and *fzd6*, and signaling from epithelial and peridermal cells to NC-derived pigment progenitors via *wnt7aa* and *atp6ap2* (**Fig 3.11 E**).

Requirements for *atp6ap2* in pigment cell development

The *atp6ap2* gene encodes a component of the v-ATPase complex, which also functions as an intracellular renin receptor and Wnt co-receptor (Cruciat et al., 2010; Ichihara & Yatabe, 2019). While mutants/knockdowns of *atp6ap2* in zebrafish have been reported to cause biliary defects (EauClaire et al., 2012), it has not been
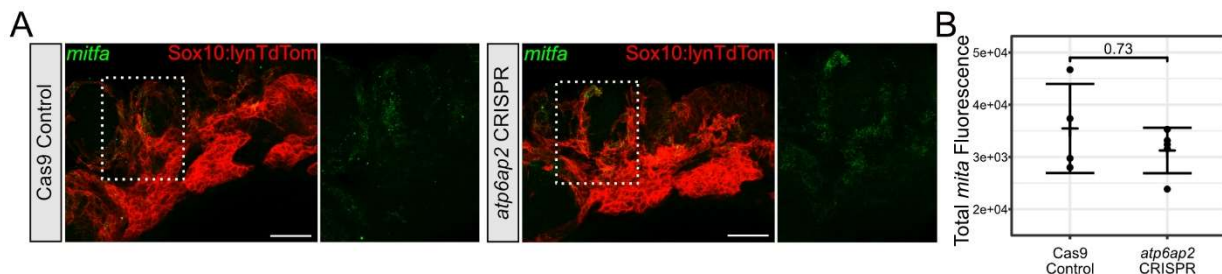
Figure 3.12: isHCR of mitfa reveals pigment progenitors are specified in atp6ap6 crispants. A) Confocal micrographs of isHCR showing *mitfa* expression (green) at 24 hpf in a Cas9-injected control (left panel) and atp6ap2 CRISPR-injected embryo (right panel) in *tg(sox10:lynTdTomato)* transgenics where NC plasma membranes are marked by tdTomato (red). **B)** Quantification of *mitfa* expression using corrected total fluorescence intensity between control embryos (n = 4 embryos, mean = 35447) and *atp6ap2* CRISPR embryos (n = 5 embryos, mean = 31242). Wilcox p-value = 0.73. Dots represent means in individual embryos. Lines represent means within conditions. Error bars represent means ±SD. For micrographs, Scale bars = 50μm.

implicated in NC development. To test *atp6ap2* function we generated targeted

mutations in zebrafish using multiplexed CRISPR-Cas9 genome editing (Wu et al,

2018). F0 embryos injected with gRNAs targeting *atp6ap2* (hereafter called "crispants")

displayed a severe loss of pigmentation compared to siblings injected with Cas9 alone

or Cas9 and gRNAs targeting the skeletal marker *phlda3* (**Fig 3.11 F**). Melanocytes

were both reduced in number and size in embryos at 48 hpf, possibly due to defects in

their specification. To test this hypothesis, we examined *mitfa* expression by isHCR, but

observed no differences in *mitfa* expression between control embryos and *atp6ap2*

crispants (**Fig 3.12 A, B**), suggesting that defects are in later differentiation or survival.

Because *atp6ap2* may act as a Wnt co-receptor, we next investigated if Wnt signaling

was affected in *atp6ap2* crispants using tg(*7xTCF:EGFP*). Crispants showed a marked

decrease in eGFP intensity in *mitfa*+ cells, as well as in other Wnt-responsive regions

(**Fig 3.11 G, H**). Together, these results indicate a previously unknown role for *atp6ap2*

in NC and maintenance of melanocytes through canonical Wnt signaling and

demonstrate the utility of predictive signaling analysis using single-cell data in

identifying novel regulatory genes.

Discussion

NC development involves coordinated fate decisions as cells are migrating to achieve proper formation of a huge variety of cell types. Previous single-cell studies have reported expression profiles of cranial NC lineages but lack detailed spatial information and rely on pseudotime to derive transitional cell states. In this study, we present a single-cell transcriptomic timeline of zebrafish cranial NC development with cells isolated from the mandibular arch migratory stream at closely spaced timepoints during their migration. Thus, the data derive from a known NC subpopulation and capture real-time gene expression dynamics to contextualize inferred lineage connections and clarify the temporal transitions that cranial NC cells undergo as they differentiate. By assimilating multiple inputs and outputs in the NC gene regulatory network using several computational pipelines, we identify a key transitional cell stage mid-migration when heterogeneity in gene expression arises and subpopulations of skeletal and pigment progenitors begin to emerge, followed by neural/glial progenitors. We uncover several novel genes that strongly mark these lineage bifurcations. In addition, through integration of bulk RNA-seq from a Wnt reporter transgenic with our single-cell data, we model lineage-specific Wnt signaling dynamics and demonstrate an essential function for the Wnt-regulatory factor, Atp6ap2, in NC-derived pigment cell development. Together, our results reveal novel transcriptional changes and cell-cell signals that coordinate NC cell migration and fate.

Premigratory cranial NC cells show no lineage bias with pigment, skeletal, and glial lineages specified mid-migration

Previous lineage-tracing studies in a variety of species have provided evidence both for multipotency and lineage restrictions in premigratory NC cells (Baggiolini et al., 2015; Krispin et al., 2010; McKinney et al., 2013; Schilling & Kimmel, 1994). Our scRNA-seq analyses show that premigratory mandibular NC cells in zebrafish have homogenous gene expression signatures. These results suggest either that premigratory NC cells are entirely unspecified toward any particular lineage or that lineage specification is not detectable at the level of transcription prior to migration. The first signs of transcriptional heterogeneity become apparent at 18 hpf, when subsets of cells express multiple markers of pigment and skeletal precursors. Notably, these two sets of precursors are already spatially segregated in the mandibular NC at this stage (**Fig 3.7**). *mitfa+* pigment progenitors are restricted to the dorsal region of the migratory stream, whereas *dlx2a+* skeletal progenitors occupy the ventral region. NC cells with signatures of glial precursors are detected slightly later in migration, at 20 hpf. This demonstrates that lineage decisions and spatial organization are coordinated early in cranial NC migration. This could be explained by the predicted changes in signaling between NC cells and their environment. Investigation into the spatial distribution of signaling molecule expression in vivo would help to elucidate the mechanics governing this apparent coordination.

The emergence of the pigment lineage followed by later specification of glial lineages is somewhat surprising, as the prevailing model describes a bifurcation between these two

lineages occurring simultaneously as the result of competing interactions between *foxd3*

and *mitfa* (Curran et al., 2010)*.* Our results support this transitional overlapping

expression of *mitfa* and *foxd3*, probably representing the interplay whereby *mitfa*

expression inhibits *foxd3* as the pigment lineage is specified (**Fig 3.7**). However, it

seems the glial lineage decision may be driven by a later re-deployment of *foxd3* in a

population of NC cells distinct from this *foxd3+/mitfa+* transitional pigment population

(**Fig 3.4 ; Fig 3.5**). These findings clarify the timing of cNC lineage decisions and

provide context to the bifurcation between pigment and glial cell fates.

Our findings are somewhat in disagreement with previously published single-cell

transcriptomics studies of the NC, which arrived at varying interpretations regarding the

temporal aspects of lineage specification. scRNA-seq data in chick cranial NC identified

a putative "neural" NC population prior to migration characterized by low expression of

NC specifiers *TFAP2A/B* and *FoxD3* and high levels of *Sox3* (Williams et al., 2019). We

found no evidence of an early neural NC progenitor population but did see a similar

expression profile in neural tube (NT) cells that we captured (**Fig 3.2**). Our findings also

differ from single-cell RNA-seq in mouse, which revealed an early bifurcation between

neural and mesenchymal lineages in trunk NC cells, with no obvious specification of

pigment cells (Soldatov et al., 2019). These discrepancies could reflect our

computational integration of the temporal data e.g. parallel analyses of pseudotime

using Monocle 3 and PAGA combined with real-time data. Alternatively, they could

reflect real regional differences in lineage specification among NC subpopulations

(trunk, cranial, and cephalic) or differences between species. A combinatorial analysis

of these currently available single-cell NC data sets with the added context of our temporal single-cell data could help to reconcile these discrepancies and build a more robust understanding of early lineage specification.

Unique deployment of Twist1 paralogs in zebrafish NC development

One striking marker of the skeletogenic lineage in our data is *twist1a*. Twist1 functions as both a NC specifier and regulator of mesenchymal NC cell lineages (Simões-Costa & Bronner, 2015). In mammals and amphibians, early Twist1 expression in the neuroepithelium marks cells that will undergo EMT and give rise to NC, with expression persisting in NC cells that later progress toward skeletal fates (Hopwood et al., 1989; Ishii et al., 2003; Soo et al., 2002). Zebrafish have two Twist1 orthologs, *twist1a* and *twist1b*, which likely arose with the whole genome duplication in teleosts and are often grouped together in the context of NC development (Germanguz et al., 2007). Here we show that they have striking mutually exclusive temporal expression profiles. While *twist1b* strongly marks early NC cells, its expression quickly drops off within the first 6 hours of migration, during which time NC cells with a signature of skeletal progenitors upregulate *twist1a* expression. This strongly suggests that in zebrafish the *twist1* duplicates diverged in function during evolution, with *twist1b* retaining one ancestral function as a NC specifier and *twist1a* another function in driving migration in skeletal lineages, consistent with the gene duplication and divergence hypothesis (Force et al., 1999). This serves to highlight species-specific strategies in NC development.

Distinct Wnt signaling components correlate with NC heterogeneity and a role for Atp6ap2

Our computational analyses of Wnt signaling synthesizes input from bulk- and scRNA-seq data allowing us to predict interesting trends in individual gene expression trajectories and cell-cell communication. This reveals evidence for several lineage- and stage-specific interactions including: 1) epithelial Wnt7aa signaling to Lrp5 in transitional NC cells, 2) epithelial Wnt4 signaling to skeletal progenitors via Fzd6, and 3) epithelial Wnt7aa signaling to pigment progenitors via Atp6ap2 (**Fig 3.11 E**). Wnt signaling has been implicated in multiple aspects of NC induction, delamination and migration, as well as multipotency and cell fate determination (reviewed in Ji et al., 2019; Rocha et al., 2020), including several of the components identified here (Willems et al., 2015). These results highlight the dynamic nature of Wnt signaling in the NC. Further investigation into individual lineage and stage-specific Wnt signals presents an exciting opportunity to uncover new mechanisms driving NC lineage specification.

As evidence of the promise of such an approach, our marker gene and predictive signaling analyses unbiasedly identified *atp6ap2* as a marker and putative regulator of pigment progenitor development potentially through canonical Wnt signaling. This gene encodes the (pro)renin receptor protein ((P)RR), which has numerous cellular functions. (P)RR is a receptor that binds prorenin to activate tissue RAS, but it is also a subunit of the v-ATPase complex and is vital for proper vesicular acidification in the kidney and heart (Ichihara & Yatabe, 2019). Atp6ap2/(P)RR also serves as a co-receptor for LRP6 in canonical Wnt signaling during anterior-posterior patterning of the central nervous

system (CNS), linking LRP6 to the v-ATPase complex to activate the Wnt-β-catenin pathway upon binding of Wnt ligand (Cruciat et al., 2010). Here we demonstrate that *atp6ap2* expression marks NC cells that adopt pigment cell fates and is vital for their survival/maturation. Further, we show that this defect correlates with a reduction in canonical Wnt signaling, supporting the notion that the function of Atp6ap2 in melanocytes resembles its previously identified role in mediating canonical Wnt signaling in the CNS. The loss of pigmentation we observe in *atp6ap2* crispants is consistent with a zebrafish mutant for *atp6ap2* called *pekin,* which exhibits biliary defects as well as hypopigmentation (EauClaire et al., 2012). Wnt signaling is required both for specification of the pigment lineage and for later differentiation of mature melanocytes (Vibert et al., 2017). We previously characterized a role for the v-ATPase associated proteins Rbc3a and Atp6v0a1 in controlling migration of pigment progenitor NC cells through modulation of canonical Wnt signaling (Tuttle et al., 2014). This role for Atp6ap2 in pigment cell development post migration further highlights the multifaceted importance of v-ATPase-mediated canonical Wnt signaling.

Materials and Methods

Zebrafish husbandry and transgenic lines

Embryos were obtained from natural breeding and staged as described in Kimmel et al., 1995. All zebrafish lines were maintained according to standard protocols (Westerfield et al, 2000). Transgenic lines used in this study include Tg(*sox10:nEOS*)[w18] (Curran et al., 2010), Tg(*sox10:lyn-tdTomato*)[ir1040] *Tg(-7.2sox10:EGFP)*[ir937],(Schilling et al., 2010), Tg(*7XTCF:EGFP*)[ia4] (Moro et al., 2012), and Tg(*7XTCF:nls-mCherry*)[ia5] (Moro et al., 2012).

Embryo dissociation and FACS

Transgenic Tg(*sox10:nEOS*) zebrafish embryos were mounted in 1% agarose and imaged on a Nikon C1 confocal. NC cells of the PA1 NC migratory stream were converted by drawing a region of interest (ROI) around the targeted cells and then exposing the ROI to 405 laser light for 10 seconds. Embryos were then dissociated using mechanical disruption and trypsin/collagenase P incubation as described in Barske et al., 2016. Dissociated cell samples were processed on a BD FACS Aria II cell sorter. Photoconverted (red) cells were separated from unconverted (green) cells based on intensity of red/green fluorescence. For Wnt reporter experiments, TCF:mCherry Sox10:LynEGFP double positive cells and Sox10:LynEGFP+ TCF:mCherry- cells were FACS sorted.

10x library construction and sequencing

Single cell suspensions were processed on a 10x Chromium platform for single-cell

library construction. Libraries were then sequenced on a HiSeq2500 (Illumina).

Read mapping and pre-processing

FASTQ files were mapped to the zebrafish transcriptome GRCz11 using CellRanger.

Mapped reads for the 12-30 hpf timeline were aggregated using CellRanger before

further computational analysis. Counts matrix normalization and scaling were performed

using Seurat v3.

Bulk RNA-seq library construction and sequencing

RNA was extracted from cell lysates using the RNEasy Micro Kit (Qiagen). cDNA

libraries were generated following the Smart Seq II protocol (Picelli et al, 2014).

Libraries were then sequenced on a NovaSeq6000 (Illumina) at a depth of ~20M reads

per sample.

Bulk RNA-seq data processing and analysis

Reads were mapped to zebrafish genome GRCz11 and quantified using STAR v2.5.2a

and RSEM v1.2.31. Differential gene expression analysis was carried out using R and

the edgeR package (Robinson et al., 2010). TCF:mCherry+ NC cells were compared to

TCF:mCherry- NC cells, and FDR <0.05 was used as a cutoff for significant differences

in expression level. Heatmaps were generated using ComplexHeatmap R package.

Standard computational analysis of single-cell RNA-seq data

Variable features identification, principal component analysis (PCA), UMAP reduction, and unbiased clustering were performed in Seurat v3. Differential expression analysis in Seurat v3 was used to identify cluster markers by Wilcoxon rank sum test. Cluster markers were then used to assign cell types. Visualizations were generated using Seurat v3 and ggplot2 with color palettes from the RColorBrewer and Viridis packages. A subset counts matrix containing TFs was analyzed with Monocle3 for trajectory and pseudotime analysis. Identification of temporally distinct TF modules was performed using the graph_test and find_gene_modules functions in Monocle3.

Multiplex CRISPR-Cas9 genome editing

Perturbations of atp6ap2 were achieved using the multiplex CRISPR technique described by Wu et al., 2018. Four unique gRNAs targeting the coding region of *atp6ap2* were synthesized using T7 MegaShortScript kit (Ambion). gRNAs were incubated with Cas9 protein (IDT) at 37°C for 5 minutes and injected into embryos at the 1-cell stage in amounts of 150 pg gRNA and 800 pg Cas9 protein per embryo. gRNA primer sequences were taken from the database provided by Wu et al, 2018.

In situ hybridization chain reaction

HCR probes were ordered from and designed by Molecular Technologies (Los Angeles, CA). Whole mount HCR was carried out as described by Choi et al., 2014.

Quantitative analyses of HCR and Wnt reporter

Raw fluorescence intensities were measured using ImageJ and corrected total cell fluorescence (CTCF) values were calculated using the formula CTCF = Integrated Density – (Area x Mean background fluorescence). Mean intensity values between timepoints and/or conditions were compared using Wilcoxon signed-rank test. For *foxd3* measurements, means were compared using Kruskal-Wallis ANOVA and then pairwise comparisons were performed using Wilcoxon signed-rank tests. Statistical tests were carried out using R, and plots were generated in R using the ggplot2 package. Other R packages used include: plyr, dplyr, reshape2, and ggpubr.

Clustering of DE genes via Dirichlet process Gaussian process (DPGP) modeling

For temporal profiling of DE genes from Wnt reporter bulk RNA-seq, genes were clustered using scRNA-seq data for cells from either the pigment branch or the non-pigment branch separately with DPGP (McDowell et al., 2018). Briefly, the DPGP model assumes that the temporal expression of a gene is generated from a mixture of Gaussian processes, such that each Gaussian process corresponds to a gene expression cluster, the parameters of which are further generated by a Dirichlet process. The concentration parameter α for the Dirichlet process was set to 0.2. To obtain temporal expression patterns for each gene, average counts of all cells in a branch for all time points were calculated (except 30 hpf for the pigment branch), which constituted a time series for that gene. DPGP was allowed to infer the optimal clustering that attained maximum a posteriori (MAP) likelihood, given those time series of average gene expression.

