

Dynamic regulation of neural crest stem cell identity during embryogenesis and oncogenesis

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

By

Debadrita Bhattacharya

August 2021

© 2021 Debadrita Bhattacharya

DYNAMIC REGULATION OF NEURAL CREST STEM CELL IDENTITY DURING EMBRYOGENESIS AND ONCOGENESIS

Debadrita Bhattacharya, PhD
Cornell University, 2021

Embryonic development involves the generation of the vast cellular diversity observed in an adult organism. This process requires spatially and temporally controlled cell state transitions through which a stem cell becomes a differentiated cell type. Thus, delineating the mechanisms that control changes in cell identity is fundamental to our understanding of organismal development. Further, since stemness can also be reactivated during oncogenesis, determining its regulation also has broad implications for cancer biology. Our lab employs the neural crest as a model system to examine the molecular control of stem cell identity in development and disease. The neural crest is a vertebrate-specific embryonic stem cell population that forms a wide range of derivatives, including neurons, chondrocytes, and melanocytes. These cells delaminate from the central nervous system and migrate through well-defined pathways within the developing embryo. Through the course of migration, they progressively lose their stem cell properties and become committed to a specific fate. Intriguingly, during tumorigenesis, neural crest stem cell identity can reemerge in adult cells, giving rise to cancers like melanoma and neuroblastoma.

In the first part of my thesis research, I combined genomic analyses with classical embryology to explore the regulation of stem cell identity during avian neural crest development. This work demonstrated how Yamanaka factors, including Sox2, Oct4, and Lin28a, control neural crest multipotency by regulating the epigenome and transcriptome of this specialized stem cell population. My findings revealed that the general pluripotency program cooperates with tissue-specific factors to modulate the gene regulatory network that endows neural crest cells with their unique features.

In the second part of my dissertation, I aimed to delineate the mechanisms that reactivate stem cell fate in neural crest derived cancers such as melanoma. By employing genomic and functional analysis,

I discovered that neural crest cells share many molecular features with cancer cells, including the metabolic adaptation Warburg effect and reliance on the same signal transduction pathway for migration. Further, by profiling melanocytic progenitors and human melanoma cells, I identified thousands of neural crest-specific epigenomic regions that reemerge during malignant transformation to establish a stem-like state within the tumor. Together, these results highlight how developmental regulatory networks are co-opted during cancer progression to promote tumor heterogeneity.

BIOGRAPHICAL SKETCH

Debadrita Bhattacharya was born in the bustling city of Kolkata, located in the eastern part of India. She completed her schooling at Carmel Convent School, where she developed a strong interest in Biology and Chemistry. During her school years, Debadrita also served as a student journalist for the national newspaper "The Statesman" and trained for over a decade as a classical Manipuri dancer. For her undergraduate studies, Debadrita pursued an Honors degree in Microbiology at the Lady Brabourne College, University of Calcutta. Her Bachelors curriculum instilled in her a fascination for Molecular Biology, and she developed an interest in pursuing a career in scientific research. With this goal in mind, she moved to Mumbai to pursue a three-year Masters by Research program at the esteemed Tata Institute of Fundamental Research. Here Debadrita joined the lab of Prof. Basuthkar J. Rao, where she studied the dynamics of DNA Damage Response in eukaryotic cells. Through her Masters' project, she became interested in studying cellular signaling systems and, in the future, wanted to interrogate how these signals were orchestrated at the tissue and organismal level. Thus, upon moving to the US to pursue her Ph.D. at Cornell University, she joined the Developmental Biology lab of Dr. Marcos Simoes-Costa. Dr. Simoes-Costa's pioneering work on neural crest development had led to the assembly of a comprehensive gene regulatory network (GRN) that controls the unique properties of this stem cell population. In the lab, Debadrita's research focused on using classical embryology and functional genomics analysis to determine how developmental signals interacted with this GRN to control neural crest stem cell fate. Upon receiving her Ph.D. Debadrita looks forward to joining the lab of Dr. Julien Sage at Stanford University, where she will be studying the molecular basis of cellular heterogeneity in Small Cell Lung Cancer.

I dedicate my thesis to my grandfather Late Mr. Amiya Kumar Bhattacharya and to my mother Dr. Swati Bhattacharya for being amazing teachers who taught me to think and ask questions.

ACKNOWLEDGMENTS

Although this thesis has my name on it, I couldn't have done it without the unrelenting support from my village of mentors, friends, family and the broader scientific community. The past 5 years in Cornell has been quite challenging, but my support system here made sure I wasn't in it alone.

First and foremost, I would like to thank my thesis advisor Dr. Marcos Simoes-Costa for being invested in every aspect of my scientific work. I will be forever indebted to him for teaching me everything that I know about Developmental Biology, for trusting me to undertake bold projects and for pushing me to strive to be the very best. Marcos's mentoring focused on bringing out my strengths and taught me how to think like a scientist.

I want to thank Dr. Andrew White for his guidance and support throughout our collaboration. I greatly benefitted from our many scientific discussions and from his inputs during postdoctoral lab search.

I am also indebted to my thesis committee members Dr. Andrew Grimson, Dr. Paul Soloway and Dr. Charles Danko for their valuable insights in my projects and for being heavily invested in my science and also my career. Their encouragement has been critical to my progress in grad school.

I would like to thank all past and present members of the Simoes-Costa for making the lab such a fun place to work in. I have learnt a lot from every member of the lab and their constant support was absolutely critical for all my projects. I would especially like to thank Dr. Ana Azambuja who was and continues to be our lab mom. Ana's kindness and her readiness to help has been my saving grace these past years. I am also thankful to Megan Rothstein, my friend, philosopher and guide in the lab and Jackie Copeland, my tissue culture buddy. It was also a lot of fun collaborating with Austin Hovland this past year on our shared project, and I am very thankful for his help with all the bioinformatic analysis. Lastly, I want to thank all the newcomers in our lab: Dr. Tatiane Kanno, Dr. Dimitrus Pramio and Fjodor Merkuri and I hope they have a wonderful time in the Simoes-Costa lab.

I am also very fortunate to have a wonderful group of friends here at Cornell and elsewhere whose love, support and kindness has been essential to my successful completion of my thesis. These include Jessica West, Mariela Nunez Santos, Porter Hall, Erik Henze, Brendan Ito, Saptarshi Chatterjee, Irma Fernandez, Ajinkya Rahane, Riasat Zaman, Jumana Badar and Saket Bagde, Mohona Chakraborty and Swagata Sen.

I am so thankful to my amazing family for their never-ending support and love for me. My parents were my first teachers and having ingrained in me the wonders of education they stood by me and trusted all my decisions. My sister was my first science teacher and I grudgingly agree that maybe I would have ended up doing a PhD if it wasn't for her.

Lastly, I want to thank my partner Amartya Mitra for being my best friend over the past four years and my cat Zelda for being an absolute joy.

Table of Contents

Chapter I: Introduction	11
Neural Crest development:	12
The regulatory basis of neural crest multipotency	14
Re-activation of stem cell identity in neural crest-derived cancers	15
Thesis conclusions:.....	17
References:.....	19
Chapter II: Control of neural crest multipotency by Wnt signaling and the Lin28/let-7 axis	22
Abstract	22
Introduction.....	23
Results	24
Discussion	40
Acknowledgments:	43
Methods	43
References	63
Chapter III: Pluripotency factors are repurposed to shape the epigenomic landscape in neural crest stem cells	67
Abstract	67
Introduction.....	68
Results	70
Discussion	89
Methods	91
References	101
Chapter IV: Metabolic reprogramming promotes neural crest migration via Yap/Tead signaling ...	104
Abstract	104
Introduction.....	104
Results	106
Discussion	121
Acknowledgments	123
Methods	123
.....	138
References	143
Chapter V: Neural crest metabolism: at the crossroads of development and disease	147

Abstract	147
Introduction.....	147
Metabolic regulation of neural crest development.....	150
Warburg Effect in other developmental contexts	157
Effects of metabolic disorders in neural crest development	162
Conclusions.....	170
Box 1. Other metabolic pathways and metabolites important for neural crest formation	173
Acknowledgements.....	175
References	176
<i>Chapter VI: Co-option of developmental gene circuits in neural crest-derived cancers</i>	<i>183</i>
Abstract:	183
Introduction:.....	184
Results:	186
Discussion	204
Methods:	206
References:	214

Chapter I: Introduction

The body of an adult vertebrate organism comprises hundreds of distinct cell types derived from a single pluripotent cell. This cellular diversity is generated by the process of cell differentiation, during which a multipotent stem cell undergoes drastic changes in its molecular identity to become an adult cell type. Thus, the balance between multipotency and differentiation is fundamental to embryo development (Young, 2011). This balance is also critical for the homeostasis of adult tissues since the reactivation of multipotency programs in differentiated cells underlies oncogenesis (Kulesa et al., 2013; Micalizzi and Ford, 2009). Despite this, we do not yet

have a clear understanding of the mechanisms that mediate these cell state transitions and those that ensure they are deployed in a spatially and temporally controlled manner. The Waddington model postulates that both cell differentiation and de-differentiation involve drastic changes in the epigenomic

landscape, which in turn activates or silences the transcriptional networks characterizing

different cell states (Fig 1). Extracellular signaling pathways have emerged as the primary players that mediate this epigenomic rewiring to dynamically regulate cellular plasticity and cell-fate commitment (Barolo and Posakony, 2002; Fagnocchi et al., 2016; Kumar et al., 2017). Yet, how extracellular signals crosstalk with cell-intrinsic factors to promote changes in cellular identity remains a central question in developmental biology. Determining how this is accomplished *in vivo* also has broad implications in cancer biology and regenerative medicine.

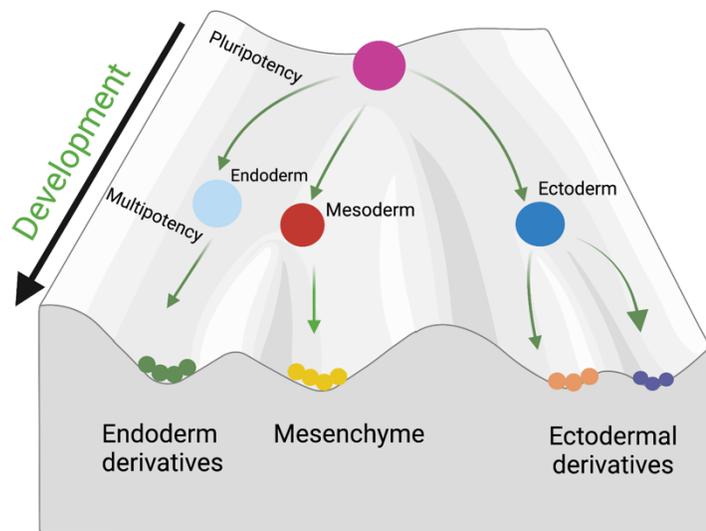


Fig 1: Waddington's epigenetic model of cell differentiation. The model shows that transition from pluripotency to lineage specific derivatives involve moving downhill on the epigenetic energy profile.

Motivated by these outstanding questions, my thesis research has centered around understanding two distinct aspects of cell state transitions: 1. How the dynamic regulation of multipotency and differentiation is implemented during embryogenesis and 2. how terminally differentiated cells re-acquire progenitor identity during oncogenesis. To investigate these questions, I used the embryonic stem cell population Neural Crest as my experimental model system. The Neural Crest (NC) is a vertebrate-specific multipotent cell-type that uniquely differentiates into ectodermal and mesenchymal derivatives and contributes to several tissues throughout the adult body (Simoes-Costa and Bronner, 2015). NC cells are also migratory, and through development, they transition from a pre-migratory stem cell to a post migratory differentiated state, which is easily tractable *in vivo* (Bronner and Simoes-Costa, 2016; Theveneau and Mayor, 2012). Further, this embryonic cell type is highly relevant in tumor biology as NC-derived cancers such as melanoma and neuroblastoma form very aggressive tumors, in part by recapitulating the developmental trajectory of this stem cell population (Johnsen et al., 2019; Kaufman et al., 2016; Kulesa et al., 2013). We also currently possess a comprehensive gene regulatory network for this cell type which provides a useful framework for studying the reorganization of the progenitors' transcriptional programs during differentiation and tumorigenesis (Hovland et al., 2020; Martik and Bronner, 2017; Sauka-Spengler and Bronner-Fraser, 2008; Simoes-Costa and Bronner, 2015). Together, these features make NC an excellent model for investigating the interplay between environmental cues and the cell-intrinsic genetic circuits, which together orchestrate cell state transitions.

Neural Crest development:

In the vertebrate embryo, NC development occurs from gastrulation to late organogenesis and comprises different stages, including induction, specification, migration, and differentiation (Fig 2). These developmental events are orchestrated by a gene regulatory network composed of several transcriptional modules that endow NC cells with their unique traits (Sauka-Spengler and Bronner-Fraser, 2008; Simoes-Costa and Bronner, 2015). NC induction occurs during early gastrulation and requires the

crosstalk between extracellular signals FGF, BMP, and WNT, which activate Neural Plate Border (NPB) genes such as *Pax3/7*, *Msx1*, and *Zic1* at the region between the neural plate and non-neural ectoderm (Simoes-Costa and Bronner, 2015) (Fig 2). This forms NPB cells that are competent to receive NC-inducing signals (Fig 2). Subsequently, during neural tube closure, the activation of the NC specification program causes a subset of NPB cells to become NC stem cells, which come to occupy the elevated dorsal folds of the neural tube.

NC specification involves the activation of early NC genes such as *FoxD3*, *Ets1*, *cMyc*, *Sox5*, which maintains the NC stem cell pool by regulating their proliferation and survival and also control NC cell state decisions and terminal differentiation.

Following the specification, the NC stem cells express genes regulating migration including *Zeb2*, *Twist1*, *Sox9* etc., which

allows these cells to undergo Epithelial to Mesenchymal Transition (EMT) and delaminate out from the Dorsal Neural Tube (DNT) (Fig 2). As the NC stem cells migrate away from the DNT along predetermined routes, they become progressively fate restricted and begin to express lineage marker genes such as *Runx2*, *NeuroG*, and *Mitf* (Hovland et al., 2020; Simoes-Costa and Bronner, 2015). Alongside activation of these cell-type-specific transcriptional circuits, NC differentiation also involves concomitant silencing of regulatory program characterizing NC stem cells (Azambuja and Simoes-Costa, 2021; Bhattacharya et al., 2018). This switch underlies the cell state transition of the multipotent NC to terminally differentiated adult cells.

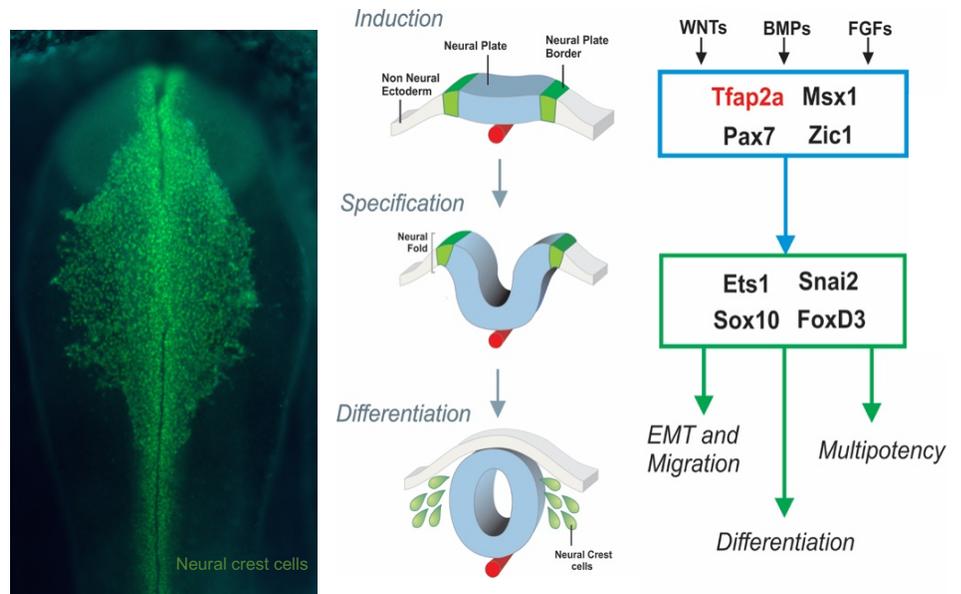


Fig 2: Top-down view of chick embryo head showing neural crest cells labelled in green. Schematic of the cross-sectional view of avian embryo head showing the distinct stages of neural crest formation. Simplified diagram of the neural crest gene regulatory network, highlighting some of the key transcription factors and signaling systems.

The regulatory basis of neural crest multipotency

By using a combination of cell transplantation and dye-labeling experiments, classical developmental biology studies have dissected the vast array of cellular derivatives of NC (Dupin et al., 2018b; Hall, 2005). These include the neurons and glia of the peripheral and enteric nervous system, craniofacial connective tissue, facial cartilage and bone, smooth muscle cells, melanocytes, and neuroendocrine cells. Indeed, neural crest progenitors have the remarkable ability to give rise to both ectodermal and mesodermal-like derivatives and hence are often referred to as the “fourth germ layer” (Hall, 2000). The scope of NC multipotency has been extensively investigated and several studies from a span of over three decades have shown that the cranial neural crest population is highly plastic. For example, *in vivo* lineage tracing of an individual dye-labelled NC cell has shown that pre-migratory cranial NC cells are truly multipotent and post-migration contribute to all NC-derived tissues (Bronner-Fraser, 1994). This has been further supported by *in vitro* clonal analysis experiments, which revealed that a single cranial NC could give rise to clonal populations comprising all differentiated derivatives, albeit in varying proportions (Dupin et al., 2018b; Ito et al., 1993). Lastly, a recent study using a “Confetti mouse” model, which allows for fluorescence-based lineage analysis of individual NC cells, verified that most NC stem cells give rise to at least two different derivatives *in vivo* (Baggiolini et al., 2015).

While the multipotent nature of NC cells has been explicitly demonstrated, the regulation of NC stem cell identity is still an area of active research. Two contrasting theories have attempted to explain the origin of NC multipotency. The first proposes a “molecular model” in which NC stem cells de-differentiate from early ectoderm and acquire multipotency due to the combined influence of WNT, BMP, and FGF signaling (Dupin et al., 2018a). According to this model, the establishment of the NPB by synergistic action of the above extracellular signals primes the cells within this region to activate the NC specification program, which allows them to re-acquire a multipotent state (LaBonne and Bronner-Fraser, 1998). This has been substantiated by the observation that increased FGF and WNT signaling is sufficient to induce the expression of NC stem cell genes in differentiated non-neural ectoderm (Leung et al., 2016).

In contrast, the opposing model suggests that NC is formed by a subset of ectodermal cells which have retained the multipotent features of the early blastula (Buitrago-Delgado et al., 2015). This model is based on the observation that NC genes such as *Sox5* and *Snail* are expressed in the pluripotent *Xenopus* blastula, and over the course of development their expression becomes restricted to the NC cells and is lost in the surrounding tissues (Buitrago-Delgado et al., 2015). This ambiguity on the origin of NC multipotency was partially resolved by recent single cell transcriptome studies in *Xenopus* and Zebrafish embryos which showed that NC stem cells are derived from the ectoderm but becomes molecularly distinct from it following the activation of NC specific transcriptional programs (Briggs et al., 2018; Farrell et al., 2018). While this evidence the former de-differentiation model, it does not inform on the molecular mechanisms by which ectodermal cells regain multipotency to become NC stem cells. Interestingly, the core pluripotency circuit composed of transcription factors *Nanog*, *Sox2*, *Klf4*, and *Oct4*, sufficient for reprogramming differentiated somatic cells into a stem cell (Young, 2011), is not active in NC stem cells. Indeed, *Nanog* and *Klf4* are lowly expressed in avian neural crest cells, while high *Sox2* levels are observed in neural plate cells (Hovland et al., 2020; Wakamatsu et al., 2004). This suggests that during development, the transition from pluripotency to multipotency requires the activation of a distinct molecular circuitry that likely cooperates with the above-mentioned extracellular signaling pathways to induce NC cells. Defining the components of this multipotency circuitry would not only inform on the early steps of neural crest formation but could also provide a general framework for understanding how specialized stem cells are established during organogenesis. Further, discerning the mechanisms that control the NC multipotent state during embryogenesis is critical for determining how this stem cell fate re-emerges during oncogenesis.

Re-activation of stem cell identity in neural crest-derived cancers

An emerging question in the field of cancer biology is how differentiated cancer cells hijack embryonic programs to regain a stem cell-like state, critical for tumor formation (Diener and Sommer, 2021; Manoranjan et al., 2012; Milla et al., 2012). Studies in several cancer types have revealed that re-

emergence of progenitor identity in tumor-initiating cells is an early event of oncogenesis. This involves the anomalous activation of developmental pathways that promote proliferation, invasion and self-renewal (Pomerantz et al., 2020; Xue et al., 2020). Thus, uncovering the upstream mechanisms that reactivate embryonic process in tumor cells has important therapeutic implications.

The stem cell population neural crest serves as an excellent model for understanding how development is recapitulated during oncogenesis. A key feature of neural crest is their extensive migratory ability, which enables these cells to colonize the distinct locations within the embryo where they will ultimately differentiate (Sauka-Spengler and Bronner-Fraser, 2008). In the course of this multi-step process, neural crest cells undergo epithelial-to-mesenchymal transition (EMT), invade and migrate through surrounding tissues, and eventually perform mesenchymal-to-epithelial transition (MET) upon reaching their destination (Theveneau and Mayor, 2012). This process is very similar to cancer invasion and is regulated by many of the same factors and signaling systems. Importantly, neural crest-derived cancers such as melanoma and neuroblastoma are especially prone to metastasis, which has led to the hypothesis that the re-activation of the progenitor migratory program underlies the aggressiveness of these tumors (Kulesa et al., 2013; Shakhova, 2014) (Johnsen et al., 2019).

Particularly, melanoma progression considerably recapitulates the developmental trajectory of neural crest. The transition from differentiated melanocytes to primary and eventually metastatic melanoma is, in part, driven by the re-activation of transcriptional programs characteristic of the stem cell population (Boiko et al., 2010; Kaufman et al., 2016). The co-existence of cancer cells at different stages of de-differentiation further contributes to the heterogeneity of melanoma tumors. Transcriptome analysis of cultured melanoma cells has revealed at least two distinct cell populations that manifest as the “proliferative phenotype,” which promotes tumor growth, and the “invasive phenotype,” which initiates metastasis (Hoek et al., 2006; Tirosh et al., 2016). Interestingly, proliferative melanoma cells are transcriptionally similar to melanocytes, while the invasive cells tend to be more de-differentiated and are neural crest-like (Arozarena and Wellbrock, 2019; Fazio et al., 2021; Restivo et al., 2017). Recent studies indicate that a switch from proliferative to invasive state underlies melanoma progression, as tumors with

predominantly invasive cells are highly metastatic and are resistant to targeted therapy (Fazio et al., 2021; Rambow et al., 2018). Intriguingly, downregulation of transcription factors controlling neural crest migration can inhibit this switch in melanoma cells (Cheng et al., 2015; Zhang et al., 2020). Alternately, transplantation of primary melanoma cells in the embryonic neural crest microenvironment is sufficient to promote cancer metastasis (Kulesa et al., 2006). These observations indicate that upstream cues controlling neural crest migration also function to reactivate this transcriptional program in melanoma cells. Further recent studies have shown that NC-specific enhancer elements are reactivated in melanoma tumors, which hints at a more broadly shared regulatory program between the embryonic cell population and its malignant derivative (Kaufman et al., 2016). This has led the field to postulate that molecular mechanisms that dictate cell state transitions during NC development are also operative at distinct stages of melanoma progression. Thus, translating lessons learned from NC development to melanoma cellular plasticity could provide new insights and novel methods for cancer treatment.

Thesis conclusions:

A defining property of developmental gene regulatory networks is that they are dynamic and can evolve by altering their topology as cells progress through the steps of fate commitment. Additionally, throughout embryogenesis, the core components of a few transcriptional networks are repurposed in multiple tissue types, which then interact with more specialized lineage-specific factors to establish very distinct cell fates. While this provides a conceptual framework for understanding how cellular diversity is attained during development, we currently lack empirical insights into the mechanisms that mediate these sequential changes in network architecture. In this dissertation, I present my work towards understanding how the transcriptional program underlying the formation of the specialized stem cell NC is established during early development, silenced upon differentiation, and is eventually reactivated during tumorigenesis.