PAGA dimensionality reduction and trajectory inference

The PAGA_tree (Wolf et al., 2019) wrapper in the Dyno (Saelens et al., 2019) package was used to infer developmental trajectories. The dataset was first processed with the default option, recipe_Zheng17 in Scanpy. The method PAGA_tree was then applied with the parameters, n_comps=50, n_neighbors=15, resolution=1, n_dcs=10. "AAACCTGAGTGGCACA" in the 12 hpf dataset was chosen as the root cell. A simplified model was obtained using Dyno based on which pseudotime was also computed in Dyno. PAGA coarse graph was used to initialize single-cell embedding in ForceAtlas2 (Jacomy et al., 2014).

Probabilistic cell type assignment

Cell Types were assigned based on known markers and novel TFs derived from TF pseudotemporal analysis (**Fig 3.2**) in a supervised approach using the package CellAssign (Zhang et al., 2019). The exact marker genes are listed in the individual csv files. The function cellassign was used with learning_rate=0.02, shrinkage=TRUE, and a size factor estimated using the computeSumFactors function in scran package (Lun et al., 2016).

Cell-cell communication analysis in scRNA-seq data

Communication between cells was predicted from our single-cell RNA-seq data using the SoptSC suite of computational tools (Wang et al., 2019). Briefly, the probability of communication between individual cells is calculated for individual ligand-receptor pairs based on their expression of these factors and expression of downstream target genes.

Analysis was performed in R using the R package for SoptSC https://github.com/mkarikom/RSoptSC/.

RNA velocity

RNA velocity for 18 hpf was computed using the Scvelo package (Bergen et al., 2020) using default parameters.

Chapter IV

Single-cell RNA-seq analysis of tailbud development reveals a hemangioblast population and a role for cell cycle regulation in midline progenitor fate specification

Introduction

In vertebrate development, many structures in the elongating body axis derive from the tailbud, a structure at the caudal-most tip of the embryo. The tailbud contains a group of self-renewing multipotent cells called neuromesodermal progenitors (NMPs), which contribute to both the neural tube (NT) and paraxial mesoderm (PM) and lateral plate mesoderm (LPM) (Griffith et al., 1992; Tzouanacou et al., 2009). NMP identity is maintained by co-expression of the transcription factors *sox2* and *brachyury* (*tbxta*), and the interaction of these factors with canonical Wnt signaling determines the bifurcation between neural and mesodermal lineages (Koch et al., 2017; Martin & Kimelman, 2012). *Tbxta* in combination with high levels of canonical Wnt signaling drives NMPs to a mesodermal identity (Martin & Kimelman, 2008, 2010). In zebrafish, the development of this lineage also depends on *tbx16* and *msgn1* (Bouldin et al., 2015; Manning & Kimelman, 2015). Less is known about the role of *sox2* in this process, but evidence indicates sustained *sox2* expression and low canonical Wnt signaling together prevent entry of NMPs into the mesoderm and instead promote a neural fate (Kinney et al., 2020; Koch et al., 2017; Takemoto et al., 2011).

Despite many years of research into the origins and derivatives of NMPs, some questions remain about their contributions to the tissues of the vertebrate body,

complicated by differences between species. In zebrafish, NMP-derived mesodermal

progenitors are further bifurcated into paraxial mesoderm and lateral plate mesoderm

(Martin & Kimelman, 2012). However, in mouse this does not appear to be the case.

Wnt3a mutants have severe loss of paraxial mesoderm due to loss of *brachyury*

expression in NMPs (Yamaguchi et al., 1999), yet lateral plate and intermediate

mesoderm populations are unaffected (Gouti et al., 2014). Further, there is evidence in

zebrafish that NMPs may contribute to vascular endothelial tissues. The zebrafish

vascular mutant *cloche* has nearly complete loss of endothelial and vascular tissues,

but maintains a population of apparent vasculature progenitors in the tail (Liao et al.,

1997; Thompson et al., 1998). It is possible therefore that in the zebrafish tailbud,

vasculature progenitors may derive from the NMPs.

Another active area of investigation in tailbud biology concerns the potential presence of

a hemangioblast population. This progenitor cell type, which gives rise to both vascular

endothelium and blood has long been theorized but difficult to identify in vivo (Lacaud &

Kouskoff, 2017; Nishikawa, 2012). Clonal expansion of cells from the primitive streak in

mouse and in vivo lineage tracing of ventral mesoderm in zebrafish reveal the

emergence of both vascular and blood lineages, leading to the conclusion that a

hemangioblast population likely derives from this mass of mesodermal progenitors

(Huber et al., 2004; Vogeli et al., 2006). Recently, using a CRISPR knock in zebrafish

line driving GFP in the *etv2* locus, it was observed that both vasculature and blood cells

were labeled (Chestnut & Sumanas, 2020). Further, single-cell RNA-seq of FACS

sorted GFP+ cells in heterozygous and homozygous *etv2* knock outs from this line

revealed an increase in red blood cells (RBCs) and LPM. They further observed by FACS an increased number of *gata1+* cells in *etv2* knockdowns (Chestnut et al., 2020). Together, these observations point to a connection between hemangioblasts and tailbud-derived LPM, but this has not been demonstrated.

Do NMP derivatives give rise to both vasculature and blood? If so, are these separate populations, or are there mechanisms driving divergence of bipotent progenitors? The lingering questions surrounding tailbud lineage relationships present an exciting opportunity to employ modern single-cell omics technologies. Recently researchers have begun to examine these questions through scRNA-seq analysis of in vitro cultured NMPs and specific FACS sorted NMP-derived lineages. These studies, however, are limited in their ability to examine lineage connections by the lack of a complete picture of the many cell types in the tailbud. Here we present a complete single-cell transcriptomic atlas of the developing tailbud in zebrafish and use it to explore the relationships between NMPs are their derivatives as well as mechanisms driving the emergence of notochord, floor plate, and hypochord lineages.

Results

Single-cell RNA-seq reveals the cellular landscape of the tailbud

To profile the many cell types and progenitor cell states in the developing tailbud comprehensively, we generated a single-cell transcriptomic atlas of the tailbud in zebrafish. 30 Tailbuds were isolated from 16 hpf embryos by manual dissection and then dissociated into single cells by trypsin and collagenase P digestion. Single-cell libraries were then constructed on the 10X Chromium platform and sequenced on an Illumina HiSeq2500 (**Fig 4.1 A**). Following filtering for low quality cells, 11,138 total single-cell transcriptomes were obtained with an average of about 3500 genes detected per cell.

Cell cycle effects were regressed, and then dimensionality reduction and clustering were performed, revealing 13 major cell types based on their expression of canonical marker genes (**Fig 4.1 B-C**). Epidermal cells were evident by their expression of *epcam* and *krt8*, distinct from peridermal cells that express *krt17* and *anxa1c*. Neural tube (NT) cells expressed *sox19a* and multiple *her* genes, with more mature neuroblasts expressing *elavl3*. Axial mesoderm (AM)/notochord cells were easily identified by expression of *noto* and *shha*. Two distinct populations of paraxial mesoderm (PM) could be identified, which we dubbed "anterior" (more mature) and "posterior" (less mature). They shared high expression of *msgn1* and *tbx16*, but the posterior PM had much higher expression of *ripply1*, and the posterior PM had high *vox* and *ved* expression. NMPs formed a cluster of cells that was located closely between the NT and PM
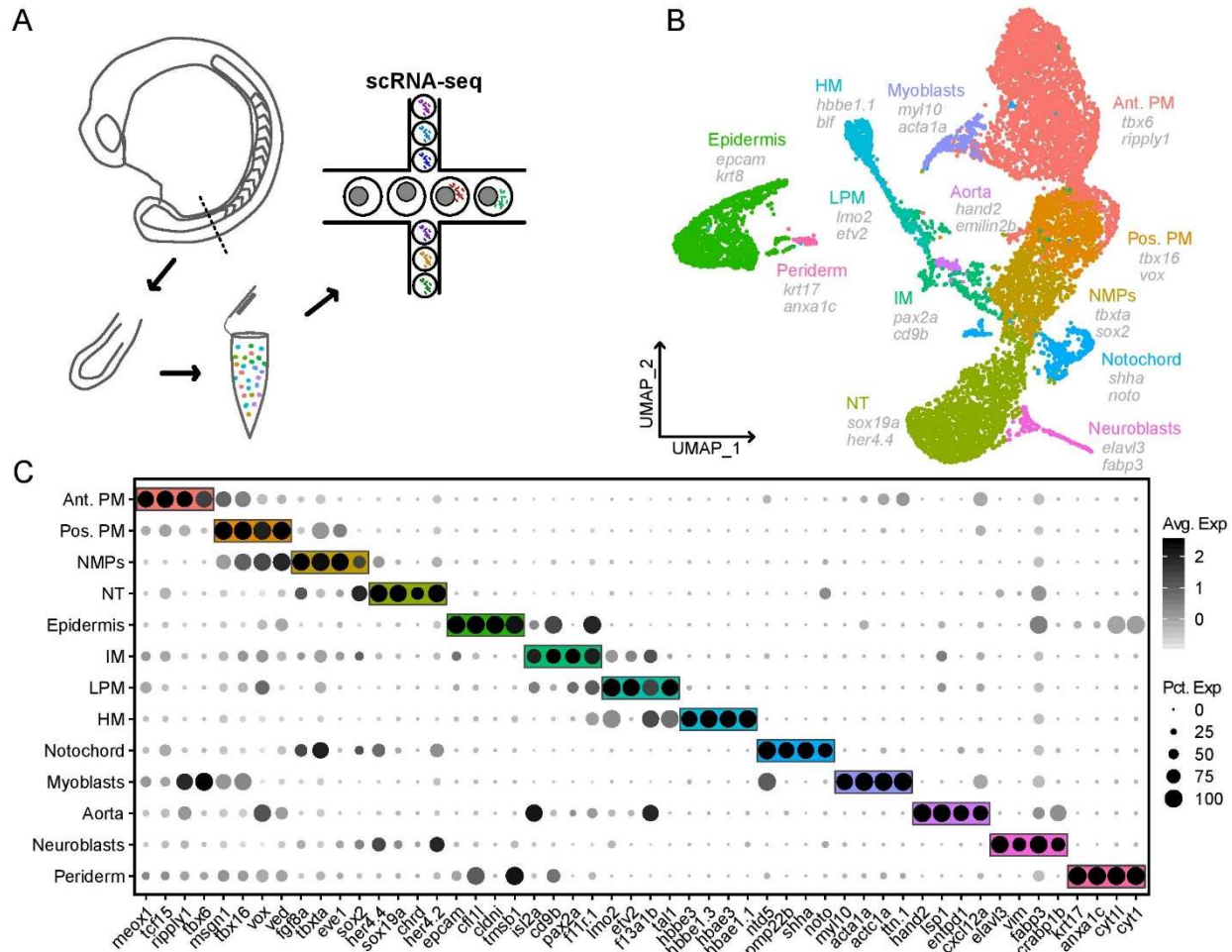
Figure 4.1: scRNA-seq of the developing vertebrate tailbud identifies major cell types. A) Schematic showing procedure for tailbud single cell isolation and sequencing. **B)** UMAP embedding of 11,138 single cells sequenced from the tailbud. Cell types are labeled, and two markers for each cell type are included below labels. **C)** Dotplot showing top 4 marker genes in each cell type by differential expression analysis. Colored boxes correspond to cell type. Greyscale indicates average expression and dot size indicates percent of expressing cells within each cell type. Marker genes were determined by Wilcoxon rank sum test.

populations in UMAP space. These cells were characterized by high expression of

sox2, tbxta, and *fgf8a*. A population of hematopoietic cells was evident by expression of

multiple *hbbe* genes. Two other populations of mesoderm were identified as

intermediate mesoderm (IM) and lateral plate mesoderm (LPM). Despite similar

expression profiles these two populations were distinguished from each other by the

LPM-specific expression of *lmo2* and *etv2* and the IM-specific expression of *pax2a*. Finally, a small population of dorsal

aorta cells were marked by unique expression of *hand2* and *emilin2b*.

Pseudotime analysis suggests lineage connections between NMPs and

hematopoietic lineages

Because epidermal lineages are distinct from the other major tailbud tissues in their

ontogeny, we subclustered all but the epidermal and peridermal lineages. The resulting

dimensionality reduction maintained the overall topology seen in the whole tailbud (**Fig**

4.2 A). Because the tail grows throughout body elongation during embryogenesis and

cells continuously differentiate from their progenitor cell niches in the tailbud, it was

likely that we captured cells at intermediate stages of differentiation. Therefore, we

performed pseudotime analysis in Monocle3 to infer differentiation trajectories (**Fig 4.2**

B). The resulting trajectories confirm the expected relationships between NMPs and

both neural and PM lineages but also predict a connection between NMPs and IM/LPM

lineages (**Fig 4.2 A-B**). We then used a comprehensive list of TFs adapted from

Vaquerizas et al., (2009) to identify sets of co-regulated TFs that correspond to the

emergence of the observed lineages. Among the top 150 TFs that differed in their

expression across pseudotime, we uncovered 7 modules of that correspond to specific

lineages (**Fig 4.2 C-D**).

The connection between NMPs and IM and LPM lineages was somewhat surprising, so

we investigated this connection further by subclustering those cell
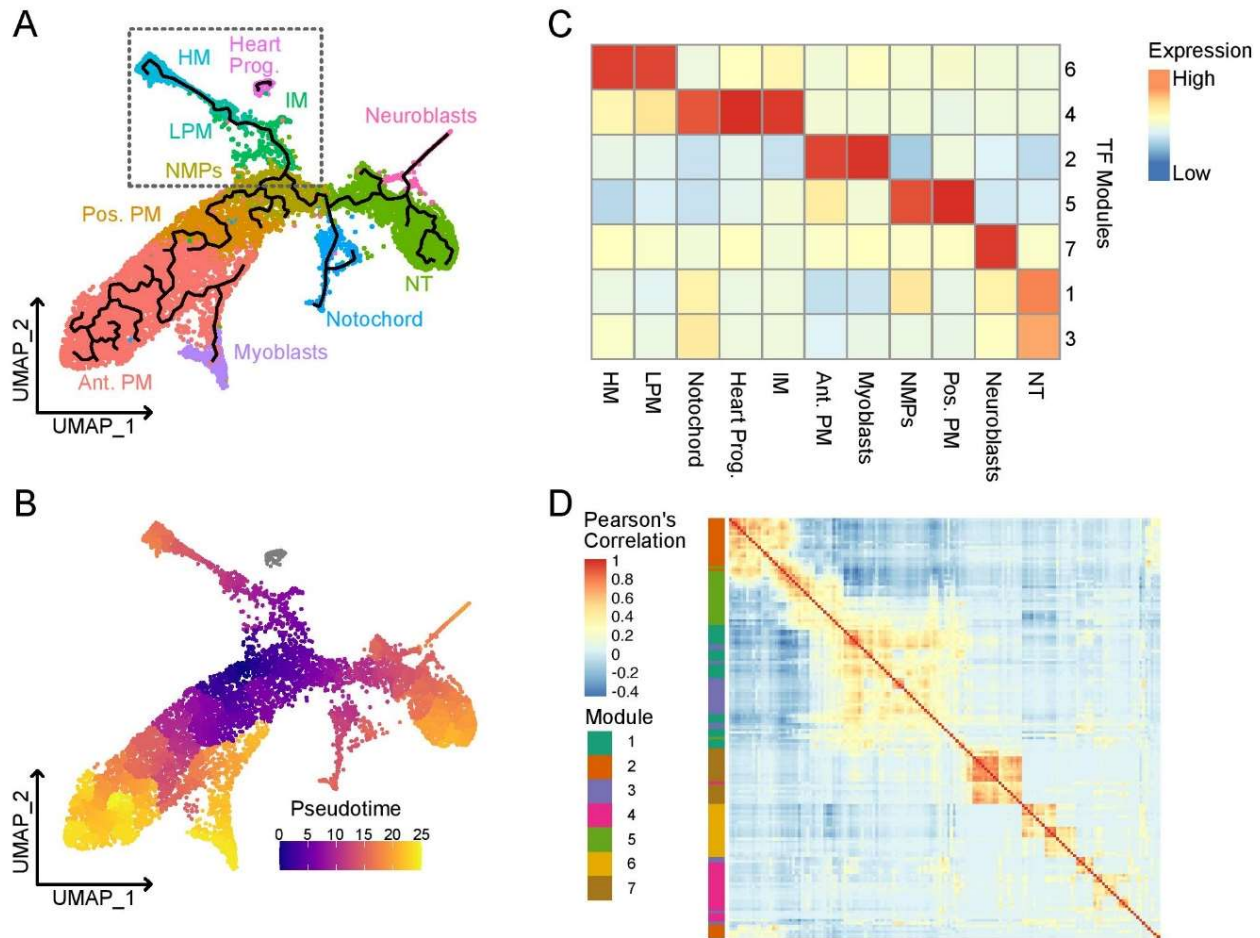
Figure 4.2: Pseudotime analysis of all non-epidermal cells uncovers connections between neuromesodermal progenitors and neural and mesodermal lineages. A) UMAP embedding of all non-epidermal cells. Trajectory computed using Monocle3 predicts connection between NMPs and neural tube, axial (notochord), paraxial, intermediate, and lateral plate mesoderm. **B)** Pseudotime values calculated in Monocle3 using differentiation trajectory with NMPs as the starting node. **C)** Heatmap showing average expression of differentially expressed transcription factors across pseudotime trajectory. 7 modules of TFs were identified with specific expression patterns in particular lineages. **D)** Heatmap showing pairwise Pearson correlation between each of the top 150 differentially expressed TFs. Several clusters of highly correlated TFs correspond to the identified modules.

populations (**Fig 4.3**). We identified two smaller subpopulations of cells that were not

recovered in the larger data set analysis: IM/pronephros, marked by *epcam*, and

Kuppfer's Vesicle, marked by *dand5* (**Fig 4.3 A-B**). Trajectory and pseudotime analysis

reproduced the connections observed in the larger data set, connecting a population of

NMPs to the IM and LPM lineages, and then to the hematopoietic cells. (**Fig 4.3 C-D**).