In the first part of my thesis work I adopted a two-pronged strategy to delineate the molecular basis of neural crest multipotency. First, taking a candidate-based approach I characterized the function of the

pluripotency factor *Lin28* in early NC stem cells where this gene is highly enriched. This led me to uncover a regulatory circuit, composed of *Lin28a* and the *let-7* miRNAs, which is activated downstream of Wnt signaling, to control the deployment and subsequent silencing of the neural crest multipotency program. My findings revealed that high levels of Wnt signalling in the dorsal neural tube induce expression of *Lin28* in pre-migratory NC, which in turn inhibits *let-7* miRNAs. As neural crest cells migrate away from the Wnt niche, their levels of *Lin28* drop and *let-7* miRNAs increase, which directly targets and inhibits NC multipotency genes. This work thus highlighted a mechanism by which the NC multipotency program is post-transcriptionally silenced upon differentiation. Next, to more comprehensively define the epigenomic and transcriptional profile of the NC stem cells and evaluate how it changes upon differentiation, in the lab, we performed time-course ATAC-seq and RNA-seq of primary avian NC cells. This analysis revealed a surprising role of the Yamanaka factors OCT4 and SOX2 in the formation of NC stem cells. I found that the OCT4-SOX2 heterodimer interacts with NC-specific pioneer factors to regulate thousands of genomic regions active in early NC cells. However, similar to *Lin28*, the regulatory targets of the OCT4-SOX2 dimer in multipotent NC cells were completely distinct from their targets in pluripotent ES cells. Together, these studies show that though components of the pluripotency network are repurposed during NC formation, the multipotent state established downstream of these ES cell factors is characterized by a distinct gene regulatory network that manifests the unique properties of cell type.

In the second part of my dissertation research, I aimed at delineating the mechanisms that reactivate stem cell fate in NC-derived cancers such as melanoma. To approach this problem from a developmental biology perspective, I first attempted to identify novel pathways that control neural crest migration since this process is most often mimicked by cancer cells. Surprisingly, this study revealed that the metabolic adaptation Warburg Effect (WE) is a crucial regulator of EMT and migration in NC cells. A characteristic feature of malignancy, WE increases the glycolytic flux of cancer cells and impacts distinct aspects of tumorigenesis. However, my finding that WE is also essential for neural crest development implicated it as an embryonic process that recurs in cancer cells. This study also uncovered that WE function to regulate the Yap/Tead signaling pathway, which transcriptionally activates the NC migratory

program to promote EMT. Excitingly, my subsequent work directly characterizing the NC-like stem cells in melanoma tumors revealed that the Yap/Tead pathway is essential for re-emergence of the NC regulatory network in cancer cells and is causal to melanoma de-differentiation. Together, these studies provided two distinct examples supporting the now long-standing postulate the tumorigenesis is essentially embryogenesis happening in reverse, in the wrong time and wrong place.

References:

- Arozarena, I., and Wellbrock, C. (2019). Phenotype plasticity as enabler of melanoma progression and therapy resistance. *Nat Rev Cancer* 19, 377-391.
- Azambuja, A.P., and Simoes-Costa, M. (2021). The connectome of neural crest enhancers reveals regulatory features of signaling systems. *Dev Cell* 56, 1268-1282 e1266.
- Baggiolini, A., Varum, S., Mateos, J.M., Bettosini, D., John, N., Bonalli, M., Ziegler, U., Dimou, L., Clevers, H., Furrer, R., *et al.* (2015). Premigratory and migratory neural crest cells are multipotent in vivo. *Cell Stem Cell* 16, 314-322.
- Barolo, S., and Posakony, J.W. (2002). Three habits of highly effective signaling pathways: principles of transcriptional control by developmental cell signaling. *Genes Dev* 16, 1167-1181.
- Bhattacharya, D., Rothstein, M., Azambuja, A.P., and Simoes-Costa, M. (2018). Control of neural crest multipotency by Wnt signaling and the Lin28/let-7 axis. *Elife* 7.
- Boiko, A.D., Razorenova, O.V., van de Rijn, M., Swetter, S.M., Johnson, D.L., Ly, D.P., Butler, P.D., Yang, G.P., Joshua, B., Kaplan, M.J., *et al.* (2010). Human melanoma-initiating cells express neural crest nerve growth factor receptor CD271. *Nature* 466, 133-137.
- Briggs, J.A., Weinreb, C., Wagner, D.E., Megason, S., Peshkin, L., Kirschner, M.W., and Klein, A.M. (2018). The dynamics of gene expression in vertebrate embryogenesis at single-cell resolution. *Science* 360.
- Bronner, M.E., and Simoes-Costa, M. (2016). The Neural Crest Migrating into the Twenty-First Century. *Curr Top Dev Biol* 116, 115-134.
- Bronner-Fraser, M. (1994). Neural crest cell formation and migration in the developing embryo. *FASEB J* 8, 699-706.
- Buitrago-Delgado, E., Nordin, K., Rao, A., Geary, L., and LaBonne, C. (2015). NEURODEVELOPMENT. Shared regulatory programs suggest retention of blastula-stage potential in neural crest cells. *Science* 348, 1332-1335.
- Cheng, P.F., Shakhova, O., Widmer, D.S., Eichhoff, O.M., Zingg, D., Frommel, S.C., Belloni, B., Raaijmakers, M.I., Goldinger, S.M., Santoro, R., *et al.* (2015). Methylation-dependent SOX9 expression mediates invasion in human melanoma cells and is a negative prognostic factor in advanced melanoma. *Genome Biol* 16, 42.
- Diener, J., and Sommer, L. (2021). Reemergence of neural crest stem cell-like states in melanoma during disease progression and treatment. *Stem Cells Transl Med* 10, 522-533.
- Dupin, E., Calloni, G.W., Coelho-Aguiar, J.M., and Le Douarin, N.M. (2018a). The issue of the multipotency of the neural crest cells. *Dev Biol*.
- Dupin, E., Calloni, G.W., Coelho-Aguiar, J.M., and Le Douarin, N.M. (2018b). The issue of the multipotency of the neural crest cells. *Dev Biol* 444 Suppl 1, S47-S59.
- Fagnocchi, L., Mazzoleni, S., and Zippo, A. (2016). Integration of Signaling Pathways with the Epigenetic Machinery in the Maintenance of Stem Cells. *Stem Cells Int* 2016, 8652748.

Farrell, J.A., Wang, Y., Riesenfeld, S.J., Shekhar, K., Regev, A., and Schier, A.F. (2018). Single-cell reconstruction of developmental trajectories during zebrafish embryogenesis. *Science* 360.

Fazio, M., van Rooijen, E., Dang, M., van de Hoek, G., Ablain, J., Mito, J.K., Yang, S., Thomas, A., Michael, J., Fabo, T., *et al.* (2021). SATB2 induction of a neural crest mesenchyme-like program drives melanoma invasion and drug resistance. *Elife* 10.

Hall, B.K. (2000). The neural crest as a fourth germ layer and vertebrates as quadroblastic not triploblastic. *Evol Dev* 2, 3-5.

Hall, B.K. (2005). Consideration of the neural crest and its skeletal derivatives in the context of novelty/innovation. *J Exp Zool B Mol Dev Evol* 304, 548-557.

Hoek, K.S., Schlegel, N.C., Brafford, P., Sucker, A., Ugurel, S., Kumar, R., Weber, B.L., Nathanson, K.L., Phillips, D.J., Herlyn, M., *et al.* (2006). Metastatic potential of melanomas defined by specific gene expression profiles with no BRAF signature. *Pigment Cell Res* 19, 290-302.

Hovland, A.S., Rothstein, M., and Simoes-Costa, M. (2020). Network architecture and regulatory logic in neural crest development. *Wiley Interdiscip Rev Syst Biol Med* 12, e1468.

Ito, K., Morita, T., and Sieber-Blum, M. (1993). In vitro clonal analysis of mouse neural crest development. *Dev Biol* 157, 517-525.

Johnsen, J.I., Dyberg, C., and Wickstrom, M. (2019). Neuroblastoma-A Neural Crest Derived Embryonal Malignancy. *Front Mol Neurosci* 12, 9.

Kaufman, C.K., Mosimann, C., Fan, Z.P., Yang, S., Thomas, A.J., Ablain, J., Tan, J.L., Fogley, R.D., van Rooijen, E., Hagedorn, E.J., *et al.* (2016). A zebrafish melanoma model reveals emergence of neural crest identity during melanoma initiation. *Science* 351, aad2197.

Kulesa, P.M., Kasemeier-Kulesa, J.C., Teddy, J.M., Margaryan, N.V., Seftor, E.A., Seftor, R.E., and Hendrix, M.J. (2006). Reprogramming metastatic melanoma cells to assume a neural crest cell-like phenotype in an embryonic microenvironment. *Proc Natl Acad Sci U S A* 103, 3752-3757.

Kulesa, P.M., Morrison, J.A., and Bailey, C.M. (2013). The neural crest and cancer: a developmental spin on melanoma. *Cells Tissues Organs* 198, 12-21.

Kumar, R., Deivendran, S., Santhoshkumar, T.R., and Pillai, M.R. (2017). Signaling coupled epigenomic regulation of gene expression. *Oncogene* 36, 5917-5926.

LaBonne, C., and Bronner-Fraser, M. (1998). Neural crest induction in *Xenopus*: evidence for a two-signal model. *Development* 125, 2403-2414.

Leung, A.W., Murdoch, B., Salem, A.F., Prasad, M.S., Gomez, G.A., and Garcia-Castro, M.I. (2016). WNT/beta-catenin signaling mediates human neural crest induction via a pre-neural border intermediate. *Development* 143, 398-410.

Manoranjan, B., Venugopal, C., McFarlane, N., Doble, B.W., Dunn, S.E., Scheinmann, K., and Singh, S.K. (2012). Medulloblastoma stem cells: where development and cancer cross pathways. *Pediatr Res* 71, 516-522.

Martik, M.L., and Bronner, M.E. (2017). Regulatory Logic Underlying Diversification of the Neural Crest. *Trends Genet* 33, 715-727.

Micalizzi, D.S., and Ford, H.L. (2009). Epithelial-mesenchymal transition in development and cancer. *Future Oncol* 5, 1129-1143.

Milla, L.A., Gonzalez-Ramirez, C.N., and Palma, V. (2012). Sonic Hedgehog in cancer stem cells: a novel link with autophagy. *Biol Res* 45, 223-230.

Pomerantz, M.M., Qiu, X., Zhu, Y., Takeda, D.Y., Pan, W., Baca, S.C., Gusev, A., Korthauer, K.D., Severson, T.M., Ha, G., *et al.* (2020). Prostate cancer reactivates developmental epigenomic programs during metastatic progression. *Nat Genet* 52, 790-799.

Rambow, F., Rogiers, A., Marin-Bejar, O., Aibar, S., Femel, J., Dewaele, M., Karras, P., Brown, D., Chang, Y.H., Debiec-Rychter, M., *et al.* (2018). Toward Minimal Residual Disease-Directed Therapy in Melanoma. *Cell* 174, 843-855 e819.

Restivo, G., Diener, J., Cheng, P.F., Kiowski, G., Bonalli, M., Biedermann, T., Reichmann, E., Levesque, M.P., Dummer, R., and Sommer, L. (2017). low neurotrophin receptor CD271 regulates phenotype switching in melanoma. *Nat Commun* 8, 1988.

Sauka-Spengler, T., and Bronner-Fraser, M. (2008). A gene regulatory network orchestrates neural crest formation. *Nat Rev Mol Cell Biol* 9, 557-568.

Shakhova, O. (2014). Neural crest stem cells in melanoma development. *Curr Opin Oncol* 26, 215-221.

Simoes-Costa, M., and Bronner, M.E. (2015). Establishing neural crest identity: a gene regulatory recipe. *Development* 142, 242-257.

Theveneau, E., and Mayor, R. (2012). Neural crest delamination and migration: from epithelium-to-mesenchyme transition to collective cell migration. *Dev Biol* 366, 34-54.

Tirosh, I., Izar, B., Prakadan, S.M., Wadsworth, M.H., 2nd, Treacy, D., Trombetta, J.J., Rotem, A., Rodman, C., Lian, C., Murphy, G., *et al.* (2016). Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science* 352, 189-196.

Wakamatsu, Y., Endo, Y., Osumi, N., and Weston, J.A. (2004). Multiple roles of Sox2, an HMG-box transcription factor in avian neural crest development. *Dev Dyn* 229, 74-86.

Xue, A.G., Chan, M., and Gujral, T.S. (2020). Pan-cancer analysis of the developmental pathways reveals non-canonical wnt signaling as a driver of mesenchymal-type tumors. *Transl Res* 224, 1-15.

Young, R.A. (2011). Control of the embryonic stem cell state. *Cell* 144, 940-954.

Zhang, X., Yang, L., Szeto, P., Abali, G.K., Zhang, Y., Kulkarni, A., Amarasinghe, K., Li, J., Vergara, I.A., Molania, R., *et al.* (2020). The Hippo pathway oncoprotein YAP promotes melanoma cell invasion and spontaneous metastasis. *Oncogene* 39, 5267-5281.

Chapter II: Control of neural crest multipotency by Wnt signaling and the *Lin28/let-7* axis

Alternate Citation¹:

D Bhattacharya, M Rothstein, AP Azambuja, M Simoes-Costa (2018) Control of neural crest multipotency by Wnt signaling and the *Lin28/let-7* axis. *Elife* 7, e40556
<https://doi.org/10.7554/eLife.40556.001>

Abstract

A crucial step in cell differentiation is the silencing of developmental programs underlying multipotency. While much is known about how lineage-specific genes are activated to generate distinct cell types, the mechanisms driving suppression of stemness are far less understood. To address this, we examined the regulation of the transcriptional network that maintains progenitor identity in avian neural crest cells. Our results show that a regulatory circuit formed by Wnt, *Lin28a*, and *let-7* miRNAs controls the deployment and the subsequent silencing of the multipotency program in a position-dependent manner. Transition from multipotency to differentiation is determined by the topological relationship between the migratory cells and the dorsal neural tube, which acts as a Wnt-producing stem cell niche. Our findings highlight a mechanism that rapidly silences complex regulatory programs and elucidate how transcriptional networks respond to positional information during cell differentiation.

This work is a published manuscript with the above citation. Debadrita Bhattacharya performed most of the experiments described in this paper. Ana P. Azambuja provided technical support for clonal analysis experiments described in Fig 2, and also performed the ChIP-qPCR and enhancer pull down experiments shown in Fig 5G and I respectively. Megan Rothstein performed the analysis shown in Fig 5j-l.

Introduction

The process of cell differentiation is characterized by major shifts in the molecular programs that control cell identity. This requires the coordination of activating and repressive regulatory mechanisms, which act together to drive overarching changes in gene expression (Davidson, 2009). In the past decades, substantial progress has been made in understanding how the activation of lineage-specific genes generates distinct cell types. In contrast, we know far less about how progenitor cell identity is silenced during cell fate commitment. Inhibitory mechanisms are crucial for terminal differentiation since silencing of multipotency networks precedes activation of lineage-specific factors and chromatin remodeling (Kalkan and Smith, 2014; Moris et al., 2016). Repression also plays an essential role in adult tissue homeostasis. Anomalous reactivation of embryonic regulatory programs in somatic tissue has been shown to underlie tumorigenesis and metastasis (Kaufman et al., 2016; Nieto, 2013). Despite their importance in development and disease, the mechanisms that suppress stemness and multipotency remain elusive.

Here, we used the cranial neural crest as a model to examine how stem cell identity is regulated during differentiation. This multipotent cell population contributes to numerous tissues and organs in vertebrate embryos, including the craniofacial skeleton, peripheral nervous system, and pigmentation of the skin (Le Douarin, 1982). Neural crest cells delaminate from the neural tube to engage in extensive migration throughout the embryo (Fig. 1a-b). As these cells move away from the neural tube, they undergo drastic changes in their transcriptional identity and transition from multipotent progenitors to committed cell types. The formation and differentiation of the neural crest are controlled by a complex gene regulatory network composed of multiple signaling pathways, transcription factors, and epigenetic modifiers (Betancur et al., 2010; Meulemans and Bronner-Fraser, 2004; Simoes-Costa and Bronner, 2015). Expression of the early components of this network defines the neural crest stem cell population that resides within the dorsal neural tube (Lignell et al., 2017). This set of genes, which includes neural

crest markers and pluripotency factors, endows the neural crest stem cells with their unique features, such as multipotency and self-renewal.

In this study, we have uncovered a regulatory circuit formed by Wnt signaling and the *Lin28a/let-7* axis that modulates the stem cell identity of avian cranial neural crest cells. This circuit functions in a position-dependent manner, promoting multipotency in the early neural crest cells and controlling the transition to differentiation in the late-migrating cells. In the premigratory neural crest, canonical Wnt signaling directly activates *Lin28a* transcription, resulting in the inhibition of *let-7* miRNA activity. As neural crest cells migrate away from the Wnt source, *let-7* levels increase, and these miRNAs target and silence multiple components of the neural crest gene regulatory network. As a result, the early network collapses, and stem cell identity is suppressed. Thus, we propose that the changes in the topological relationship between the neural crest and the dorsal neural tube drive the transition from multipotent stem cell to differentiated cell type. Our model integrates signaling, transcription, and post-transcriptional regulation to clarify how positional information ultimately affects gene network topology and stem cell identity.

Results

Dynamic expression of *Lin28a* during neural crest development

The developmental program controlling the formation of neural crest cells is encoded in a modular transcriptional gene regulatory network (Betancur et al., 2010; Meulemans and Bronner-Fraser, 2004; Simoes-Costa and Bronner, 2015) (Fig. 1c), in which the early modules mediate neural crest induction and specification, while the later modules regulate their ability to migrate and differentiate into diverse cell types (Meulemans and Bronner-Fraser, 2004; Simoes-Costa and Bronner, 2015). A large part of the early gene regulatory network (Fig. 1c) consists of factors that maintain neural crest cells in a multipotent state, such as *Pax7*, *FoxD3*, *Ets1*, *Myc*, and *Sox5* (Buitrago-Delgado et al., 2015; Le Douarin and Dupin, 2016). These factors define the neural crest stem cell pool and are co-expressed in pre-migratory and

early migrating cells (Lignell et al., 2017). To examine how stem cell identity changes during cell commitment, we quantified the expression levels of these factors during different stages of embryonic development. We labeled cranial neural crest cells with an enhancer of the *Tfap2a* gene (here referred to as *Tfap2aE1*) (Attanasio et al., 2013), which drives specific expression of reporter genes in the avian neural crest lineage, starting at HH6 and persisting until late migratory stages (HH16). We generated a *Tfap2aE1*-eGFP construct to FACS-sort pure populations of neural crest cells from different embryonic stages for RT-PCR analysis. At the stages corresponding to progenitor specification, we observed an increase in the mRNA levels of genes that are part of the neural crest stem cell signature, like *Pax7*, *FoxD3*, *Ets1*, *Myc*, and *Sox5* (Fig. 1d). While the expression of these factors peaked at different time points, we observed that they were simultaneously downregulated during the late stages of migration. Such striking changes in the neural crest transcriptional network raised the intriguing possibility that a regulatory mechanism exists to silence stem cell identity at the onset of differentiation.

Our transcriptomic analyses of cranial neural crest cells (Simoes-Costa and Bronner, 2016; Simoes-Costa et al., 2014) had previously shown that pluripotency factor *Lin28a* is strongly enriched in this cell population. To test if *Lin28a* is involved in the transition from multipotency to differentiation, we first examined its expression levels during neural crest development. *In situ* hybridization and immunohistochemistry analysis revealed that *Lin28a* was robustly expressed in the neural folds, premigratory, and early migrating neural crest of the chick embryo (Fig S1a-p). Isolation of neural crest with the *Tfap2aE1* and RT-PCR quantitation showed strong enrichment of *Lin28a* mRNA during specification (Fig S1q). This was not observed for the other *Lin28* gene present in the chick genome, *Lin28b* (Tzialikas and Romer-Seibert, 2015). The paralog was expressed at much lower levels and was not enriched in neural crest cells (Fig S1q-r). Interestingly, there was a marked reduction in *Lin28a* mRNA levels at later stages, concomitant with silencing of the neural crest stem cell genes (Fig. 1d). These results show that *Lin28a* is dynamically expressed during neural crest development and that it is most abundant in premigratory cells that reside within the neural tube.

Figure 1

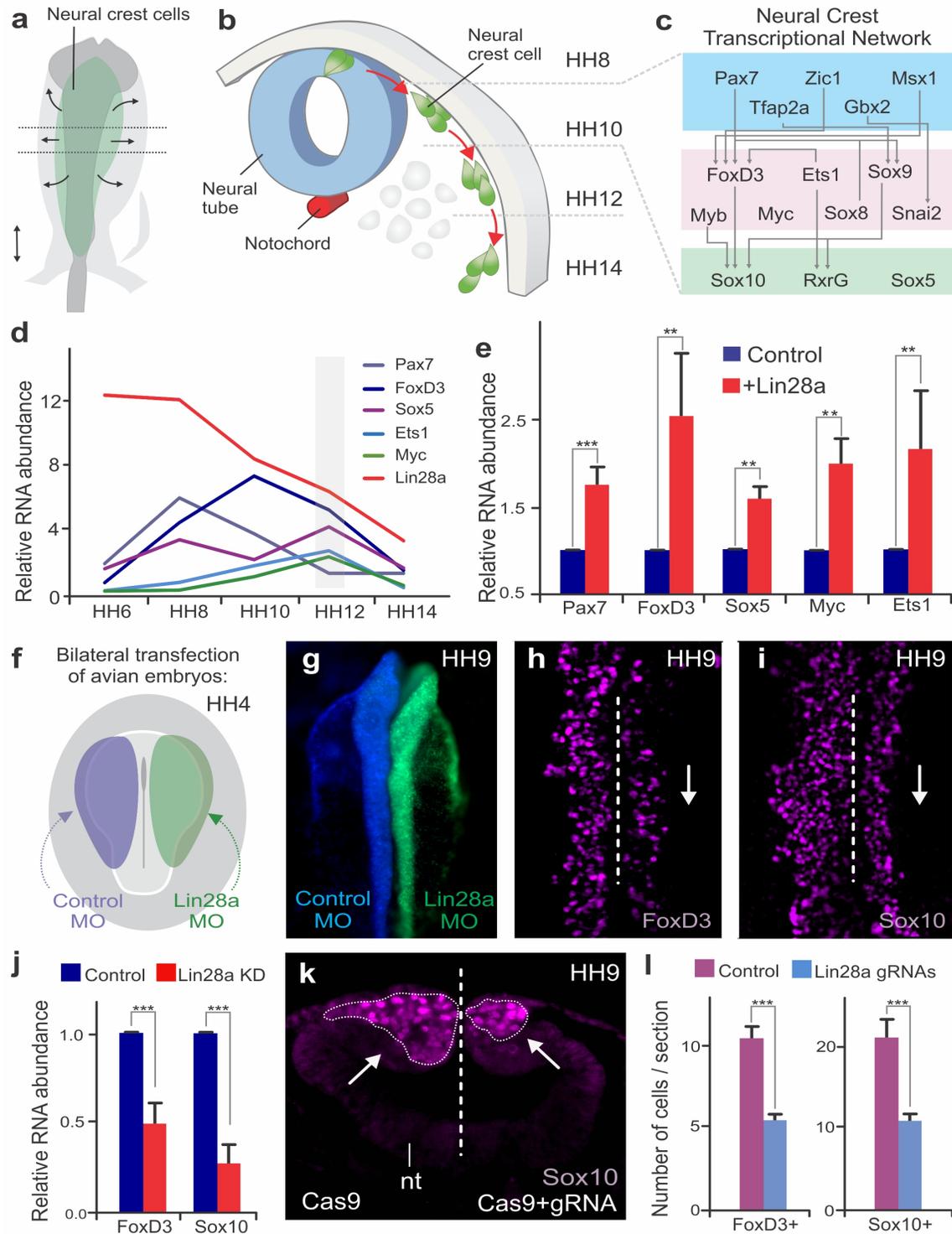


Figure 1. Changes in *Lin28a* levels impact neural crest development *in vivo*.

(a-b) Neural crest migration during avian development. (a) Neural crest progenitor cells (green) are specified on dorsal folds of the neural tube (grey) during early development. (b) Transverse section of the neural tube showing the position of neural crest cells through development as they progressively move away from the neural tube to differentiate. HH8 and HH14 are the earliest and latest developmental stages shown in the diagram, respectively.

(c) A schematic of the early gene regulatory network composed of transcription factors involved in neural crest cells formation. (d) Expression levels of *Lin28a* and transcription factors of the early gene regulatory circuit in sorted neural crest cells obtained from different stages. (e) Constitutive expression of *Lin28a* results in the maintenance of multipotency genes in late neural crest cells. RT-RT-PCR for *Pax7*, *FoxD3*, *Sox5*, *Myc*, and *Ets1* comparing the expression of these genes in control vs. *Lin28a* overexpressing migratory neural crest cells. (f) Electroporation scheme for loss-of-function assays in which control reagent (blue) and targeted reagent (green) were injected in different sides of a HH4 chick embryo. (g) Dorsal whole-mount view of HH9 embryo with Control MO on the left and *Lin28a* MO on the right. Immunohistochemistry for neural crest markers *FoxD3* (h) and *Sox10* (i) on *Lin28a* knockdown. The dotted line represents embryo midline (j) RT-RT-PCR for *FoxD3*, and *Sox10* transcripts in control vs. *Lin28a* MO treated neural folds. (k-l) CRISPR-Cas9 mediated knockdown of *Lin28a* recapitulates the MO phenotype. (k) Transverse section showing *Sox10* positive cells in control and knockdown sides of the embryo head, showing a reduction in the number of neural crest cells (arrow). (l) Quantification of *FoxD3*⁺ and *Sox10*⁺ cells following CRISPR-Cas9 mediated knockdown of *Lin28a*. Error bars in (e), (j), and (l) represent standard error. H.H.: Hamburger and Hamilton developmental stages, MO: Morpholino.