We analyzed gene expression across pseudotime and discovered 6

Figure 4.3: Subclustering of NMPs and non-axial or paraxial mesodermal lineages uncovers pronephros and Kupffer's Vesicle cells and characterizes novel lineage relationship between NMPs and hematopoiesis. A) UMAP showing dimensionality reduction on a subset of cells including NMPs, all non-paraxial mesoderm, and the aorta. **B)** Dotplot showing top 4 marker genes in each cell type by differential expression analysis. Colored boxes correspond to cell type. Greyscale indicates average expression and dot size indicates percent of expressing cells within each cell type. **C)** Trajectory computed in Monocle3 overlayed with UMAP colored by unbiased clustering. **D)** Pseudotime values computed in Monocle3 using the differentiation trajectory with NMP cluster 4 as the starting node. **E)** Heatmap showing average expression in each unbiased cluster for differentially expressed gene modules identified in Monocle3. **F)** Gene signature scores for gene modules overlayed with UMAP embedding. **G)** Local regression plots showing moving average expression for *id3*, *notch1a*, *gata1a*, and *pax2a* across pseudotime for clusters 4, 6, 5, 3, and 1, which represent the trajectory leading from NMPs to HM through IM and LPM. **H)** Heatmap with cells from NMP to HM trajectory ordered by pseudotime showing expression of genes that mark a smooth transition between cell types. Module identities for each gene are displayed on the left.

modules of genes that correspond to stages across the supposed differentiation

trajectory (**Fig 4.3 E-F**). Among these genes were several factors that have known roles

in specification of mesodermal lineages. Expression of *id3* and *notch1a* both drop off as

cells progress through the IM and LPM phases of the trajectory. *Pax2a* increases in

early IM cells and then decreases in later IM and then further decreases in LPM and

hematopoietic cells as *gata1a* increases drastically (**Fig 4.3 G**). When cells are

arranged by pseudotime, expression of key markers for each lineage form a smooth

gradient that

implies a transition from NMP through IM and LPM to an eventual hematopoietic fate

(**Fig 4.3 H**).

To address outstanding questions regarding the fate of the MPCs that give rise to

notochord, hypochord, and floor plate, we performed analysis on the population of cells
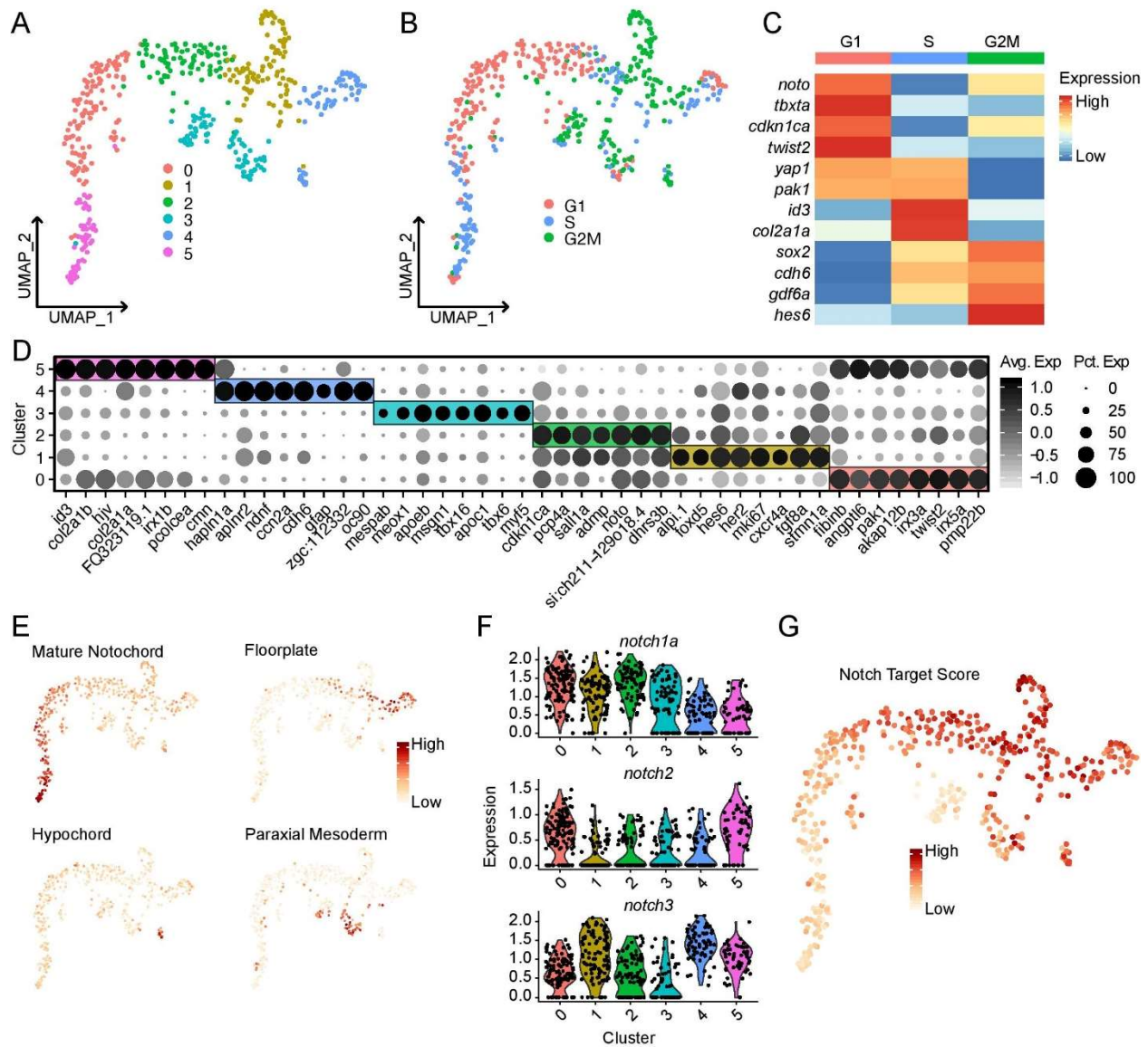
Figure 4.4: Analysis of notochord population reveals hypochord and floor plate progenitor populations and identifies lineage-specific cell cycle states. A) UMAP embedding of all cells assigned a notochord identity in the full tailbud analysis. Unbiased clustering resulted in 5 distinct clusters. **B)** UMAP colored by cell cycle states determined by cell cycle scoring. G1, S, and G2/M stages are not evenly distributed among clusters. **C)** Heatmap showing average expression among cells in each cell cycle stage for several factors driving development of MPC-derived tissues. **D)** Dotplot showing top 4 marker genes in each cell type by differential expression analysis. Colored boxes correspond to cell type. Greyscale indicates average expression and dot size indicates percent of expressing cells within each cell type. **E)** Gene scoring for canonical markers for each MPC-derived cell type and paraxial mesoderm. **F)** Violin plots showing expression distribution among cells in each cluster for *notch* genes. **G)** Gene scoring for Notch signaling targets

assigned the notochord (AM) cell type identity. Following dimensionality reduction and clustering, we found 5 clusters of cells with distinct expression profiles (**Fig 4.4 A,D**). Interestingly, while the clusters separated somewhat according to cell cycle state (**Fig 4.4 B**), based on distinct patterns of gene expression, these cell cycle signatures also corresponded to expression of notochord and floor plate specifier genes (**Fig 4.4 C**). Cells in G1 phase expressed genes associated with an immature notochord or notochord progenitor state including *noto* and *tbxta*. Cells in S phase strongly resembled mature notochord cells, with expression of *col2a1a* and *id3*. Cells in G2/M phase showed expression of floor plate markers including *cdh6* and *hes6*, which also imply active Notch signaling. These expression profiles suggested the cell cycle phases may be indicative of differential acquisition of a notochord or floor plate identity, so we examined this directly by computing gene scores (Tirosh et al., 2016) across the data set for notochord, floor plate, and hypochord cell type signatures using a list of canonical markers for each (**Fig 4.4 E**). Clusters 0 and 4 showed enrichment for markers of mature notochord, while parts of clusters 1 and 4 were enriched for floor plate markers. A small population assigned to cluster 4 was enriched for a hypochord signature. Cluster 3 showed enrichment for paraxial mesoderm, suggesting these cells may have been mis-identified in the larger clustering analysis. To further investigate the identity of the putative floor plate subpopulation, we examined Notch signaling. Expression of *notch1a* was relatively broad across the cell types, and *notch2* expression was low in the floor plate population. However, *notch3* was expressed highly in the floor plate population (**Fig 4.4 F**). Further, gene scoring for canonical Notch signaling targets, revealed strong enrichment in the putative floor plate subpopulation. Taken together,
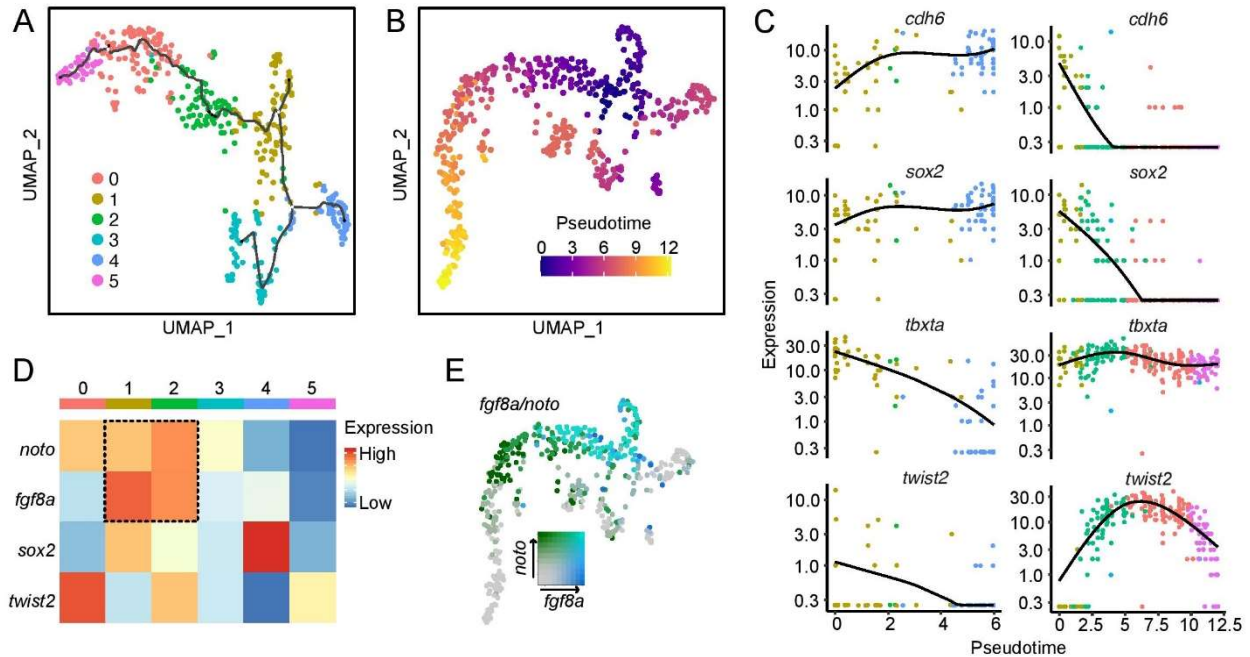
Figure 4.5: Pseudotime analysis details emergence of notochord and floor plate and identifies a molecular signature of MPCs. A) UMAP embedding and trajectory computed in Monocle3. **B)** Pseudotime values computed from differentiation trajectory with putative MPC population (cluster 1) as the starting node overlayed with Seurat UMAP. **C)** Local regression for markers of floor plate and notochord markers across the floor plate trajectory branch (cluster 1 to cluster 4) and notochord trajectory branch (cluster 1 to cluster 5). Line indicates moving average expression. **D)** Heatmap showing average expression within clusters for floor plate and notochord markers as well as *fgf8a* and *sox2*. **E)** UMAP overlayed with expression for both *fgf8a* and *noto*. Overlapping expression observed in cluster 1.

these analyses indicate that despite all these cells initially being assigned a notochord identity in the full tailbud analysis, we did indeed capture all three MPC-derived tissues.

To better understand the connections between these MPC-derived cell types, we next performed trajectory and pseudotime analyses. Dimensionality reduction and trajectory inference in Monocle3 predicted the expected connections between clusters (**Fig 4.5 A**). The trajectory also implied a branching point between floor plate and notochord fates originating from cluster 3, so pseudotime was calculated with this cluster as the starting node (**Fig 4.5 B**). Floor plate markers *cdh6* and *sox2* increased along the trajectory

103

branch leading to cluster 4, while in the trajectory leading to cluster 5, expression of these genes drops off rapidly. In contrast, *cdh6* and *sox2* decrease in the cluster 5 branch, while *twist2* increases and decreases by the end of the trajectory (**Fig 4.5C**). Next we investigated gene expression within the putative MPC-population which gives rise to both of these branches. Strikingly, this population is uniquely marked by overlapping expression of *fgf8a* and *noto,* representing a potential novel signature of the MPCs.

Discussion

Possible mechanistic requirement for cell cycle control in the specification of MPC-derived tissues

It has been well established that notochord, hypochord, and floor plate arise from a common precursor cell in the tailbud called the midline progenitor cell (MPC) (Catala et al., 1995, 1996; Selleck & Stern, 1991). Recently, in zebrafish it was shown that the emergence of these tissues involves sequential lineage bifurcations, with two MPC-derived bipotent progenitors giving rise to notochord/floor plate and
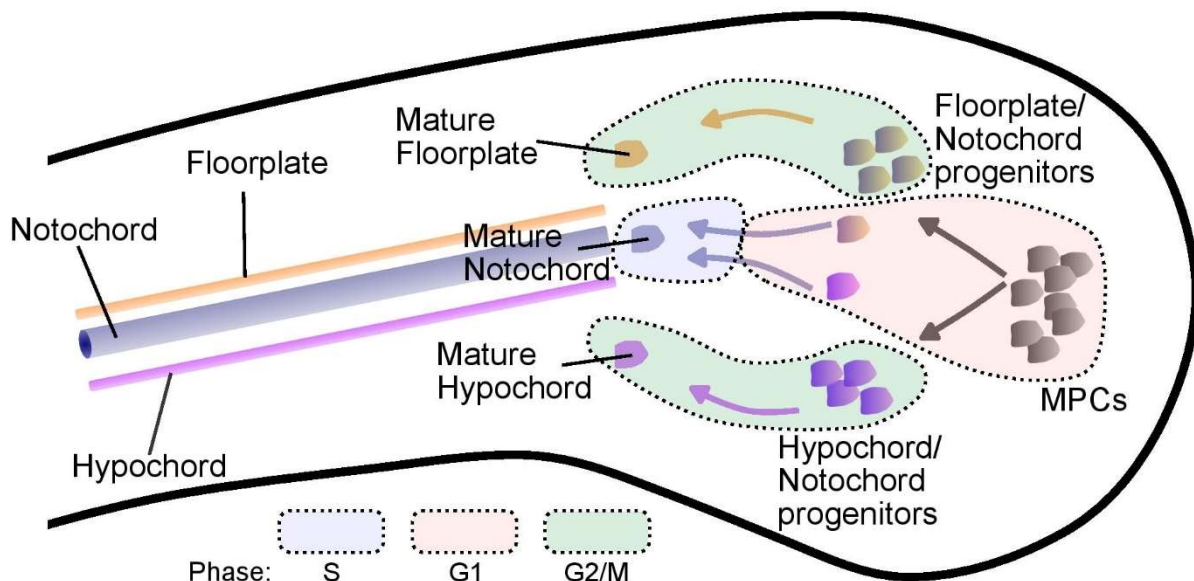
Figure 4.6: Lineage relationships between tailbud midline progenitor cells (MPCs) derivatives and corresponding cell cycle states. Diagram showing the lineage bifurcations leading to emergence of notochord, hypochord, and floorplate from MPCs. MPCs give rise to two bipotent populations of cells which then bifurcate further. Dashed outlined shape fill colors correspond to cell cycle states. The trajectory giving rise to notochord is characterized by an extended cell cycle arrest in G1 phase before late entry into S phase. Trajectories giving rise to hypochord or floorplate quickly enter G2/M phase.

notochord/hypochord, respectively (Row et al., 2016). The lineage decisions of these progenitors are driven by differential levels of Wnt and Notch signaling. Consistent with

this, our data indicate that Notch signaling precludes a notochord fate in this population, with cells enriched for Notch targets progressing to either hypochord or floor plate (**Fig 4.4**). Another interesting aspect of this process is that midline progenitors appear to undergo a cell cycle arrest which keeps them in G1 phase before they progress to a notochord fate (Sugiyama et al., 2009, 2014). While this has been observed in vivo in zebrafish, how this arrest is related to the transcriptional changes driving specification of MPC-derived lineages is not known. Our data show that the putative MPC population is indeed almost exclusively in G1 phase. Further, the trajectory leading to a notochord fate remains in G1, while the trajectory leading to floor plate/hypochord quickly enter G2/M phase (**Fig 4.6**). While these findings alone do not offer a mechanistic explanation for the role of cell cycle arrest in MPC lineage specification, they do indicate that these processes are tightly correlated. Further, they may imply that Notch signaling plays some role in the cell cycle progression that drives the non-notochord fates. In vivo perturbation of cell cycle regulation in MPCs will help to deconvolve the relationships between the cell cycle, Notch, and lineage specification.

Evidence for an NMP-derived hemangioblast population in the zebrafish tailbud

For many years, researchers have studied the possibility of an embryonic hemangioblast population of cells capable of giving rise to both vasculature and blood cells (Lacaud & Kouskoff, 2017; Nishikawa, 2012). However, the hemangioblast has been difficult to characterize in vivo, and the ontology of this population is not clear. Our data show a clear connection between NMP-derived vascular progenitors (LPM) and a population of hematopoietic cells marked by multiple hemoglobin genes. Indeed, the

106

progression from NMP to hematopoiesis in our data is marked by an initial upregulation of *etv2* followed by a subsequent downregulation and an increase in *gata1a* expression. This progression is also characterized by increased BMP signaling, as marked by *id3* expression, which then decreases (**Fig 4.3**). Therefore, cells in the LPM progenitor pool that downregulate *etv2* would go on to give rise to red blood cells, while those maintaining *etv2* expression would likely give rise to vascular endothelium.