***Lin28a* promotes maintenance of neural crest stem cell identity**

To verify whether *Lin28a* is involved in the regulation of neural crest stem cell identity, we used transient expression vectors to manipulate its levels in developing chick embryos. First, we investigated how sustained expression of *Lin28a* affects the silencing of the early neural crest genes at later stages of migration. Neural crest cells transfected with a *Lin28a* expression vector were sorted through FACS, and their expression profile was compared to wild-type neural crest through RT-PCR. This analysis revealed that migratory cells constitutively expressing *Lin28a* maintained expression of neural crest stem cell factors even at stages when these genes would normally be downregulated (Fig. 1e). We confirmed these results by performing immunohistochemistry for *Pax7* and *FoxD3* following *Lin28a* gain-of-function (Fig S2a-h). This maintenance of early neural crest factors resulted in a delay in differentiation, as evidenced by a decrease in the expression of ectomesenchymal, neuronal and glial markers in late migrating cells (Fig S2i). This is consistent with the possibility that *Lin28a* acts to prevent the silencing of the multipotency network that precedes cell differentiation.

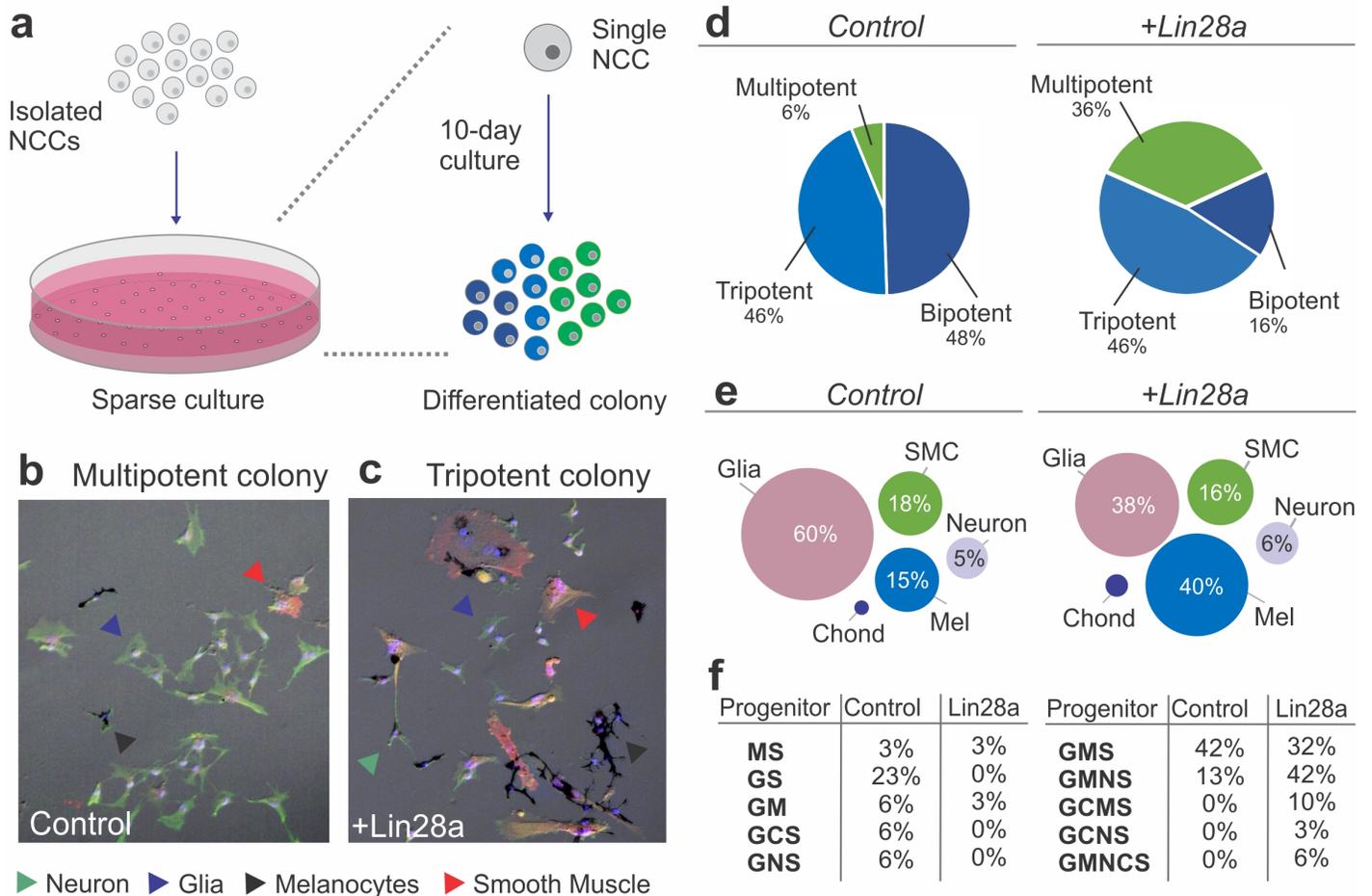
In the reciprocal experiment, we tested the effects of premature downregulation of *Lin28a* in neural crest stem cells. To this end, we performed bilateral electroporations to transfect embryos with a *Lin28a* translation-blocking morpholino on the right side (green) and a control morpholino (blue) on the left side (Fig. 1f). This treatment resulted in a strong knockdown of *Lin28a* protein in the morpholino-

transfected side of the embryo (FigS3 3a-d). Loss of *Lin28a* resulted in decreased expression of neural crest markers *FoxD3* and *Sox10*, as shown by immunohistochemistry (Fig. 1g-i) and *in situ* hybridization (Fig. S3e-h). Immunohistochemistry for phospho-histone H3 and Caspase-3 confirmed that this phenotype was not due to changes in cell cycle exit or cell death (Fig S4a-c,f-g). Microdissection of dorsal neural folds of targeted embryos allowed for the quantification of the phenotype with RT-PCR, which revealed a significant decrease in the expression of the neural crest markers (Fig. 1j). To confirm this phenotype, we employed CRISPR/Cas9 genome editing and RNA interference (using Dicer-substrate siRNAs, DisRNAs) (Kim et al., 2005) to disrupt *Lin28a* expression (Fig. 1k-l, Fig S3 i-m). Consistent with the morpholino experiments, targeting the first exon of *Lin28a* with a pair of gRNAs resulted in a significant reduction in the number of *FoxD3* and *Sox10* positive cells (Fig1k-l). Finally, DsiRNA-mediated knockdown of *Lin28a*, but not of *Lin28b*, resulted in a significant decrease in the expression of *FoxD3* and *Sox10* (Fig S3n-o).

To evaluate possible long-term consequences of *Lin28a* knockdown in neural crest derivatives, we employed electroporation of morpholinos combined with a cornish pasty culture system (Nagai et al., 2011), which allows for long-term incubation of electroporated embryos *ex ovo*. Bilaterally-injected morphant embryos were cultured until stage HH15 so that we could observe the formation of cranial ganglia. Immunohistochemistry with neuronal marker Tuj1 revealed that knockdown of *Lin28a* resulted in dispersed ganglia with abnormal condensation. In particular, we observed a marked reduction in the maxillomandibular lobe of the trigeminal ganglion (Fig S4j-o). This phenotype is consistent with previous studies that highlight the requirement of neural crest cells for timely ganglion condensation and accurate establishment of neuronal connections (Hamburger, 1961; Shiau et al., 2008; Stark et al., 1997). Taken together, these functional experiments indicate that high levels of *Lin28a* during early development are necessary for neural crest specification, while its subsequent downregulation is required for silencing of progenitor identity.

Lin28a regulates neural crest multipotency

Lin28a is a *bona fide* pluripotency factor that has been shown to underlie stemness and to drive reprogramming of somatic cells (Yu et al., 2007; Zhang et al., 2016). Our observation that *Lin28a* manipulation impacts the neural crest transcriptional network suggests that this factor may regulate some of the stem cell properties that characterize this cell population. To directly test if *Lin28a* promotes neural crest multipotency, we examined how manipulating its expression affects the developmental potential of individual progenitors, using single-cell lineage analysis (Sieber-Blum and Cohen, 1980; Trentin et al., 2004). Neural crest cells from quail embryos transfected with *Lin28a* (or the empty PCI-H2b: RFP vector) were cultured in sparse conditions, such that a single progenitor could form clonal colonies of differentiated cells (Fig. 2a). After a ten-day incubation period in a standard culture medium, individual cells gave rise to colonies that were composed of multiple differentiated derivatives, which were identified based on cell morphology and molecular markers (Fig. 2b-c). The somatic cell types observed included the typical neural crest derivatives, such as neurons (N), glia (G), melanocytes (M), chondroblasts (C), and smooth muscle cells (S). Quantification of cell types in colonies from the control and experimental groups revealed a six-fold increase in multipotent (that generated four or more cell types) neural crest cells after transfection with a *Lin28a* expression vector (Fig. 2d). Furthermore, while mock-transfected neural crest cells were biased towards a glial fate (60% of the total progeny), the experimental group had a more uniform distribution of derivatives (Fig. 2e). Consistent with this, classification of neural crest progenitors according to their developmental potentials revealed a larger percentage of multipotent progenitors (like GCMS progenitors, which give rise to glia, chondroblasts, melanocytes, and smooth muscles) when compared to the control group (Fig. 2f). The results from this single-cell clonal analysis indicate that the changes in the transcriptional network observed following manipulation in the *Lin28a* levels affect specific properties of neural crest cells, such as their ability to give rise to multiple cell types.

Figure 2**Figure 2. *Lin28a* modulates the developmental potential of neural crest cells.**

(a) Diagram showing the experimental design for neural crest single-cell clonal analysis. Quail neural crest cells transfected with control or *Lin28a* expression construct were isolated and plated sparsely to ensure the formation of clonal colonies. Cells were cultured for up to 10 days, after which the different cell types in each colony were identified by immunofluorescence. (b-d) Neural crest cells expressing higher levels of *Lin28a* formed increased number of multipotent colonies. Representative images showing a tripotent colony derived from a neural crest cell transfected with control construct (b) and a multipotent colony derived from a single neural crest cell transfected with a *Lin28a* expression construct (+*Lin28a*) (c). (d) Pie-charts showing the percentage of bi-, tri-, and multipotent colonies formed by control and +*Lin28a* neural crest cells. (e) Representation of the frequency of different cell types observed in all colonies formed by control and +*Lin28a* neural crest cells. (f) Table listing the percentages of the different neural crest progenitor cells observed in control and +*Lin28a* conditions. NCC: neural crest cell.

The Lin28a/let-7 axis modulates neural crest differentiation

Lin28a has been shown to regulate genes post-transcriptionally both directly by binding to mRNAs or indirectly by inhibiting maturation of the *let-7* family of microRNAs, which are potent post-transcriptional

repressors (Newman et al., 2008). To investigate if Lin28a function in neural crest is *let-7* dependent, we employed a *let-7* sensor to survey microRNA activity during different stages of development (Fig. 3a). In this construct, a reporter gene with a destabilization domain (mCherry-PEST) was placed upstream of multiple *let-7* target sites, such that increased *let-7* activity results in decreased mCherry fluorescence. By transfecting embryos with the sensor, we observed that although neural crest stem cells have low *let-7* activity, it continuously increases as the cells migrate away from the neural tube (Fig. 3b). Consistent with this, RT-PCR analysis revealed a significant increase in the expression levels of *let-7* miRNAs from the premigratory (HH8) to the late-migrating (HH12) neural crest (Fig. 3c). Since *let-7* activity and expression were inversely correlated with *Lin28a* expression levels (Fig. 1d), we tested the effect of Lin28a knockdown on *let-7* activity in the neural crest. The results show that in the absence of Lin28a, *let-7* activity increased, as evidenced by a reduction in *let-7* sensor expression (Fig. 3d-e) and direct measurement of multiple mature *let-7*s through RT-PCR (Fig 3f). Furthermore, transfection with a *let-7a* mimic molecule resulted in the loss of FoxD3+ cells (Fig 3g-h) and reduced the expression of neural crest markers (Fig. 3i), recapitulating the phenotype of *Lin28a* knockdown. This treatment did not alter cell death or proliferation, as evidenced by immunostaining with pH3 and Caspase-3 antibodies (Fig S4d-e,h-i).