This interpretation is consistent with both our data and the recent observations by Chestnut et al and strongly supports the notion that NMP-derived cells give rise to both vascular endothelium and red blood cells. It seems, therefore that there is a hemangioblast population in zebrafish which is in fact a common progenitor pool derived from NMPs that progresses through an LPM-like identity before acquiring either a vascular endothelial or hematopoietic cell fate based at least partially on expression of *etv2*. This population would be distinct from hemangioblasts that derive from gastrulation-specified LPM. Our data also suggest that this decision is likely controlled by BMP signaling. Experiments in vivo should be performed to verify the validity of this finding by directly testing the ability of NMP-derived LPM progenitors to form blood with the proper levels of BMP signaling and repression of *etv2*.

Materials and Methods

Tailbud dissociation and 10x library construction and sequencing

Wild type AB zebrafish embryos were obtained by natural breeding. Tailbuds were

obtained by manual dissection at 16 hpf. Tailbuds were then dissociated into single cell

suspension using mechanical disruption and trypsin/collagenase P incubation as

described in Barske et al., 2016.

10x library construction and sequencing

Single cell suspensions were processed on a 10x Chromium platform for single-cell

library construction. Libraries were then sequenced on a HiSeq2500 (Illumina).

Read mapping and preprocessing

FASTQ files were mapped to the zebrafish transcriptome GRCz11 using CellRanger.

Mapped reads for the 12-32 hpf timeline were aggregated using CellRanger before

further computational analysis. Counts matrix normalization and scaling were performed

using Seurat v3.

Computational analysis and pseudotime computation of single-cell RNA-seq data

Variable features identification, principal component analysis (PCA), UMAP reduction,

and unbiased clustering were performed in Seurat v3. Differential expression analysis in

Seurat v3 was used to identify cluster markers by Wilcoxon rank sum test. Cluster

markers were then used to assign cell types. Visualizations were generated using

Seurat v3 and ggplot2 with color palettes from the RColorBrewer and Viridis packages.

A subset counts matrix containing TFs was analyzed with Monocle3 for trajectory and pseudotime analysis. Identification of temporally distinct TF modules was performed using the graph_test and find_gene_modules functions in Monocle3.

Chapter V

Conclusion and Future Directions

Cell fate determination is a fundamental aspect of embryonic development. The decision of a multipotent progenitor cell to adopt one fate over others can involve many layers of transcriptional regulation and signaling dynamics. Moreover, the development of complex tissues involves lineage decisions made not only on the level of individual cells, but also on the level of large groups of cells that must simultaneously coordinate intercellular signals and dynamic physical rearrangements. The incredible complexity of this process makes cell fate determination exceedingly difficult to study comprehensively. To build a more complete understanding, a multifaceted approach is needed, utilizing both traditional genetic and molecular techniques combined with the unbiased approaches of modern -omics technologies. Such investigations allow for the examination of entire aspects of cellular biology (eg: transcriptome, proteome, epigenome), which can reveal previously-unseen heterogeneity and/or plasticity in lineage specification that would almost certainly be missed by investigation driven entirely by *a priori* knowledge. And further, these unbiased approaches are at their most powerful when used to inform new directed investigations.

In this thesis, I have presented my work using such an approach to address long-standing questions regarding cell fate determination in neural crest (NC) biology as well as an entirely data-driven analysis of lineage specification in the tailbud. In chapter II, my functional characterization of Lmo7a in NC migration and cell fate decisions links focal adhesion dynamics with canonical Wnt signaling, presenting a novel mechanism

for coordination of cell fate decisions and migration. In chapter III, I leveraged in vivo imaging along with the integration of bulk and temporal single-cell RNA-sequencing to uncover the precise timing of skeletal, pigment, and neural/glial fate decisions in the cranial NC. These analyses also uncovered a novel role in NC canonical Wnt signaling and pigment cell development for the (pro)renin receptor (*atp6ap2*). In chapter IV, I generated the first complete single-cell transcriptomic map of cellular diversity in the tailbud and investigate the mechanisms regulating lineage connections between several cell types. Each of these studies on their own provide novel insights into mechanisms driving lineage specification in these embryonic multipotent progenitors and present exciting opportunities for further investigation.

Together, these studies highlight the strength of applying modern -omics techniques combined with genetics and molecular biology to questions of lineage specification. Both the NC and tailbud multipotent progenitors are examples of cells that must coordinate cell fate decisions within a rapidly changing environment as the embryo develops and in the case of NC cells, as they migrate. In both cases, traditional concepts of progressive lineage bifurcations appear to require a more nuanced interpretation. My work uncovering the role of Lmo7a in NC migration and lineage specification suggests the decisions between neural/glial and pigment lineages involve interactions between adhesion factors and Wnt signaling components, the loss of which can result in aberrant bipotent cell states. My scRNA-seq timeline of NC development uncovered putative transitional states that precede lineage bifurcation and correlate with spatial reorganization within the migratory stream. Finally, my work in constructing a

single-cell atlas of transcriptional diversity in the tailbud revealed multiple transitional

states whose ultimate fate decisions require input from major signaling pathways like

BMP as well as regulation of the cell cycle. In each of these instances, the unbiased

large-scale investigation afforded by modern -omics combined with directed

investigation through traditional techniques reveal that regulation of lineage specification

involves integration of many factors to achieve robust development. Further, they

suggest that there is more heterogeneity and a greater number of transitional cell states

present within these tissues than was previously observed. Together, these

observations could suggest greater plasticity and more complex regulation of lineage

acquisition than current models account for.

A detailed picture of the emergence of lineages in the cranial NC

A long-standing challenge in the study of NC development has been pinpointing the

timing of lineage decisions and the capacity of individual NC cells to contribute to each

lineage. The difficulty in addressing these particular issues comes in part from the rapid

and dynamic nature of the early stages of NC development. NC cells rapidly migrate

away from the neuroepithelium, covering sometimes great distances within the embryo

and encountering many environments. Examining this kind of dynamic heterogeneity

within a cell population is a task particularly well-suited to single-cell transcriptomics,

which allows for profiling the entire transcriptome of thousands of cells simultaneously.

In the last 4 years, a number of papers have begun to examine NC lineage specification

using single-cell transcriptomics, but have not produced consistent conclusions

(Lukoseviciute et al., 2018; Morrison et al., 2017; Soldatov et al., 2019; Williams et al.,

2019). The issues surrounding the interpretation of these studies were discussed in detail in chapter I: lack of pure NC populations, reliance on pseudotime alone, differences between trunk and cranial NC, etc. I have attempted to reconcile outstanding questions in neural crest lineage specification through careful analysis and follow up studies of single-cell RNA-seq data that maintains a degree of spatial and temporal information.

First, I circumvent difficulties in interpreting computationally derived lineage connections from pseudotime and RNA velocity by constructing a single-cell RNA-seq timeline with fine temporal resolution. My data includes 6 time points that correspond to premigratory and several migratory stages, resulting in a comprehensive view into the early and middle stages of NC development. By including this temporal information, I confirm that the pseudotemporal connections describing the emergence of lineages comports with real-world time, greatly increasing the reliability of this information. Using this approach, I found that the first truly distinct lineage specified in the cranial NC is pigment. These cells become identifiable by expression of *mitfa*, *pax3/7*, and several other canonical pigment specifiers by 18 hpf, but not 14 hpf. This demonstrates that this lineage is indeed specified in the middle stages of migration, but not before. In addition, I identified a later small neural/glial population expressing *foxd3* that branches off from a larger group of mesenchymal-like NC cells that presumably adopt a skeletal fate based on their expression of *dlx* genes and *col2a1b*. I confirmed skeletal and pigment lineage divergence in vivo at 18 hpf through in situ Hybridization Chain Reaction (isHCR), further bolstering the notion that this timing is accurate.

My findings are in contrast to all of the previous NC single-cell studies (Lukoseviciute et al., 2018; Morrison et al., 2017; Soldatov et al., 2019; Williams et al., 2019), which never identified pigment progenitors through sequencing. Importantly, my data also strongly indicate that the premigratory NC is indeed homogenous in gene expression, with no obvious signs of lineage specification until mid-migration. This observation is consistent with the findings from Soldatov et al, (2019) which similarly show emergence of lineages after migration onset. This, however, does not conclusively show there is no biasing of lineages at this stage, merely that any such biasing is not present at the level of transcription. Further, through signaling probability analysis, I show that the emergence of lineages corresponds to shifts in Wnt signaling dynamics, most of which involve signals from surrounding cell type. Together, my single-cell analyses strongly support a hybrid model for NC lineage specification (**Fig 1.1**) where cell fate decisions occur dynamically during migration.

This model could be further investigated through emerging technologies in spatial transcriptomics. Multiplexed in situ techniques including HCR and RNAscope allow investigation of overlapping expression of many genes at once and could help to address the spatial separation of lineages more conclusively in the migratory stream. One study attempted a version of this using HCR, and spatially resolved a stem-like niche in the chick NC, but this study was limited to a few dozen genes (Lignell et al., 2017). Alternatively, in situ sequencing techniques like the 10X Genomics Visium platform allow for whole transcriptome sequencing with preserved spatial information

down to the level of a few cells. Together with the knowledge presented here and in other sequencing studies, these techniques would allow us to build a truly complete understanding of how lineage-specific expression profiles emerge and segregate within NC migratory streams. This would conclusively answer the questions of when and where NC cell fate decisions occur.

Several of the novel markers I predicted from the single-cell RNA-seq timeline were validated in vivo through isHCR. Each of these genes had lineage-specific expression in the NC, and one gene, *atp6ap2*, has a functional role in melanocyte development. It is likely that the other genes I identified computationally have unknown functional roles as well. Furthermore, the signaling analysis I performed for Wnt signaling dynamics predicts signaling between the NC and surrounding tissues through ligand-receptor pairs that are specific to particular lineages. Both these novel lineage markers and lineage-specific signaling molecules provide a wealth of potential for follow up functional characterization studies. Rapid mutagenesis screens in zebrafish have been made easy and efficient by advances in CRISPR-Cas9 techniques (Wu et al., 2018). Such a screen guided by the results from my single-cell analysis would likely reveal previously unknown functions for many candidate genes.

Characterization of Lmo7a function reveals coordination of NC migration and cell fate through focal adhesion and Wnt signaling dynamics

Lmo7 has been studied in multiple cellular contexts including cardiac muscle, skeletal muscle, and epithelial cells, where it is known to play roles in both cell adhesion and

transcriptional regulation (Dedeic et al., 2011; Holaska et al., 2006; Ooshio et al., 2004; Ott et al., 2008; Wozniak et al., 2013). It has also been implicated in Emery-Dreifuss muscular dystrophy and multiple cancers including lung adenocarcinoma (Hu et al., 2011; Mull et al., 2015; Nakamura et al., 2011). My work describes the first identified role for Lmo7 in NC biology, with a hypothesized mechanism that is similar to but distinct from its mechanistic requirements in other cell types. I show that focal adhesion (FA) dynamics are dysfunctional in Lmo7a-deficient NC cells, which draws an interesting parallel to the described role for Lmo7 in muscle, where it has been shown to localize to FAs (Wozniak et al., 2013). Unlike in these cells however, I show that Lmo7a also has an effect of canonical Wnt signaling in NC cells. This finding suggests a previously unknown mechanistic relationship between FA dynamics and Wnt signaling in NC cells.

Focal adhesions and Wnt signaling are known to be connected in cancer biology and in several developmental contexts (Fonar & Frank, 2011; Wörthmüller & Rüegg, 2020). In mouse osteoblasts, Wnt signaling is enhanced by FAK (focal adhesion kinase)-mediated stabilization of β-catenin, which enhances its nuclear localization and increases downstream Wnt target expression (Fonar & Frank, 2011). This is strikingly similar to the model I propose for the connection between focal adhesion complexes and Wnt signaling through Lmo7a in NC cells. While much of the molecular biology of this model is yet unproven, my work presents the intriguing possibility that Wnt signaling in the NC may be regulated in part by FAK. This would open up an entirely new

paradigm for the coordination between NC migration and lineage specification, which both involve canonical Wnt signaling.

 Further investigation into mechanistic requirements for FAs in NC cells could further illuminate these potential connections. Perturbation of FAK would help to test this directly. However, the essential nature of FAK in embryonic development makes whole embryo mutants confounding for the analysis of NC-specific defects. Instead, targeted mutagenesis in NC cells through UAS or other cell-type specific genome editing methodologies would circumvent this potential complication. Alternatively, transient chemical inhibition of FAK at the essential timepoints identified by my temporal single-cell RNA-seq analysis (Chapter III) could aid in specificity of functional characterization. Analyses of NC lineage specification and migration in such mutants would help to characterize the crosstalk between FAs and Wnt signaling in NC biology and illuminate coordination between migration and cell fate decisions.

Lineage-specific and temporal requirements for Wnt regulatory factors in the NC

Wnt signaling is among the most important signaling pathways in NC biology, participating in almost every step of development from induction to EMT and migration to lineage specification (De Calisto et al., 2005; Hari et al., 2012; Ji et al., 2019). Temporal control of relative Wnt signaling levels is critical for proper NC development (Hutchins & Bronner, 2018; Maj et al., 2016), highlighting the dynamic nature of Wnt requirements. Previously, our lab characterized the roles of two novel Wnt regulatory factors: Rbc3a and Ovol1a (Piloto & Schilling, 2010; Tuttle et al., 2014). These genes

are striking both in their roles in promoting migration, and in the lineage-specific requirement of these roles. Both Rbc3a and Ovol1a affected only the pigment lineage, with NC cells that failed to migrate in their absence adopting exclusively pigment identities. This is in contrast to Lmo7a, which appears to affect the migration of both pigment and glial progenitors. One striking implication of this difference is that it implies final location alone is insufficient to determine NC cell fate. Rather, it might suggest that these lineages have differential requirements for particular Wnt-regulatory mechanisms. Ovol1a acts downstream of Wnt signaling, activating downstream target gene expression. Rbc3a and Lmo7a both appear to act upstream. Rbc3a participates in the trafficking of Fzd molecules, and Lmo7a influences the stabilization/localization of β-catenin. These mechanistic differences do not in and of themselves explain the lineage-specific requirements, and further comparative analysis is needed.

All three of the above-described proteins are expressed in premigratory and early-migrating NC cells. Conversely, *atp6ap2*, is expressed as the pigment lineage becomes specified. Loss of this gene does not affect migration or pigment specification, but rather later maturation or survival. My computational analyses and follow up functional characterization do suggest however that *atp6ap2* is required for proper Wnt signaling. This therefore potentially represents another Wnt-regulatory factor that has a distinct developmental role, but with a temporal-specific requirement that differs from Ovol1a, Rbc3a, or Lmo7a. Further characterization of the specific mechanistic role this gene plays would offer insights into pigment cell development. But even more exciting is the potential afforded by all of these Wnt-regulatory genes to study the temporal and

lineage-specific requirement for Wnt signaling in the NC as a whole. In chapter II, I present an initial bulk transcriptomic comparison of *lmo7a-* and o*vol1a*-deficient NC cells. This analysis indicates strikingly different expression profiles and could be extended to r*bc3a* and *atp6ap2*. Further, a single-cell transcriptomics comparison of embryos deficient in each of these genes could illuminate the heterogeneous requirements across NC lineages and developmental time.

New perspectives on tailbud multipotent progenitors using single-cell RNA-seq

The tailbud contains multipotent progenitor cell populations that give rise to numerous tissues in the posterior segments of the adult body plan. These include the neuromesodermal progenitors (NMPs) and the midline progenitor cells (MPCs) which contribute to neural tube and paraxial/lateral mesoderm and notochord, floor plate, and hypochord, respectively. The relationships between these progenitors and their derivatives have been studied extensively, but such studies would be aided greatly by a complete transcriptomic profile of all these cells. Current single-cell studies are limited in their scope, only profiling a handful of cell types (Chestnut et al., 2020; Gouti et al., 2017). In chapter IV, I present the first comprehensive transcriptomic atlas of the zebrafish tailbud which provides insights into the biology of NMP and MPC lineage specification.

Our data shows a clear connection between paraxial mesoderm (PM) and neural tube (NT) through an apparent NMP population, as is expected. Surprisingly, however, we

also find an apparent relationship between intermediate mesoderm (IM), lateral plate mesoderm (LPM) and hematopoietic lineages all stemming from the same NMP population. While it is known that LPM derives from NMPs, the notion that blood derivatives may share a common origin with these cells has not been demonstrated. Recent studies using an *etv2* knock in line of zebrafish have also implied such a connection, however (Chestnut et al., 2020; Chestnut & Sumanas, 2020). Together with these recent findings, our computationally-derived lineage relationships between these cell types may in fact point to there being a hemangioblast population in the zebrafish tailbud that is NMP-derived. This would answer long-standing questions regarding this cell type, which has been very difficult to profile in vivo (Lacaud & Kouskoff, 2017; Nishikawa, 2012). The transcriptional profile of the putative hemangioblast population is characterized by expression of LPM, IM, and hematopoietic markers. In vivo validation of such an expression profile using isHCR or other high resolution in situ techniques would help to verify their existence as well as characterize their spatial segregation in the tailbud. Further, because our data suggest temporal requirements for BMP signaling in the lineage split between lateral mesoderm and hematopoiesis, in vivo experiments using BMP agonists or antagonists could shed further light on these bipotent progenitors.