Figure 3

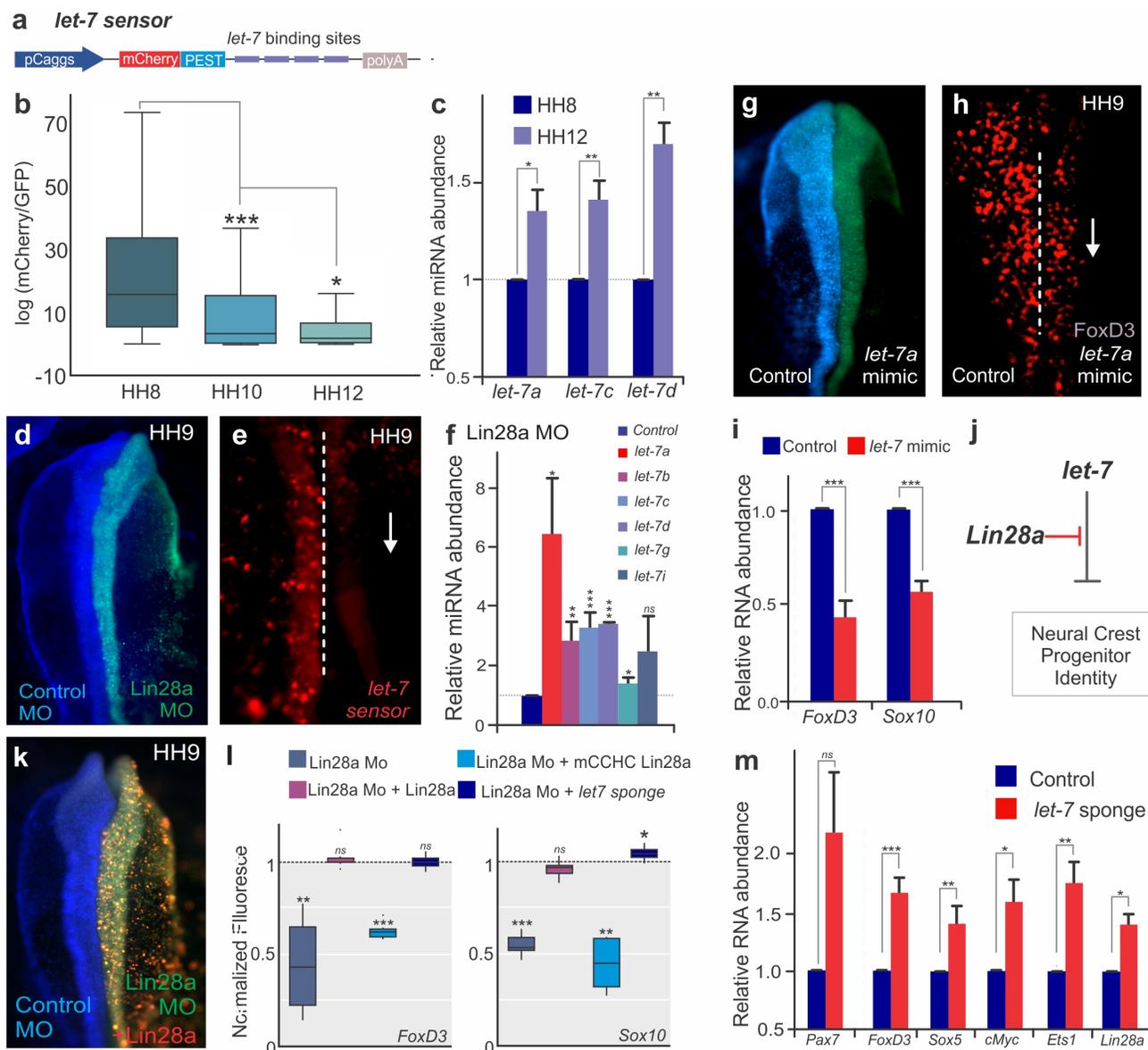


Figure 3. The *Lin28a/let-7* axis modulates neural crest progenitor identity *in vivo*.

(a) A schematic representation of the *let-7* sensor, which consists of several *let-7* binding sites downstream of destabilized mCherry fluorescent protein. (b-c) The activity of mature *let-7* miRNAs increases through neural crest development. (b) Boxplots showing mCherry/GFP fluorescence ratio, a readout of *let-7* sensor activity, in neural crest cells at different developmental stages. (c) RT-RT-PCR for mature *let-7* family miRNAs comparing their levels in neural crest cells sorted from HH8 and HH12 embryos. (d-f) Loss of *Lin28a* results in increased activity of mature *let-7* miRNAs. (d) Whole-mount view of an embryo bilaterally injected with control and *Lin28a* Mo. (e) Representative image showing *let-7* sensor fluorescence in control vs. *Lin28a* Mo side of an embryo. The dotted line represents the embryo midline. (f) RT-PCR for mature *let-7* family miRNAs, in the background of *Lin28a* knockdown. (g) Whole-mount view of an embryo electroporated with control and *let-7a* mimic. (h) Immunohistochemistry for FoxD3 positive neural crest cells in the presence of *let-7a* mimic. Dotted line represents the embryo midline. (i) Quantification of transcript levels of *FoxD3* and *Sox10* in the presence of increased *let-7a*.

(j) Model for modulation of neural crest identity by the *Lin28a/let-7* axis. (k) Representative dorsal view of an embryo electroporated with control MO (blue) on the left and *Lin28a* MO (green) co-injected with a *Lin28a* expression vector (red) on the right. (l) Boxplots showing the quantification of FoxD3 and Sox10 fluorescence in epistatic experiments, in which *Lin28a* Mo was co-electroporated with *Lin28a* expression vector, mCCHC *Lin28a*, and a *let-7* sponge construct. (m) Loss of *let-7* activity results in maintenance of multipotency genes in late neural crest cells. RT-RT-PCR for *Pax7*, *FoxD3*, *Sox5*, *Myc*, *Ets1*, and *Lin28a*. comparing the expression of these genes in control vs. late migratory neural crest cells expressing *let-7* sponge construct. Error bars in (c), (f), (i) and (m) represent standard error. HH: Hamburger and Hamilton developmental stages, MO: Morpholino.

To confirm that *Lin28a* regulates neural crest development predominantly via *let-7* inhibition (Fig. 3j), we performed *Lin28a* knockdown and attempted to rescue the phenotype with a wild-type *Lin28a* construct, a *Lin28a* mutant unable to bind to *let-7* miRNAs (mCCHC *Lin28a*) (Heo et al., 2008) or a *let-7* sponge construct, which sequesters *let-7* molecules to decrease its activity (Kumar et al., 2008) (Fig. 3k-l). The wild-type *Lin28a* protein was able to rescue the loss of neural crest; however, the mCCHC *Lin28a* mutant protein could not restore *FoxD3* or *Sox10* expression. Furthermore, transfection with a *let-7* sponge construct, which reduces the levels of free mature *let-7*, recapitulated the wild-type *Lin28a* rescue (Fig. 3k-l). These results indicate that *Lin28a* regulates neural crest genes via the *let-7* dependent pathway and suggest that *let-7* miRNAs mediate the silencing of neural crest genes observed during late migration (Fig. 1d). To test if the increase in *let-7* levels (Fig. 3b-c) results in silencing of early neural crest genes, we reduced the levels of mature *let-7s* using the sponge construct and quantified the expression of multipotency genes in late neural crest cells. RT-PCR analysis revealed that this inhibition of *let-7* activity results in maintenance of *Pax7*, *FoxD3*, *Sox5*, *cMyc*, and *Ets1*, recapitulating the effect of *Lin28a* overexpression (Fig. 1e, Fig. 3m). Consistent with published data showing that *Lin28a* is itself a *let-7* target (Rybak et al., 2008), and that these factors form a double-negative feedback loop, we also detected higher levels of the pluripotency factor in these cells (Fig. 3m). Taken together, these experiments indicate that neural crest stem cell identity is regulated by the *Lin28a/let-7* axis, consistent with the possibility that the balance between these two factors underlies the collapse of the multipotency network observed during differentiation.

Post-transcriptional silencing of the neural crest gene regulatory network

To test the global transcriptional effects of the disruption of *Lin28a/let-7* balance, we employed single embryo Nanostring analysis to assay the effects of *Lin28a* knockdown and *let-7* gain-of-function on the expression of ~100 genes involved in neural, placodal, and neural crest development (Simoes-Costa et al., 2015) (Fig. 4a-d). Both *Lin28a*-loss- and *let-7* gain-of-function recapitulated the changes that occur in late migratory neural crest cells (Fig 1d). By comparing control and targeted cells from individual embryos, we found that neural crest genes were strongly down-regulated in both treatments, suggesting that premature *let-7* activity has a systemic effect on the neural crest gene regulatory network (Fig 4d). Furthermore, genes that have been reported to modulate neural crest multipotency, such as *Sox10*, *FoxD3*, *Sox5*, *Ets1*, and *cMyc*, were particularly susceptible to changes in the *Lin28a/let-7* balance (Fig 4d). We validated the Nanostring results by single-embryo RT-PCR (Fig 4e), which additionally showed a strong downregulation of *Pax7* and *Tfap2b* genes (for which we lacked functional Nanostring probes) following manipulation of *let-7* levels. These results indicate the *Lin28a/let-7* axis regulates neural crest development by modulating the entire transcriptional network. Furthermore, the striking loss of stem cell genes (Lignell et al., 2017) observed in these experiments indicates that increased *let-7* activity suppresses multipotency and stemness in neural crest progenitors.

Next, we employed UTR-reporter assays to identify direct targets of *let-7* microRNAs in the early neural crest transcriptional network. 3' UTRs of six neural crest genes (Fig S5a) that were robustly affected in our functional assays (Fig. 4d-e) were cloned downstream of a destabilized reporter gene (mCherry-PEST) and bilaterally transfected in chick embryos with or without a *let-7* mimic (Fig. 4f-h). A similar construct driving GFP expression but lacking the 3' UTR was used as a transfection control. Flow cytometry analysis was used to compare reporter activity on each side of the same embryo (Fig. 4i). The 3' UTR reporters for *FoxD3*, *Pax7*, and *Myc* (Fig 4i-j; Fig S5c,f), which have *let-7* target sites (Fig S5a), showed decreased activity when they were co-transfected with the *let-7* mimic. In contrast, the reporters for *Sox10*, *Zic1*, and *Sox8* UTRs, which lack target sites for the microRNAs, were unaffected by the gain-

of-function assay (Fig 4j, Fig.S5b,d-e). To further test the importance of post-transcriptional regulation in the silencing of early neural crest genes, we compared endogenous FoxD3 protein expression with the activity of the enhancer that controls its expression in migratory cells (*FoxD3NC2*) (Simoes-Costa et al., 2012). We found that FoxD3 protein expression rapidly decreased during migration, while enhancer activity did not significantly change between neural crest cells close to the neural tube and those that had further migrated out (Fig S5g-j). This indicates that post-transcriptional regulation, which is absent from the enhancer construct, is necessary for timely silencing of components of the gene regulatory network.

To demonstrate the importance of *let-7* target sites for the endogenous regulation of neural crest genes, we employed CRISPR/Cas9 genome editing. A Cas9/eGFP expression vector containing gRNAs targeted to *let-7* sites within the 3'UTR of *Pax7* and *FoxD3* were transfected in gastrula stage chick embryos with bilateral electroporation (the control side was transfected with Cas9/eGFP only). Transfected cells from the control and experimental sides of single embryos were isolated with FACS at HH12 and analyzed with RT-PCR (Fig S5k). Targeting of individual *let-7* sites in both the *Pax7* and *FoxD3* loci resulted in a mild but consistent increase in the expression of these genes (Fig. 4k), at the stages that they are normally downregulated; a control gRNA (gRNA4) targeting a region of the *FoxD3* UTR devoid of *let-7* miRNAs sites had no effect on gene expression. These results indicate that *let-7* miRNAs orchestrate the silencing of progenitor cell identity by directly repressing critical network nodes (Fig. 4i). We speculate that these inhibitory interactions propagate in a domino-like fashion throughout the network, resulting in its collapse (Fig. 4l).