In this data set, we also captured all of the MPC-derived tissues: notochord, hypochord, and floor plate. In my analysis, I identify a putative MPC population, which is connected to both floor plate and notochord, representing one of the recently-described bipotent progenitors in the zebrafish tailbud (Row et al., 2016). This cell population is marked

uniquely by overlapping expression of *fgf8a* and *noto*. This expression profile should be verified in vivo using isHCR or other in situ techniques. Strikingly, this population is also almost exclusively characterized by G1 cell cycle state. With cells that remain in G1 progressing toward a notochord fate before entering the cell cycle, and those exiting G1 immediately progressing toward a floor plate identity. This observation comports with recent in vivo observations that notochord progenitors arrest in G1 phase (Sugiyama et al., 2009, 2014). Further, they show that cell cycle states in MPCs directly relate to their acquisition of lineage-specific transcriptional programs. This could imply a function role for cell cycle arrest or a mechanism that regulates both cell cycle state and lineage specification in concert. In vivo perturbation of cell cycle regulatory mechanisms in these progenitor pools could help to disentangle these possibilities and further our understanding of the mechanisms governing the development of MPC-derived tissues.

As with any scientific pursuit, new advances in technologies offer us new perspectives and means to investigate old questions. Over the last decade, one of the most significant advances driving discoveries in the field of cell fate and lineage determination is undoubtedly the rise of next generation sequencing techniques. In this thesis I have demonstrated how careful analysis of large transcriptomic data sets can expand our understanding of developmental systems. This type of investigation, however, is at its most powerful when paired with in vivo experimentation and when informed by the wealth of knowledge generated over many decades by the investigators that laid the foundations of embryology. Over the next years, current methodologies will undoubtedly be improved upon and bolstered by new technological breakthroughs. And with each

new window into the complex processes driving cell fate decisions, we move ever closer

to a complete answer to the age-old question: "how does a fully formed animal grow

from a single cell?"

References

Ahsan, K., Singh, N., Rocha, M., Huang, C., & Prince, V. E. (2019). Prickle1 is required for EMT and migration of zebrafish cranial neural crest. *Developmental Biology*, *448*(1), 16–35. https://doi.org/10.1016/j.ydbio.2019.01.018

Axelrod, D. (1979). Carbocyanine dye orientation in red cell membrane studied by microscopic fluorescence polarization. *Biophysical Journal*, *26*(3), 557–573. https://doi.org/10.1016/S0006-3495(79)85271-6

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*(3), 314–322. https://doi.org/10.1016/j.stem.2015.02.017

Balakier, H., & Pedersen, R. A. (1982). Allocation of cells to inner cell mass and trophectoderm lineages in preimplantation mouse embryos. *Developmental Biology*, *90*(2), 352–362. https://doi.org/10.1016/0012-1606(82)90384-0

Ballard, W. W. (1973). A new fate map forSalmo gairdneri. *Journal of Experimental Zoology*, *184*(1), 49–73. https://doi.org/10.1002/jez.1401840105

Ballard, W. W. (1982). Morphogenetic movements and fate map of the cypriniform teleost, Catostomus commersoni (lacepede). *Journal of Experimental Zoology*, *219*(3), 301–321. https://doi.org/10.1002/jez.1402190306

Ballard, W. W. (1986). Morphogenetic movements and a provisional fate map of development in the holostean fishAmia calva. *Journal of Experimental Zoology*, *238*(3), 355–372. https://doi.org/10.1002/jez.1402380309

Barske, L., Askary, A., Zuniga, E., Balczerski, B., Bump, P., Nichols, J. T., & Crump, J.

G. (2016). Competition between Jagged-Notch and Endothelin1 Signaling Selectively Restricts Cartilage Formation in the Zebrafish Upper Face. *PLoS Genetics*, *12*(4), 1005967. https://doi.org/10.1371/journal.pgen.1005967

Beck, C. W. (2015). Development of the vertebrate tailbud. *Wiley Interdisciplinary Reviews: Developmental Biology*, *4*(1), 33–44. https://doi.org/10.1002/wdev.163

Bergen, V., Lange, M., Peidli, S., Wolf, F. A., & Theis, F. J. (2020). Generalizing RNA velocity to transient cell states through dynamical modeling. *Nature Biotechnology*, *38*(12), 1408–1414. https://doi.org/10.1038/s41587-020-0591-3

Boer, E. F., Howell, E. D., Schilling, T. F., Jette, C. A., & Stewart, R. A. (2015). Fascin1-Dependent Filopodia are Required for Directional Migration of a Subset of Neural Crest Cells. *PLOS Genetics*, *11*(1), e1004946. https://doi.org/10.1371/journal.pgen.1004946

Bouldin, C. M., Manning, A. J., Peng, Y. H., Farr, G. H., Hung, K. L., Dong, A., & Kimelman, D. (2015). Wnt signaling and tbx16 form a bistable switch to commit bipotential progenitors to mesoderm. *Development (Cambridge)*, *142*(14), 2499–2507. https://doi.org/10.1242/dev.124024

Briggs, J. A., Weinreb, C., Wagner, D. E., Megason, S., Peshkin, L., Kirschner, M. W., & Klein, A. M. (2018). The dynamics of gene expression in vertebrate embryogenesis at single-cell resolution. *Science*, *360*(6392). https://doi.org/10.1126/science.aar5780

Bronner-Fraser, M., Stern, C. D., & Fraser, S. (1991). Analysis of neural crest cell lineage and migration. *Journal of Craniofacial Genetics and Developmental Biology*, *11*(4), 214–222.

Bronner-Fraser, Marianne, & Fraser, S. E. (1988). Cell lineage analysis reveals

multipotency of some avian neural crest cells. *Nature*, *335*(6186), 161–164.

https://doi.org/10.1038/335161a0

Bronner-Fraser, M., Sieber-blum, M., & Cohen, A. M. (1980). Clonal analysis of the

avian neural crest: Migration and maturation of mixed neural crest clones injected

into host chicken embryos. *Journal of Comparative Neurology*, *193*(2), 423–434.

https://doi.org/10.1002/cne.901930209

Burstyn-Cohen, T., Stanleigh, J., Sela-Donenfeld, D., & Kalcheim, C. (2004). Canonical

Wnt activity regulates trunk neural crest delamination linking BMP/noggin signaling

with G1/S transition. *Development*, *131*(21), 5327–5339.

https://doi.org/10.1242/dev.01424

Cambray, Noemi, & Wilson, V. (2007). Two distinct sources for a population of maturing

axial progenitors. *Development*, *134*(15), 2829–2840.

https://doi.org/10.1242/dev.02877

Cambray, Noemí, & Wilson, V. (2002). Axial progenitors with extensive potency are

localised to the mouse chordoneural hinge. *Development*, *129*(20), 4855–4866.

Cannoodt, R., Saelens, W., & Saeys, Y. (2016). Computational methods for trajectory

inference from single-cell transcriptomics. In *European Journal of Immunology* (Vol.

46, Issue 11, pp. 2496–2506). Wiley-VCH Verlag.

https://doi.org/10.1002/eji.201646347

Cao, J., Spielmann, M., Qiu, X., Huang, X., Ibrahim, D. M., Hill, A. J., Zhang, F.,

Mundlos, S., Christiansen, L., Steemers, F. J., Trapnell, C., & Shendure, J. (2019).

The single-cell transcriptional landscape of mammalian organogenesis. *Nature*,

566(7745), 496–502. https://doi.org/10.1038/s41586-019-0969-x

Carmona-Fontaine, C., Matthews, H. K., Kuriyama, S., Moreno, M., Dunn, G. A., Parsons, M., Stern, C. D., & Mayor, R. (2008). Contact inhibition of locomotion in vivo controls neural crest directional migration. *Nature*, *456*(7224), 957–961. https://doi.org/10.1038/nature07441

Carmona-Fontaine, C., Theveneau, E., Tzekou, A., Tada, M., Woods, M., Page, K. M., Parsons, M., Lambris, J. D., & Mayor, R. (2011). Complement Fragment C3a Controls Mutual Cell Attraction during Collective Cell Migration. *Developmental Cell*, *21*(6), 1026–1037. https://doi.org/10.1016/j.devcel.2011.10.012

Catala, M., Teillet, M. A., De Robertis, E. M., & Le Douarin, N. M. (1996). A spinal cord fate map in the avian embryo: While regressing, Hensen's node lays down the notochord and floor plate thus joining the spinal cord lateral walls. *Development*, *122*(9), 2599–2610.

Catala, M., Teillet, M. A., & Le Douarin, N. M. (1995). Organization and development of the tail bud analyzed with the quail-chick chimaera system. *Mechanisms of Development*, *51*(1), 51–65. https://doi.org/10.1016/0925-4773(95)00350-A

Chalpe, A. J., Prasad, M., Henke, A. J., & Paulson, A. F. (2010). Regulation of cadherin expression in the chicken neural crest by the Wnt/β-catenin signaling pathway. *Cell Adhesion and Migration*, *4*(3), 431–438. https://doi.org/10.4161/cam.4.3.12138

Chestnut, B., Casie Chetty, S., Koenig, A. L., & Sumanas, S. (2020). Single-cell transcriptomic analysis identifies the conversion of zebrafish Etv2-deficient vascular progenitors into skeletal muscle. *Nature Communications*, *11*(1), 1–16. https://doi.org/10.1038/s41467-020-16515-y

Chestnut, B., & Sumanas, S. (2020). Zebrafish etv2 knock-in line labels vascular

    endothelial and blood progenitor cells. *Developmental Dynamics*, *249*(2), 245–261.

    https://doi.org/10.1002/dvdy.130

Choi, H. M. T., Beck, V. A., & Pierce, N. A. (2014). Next-generation in situ hybridization

    chain reaction: Higher gain, lower cost, greater durability. *ACS Nano*, *8*(5), 4284–

    4294. https://doi.org/10.1021/nn405717p

Cruciat, C. M., Ohkawara, B., Acebron, S. P., Karaulanov, E., Reinhard, C., Ingelfinger,

    D., Boutros, M., & Niehrs, C. (2010a). Requirement of prorenin receptor and

    vacuolar H+-ATPase-mediated acidification for Wnt signaling. *Science*, *327*(5964),

    459–463. https://doi.org/10.1126/science.1179802

Cruciat, C. M., Ohkawara, B., Acebron, S. P., Karaulanov, E., Reinhard, C., Ingelfinger,

    D., Boutros, M., & Niehrs, C. (2010b). Requirement of prorenin receptor and

    vacuolar H+-ATPase-mediated acidification for Wnt signaling. *Science*, *327*(5964),

    459–463. https://doi.org/10.1126/science.1179802

Curran, K., Lister, J. A., Kunkel, G. R., Prendergast, A., Parichy, D. M., & Raible, D. W.

    (2010). Interplay between Foxd3 and Mitf regulates cell fate plasticity in the

    zebrafish neural crest. *Developmental Biology*, *344*(1), 107–118.

    https://doi.org/10.1016/j.ydbio.2010.04.023

Dai, X., Jiang, W., Zhang, Q., Xu, L., Geng, P., Zhuang, S., Petrich, B. G., Jiang, C.,

    Peng, L., Bhattacharya, S., Evans, S. M., Sun, Y., Chen, J., & Liang, X. (2013).

    Requirement for integrin-linked kinase in neural crest migration and differentiation

    and outflow tract morphogenesis. *BMC Biology*, *11*. https://doi.org/10.1186/1741-

    7007-11-107

De Calisto, J., Araya, C., Marchant, L., Riaz, C. F., & Mayor, R. (2005). Essential role of

    non-canonical Wnt signalling in neural crest migration. *Development*, *132*(11),

    2587–2597. https://doi.org/10.1242/dev.01857

Dedeic, Z., Cetera, M., Cohen, T. V., & Holaska, J. M. (2011). Emerin inhibits Lmo7

    binding to the Pax3 and MyoD promoters and expression of myoblast proliferation

    genes. *Journal of Cell Science*, *124*(10), 1691–1702.

    https://doi.org/10.1242/jcs.080259

Deppe, U., Schierenberg, E., Cole, T., Krieg, C., Schmitt, D., Yoder, B., & von

    Ehrenstein, G. (1978). Cell lineages of the embryo of the nematode Caenorhabditis

    elegans. *Proceedings of the National Academy of Sciences of the United States of*

    America, *75*(1), 376–380. https://doi.org/10.1073/pnas.75.1.376

Dorsky, R. I., Moon, R. T., & Raible, D. W. (1998). Control of neural crest cell fate by the

    Wnt signalling pathway. *Nature*, *396*(6709), 370–372. https://doi.org/10.1038/24620

Dottori, M., Gross, M. K., Labosky, P., & Goulding, M. (2001). The winged-helix

    transcription factor Foxd3 suppresses interneuron differentiation and promotes

    neural crest cell fate. *Development*, *128*(21), 4127–4138.

Du, T. T., Dewey, J. B., Wagner, E. L., Cui, R., Heo, J., Park, J. J., Francis, S. P.,

    Perez-Reyes, E., Guillot, S. J., Sherman, N. E., Xu, W., Oghalai, J. S., Kachar, B.,

    & Shin, J. B. (2019). LMO7 deficiency reveals the significance of the cuticular plate

    for hearing function. *Nature Communications*, *10*(1), 1–15.

    https://doi.org/10.1038/s41467-019-09074-4

Dupin, E., Calloni, G. W., Coelho-Aguiar, J. M., & Le Douarin, N. M. (2018). The issue

    of the multipotency of the neural crest cells. In *Developmental Biology* (Vol. 444,

pp. S47–S59). Elsevier Inc. https://doi.org/10.1016/j.ydbio.2018.03.024

Dupin, E., Calloni, G. W., & Le Douarin, N. M. (2010). The cephalic neural crest of
amniote vertebrates is composed of a large majority of precursors endowed with
neural, melanocytic, chondrogenic and osteogenic potentialities. In *Cell Cycle* (Vol.
9, Issue 2, pp. 238–249). Taylor and Francis Inc.
https://doi.org/10.4161/cc.9.2.10491

Eagleson, G., Ferreiro, B., & Harris, W. A. (1995). Fate of the anterior neural ridge and
the morphogenesis of the xenopus forebrain. *Journal of Neurobiology*, *28*(2), 146–
158.

Eagleson, G. W., & Harris, W. A. (1990). Mapping of the presumptive brain regions in
the neural plate of Xenopus laevis. *Journal of Neurobiology*, *21*(3), 427–440.
https://doi.org/10.1002/neu.480210305

EauClaire, S. F., Cui, S., Ma, L., Matous, J., Marlow, F. L., Gupta, T., Burgess, H. A.,
Abrams, E. W., Kapp, L. D., Granato, M., Mullins, M. C., & Matthews, R. P. (2012).
Mutations in vacuolar H+-ATPase subunits lead to biliary developmental defects in
zebrafish. *Developmental Biology*, *365*(2), 434–444.
https://doi.org/10.1016/j.ydbio.2012.03.009

Eriksson, J., & Löfberg, J. (2000). Development of the hypochord and dorsal aorta in the
zebrafish embryo (Danio rerio). *Journal of Morphology*, *244*(3), 167–176.
https://doi.org/10.1002/(SICI)1097-4687(200006)244:3<167::AID-
JMOR2>3.0.CO;2-J

Farrell, J. A., Wang, Y., Riesenfeld, S. J., Shekhar, K., Regev, A., & Schier, A. F.
(2018). Single-cell reconstruction of developmental trajectories during zebrafish

embryogenesis. *Science*, *360*(6392). https://doi.org/10.1126/science.aar3131

Ferronha, T., Angeles Rabadán, M., Gil-Guiñon, E., Le Dréau, G., de Torres, C., &
Martí, E. (2013). LMO4 is an essential cofactor in the Snail2-mediated epithelial-to-
mesenchymal transition of neuroblastoma and neural crest cells. *Journal of
Neuroscience*, *33*(7), 2773–2783. https://doi.org/10.1523/JNEUROSCI.4511-
12.2013

Fonar, Y., & Frank, D. (2011). FAK and WNT Signaling: The Meeting of Two Pathways
in Cancer and Development. *Anti-Cancer Agents in Medicinal Chemistry*, *11*(7),
600–606. https://doi.org/10.2174/187152011796817673

Furuya, M., Tsuji, N., Endoh, T., Moriai, R., Kobayashi, D., Yagihashi, A., & Watanabe,
N. (2002). A novel gene containing PDZ and LIM domains, PCD1, is overexpressed
in human colorectal cancer. *Anticancer Research*, *22*(6 C), 4183–4186.

Gans, C., & Northcutt, R. G. (1983). Neural crest and the origin of vertebrates: A new
head. In *Science* (Vol. 220, Issue 4594, pp. 268–274). American Association for the
Advancement of Science. https://doi.org/10.1126/science.220.4594.268

Garcia-Lopez, R., Pombero, A., & Martinez, S. (2009). Fate map of the chick embryo
neural tube. In *Development Growth and Differentiation* (Vol. 51, Issue 3, pp. 145–
165). https://doi.org/10.1111/j.1440-169X.2009.01096.x

Garnett, A. T., Square, T. A., & Medeiros, D. M. (2012). BMP, wnt and FGF signals are
integrated through evolutionarily conserved enhancers to achieve robust
expression of Pax3 and Zic genes at the zebrafish neural plate border.
Development (Cambridge), *139*(22), 4220–4231.
https://doi.org/10.1242/dev.081497

Garriock, R. J., Chalamalasetty, R. B., Kennedy, M. W., Canizales, L. C., Lewandoski, M., & Yamaguchi, T. P. (2015). Lineage tracing of neuromesodermal progenitors reveals novel wnt-dependent roles in trunk progenitor cell maintenance and differentiation. *Development (Cambridge)*, *142*(9), 1628–1638. https://doi.org/10.1242/dev.111922

Germanguz, I., Lev, D., Waisman, T., Kim, C.-H., & Gitelman, I. (2007). Four *twist* genes in zebrafish, four expression patterns. *Developmental Dynamics*, *236*(9), 2615–2626. https://doi.org/10.1002/dvdy.21267

Gierahn, T. M., Wadsworth, M. H., Hughes, T. K., Bryson, B. D., Butler, A., Satija, R., Fortune, S., Christopher Love, J., & Shalek, A. K. (2017). Seq-Well: Portable, low-cost rna sequencing of single cells at high throughput. *Nature Methods*, *14*(4), 395–398. https://doi.org/10.1038/nmeth.4179

Gimlich, R. L., & Braun, J. (1985). Improved fluorescent compounds for tracing cell lineage. *Developmental Biology*, *109*(2), 509–514. https://doi.org/10.1016/0012-1606(85)90476-2

Goody, M. F., Kelly, M. W., Lessard, K. N., Khalil, A., & Henry, C. A. (2010). Nrk2b-mediated NAD+ production regulates cell adhesion and is required for muscle morphogenesis in vivo. Nrk2b and NAD+ in muscle morphogenesis. *Developmental Biology*, *344*(2), 809–826. https://doi.org/10.1016/j.ydbio.2010.05.513

Gouti, M., Delile, J., Stamataki, D., Wymeersch, F. J., Huang, Y., Kleinjung, J., Wilson, V., & Briscoe, J. (2017). A Gene Regulatory Network Balances Neural and Mesoderm Specification during Vertebrate Trunk Development. *Developmental Cell*, *41*(3), 243-261.e7. https://doi.org/10.1016/j.devcel.2017.04.002

131

Gouti, M., Tsakiridis, A., Wymeersch, F. J., Huang, Y., Kleinjung, J., Wilson, V., & Briscoe, J. (2014). In vitro generation of neuromesodermal progenitors reveals distinct roles for wnt signalling in the specification of spinal cord and paraxial mesoderm identity. *PLoS Biology*, *12*(8). https://doi.org/10.1371/journal.pbio.1001937

Griffith, C. M., Wiley, M. J., & Sanders, E. J. (1992). The vertebrate tail bud: three germ layers from one tissue. In *Anatomy and Embryology* (Vol. 185, Issue 2, pp. 101–113). Springer-Verlag. https://doi.org/10.1007/BF00185911

Haber, A. L., Biton, M., Rogel, N., Herbst, R. H., Shekhar, K., Smillie, C., Burgin, G., Delorey, T. M., Howitt, M. R., Katz, Y., Tirosh, I., Beyaz, S., Dionne, D., Zhang, M., Raychowdhury, R., Garrett, W. S., Rozenblatt-Rosen, O., Shi, H. N., Yilmaz, O., … Regev, A. (2017). A single-cell survey of the small intestinal epithelium. *Nature*, *551*(7680), 333–339. https://doi.org/10.1038/nature24489

Hamidouche, Z., Haÿ, E., Vaudin, P., Charbord, P., Schüle, R., Marie, P. J., & Fromigué, O. (2008). FHL2 mediates dexamethasone-induced mesenchymal cell differentiation into osteoblasts by activating Wnt/β-catenin signaling-dependent Runx2 expression. *The FASEB Journal*, *22*(11), 3813–3822. https://doi.org/10.1096/fj.08-106302

Hampel, S., Chung, P., McKellar, C. E., Hall, D., Looger, L. L., & Simpson, J. H. (2011). Drosophila Brainbow: A recombinase-based fluorescence labeling technique to subdivide neural expression patterns. *Nature Methods*, *8*(3), 253–259. https://doi.org/10.1038/nmeth.1566

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/β-catenin signaling. *Development (Cambridge)*, *139*(12), 2107–2117. https://doi.org/10.1242/dev.073064

Harrison, D. A., & Perrimon, N. (1993). Simple and efficient generation of marked clones in Drosophila. *Current Biology*, *3*(7), 424–433. https://doi.org/10.1016/0960-9822(93)90349-S

He, J., Zhang, N., Zhang, J., Jiang, B., & Wu, F. (2018). Migration critically meditates osteoblastic differentiation of bone mesenchymal stem cells through activating canonical Wnt signal pathway. *Colloids and Surfaces B: Biointerfaces*, *171*, 205–213. https://doi.org/10.1016/j.colsurfb.2018.07.017

Heuberger, J., & Birchmeier, W. (2010). Interplay of cadherin-mediated cell adhesion and canonical Wnt signaling. In *Cold Spring Harbor perspectives in biology* (Vol. 2, Issue 2, pp. 2915–2916). Cold Spring Harbor Laboratory Press. https://doi.org/10.1101/cshperspect.a002915

His, W. (1868). Untersuchungen über die erste Anlage des Wirbelthierleibes : die erste Entwickelung des Hühnchens im Ei / von Wilhelm His. *Untersuchungen Über Die Erste Anlage Des Wirbelthierleibes : Die Erste Entwickelung Des Hühnchens Im Ei / von Wilhelm His.* https://doi.org/10.5962/bhl.title.15288

Hoffman, T. L., Javier, A. L., Campeau, S. A., Knight, R. D., & Schilling, T. F. (2007). Tfap2 transcription factors in zebrafish neural crest development and ectodermal evolution. *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution*, *308*(5), 679–691. https://doi.org/10.1002/jez.b.21189

Holaska, J. M., Rais-Bahrami, S., & Wilson, K. L. (2006). Lmo7 is an emerin-binding

protein that regulates the transcription of emerin and many other muscle-relevant

genes. *Human Molecular Genetics*, *15*(23), 3459–3472.

https://doi.org/10.1093/hmg/ddl423

Hopwood, N. D., Pluck, A., & Gurdon, J. B. (1989). A Xenopus mRNA related to

Drosophila twist is expressed in response to induction in the mesoderm and the

neural crest. *Cell*, *59*(5), 893–903. https://doi.org/10.1016/0092-8674(89)90612-0

Hu, Q., Guo, C., Li, Y., Aronow, B. J., & Zhang, J. (2011). LMO7 Mediates Cell-Specific

Activation of the Rho-Myocardin-Related Transcription Factor-Serum Response

Factor Pathway and Plays an Important Role in Breast Cancer Cell Migration.

Molecular and Cellular Biology, *31*(16), 3223–3240.

https://doi.org/10.1128/mcb.01365-10

Huber, T. L., Kouskoff, V., Fehling, H. J., Palis, J., & Keller, G. (2004). Haemangioblast

commitment is initiated in the primitive streak of the mouse embryo. *Nature*,

432(7017), 625–630. https://doi.org/10.1038/nature03122

Hutchins, E. J., & Bronner, M. E. (2018). Draxin acts as a molecular rheostat of

canonical Wnt signaling to control cranial neural crest EMT. *Journal of Cell Biology*,

217(10), 3683–3697. https://doi.org/10.1083/JCB.201709149

Ichihara, A., & Yatabe, M. S. (2019). The (pro)renin receptor in health and disease. In

Nature Reviews Nephrology (Vol. 15, Issue 11, pp. 693–712). Nature Publishing

Group. https://doi.org/10.1038/s41581-019-0160-5

Ishii, M., Merrill, A. E., Chan, Y. S., Gitelman, I., Rice, D. P. C., Sucov, H. M., & Maxson,

R. E. (2003). Msx2 and Twist cooperatively control the development of the neural

crest-derived skeletogenic mesenchyme of the murine skull vault. *Development*,

130(24), 6131–6142. https://doi.org/10.1242/dev.00793

Jaitin, D. A., Kenigsberg, E., Keren-Shaul, H., Elefant, N., Paul, F., Zaretsky, I., Mildner,

A., Cohen, N., Jung, S., Tanay, A., & Amit, I. (2014). Massively parallel single-cell

RNA-seq for marker-free decomposition of tissues into cell types. *Science*,

343(6172), 776–779. https://doi.org/10.1126/science.1247651

Jeong, Y., & Epstein, D. J. (2003). Distinct regulators of Shh transcription in the floor

plate and notochord indicate separate origins for these tissues in the mouse node.

In *Development* (Vol. 130, Issue 16, pp. 3891–3902). The Company of Biologists

Ltd. https://doi.org/10.1242/dev.00590

Ji, Y., Hao, H., Reynolds, K., McMahon, M., & Zhou, C. J. (2019). Wnt Signaling in

Neural Crest Ontogenesis and Oncogenesis. In *Cells* (Vol. 8, Issue 10). NLM

(Medline). https://doi.org/10.3390/cells8101173

Johnson, M. H., & Ziomek, C. A. (1981). The foundation of two distinct cell lineages

within the mouse morula. *Cell*, *24*(1), 71–80. https://doi.org/10.1016/0092-

8674(81)90502-X

Kalcheim, C., & Kumar, D. (2017). Cell fate decisions during neural crest ontogeny. *The

International Journal of Developmental Biology*, *61*(3-4–5), 195–203.

https://doi.org/10.1387/ijdb.160196ck

Kang, S., Xu, H., Duan, X., Liu, J.-J., He, Z., Yu, F., Zhou, S., Meng, X.-Q., Cao, M., &

Kennedy, G. C. (2000). PCD1, a Novel Gene Containing PDZ and LIM Domains, Is

Overexpressed in Several Human Cancers. *Cancer Research*, *60*(18), 5296–5302.

Kardon, G., Campbell, J. K., & Tabin, C. J. (2002). Local extrinsic signals determine

muscle and endothelial cell fate and patterning in the vertebrate limb.

Developmental Cell, *3*(4), 533–545. https://doi.org/10.1016/S1534-5807(02)00291-5

Karlsson, T., Kvarnbrink, S., Holmlund, C., Botling, J., Micke, P., Henriksson, R., Johansson, M., & Hedman, H. (2018). LMO7 and LIMCH1 interact with LRIG proteins in lung cancer, with prognostic implications for early-stage disease. *Lung Cancer*, *125*, 174–184. https://doi.org/10.1016/j.lungcan.2018.09.017

Kerosuo, L., & Bronner-Fraser, M. (2012). What is bad in cancer is good in the embryo: Importance of EMT in neural crest development. In *Seminars in Cell and Developmental Biology* (Vol. 23, Issue 3, pp. 320–332). Elsevier Ltd. https://doi.org/10.1016/j.semcdb.2012.03.010

Khang, D., Choi, J., Im, Y. M., Kim, Y. J., Jang, J. H., Kang, S. S., Nam, T. H., Song, J., & Park, J. W. (2012). Role of subnano-, nano- and submicron-surface features on osteoblast differentiation of bone marrow mesenchymal stem cells. *Biomaterials*, *33*(26), 5997–6007. https://doi.org/10.1016/j.biomaterials.2012.05.005

Kimmel, C.B., Warga, R. M., & Schilling, T. F. (1990). Origin and organization of the zebrafish fate map. *Development*, *108*(4).

Kimmel, Charles B., Ballard, W. W., Kimmel, S. R., Ullmann, B., & Schilling, T. F. (1995). Stages of embryonic development of the zebrafish. *Developmental Dynamics*, *203*(3), 253–310. https://doi.org/10.1002/aja.1002030302

Kinney, B. A., Al Anber, A., Row, R. H., Tseng, Y. J., Weidmann, M. D., Knaut, H., & Martin, B. L. (2020). Sox2 and Canonical Wnt Signaling Interact to Activate a Developmental Checkpoint Coordinating Morphogenesis with Mesoderm Fate Acquisition. *Cell Reports*, *33*(4), 108311.

https://doi.org/10.1016/j.celrep.2020.108311

Klein, A. M., Mazutis, L., Akartuna, I., Tallapragada, N., Veres, A., Li, V., Peshkin, L.,

Weitz, D. A., & Kirschner, M. W. (2015). Droplet barcoding for single-cell

transcriptomics applied to embryonic stem cells. *Cell*, *161*(5), 1187–1201.

https://doi.org/10.1016/j.cell.2015.04.044

Koch, F., Scholze, M., Wittler, L., Schifferl, D., Sudheer, S., Grote, P., Timmermann, B.,

Macura, K., & Herrmann, B. G. (2017). Antagonistic Activities of Sox2 and

Brachyury Control the Fate Choice of Neuro-Mesodermal Progenitors.

Developmental Cell, *42*(5), 514-526.e7.

https://doi.org/10.1016/j.devcel.2017.07.021

Korenbaum, E., & Rivero, F. (2002). Calponin homology domains at a glance. *Journal of

Cell Science*, *115*(18), 3543–3545. https://doi.org/10.1242/jcs.00003

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*(4), 585–595. https://doi.org/10.1242/dev.041509

La Manno, G., Soldatov, R., Zeisel, A., Braun, E., Hochgerner, H., Petukhov, V.,

Lidschreiber, K., Kastriti, M. E., Lönnerberg, P., Furlan, A., Fan, J., Borm, L. E., Liu,

Z., van Bruggen, D., Guo, J., He, X., Barker, R., Sundström, E., Castelo-Branco,

G., … Kharchenko, P. V. (2018). RNA velocity of single cells. *Nature*, *560*(7719),

494–498. https://doi.org/10.1038/s41586-018-0414-6

LaBonne, C., & Bronner-Fraser, M. (1998). Neural crest induction in Xenopus: Evidence

for a two-signal model. *Development*, *125*(13), 2403–2414.

Lacaud, G., & Kouskoff, V. (2017). Hemangioblast, hemogenic endothelium, and

primitive versus definitive hematopoiesis. *Experimental Hematology*, *49*, 19–24. https://doi.org/10.1016/j.exphem.2016.12.009

Lander, R., Nasr, T., Ochoa, S. D., Nordin, K., Prasad, M. S., & Labonne, C. (2013). Interactions between Twist and other core epithelial-mesenchymal transition factors are controlled by GSK3-mediated phosphorylation. *Nature Communications*, *4*. https://doi.org/10.1038/ncomms2543

Langhe, R. P., Gudzenko, T., Bachmann, M., Becker, S. F., Gonnermann, C., Winter, C., Abbruzzese, G., Alfandari, D., Kratzer, M. C., Franz, C. M., & Kashef, J. (2016). Cadherin-11 localizes to focal adhesions and promotes cell-substrate adhesion. *Nature Communications*, *7*(1), 1–10. https://doi.org/10.1038/ncomms10909

Lawson, K. A., Meneses, J. J., & Pedersen, R. A. (1986). Cell fate and cell lineage in the endoderm of the presomite mouse embryo, studied with an intracellular tracer. *Developmental Biology*, *115*(2), 325–339. https://doi.org/10.1016/0012-1606(86)90253-8

Le Douarin, N. M., & Teillet, M. A. (1973). The migration of neural crest cells to the wall of the digestive tract in avian embryo. *Journal of Embryology and Experimental Morphology*, *30*(1), 31–48.

LeDouarin, N. (1980). Migration and differentiation of neural crest cells. *Current Topics in Developmental Biology*, *16*(C), 31–85. https://doi.org/10.1016/S0070-2153(08)60153-2

Li, B., Mackay, D. R., Dai, Q., Li, T. W. H., Nair, M., Fallahi, M., Schonbaum, C. P., Fantes, J., Mahowald, A. P., Waterman, M. L., Fuchs, E., & Dai, X. (2002). The LEF1/β-catenin complex activates movo1, a mouse homolog of Drosophila ovo

required for epidermal appendage differentiation. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(9), 6064–6069. https://doi.org/10.1073/pnas.092137099

Liao, W., Bisgrove, B. W., Sawyer, H., Hug, B., Bell, B., Peters, K., Grunwald, D. J., & Stainier, D. Y. R. (1997). The zebrafish gene cloche acts upstream of a flk-1 homologue to regulate endothelial cell differentiation. *Development*, *124*(2), 381–389.

Lignell, A., Kerosuo, L., Streichan, S. J., Cai, L., & Bronner, M. E. (2017). Identification of a neural crest stem cell niche by Spatial Genomic Analysis. *Nature Communications*, *8*(1). https://doi.org/10.1038/s41467-017-01561-w

Lin, Y. H., Zhen, Y. Y., Chien, K. Y., Lee, I. C., Lin, W. C., Chen, M. Y., & Pai, L. M. (2017). LIMCH1 regulates nonmuscle myosin-II activity and suppresses cell migration. *Molecular Biology of the Cell*, *28*(8), 1054–1065. https://doi.org/10.1091/mbc.E15-04-0218

Lister, J. A., Cooper, C., Nguyen, K., Modrell, M., Grant, K., & Raible, D. W. (2006). Zebrafish Foxd3 is required for development of a subset of neural crest derivatives. *Developmental Biology*, *290*(1), 92–104. https://doi.org/10.1016/j.ydbio.2005.11.014

Lister, R., O'Malley, R. C., Tonti-Filippini, J., Gregory, B. D., Berry, C. C., Millar, A. H., & Ecker, J. R. (2008). Highly Integrated Single-Base Resolution Maps of the Epigenome in Arabidopsis. *Cell*, *133*(3), 523–536. https://doi.org/10.1016/j.cell.2008.03.029

Livet, J., Weissman, T. A., Kang, H., Draft, R. W., Lu, J., Bennis, R. A., Sanes, J. R., &

Lichtman, J. W. (2007). Transgenic strategies for combinatorial expression of fluorescent proteins in the nervous system. *Nature*, *450*(7166), 56–62. https://doi.org/10.1038/nature06293

Löfberg, J., & Collazo, A. (1997). Hypochord, an enigmatic embryonic structure: Study of the axolotl embryo. *Journal of Morphology*, *232*(1), 57–66. https://doi.org/10.1002/(SICI)1097-4687(199704)232:1<57::AID-JMOR3>3.0.CO;2-L

Lukoseviciute, M., Gavriouchkina, D., Williams, R. M., Hochgreb-Hagele, T., Senanayake, U., Chong-Morrison, V., Thongjuea, S., Repapi, E., Mead, A., & Sauka-Spengler, T. (2018). From Pioneer to Repressor: Bimodal foxd3 Activity Dynamically Remodels Neural Crest Regulatory Landscape In Vivo. *Developmental Cell*, *47*(5), 608-628.e6. https://doi.org/10.1016/j.devcel.2018.11.009

Lun, A. T. L., McCarthy, D. J., & Marioni, J. C. (2016). A step-by-step workflow for low-level analysis of single-cell RNA-seq data with Bioconductor. *F1000Research*, *5*. https://doi.org/10.12688/f1000research.9501.2

Luo, T., Lee, Y. H., Saint-Jeannet, J. P., & Sargent, T. D. (2003). Induction of neural crest in Xenopus by transcription factor AP2α. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(2), 532–537. https://doi.org/10.1073/pnas.0237226100

Macosko, E. Z., Basu, A., Satija, R., Nemesh, J., Shekhar, K., Goldman, M., Tirosh, I., Bialas, A. R., Kamitaki, N., Martersteck, E. M., Trombetta, J. J., Weitz, D. A., Sanes, J. R., Shalek, A. K., Regev, A., & McCarroll, S. A. (2015). Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell*,

161(5), 1202–1214. https://doi.org/10.1016/j.cell.2015.05.002

Maj, E., Künneke, L., Loresch, E., Grund, A., Melchert, J., Pieler, T., Aspelmeier, T., &

Borchers, A. (2016). Controlled levels of canonical Wnt signaling are required for

neural crest migration. *Developmental Biology*, *417*(1), 77–90.

https://doi.org/10.1016/j.ydbio.2016.06.022

Manning, A. J., & Kimelman, D. (2015). Tbx16 and Msgn1 are required to establish

directional cell migration of zebrafish mesodermal progenitors. *Developmental*

Biology, *406*(2), 172–185. https://doi.org/10.1016/j.ydbio.2015.09.001

Mao, X., Fujiwara, Y., Chapdelaine, A., Yang, H., & Orkin, S. H. (2001). Activation of

EGFP expression by Cre-mediated excision in a new ROSA26 reporter mouse

strain. *Blood*, *97*(1), 324–326. https://doi.org/10.1182/blood.V97.1.324

Martin, B. L., & Kimelman, D. (2008). Regulation of Canonical Wnt Signaling by

Brachyury Is Essential for Posterior Mesoderm Formation. *Developmental Cell*,

15(1), 121–133. https://doi.org/10.1016/j.devcel.2008.04.013

Martin, B. L., & Kimelman, D. (2010). Brachyury establishes the embryonic mesodermal

progenitor niche. *Genes and Development*, *24*(24), 2778–2783.

https://doi.org/10.1101/gad.1962910

Martin, B. L., & Kimelman, D. (2012). Canonical Wnt Signaling Dynamically Controls

Multiple Stem Cell Fate Decisions during Vertebrate Body Formation.