Figure 4

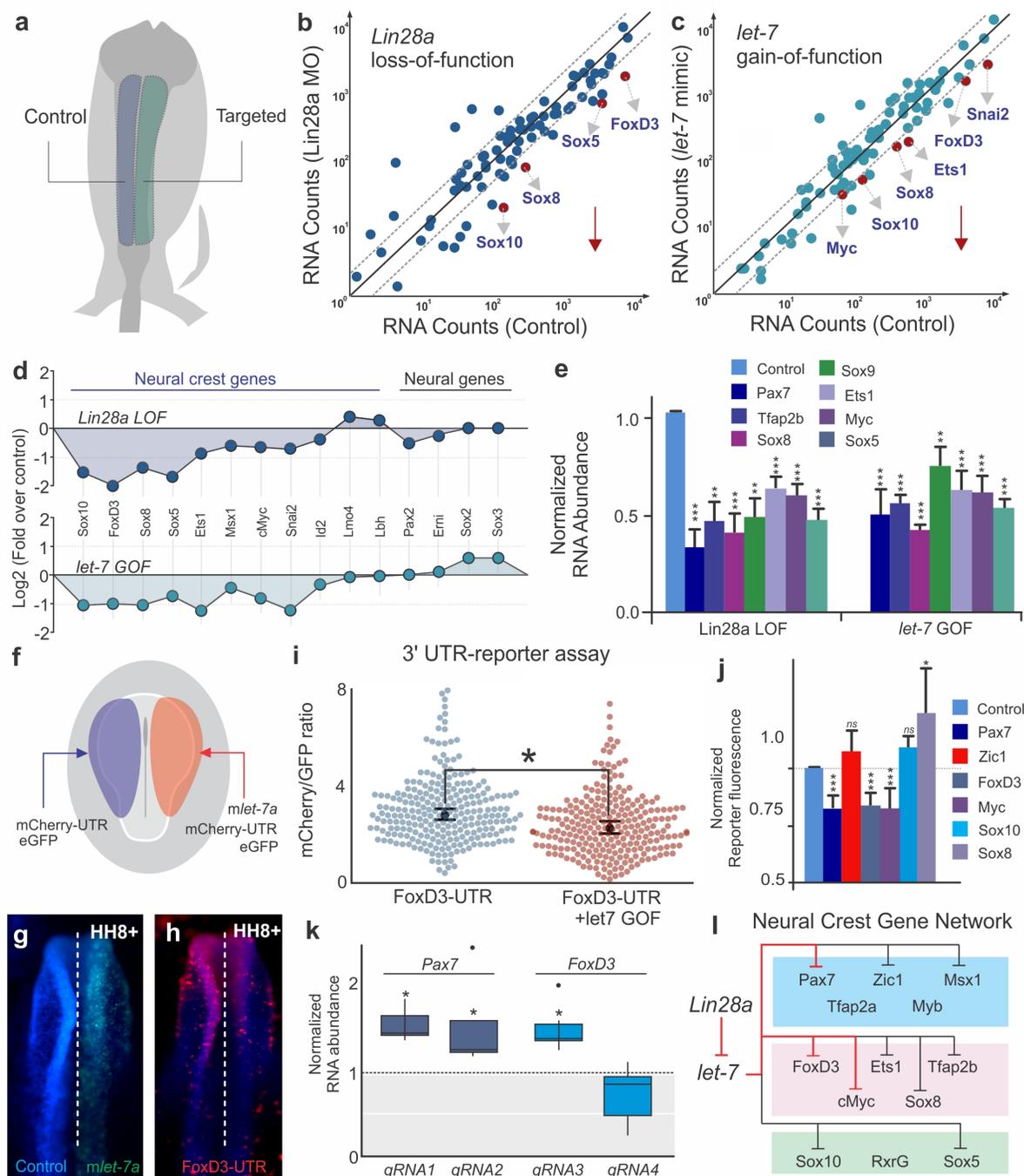


Figure 4. *let-7* directly targets multipotency circuits in the neural crest transcriptional network.

(a) Control and targeted neural folds were dissected from the same embryo for Nanostring analysis. Comparison of transcript levels between (b) Control MO and *Lin28a* MO electroporated cells and (c) between control and *let-7* mimic injected cells. Genes below the diagonal dotted line were significantly downregulated (downward red arrow) in each condition. (d) Neural crest genes were similarly affected in *Lin28a* loss of function (LOF) and *let-7* gain-of-function (GOF) assays. (e) RT-PCR was performed with control and targeted neural folds to validate Nanostring results. (f) Electroporation scheme for *in vivo* 3'UTR reporter assay. Individual 3'UTR reporter constructs were co-injected with a control oligo (left) and a *let-7* mimic (right). Flow cytometry analysis was performed to measure

mCherry and eGFP fluorescence of individual cells. (g-h) whole-mount view of an embryo showing FoxD3-UTR reporter fluorescence in control vs. *let-7* mimic transfected side of the embryo. (i) Representative scatter plots of FoxD3 UTR-reporter assay, showing the mCherry/GFP intensity ratio in cells analyzed from the control (gene-UTR) and *let-7a* mimic transfected (gene-UTR+*let7* GOF) sides of the same embryo. Each dot in the plot represents a single cell. (j) Average fold change in the ratio of mCherry/GFP intensity for each 3'UTR analyzed. (k) Quantification of fold change in *Pax7* and *Foxd3* transcript levels in late migratory neural crest cells when the *let-7* binding site on the 3'-UTR of these genes are targeted with specific gRNAs. *gRNA1* and *gRNA2* against *Pax7* 3'-UTR specifically target the two *let-7* binding sites, while *gRNA3* and *gRNA4* for *FoxD3* 3'-UTR target a *let-7* binding site and another control region on the UTR, respectively. (l) *Lin28a/let-7* targets in the early neural crest transcriptional network, showing genes that are directly (red inhibitory lines) or indirectly (black inhibitory lines) affected by *let-7*. Error bars in (e) and (j) represent standard error and standard deviation, respectively. MO: Morpholino, LOF: loss-of-function, GOF: gain-of-function.

Wnt signaling regulates multipotency in a position-dependent manner

The above results show that the dynamics of *Lin28a/let-7* activity regulate multiple targets of the early neural crest transcriptional network. To explore the upstream regulators of the *Lin28a/let-7* axis, we investigated the transcriptional regulation of *Lin28a* in neural crest cells. Assay for Transposase-Accessible Chromatin (ATAC-seq) performed in neural crest cells sorted from HH8 embryos revealed eleven non-coding regions of open chromatin in the *Lin28a* locus (Fig. S6a). Transient transgenesis experiments in chick embryos showed that only one of these regions, located in the second intron of *Lin28a* (*Lin28E1*) (Fig. 5a, Fig.5-fig supplement 1b), was able to drive reporter activity in the neural crest and ectoderm (Fig. 5b). The sequence of this enhancer, which is conserved in amniotes, contains four TCF/LEF binding sites (Fig S6c), suggesting regulation by canonical Wnt signaling. Consistent with this, *Lin28a* protein expression was strongest in the dorsal neural tube, a known source of Wnt ligands (Hollyday et al., 1995; Simoes-Costa et al., 2015) (Fig. 5c-d). To test whether Wnt-signaling directly regulates *Lin28a* in neural crest cells, we first mutated the four TCF/LEF binding sites in *Lin28E1*. This resulted in complete loss of enhancer activity, specifically in neural crest cells (Fig. 5e-f), and also prevented enhancer association with Wnt-effector Lef1 in enhancer pull-down experiments (Fig. 5g). Second, we conducted loss-of-function experiments by disrupting Wnt signaling using morpholinos targeting two Wnt ligands expressed in the dorsal neural folds, *Wnt1* and *Wnt4* (Simoes-Costa et al., 2015). This knockdown reduced the expression of both endogenous *Lin28a* and *Lin28E1* activity (Fig.

5h). Moreover, loss of Wnt signaling resulted in increased levels of mature *let-7s* (Fig 5h). Finally, chromatin immunoprecipitation (ChIP) revealed that Lef1 and nuclear β -catenin (*Ctnnb1*) are associated with *Lin28E1* in neural crest cells, indicating that canonical Wnts are directly regulating *Lin28a*. This interaction is dependent on Wnt activity, as binding to the enhancer was lost in embryos transfected with a *Wnt1* dominant-negative construct (Fig 5i).

Taken together, these results indicate that a Wnt-*Lin28a/let-7* regulatory circuit controls neural crest stem cell identity during differentiation. We hypothesize that a Wnt-mediated stem cell niche in the dorsal neural tube activates *Lin28a* expression in neural crest cells, thereby protecting the stem cell regulatory network from *let-7* mediated repression. If this assumption is correct, neural crest cells should exhibit a reduction in Wnt-activation and an increase in *let-7* activity as they migrate away from the neural tube. To test this, we measured nuclear β -catenin (*Ctnnb1*) and *let-7* sensor fluorescence as a function of distance from the neural tube in single migratory neural crest cells (Fig 5j). Consistent with our prediction, the analysis revealed a decrease in Wnt activity and increased *let-7* mediated repression during migration (Fig 5k-l). Based on these results, we propose that the topological relationship between a neural crest cell and the Wnt niche determines the balance of *Lin28a/let-7* activity, which in turn modulates the early transcriptional network. Thus, we expanded the Wnt niche beyond the dorsal neural tube by ectopically expressing Wnt1 in migratory neural crest cells (Fig. 5m, increased *Axin2* transcript levels confirmed over-activation of the pathway) (Jho et al., 2002). As predicted, we found that migrating neural crest cells constitutively expressing Wnt1 have higher levels of *Lin28a* and lower *let-7* activity. The expansion of the Wnt niche also prevents silencing of early neural crest factors *Pax7* and *FoxD3* (Fig 5m). Furthermore, this maintenance of stem identity resulted in suppression of differentiation, as Wnt gain-of-function resulted in lower expression levels of drivers of ectomenchymal differentiation *Runx2*, *Alx1*, and *Barx2* (Fig S7a-b). These findings show that manipulation of the Wnt-*Lin28a/let-7* regulatory circuit impacts both progenitor identity and the onset of differentiation, indicating that this mechanism controls the transition between these two states.

Figure 5

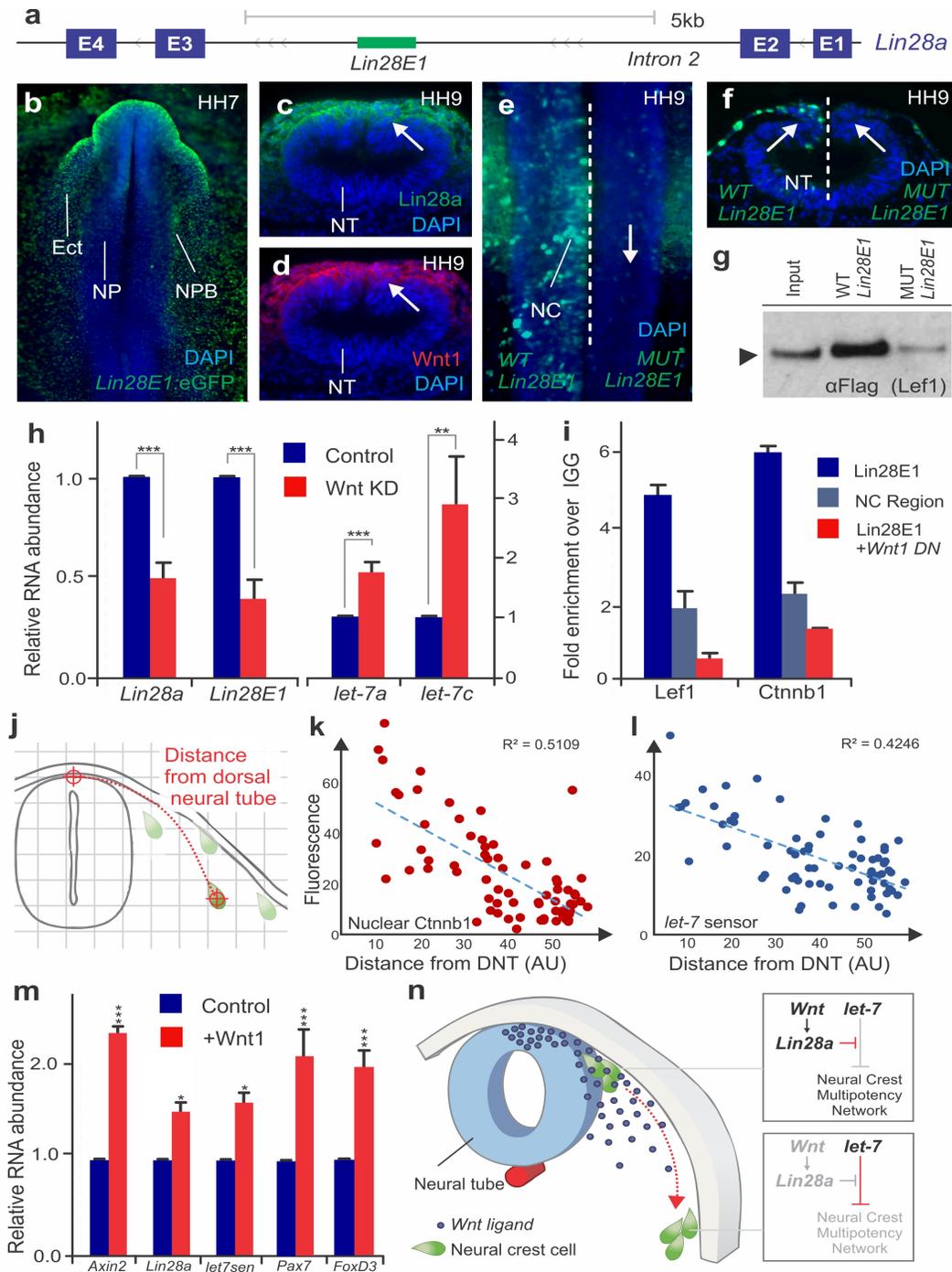


Figure 5. Positional information determines silencing of multipotency during neural crest migration