Developmental Cell, *22*(1), 223–232. https://doi.org/10.1016/j.devcel.2011.11.001

Martin, B., Schneider, R., Janetzky, S., Waibler, Z., Pandur, P., Kühl, M., Behrens, J.,

Von Mark, K. Der, Starzinski-Powitz, A., & Wixler, V. (2002). The LIM-only protein

FHL2 interacts with β-catenin and promotes differentiation of mouse myoblasts.

Journal of Cell Biology, *159*(1), 113–122. https://doi.org/10.1083/jcb.200202075

Matthews, J. M., Lester, K., Joseph, S., & Curtis, D. J. (2013). LIM-domain-only proteins in cancer. In *Nature Reviews Cancer* (Vol. 13, Issue 2, pp. 111–122). Nature Publishing Group. https://doi.org/10.1038/nrc3418

Mayor, R., & Theveneau, E. (2012). The neural crest. *Development (Cambridge)*, *140*(11), 2247–2251. https://doi.org/10.1242/dev.091751

McBeath, R., Pirone, D. M., Nelson, C. M., Bhadriraju, K., & Chen, C. S. (2004). Cell shape, cytoskeletal tension, and RhoA regulate stem cell lineage commitment. *Developmental Cell*, *6*(4), 483–495. https://doi.org/10.1016/S1534-5807(04)00075-9

McDowell, I. C., Manandhar, D., Vockley, C. M., Schmid, A. K., Reddy, T. E., & Engelhardt, B. E. (2018). Clustering gene expression time series data using an infinite Gaussian process mixture model. *PLoS Computational Biology*, *14*(1), e1005896. https://doi.org/10.1371/journal.pcbi.1005896

McGrew, M. J., Sherman, A., Lillico, S. G., Ellard, F. M., Radcliffe, P. A., Gilhooley, H. J., Mitrophanous, K. A., Cambray, N., Wilson, V., & Sang, H. (2008). Localised axial progenitor cell populations in the avian tail bud are not committed to a posterior Hox identity. *Development*, *135*(13), 2289–2299. https://doi.org/10.1242/dev.022020

McKinney, M. C., Fukatsu, K., Morrison, J., McLennan, R., Bronner, M. E., & Kulesa, P. M. (2013). Evidence for dynamic rearrangements but lack of fate or position restrictions in premigratory avian trunk neural crest. *Development (Cambridge)*, *140*(4), 820–830. https://doi.org/10.1242/dev.083725

McLennan, R., Dyson, L., Prather, K. W., Morrison, J. A., Baker, R. E., Maini, P. K., & Kulesa, P. M. (2012). Multiscale mechanisms of cell migration during development: Theory and experiment. *Development (Cambridge)*, *139*(16), 2935–2944. https://doi.org/10.1242/dev.081471

McLennan, R., Schumacher, L. J., Morrison, J. A., Teddy, J. M., Ridenour, D. A., Box, A. C., Semerad, C. L., Li, H., McDowell, W., Kay, D., Maini, P. K., Baker, R. E., & Kulesa, P. M. (2015). Neural crest migration is driven by a few trailblazer cells with a unique molecular signature narrowly confined to the invasive front. *Development (Cambridge)*, *142*(11), 2014–2025. https://doi.org/10.1242/dev.117507

McLennan, R., Teddy, J. M., Kasemeier-Kulesa, J. C., Romine, M. H., & Kulesa, P. M. (2010). Vascular endothelial growth factor (VEGF) regulates cranial neural crest migration in vivo. *Developmental Biology*, *339*(1), 114–125. https://doi.org/10.1016/j.ydbio.2009.12.022

Melby, A. E., Warga, R. M., & Kimmel, C. B. (1996). Specification of cell fates at the dorsal margin of the zebrafish gastrula. *Development*, *122*(7), 2225–2237.

Meulemans, D., & Bronner-Fraser, M. (2004). Gene-regulatory interactions in neural crest evolution and development. In *Developmental Cell* (Vol. 7, Issue 3, pp. 291–299). Elsevier. https://doi.org/10.1016/j.devcel.2004.08.007

Minchin, J. E. N., & Hughes, S. M. (2008). Sequential actions of Pax3 and Pax7 drive xanthophore development in zebrafish neural crest. *Developmental Biology*, *317*(2), 508–522. https://doi.org/10.1016/j.ydbio.2008.02.058

Montero-Balaguer, M., Lang, M. R., Sachdev, S. W., Knappmeyer, C., Stewart, R. A., De La Guardia, A., Hatzopoulos, A. K., & Knapik, E. W. (2006). The mother

superior mutation ablates foxd3 activity in neural crest progenitor cells and depletes neural crest derivatives in zebrafish. *Developmental Dynamics*, *235*(12), 3199–3212. https://doi.org/10.1002/dvdy.20959

Moro, E., Ozhan-Kizil, G., Mongera, A., Beis, D., Wierzbicki, C., Young, R. M., Bournele, D., Domenichini, A., Valdivia, L. E., Lum, L., Chen, C., Amatruda, J. F., Tiso, N., Weidinger, G., & Argenton, F. (2012). In vivo Wnt signaling tracing through a transgenic biosensor fish reveals novel activity domains. *Developmental Biology*, *366*(2), 327–340. https://doi.org/10.1016/j.ydbio.2012.03.023

Morrison, J. A., McLennan, R., Wolfe, L. A., Gogol, M. M., Meier, S., McKinney, M. C., Teddy, J. M., Holmes, L., Semerad, C. L., Box, A. C., Li, H., Hall, K. E., Perera, A. G., & Kulesa, P. M. (2017). Single-cell transcriptome analysis of avian neural crest migration reveals signatures of invasion and molecular transitions. *ELife*, *6*. https://doi.org/10.7554/eLife.28415

Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L., & Wold, B. (2008). Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature Methods*, *5*(7), 621–628. https://doi.org/10.1038/nmeth.1226

Mull, A., Kim, G., & Holaska, J. M. (2015). LMO7-null mice exhibit phenotypes consistent with emery-dreifuss muscular dystrophy. *Muscle and Nerve*, *51*(2), 222–228. https://doi.org/10.1002/mus.24286

Nagalakshmi, U., Wang, Z., Waern, K., Shou, C., Raha, D., Gerstein, M., & Snyder, M. (2008). The transcriptional landscape of the yeast genome defined by RNA sequencing. *Science*, *320*(5881), 1344–1349. https://doi.org/10.1126/science.1158441

Nakagawa, & Takeichi. (1995). Neural crest cell-cell adhesion controlled by sequential and subpopulation-specific expression of novel cadherins. *Development*, *121*(5), 1321–1332. https://doi.org/10.1016/0168-9525(95)90550-2

Nakagawa, & Takeichi. (1998). Neural crest emigration from the neural tube depends on regulated cadherin expression. *Development*, *125*(15), 2963–2971.

Nakamura, H., Hori, K., Tanaka-Okamoto, M., Higashiyama, M., Itoh, Y., Inoue, M., Morinaka, S., & Miyoshi, J. (2011). Decreased expression of LMO7 and its clinicopathological significance in human lung adenocarcinoma. *Experimental and Therapeutic Medicine*, *2*(6), 1053–1057. https://doi.org/10.3892/etm.2011.329

Nakamura, H., Mukai, M., Komatsu, K., Tanaka-Okamoto, M., Itoh, Y., Ishizaki, H., Tatsuta, M., Inoue, M., & Miyoshi, J. (2005). Transforming growth factor-β1 induces LMO7 while enhancing the invasiveness of rat ascites hepatoma cells. *Cancer Letters*, *220*(1), 95–99. https://doi.org/10.1016/j.canlet.2004.07.023

Nishikawa, S. (2012). Hemangioblast: An in vitro phantom. *Wiley Interdisciplinary Reviews: Developmental Biology*, *1*(4), 603–608. https://doi.org/10.1002/wdev.38

Northcutt, R. G., & Gans, C. (1983). The genesis of neural crest and epidermal placodes: a reinterpretation of vertebrate origins. In *The Quarterly review of biology* (Vol. 58, Issue 1, pp. 1–28). Q Rev Biol. https://doi.org/10.1086/413055

Ochoa, S. D., Salvador, S., & LaBonne, C. (2012). The LIM adaptor protein LMO4 is an essential regulator of neural crest development. *Developmental Biology*, *361*(2), 313–325. https://doi.org/10.1016/j.ydbio.2011.10.034

Olesnicky Killian, E. C., Birkholz, D. A., & Artinger, K. B. (2009). A role for chemokine signaling in neural crest cell migration and craniofacial development.

Developmental Biology, *333*(1), 161–172.

https://doi.org/10.1016/j.ydbio.2009.06.031

Ooshio, T., Irie, K., Morimoto, K., Fukuhara, A., Imai, T., & Takai, Y. (2004).

Involvement of LMO7 in the association of two cell-cell adhesion molecules, nectin

and E-cadherin, through afadin and α-actinin in epithelial cells. *Journal of Biological*

Chemistry, *279*(30), 31365–31373. https://doi.org/10.1074/jbc.M401957200

Ott, E. B., van den Akker, N. M. S., Sakalis, P. A., Gittenberger-de Groot, A. C., Te

Velthuis, A. J. W., & Bagowski, C. P. (2008). The lim domain only protein 7 is

important in zebrafish heart development. *Developmental Dynamics*, *237*(12),

3940–3952. https://doi.org/10.1002/dvdy.21807

Pan, Y., Freundlich, T., Weissman, T. A., Schoppik, D., Cindy Wang, X., Zimmerman,

S., Ciruna, B., Sanes, J. R., Lichtman, J. W., & Schier, A. F. (2013). Zebrabow:

Multispectral cell labeling for cell tracing and lineage analysis in zebrafish.

Development (Cambridge), *140*(13), 2835–2846.

https://doi.org/10.1242/dev.094631

Picelli, S., Faridani, O. R., Björklund, Å. K., Winberg, G., Sagasser, S., & Sandberg, R.

(2014). Full-length RNA-seq from single cells using Smart-seq2. *Nature Protocols*,

9(1), 171–181. https://doi.org/10.1038/nprot.2014.006

Piloto, S., & Schilling, T. F. (2010). Ovo1 links Wnt signaling with N-cadherin localization

during neural crest migration. *Development*, *137*(12), 1981–1990.

https://doi.org/10.1242/dev.048439

Placzek, M., & Briscoe, J. (2005). The floor plate: Multiple cells, multiple signals. In

Nature Reviews Neuroscience (Vol. 6, Issue 3, pp. 230–240). Nat Rev Neurosci.

146

https://doi.org/10.1038/nrn1628

Prasad, M. S., Charney, R. M., & García-Castro, M. I. (2019). Specification and formation of the neural crest: Perspectives on lineage segregation. In *Genesis* (Vol. 57, Issue 1, p. e23276). John Wiley and Sons Inc. https://doi.org/10.1002/dvg.23276

Raible, D. W., & Eisen, J. S. (1994). Restriction of neural crest cell fate in the trunk of the embryonic zebrafish. *Development*, *120*(3).

Richardson, J., Gauert, A., Briones Montecinos, L., Fanlo, L., Alhashem, Z. M., Assar, R., Marti, E., Kabla, A., Härtel, S., & Linker, C. (2016). Leader Cells Define Directionality of Trunk, but Not Cranial, Neural Crest Cell Migration. *Cell Reports*, *15*(9), 2076–2088. https://doi.org/10.1016/j.celrep.2016.04.067

Rosenquist, G. C. (1971). The location of the pregut endoderm in the chick embryo at the primitive streak stage as determined by radioautographic mapping. *Developmental Biology*, *26*(2), 323–335. https://doi.org/10.1016/0012-1606(71)90131-X

Row, R. H., Pegg, A., Kinney, B. A., Farr, G. H., Maves, L., Lowell, S., Wilson, V., & Martin, B. L. (2018). BMP and FGF signaling interact to pattern mesoderm by controlling basic helix-loop-helix transcription factor activity. *ELife*, *7*. https://doi.org/10.7554/eLife.31018

Row, R. H., Tsotras, S. R., Goto, H., & Martin, B. L. (2016). The zebrafish tailbud contains two independent populations of midline progenitor cells that maintain long-term germ layer plasticity and differentiate in response to local signaling cues. *Development (Cambridge)*, *143*(2), 244–254. https://doi.org/10.1242/dev.129015

Rozenblatt-Rosen, O., Stubbington, M. J. T., Regev, A., & Teichmann, S. A. (2017). The

    Human Cell Atlas: From vision to reality. In *Nature* (Vol. 550, Issue 7677, pp. 451–

    453). Nature Publishing Group. https://doi.org/10.1038/550451a

Saelens, W., Cannoodt, R., Todorov, H., & Saeys, Y. (2019). A comparison of single-

    cell trajectory inference methods. *Nature Biotechnology*, *37*(5), 547–554.

    https://doi.org/10.1038/s41587-019-0071-9

Sang, M., Ma, L., Sang, M., Zhou, X., Gao, W., & Geng, C. (2014). LIM-domain-only

    proteins: Multifunctional nuclear transcription coregulators that interacts with

    diverse proteins. In *Molecular Biology Reports* (Vol. 41, Issue 2, pp. 1067–1073).

    Kluwer Academic Publishers. https://doi.org/10.1007/s11033-013-2952-1

Santiago, A., & Erickson, C. A. (2002). Ephrin-B ligands play a dual role in the control of

    neural crest cell migration. *Development*, *129*(15), 3621–3632.

Sasai, N., Mizuseki, K., & Sasai, Y. (2001). Requirement of FoxD3-class signaling for

    neural crest determination in Xenopus. *Development*, *128*(13), 2525–2536.

Sasaki, M., Tsuji, N., Furuya, M., Kondoh, K., Kamagata, C., Kobayashi, D., Yagihashi,

    A., & Watanabe, N. (2003). PCD1, a novel gene containing PDZ and LIM domains,

    is overexpressed in human breast cancer and linked to lymph node metastasis.

    Anticancer Research, *23*(3 B), 2717–2721.

Schaum, N., Karkanias, J., Neff, N. F., May, A. P., Quake, S. R., Wyss-Coray, T.,

    Darmanis, S., Batson, J., Botvinnik, O., Chen, M. B., Chen, S., Green, F., Jones, R.

    C., Maynard, A., Penland, L., Pisco, A. O., Sit, R. V., Stanley, G. M., Webber, J. T.,

    … Weissman, I. L. (2018). Single-cell transcriptomics of 20 mouse organs creates

    a Tabula Muris. *Nature*, *562*(7727), 367–372. https://doi.org/10.1038/s41586-018-

0590-4

Schilling, T. F., & Kimmel, C. B. (1994). Segment and cell type lineage restrictions
during pharyngeal arch development in the zebrafish embryo. *Development*,
120(3).

Schilling, T. F., Le Pabic, P., & Hoffman, T. L. (2010). Using transgenic zebrafish (
Danio rerio) to study development of the craniofacial skeleton. *Journal of Applied
Ichthyology*, *26*(2), 183–186. https://doi.org/10.1111/j.1439-0426.2010.01401.x

Selleck, M. A. J., & Stern, C. D. (1991). Fate mapping and cell lineage analysis of
Hensen's node in the chick embryo. *Development*, *112*(2), 615–626.

Serbedzija, G. N., Bronner-Fraser, M., & Fraser, S. E. (1989). A vital dye analysis of the
timing and pathways of avian trunk neural crest cell migration. *Development*,
106(4), 809–816.

Serbedzija, G. N., Bronner-Fraser, M., & Fraser, S. E. (1992). Vital dye analysis of
cranial neural crest cell migration in the mouse embryo. *Development*, *116*(2), 297–
307.

Serbedzija, G. N., Fraser, S. E., & Bronner-Fraser, M. (1990). Pathways of trunk neural
crest cell migration in the mouse embryo as revealed by vital dye labelling.
Development, *108*(4), 605–612.

Simões-Costa, M., & Bronner, M. E. (2015). Establishing neural crest identity: a gene
regulatory recipe. In *Development (Cambridge, England)* (Vol. 142, Issue 2, pp.
242–257). Oxford University Press for The Company of Biologists Limited.
https://doi.org/10.1242/dev.105445

Sjöblom, B., Ylänne, J., & Djinović-Carugo, K. (2008). Novel structural insights into F-

actin-binding and novel functions of calponin homology domains. In *Current*

Opinion in Structural Biology (Vol. 18, Issue 6, pp. 702–708). Curr Opin Struct Biol.

https://doi.org/10.1016/j.sbi.2008.10.003

Smith, A., Robinson, V., Patel, K., & Wilkinson, D. G. (1997). The EphA4 and EphB1

receptor tyrosine kinases and ephrin-B2 ligand regulate targeted of branchial neural

crest cells. *Current Biology*, *7*(8), 561–570. https://doi.org/10.1016/s0960-

9822(06)00255-7

Soldatov, R., Kaucka, M., Kastriti, M. E., Petersen, J., Chontorotzea, T., Englmaier, L.,

Akkuratova, N., Yang, Y., Häring, M., Dyachuk, V., Bock, C., Farlik, M., Piacentino,

M. L., Boismoreau, F., Hilscher, M. M., Yokota, C., Qian, X., Nilsson, M., Bronner,

M. E., … Adameyko, I. (2019). Spatiotemporal structure of cell fate decisions in

murine neural crest. *Science*, *364*(6444). https://doi.org/10.1126/science.aas9536

Soo, K., O'Rourke, M. P., Khoo, P. L., Steiner, K. A., Wong, N., Behringer, R. R., &

Tam, P. P. L. (2002). Twist function is required for the morphogenesis of the

cephalic neural tube and the differentiation of the cranial neural crest cells in the

mouse embryo. *Developmental Biology*, *247*(2), 251–270.

https://doi.org/10.1006/dbio.2002.0699

Soriano, P. (1999). Generalized lacZ expression with the ROSA26 Cre reporter strain

[1]. In *Nature Genetics* (Vol. 21, Issue 1, pp. 70–71). Nat Genet.

https://doi.org/10.1038/5007

Spemann, H., & Mangold, H. (1924). über Induktion von Embryonalanlagen durch

Implantation artfremder Organisatoren. *Archiv Für Mikroskopische Anatomie Und*

Entwicklungsmechanik, *100*(3–4), 599–638. https://doi.org/10.1007/BF02108133

Stemple, D. L. (2005). Structure and function of the notochord: An essential organ for

    chordate development. In *Development* (Vol. 132, Issue 11, pp. 2503–2512). The

    Company of Biologists Ltd. https://doi.org/10.1242/dev.01812

Stemple, D. L., & Anderson, D. J. (1992). Isolation of a stem cell for neurons and glia

    from the mammalian neural crest. *Cell*, *71*(6), 973–985.

    https://doi.org/10.1016/0092-8674(92)90393-Q

Stewart, R. A., Arduini, B. L., Berghmans, S., George, R. E., Kanki, J. P., Henion, P. D.,

    & Look, A. T. (2006). Zebrafish foxd3 is selectively required for neural crest

    specification, migration and survival. *Developmental Biology*, *292*(1), 174–188.

    https://doi.org/10.1016/j.ydbio.2005.12.035

Stuhlmiller, T. J., & García-Castro, M. I. (2012). Current perspectives of the signaling

    pathways directing neural crest induction. In *Cellular and Molecular Life Sciences*

    (Vol. 69, Issue 22, pp. 3715–3737). Cell Mol Life Sci.

    https://doi.org/10.1007/s00018-012-0991-8

Sugiyama, M., Saitou, T., Kurokawa, H., Sakaue-Sawano, A., Imamura, T., Miyawaki,

    A., & Iimura, T. (2014). Live Imaging-Based Model Selection Reveals Periodic

    Regulation of the Stochastic G1/S Phase Transition in Vertebrate Axial

    Development. *PLoS Computational Biology*, *10*(12).

    https://doi.org/10.1371/journal.pcbi.1003957

Sugiyama, M., Sakaue-Sawano, A., Iimura, T., Fukami, K., Kitaguchi, T., Kawakami, K.,

    Okamoto, H., Higashijima, S. I., & Miyawaki, A. (2009). Illuminating cell-cycle

    progression in the developing zebrafish embryo. *Proceedings of the National*

    Academy of Sciences of the United States of America, *106*(49), 20812–20817.

https://doi.org/10.1073/pnas.0906464106

Sulston, J. E., Schierenberg, E., White, J. G., & Thomson, J. N. (1983). The embryonic

cell lineage of the nematode Caenorhabditis elegans. In *Developmental Biology*

(Vol. 100, Issue 1, pp. 64–119). Dev Biol. https://doi.org/10.1016/0012-

1606(83)90201-4

Szabó, A., & Mayor, R. (2018). Mechanisms of Neural Crest Migration. *Annual Review

of Genetics*, *52*(1), 43–63. https://doi.org/10.1146/annurev-genet-120417-031559

Takahashi, K., & Yamanaka, S. (2006). Induction of Pluripotent Stem Cells from Mouse

Embryonic and Adult Fibroblast Cultures by Defined Factors. *Cell*, *126*(4), 663–

676. https://doi.org/10.1016/j.cell.2006.07.024

Takemoto, T., Uchikawa, M., Yoshida, M., Bell, D. M., Lovell-Badge, R., Papaioannou,

V. E., & Kondoh, H. (2011). Tbx6-dependent Sox2 regulation determines neural or

mesodermal fate in axial stem cells. *Nature*, *470*(7334), 394–398.

https://doi.org/10.1038/nature09729

Teddy, J. M., & Kulesa, P. M. (2004). In vivo evidence for short- and long-range cell

communication in cranial neural crest cells. *Development*, *131*(24), 6141–6151.

https://doi.org/10.1242/dev.01534

Teixeira, V., Lourenco, S., Falzon, M., Capitanio, A., Bottoms, S., Carroll, B., Brown, J.,

George, J., & Janes, S. (2014). S112 Mmp12 And Lmo7 Are Key Genes Involved

In The Early Pathogenesis Of Squamous Cell Carcinoma Of The Lung. *Thorax*,

69(Suppl 2), A59–A60. https://doi.org/10.1136/thoraxjnl-2014-206260.118

Theveneau, E., & Linker, C. (2017). Leaders in collective migration: Are front cells really

endowed with a particular set of skills? In *F1000Research* (Vol. 6). Faculty of 1000

Ltd. https://doi.org/10.12688/f1000research.11889.1

Theveneau, E., Marchant, L., Kuriyama, S., Gull, M., Moepps, B., Parsons, M., & Mayor, R. (2010). Collective Chemotaxis Requires Contact-Dependent Cell Polarity. *Developmental Cell*, *19*(1), 39–53. https://doi.org/10.1016/j.devcel.2010.06.012

Thompson, M. A., Ransom, D. G., Pratt, S. J., MacLennan, H., Kieran, M. W., Detrich, H. W., Vail, B., Huber, T. L., Paw, B., Brownlie, A. J., Oates, A. C., Fritz, A., Gates, M. A., Amores, A., Bahary, N., Talbot, W. S., Her, H., Beier, D. R., Postlethwait, J. H., & Zon, L. I. (1998). The cloche and spadetail genes differentially affect hematopoiesis and vasculogenesis. *Developmental Biology*, *197*(2), 248–269. https://doi.org/10.1006/dbio.1998.8887

Tirosh, I., Izar, B., Prakadan, S. M., Wadsworth, M. H., Treacy, D., Trombetta, J. J., Rotem, A., Rodman, C., Lian, C., Murphy, G., Fallahi-Sichani, M., Dutton-Regester, K., Lin, J. R., Cohen, O., Shah, P., Lu, D., Genshaft, A. S., Hughes, T. K., Ziegler, C. G. K., … Garraway, L. A. (2016). Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science*, *352*(6282), 189–196. https://doi.org/10.1126/science.aad0501

Toro-Tapia, G., Villaseca, S., Beyer, A., Roycroft, A., Marcellini, S., Mayor, R., & Torrejoń, M. (2018). The Ric-8A/Gα13/FAK signalling cascade controls focal adhesion formation during neural crest cell migration in xenopus. *Development (Cambridge)*, *145*(22). https://doi.org/10.1242/dev.164269

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. *Nature*

Biotechnology, *32*(4), 381–386. https://doi.org/10.1038/nbt.2859

Trokovic, N., Trokovic, R., & Partanen, J. (2005). Fibroblast growth factor signalling and regional specification of the pharyngeal ectoderm. *International Journal of Developmental Biology*, *49*(7), 797–805. https://doi.org/10.1387/ijdb.051976nt

Tuttle, A. M., Hoffman, T. L., & Schilling, T. F. (2014). Rabconnectin-3a Regulates Vesicle Endocytosis and Canonical Wnt Signaling in Zebrafish Neural Crest Migration. *PLoS Biology*, *12*(5), e1001852. https://doi.org/10.1371/journal.pbio.1001852

Tzouanacou, E., Wegener, A., Wymeersch, F. J., Wilson, V., & Nicolas, J. F. (2009). Redefining the Progression of Lineage Segregations during Mammalian Embryogenesis by Clonal Analysis. *Developmental Cell*, *17*(3), 365–376. https://doi.org/10.1016/j.devcel.2009.08.002

Vaquerizas, J. M., Kummerfeld, S. K., Teichmann, S. A., & Luscombe, N. M. (2009). A census of human transcription factors: Function, expression and evolution. In *Nature Reviews Genetics* (Vol. 10, Issue 4, pp. 252–263). Nature Publishing Group. https://doi.org/10.1038/nrg2538

Vega-Lopez, G. A., Cerrizuela, S., & Aybar, M. J. (2017). Trunk neural crest cells: formation, migration and beyond. *The International Journal of Developmental Biology*, *61*(1–2), 5–15. https://doi.org/10.1387/ijdb.160408gv

Vibert, L., Aquino, G., Gehring, I., Subkankulova, T., Schilling, T. F., Rocco, A., & Kelsh, R. N. (2017). An ongoing role for Wnt signaling in differentiating melanocytes in vivo. *Pigment Cell and Melanoma Research*, *30*(2), 219–232. https://doi.org/10.1111/pcmr.12568

Vogeli, K. M., Jin, S. W., Martin, G. R., & Stainier, D. Y. R. (2006). A common progenitor

for haematopoietic and endothelial lineages in the zebrafish gastrula. *Nature*,

443(7109), 337–339. https://doi.org/10.1038/nature05045

Vogt, W. (1929). Gestaltungsanalyse am Amphibienkeim mit Örtlicher Vitalfärbung - II.

Teil. Gastrulation und Mesodermbildung bei Urodelen und Anuren. *Wilhelm Roux'*

Archiv Für Entwicklungsmechanik Der Organismen, *120*(1), 384–706.

https://doi.org/10.1007/BF02109667

Wada, N., Javidan, Y., Nelson, S., Carney, T. J., Kelsh, R. N., & Schilling, T. F. (2005).

Hedgehog signaling is required for cranial neural crest morphogenesis and

chondrogenesis at the midline in the zebrafish skull. *Development*, *132*(17), 3977–

3988. https://doi.org/10.1242/dev.01943

Waddington, C. H. (1932). III. Experiments on the development of chick and duck

embryos, cultivated in vitro. *Philosophical Transactions of the Royal Society of*

London. Series B, Containing Papers of a Biological Character, *221*(474–482),

179–230. https://doi.org/10.1098/rstb.1932.0003

Waddington, C. H. (1957). *The Strategy of the Genes; a Discussion of Some Aspects of*

Theoretical Biology. Allen & Unwin.

Wagner, D. E., Weinreb, C., Collins, Z. M., Briggs, J. A., Megason, S. G., & Klein, A. M.

(2018). Single-cell mapping of gene expression landscapes and lineage in the

zebrafish embryo. *Science*, *360*(6392), 981–987.

https://doi.org/10.1126/science.aar4362

Wang, H. U., & Anderson, D. J. (1997). Eph family transmembrane ligands can mediate

repulsive guidance of trunk neural crest migration and motor axon outgrowth.

Neuron, *18*(3), 383–396. https://doi.org/10.1016/S0896-6273(00)81240-4

Wang, Melville, D. B., Montero-Balaguer, M., Hatzopoulos, A. K., & Knapik, E. W.
(2011). Tfap2a and Foxd3 regulate early steps in the development of the neural
crest progenitor population. *Developmental Biology*, *360*(1), 173–185.
https://doi.org/10.1016/j.ydbio.2011.09.019

Wang, S., Karikomi, M., Maclean, A. L., & Nie, Q. (2019). Cell lineage and
communication network inference via optimization for single-cell transcriptomics.
Nucleic Acids Research, *47*(11), 66. https://doi.org/10.1093/nar/gkz204

Weisblat, D. A., & Shankland, M. (1985). Cell lineage and segmentation in the leech.
*Philosophical Transactions of the Royal Society of London. Series B, Biological
Sciences*, *312*(1153), 39–56. https://doi.org/10.1098/rstb.1985.0176

Weisblat, David A., Sawyer, R. T., & Stent, G. S. (1978). Cell lineage analysis by
intracellular injection of a tracer enzyme. *Science*, *202*(4374), 1295–1298.
https://doi.org/10.1126/science.725606

Westerfield. (2000). *The zebrafish book. A guide for the laboratory use of zebrafish
(Danio rerio)*. University of Oregon Press. https://zfin.org/zf_info/zfbook/zfbk.html

Weston, J. A. (1970). The migration and differentiation of neural crest cells. In *Advances
in morphogenesis* (Vol. 8, pp. 41–114). Adv Morphog. https://doi.org/10.1016/b978-
0-12-028608-9.50006-5

Wetts, R., & Fraser, S. E. (1988). Multipotent precursors can give rise to all major cell
types of the frog retina. *Science*, *239*(4844), 1142–1145.
https://doi.org/10.1126/science.2449732

Williams, R. M., Candido-Ferreira, I., Repapi, E., Gavriouchkina, D., Senanayake, U.,

Ling, I. T. C., Telenius, J., Taylor, S., Hughes, J., & Sauka-Spengler, T. (2019). Reconstruction of the Global Neural Crest Gene Regulatory Network In Vivo. *Developmental Cell*, *51*(2), 255-276.e7. https://doi.org/10.1016/j.devcel.2019.10.003

Wolf, F. A., Hamey, F. K., Plass, M., Solana, J., Dahlin, J. S., Göttgens, B., Rajewsky, N., Simon, L., & Theis, F. J. (2019). PAGA: graph abstraction reconciles clustering with trajectory inference through a topology preserving map of single cells. *Genome Biology*, *20*(1), 1–9. https://doi.org/10.1186/s13059-019-1663-x

Wörthmüller, J., & Rüegg, C. (2020). The crosstalk between FAK and Wnt signaling pathways in cancer and its therapeutic implication. *International Journal of Molecular Sciences*, *21*(23), 1–23. https://doi.org/10.3390/ijms21239107

Wozniak, M. A., Baker, B. M., Chen, C. S., & Wilson, K. L. (2013). The emerin-binding transcription factor Lmo7 is regulated by association with p130cas at focal adhesions. *PeerJ*, *2013*(1). https://doi.org/10.7717/peerj.134

Wu, J., Saint-Jeannet, J. P., & Klein, P. S. (2003). Wnt-frizzled signaling in neural crest formation. In *Trends in Neurosciences* (Vol. 26, Issue 1, pp. 40–45). Elsevier. https://doi.org/10.1016/S0166-2236(02)00011-5

Wu, R. S., Lam, I. I., Clay, H., Duong, D. N., Deo, R. C., & Coughlin Correspondence, S. R. (2018). A Rapid Method for Directed Gene Knockout for Screening in G0 Zebrafish. *Developmental Cell*, *46*, 112–125. https://doi.org/10.1016/j.devcel.2018.06.003

Wu, R. S., Lam, I. I., Clay, H., Duong, D. N., Deo, R. C., & Coughlin, S. R. (2018). A Rapid Method for Directed Gene Knockout for Screening in G0 Zebrafish.

Developmental Cell, *46*(1), 112-125.e4.

https://doi.org/10.1016/j.devcel.2018.06.003

Yamaguchi, T. P., Takada, S., Yoshikawa, Y., Wu, N., & McMahon, A. P. (1999). T

(Brachyury) is a direct target of Wnt3a during paraxial mesoderm specification.

Genes and Development, *13*(24), 3185–3190.

https://doi.org/10.1101/gad.13.24.3185

Zeisel, A., M͡oz-Manchado, A. B., Codeluppi, S., Lönnerberg, P., Manno, G. La, Juréus,

A., Marques, S., Munguba, H., He, L., Betsholtz, C., Rolny, C., Castelo-Branco, G.,

Hjerling-Leffler, J., & Linnarsson, S. (2015). Cell types in the mouse cortex and

hippocampus revealed by single-cell RNA-seq. *Science*, *347*(6226), 1138–1142.

https://doi.org/10.1126/science.aaa1934

Zhang, A. W., O'Flanagan, C., Chavez, E. A., Lim, J. L. P., Ceglia, N., McPherson, A.,

Wiens, M., Walters, P., Chan, T., Hewitson, B., Lai, D., Mottok, A., Sarkozy, C.,

Chong, L., Aoki, T., Wang, X., Weng, A. P., McAlpine, J. N., Aparicio, S., … Shah,

S. P. (2019). Probabilistic cell-type assignment of single-cell RNA-seq for tumor

microenvironment profiling. *Nature Methods*, *16*(10), 1007–1015.

https://doi.org/10.1038/s41592-019-0529-1