(a) The avian *Lin28a* gene locus, showing the exonic (E1-E4) and intronic regions. *Lin28E1* is a 698 bp region within the second intron, which was identified as a *Lin28a* enhancer. (b) Expression pattern of a *Lin28E1* driven eGFP construct, which is active in the ectoderm (Ect) and at the neural plate border (NPB). Immunohistochemistry for *Lin28a* (c) and *Wnt1* (d) on transverse sections of HH9 embryo (arrows point to the dorsal tube). (e-f) Comparison of *Lin28E1* (left) and *MUT Lin28E1* (right) reporter activity in a bilaterally electroporated embryo. Dorsal view of the head of a whole-mount embryo (e) and transverse section (f, arrows point to neural crest cells). (g) Western blot for flag-Lef1 following enhancer pull-down experiment with wild type and mutant *Lin28E1*. (h) RT-PCR for endogenous

Lin28a mRNA, *Lin28E1* reporter, and mature *let-7* miRNAs on combinatorial knockdown of Wnt1 and Wnt4 (i) Chromatin immuno-precipitation for Lef-1 and β -catenin (Ctnnb1), performed with neural folds of W.T. embryos, and embryos electroporated with Wnt dominant-negative construct. (j) Diagram outlining the parameters for single-cell measurements of *let-7* sensor fluorescence and nuclear β -catenin (Ctnnb1), as a function of the distance to the dorsal neural tube. Fluorescence intensity of nuclear β -catenin (k) and *let-7* sensor (l) is inversely correlated with the distance of the neural crest cell from the neural tube. In both graphs, each dot represents a single cell. (m) Quantitative comparison of transcript levels of *Axin2*, *Lin28a*, *let-7* sensor, *Pax7*, and *FoxD3* in control vs. Wnt1 overexpressing neural crest cells (n) Model summarizing the results. A.U.: arbitrary units, D.N.: dominant-negative, DNT: dorsal neural tube, IGG: Immunoglobulin G, *Lin28E1*: *Lin28a* Enhancer 1, Neg region: negative control region. Error bars in (h) and (m) represent standard error and error bars in (i) reflect standard deviation between technical replicates.

Discussion

The neural crest is a migratory and multipotent cell type that undergoes extensive regulatory changes during differentiation. Hence, it is a powerful *in vivo* model to explore how environmental cues and transcriptional identity are integrated during cell state transitions. In this study, we examine neural crest development to characterize a mechanism linking positional information and the remodeling of gene regulatory networks that control multipotency. According to our model (Fig 5n), high levels of Wnt ligands produced by the dorsal neural tube activate *Lin28a* transcription in neural crest stem cells. The high levels of the *Lin28a*, in turn, inhibit *let-7* activity, protecting the neural crest transcriptional network from repression by these microRNAs. As neural crest cells migrate away from the Wnt source, *Lin28a* levels are significantly reduced, resulting in an increase of mature *let-7* levels and subsequent repression of multipotency factors. This inhibition of crucial network nodes results in silencing of the neural crest gene regulatory network and loss of stem cell identity.

Wnt is a major modulator of neural crest identity, acting reiteratively during the formation and differentiation of this cell type (Raible and Ragland, 2005). Our results indicate that canonical Wnt signaling acts via the *Lin28a/let-7* axis to promote neural crest multipotency. Our model elucidates how this signaling system provides developing cells with spatial information during cell fate restriction (Loh et al., 2016). As neural crest cells migrate away from the dorsal neural tube, we observe a gradual decrease in the activation of the pathway (Fig 5k), which underlies the silencing of neural crest progenitor identity

(Fig 1d). These spatial dynamics suggest an interesting parallel to the classic niche models in stem cell biology, in which the microenvironment provides signals that endow progenitor cells with broad potential (Morrison and Spradling, 2008). Our results indicate that the dorsal neural tube is an important signaling center in the embryo, acting to maintain neural crest stem cells in a multipotent state. The *Lin28a/let-7* axis is a crucial component of this mechanism. Little was known about the role of *Lin28a* in neural crest development before our analysis; previous work showed that *Lin28a* and its paralog, *Lin28b*, are dynamically expressed during amniote embryonic development (Yokoyama et al., 2008), and *in vitro* experiments suggest that these factors play a role in neurogliogenesis (Balzer et al., 2010). Here we identify an essential function of *Lin28a* factor during the early stages of neural crest development, which is consistent with its role as a regulator of pluripotency in stem cells and cancer (Shyh-Chang and Daley, 2013).

We postulate that this regulatory mechanism is independent of the role of Wnts in neural crest induction and specification (Garcia-Castro et al., 2002; Simoes-Costa et al., 2015), and its later function as a driver of differentiation into melanocytes (Dorsky et al., 1998) and sensory neurons (Lee et al., 2004). Neural crest cells display specific responses to Wnt manipulations performed at distinct stages of development, both before (Garcia-Castro et al., 2002; Simoes-Costa et al., 2015) and after specification (Hari et al., 2012). Thus, we believe our model is compatible with the previous studies that show a requirement of canonical Wnts in cell fate decisions. While we still have a limited understanding of the mechanisms that compartmentalize the distinct functions of the pathway, time-controlled studies have shown that timing of intracellular response is crucial for specificity (Hari et al., 2012). Our results suggest that *Lin28* activation by Wnts is established very early in neural crest progenitors and that this interaction is lost as the response to the pathway is attenuated during migration. In birds, neural crest cells form two waves of migration, with the melanocytic subpopulation delaminating later than the chondrocytic and neural progenitors (Erickson and Goins, 1995; Le Douarin and Kalcheim, 1999). While we hypothesize that the *Wnt-Lin28/let-7* circuit operates in all neural crest cells, further experiments will be necessary to

clarify how the *Lin28/let-7* axis operates in these subpopulations, which can differentiate at distinct positions relative to the dorsal neural tube.

Our findings also clarify how intricate regulatory programs can be rapidly silenced during cell state transitions. Studies in the neural crest gene network have revealed numerous positive interactions that act to stabilize progenitor cell identity (Simoës-Costa and Bronner, 2015). Indeed, the presence of positive regulatory loops is a common feature of developmental regulatory networks (Davidson, 2010; Davidson and Levine, 2008). Nevertheless, the identity of a progenitor cell is transient by definition, and thus the network has to collapse as cells differentiate. This rapid silencing of progenitor identity during differentiation is evident in our quantitative analysis (Fig. 1d), which shows that the majority of the components of the network are only transiently expressed, being downregulated as the cells migrate away from the neural tube. Genes that are interlinked by positive interactions are still progressively lost as cells migrate ventrally. For instance, *Pax3/7* has been shown to activate itself to stabilize and maintain neural crest identity (Plouhinec et al., 2014). Similarly, *Tfap2a*, *Sox9*, and *Snai2* form a positive feedback loop in neural crest cells (Luo et al., 2003) which may explain the rapid increase in expression of these factors during neural crest specification. *SoxE* factors are also known to partake and auto-regulatory loops in multiple contexts (Honore et al., 2003; Mead et al., 2013; O'Donnell et al., 2006). Thus, there is a disconnect between the logic encoded in network architecture (Simoës-Costa and Bronner, 2015) and plasticity of cell identity observed *in vivo*.

We address this by identifying a post-transcriptional mechanism that is able to override positive regulatory interactions, silencing the neural crest gene regulatory network. Our experiments indicate that in the absence of *Lin28a*, the expression of mature *let-7* microRNAs increases dramatically (Fig. 3c-e), reaching levels that are sufficient to inhibit neural crest stem cell identity. This is due to the fact that the *let-7* target genes identified in our UTR-reporter assay represent critical nodes of the network that are involved in the regulation of multiple neural crest genes (Simoës-Costa and Bronner, 2015). *Pax3/7*, *FoxD3*, and *cMyc* are stem cell factors and *bona fide* markers of neural crest cells, which play central

roles in induction and specification (Basch et al., 2006; Dottori et al., 2001; Kerosuo and Bronner, 2016; Krishnakumar et al., 2016). These factors are also part of a group of genes that define neural crest stem cell identity (Lignell et al., 2017) located within the Wnt niche. Our demonstration that a Wnt-Lin28a/*let-7* regulatory circuit is able to modulate this regulatory program highlights how gene networks can be extensively remodeled during cell state transitions.

Acknowledgments:

We are indebted to Carol Bayles and Adam Paul Wojno for cell-sorting assistance at the BRC Flow Cytometry Cell Sorting Facility at Cornell University. We thank Dr. Rebecca M. Williams for support in imaging acquisition through the BRC Imaging Facility and Carolina Purcell for assistance with the long-term culture system for avian embryos. This work was supported by NIH grant R00DE024232, the Meinig Family Foundation, and a Basil O'Connor Starter Scholar Award to M.S.-C.

Methods

Embryo collection and fixation

Fertilized chicken eggs (Leghorn White) were purchased from University of Connecticut (Department of Animal Science). Eggs were incubated at 37°C until the embryos reached the desired developmental stage. Embryos were collected and cultured according to the E.C. protocol (Chapman et al., 2001) and staged based on Hamburger and Hamilton (Hamburger and Hamilton, 1951). For immunohistochemistry, embryos were fixed with phosphate buffer (P.B.) containing 4% PFA for 20 minutes at room temperature (R.T.) and processed immediately. For *in situ* hybridization, embryos were fixed in phosphate buffer saline (PBS) containing 4% paraformaldehyde (PFA) for 2 hours at R.T. or overnight at 4°C. Following fixation, embryos were dissected, washed with PBST, dehydrated, and stored in methanol at -20°C. Whole-mount *in-situ* hybridization was performed as previously described (Wilkinson, 1992). For double *in-situ* hybridization, we used the Tyramide TSA system from Perkin

Elmer (TSA Plus Cyanine 5 & Fluorescein, NEL754001KT) as previously described (Denkers et al., 2004).

Embryo transfection and perturbation experiments

Chick embryos at HH4-5 were transfected with morpholinos, DsiRNAs, and DNA constructs by *ex ovo electroporation*, as previously described (Simoes-Costa et al., 2015). Briefly, morpholinos/DsiRNAs or DNA expression vectors were injected between the epiblast and vitelline membrane of dissected embryos and electroporated with platinum electrodes (five 50ms pulses of 5.1V, with an interval of 100ms between pulses) (Sauka-Spengler and Barembaum, 2008). In all gene knockdown and overexpression experiments, the embryos were injected bilaterally with the control reagent on the left side and the targeted reagent on the right side. Whole embryo injections were performed for enhancer analysis and neural crest sorting experiments. Following electroporation, embryos were cultured in albumin at 37°C until they reached the desired developmental stages. Embryo survival was >90%, and all embryos were screened to ensure that only uniformly electroporated, healthy embryos were used for further analysis. *Lin28a* knockdown was performed using FITC labeled translation-blocking morpholino (5'-AAACAGACCCCATCCCGACTCGC-3') (GeneTools). Both control and *Lin28a* morpholinos were injected at a final concentration of 1.5 mM, supplemented with 1µg/µl of carrier DNA and 10mM Tris pH 8.0. For *Lin28a* and *Lin28b* loss of function experiments using DsiRNAs, the control and targeted DsiRNAs (IDT) were injected at a final concentration 20 µM. The sequence of the DsiRNAs used are as follows:

Lin28a DsiRNA1: 5'GCCGUUGAAUUCACCUUCAAGAAAT-3'

Lin28a DsiRNA2: 5'-GGGGUCUGUUUCCAACCAGCAGUTT-3'

Lin28b DsiRNA1: 5'-GUGGAAUUUACUUACAAGAAAUCTT-3'

Lin28b DsiRNA2: 5'-AAGCUUACAUGGAAGGAUUUAGAA-3'.

The *gga-let-7-a-5p* mimic (miScript miRNA mimic, Qiagen, MSY0001101) was electroporated at a concentration of 100 μ M, with 1 μ g/ μ l of carrier DNA and 10mM Tris pH8.0, to facilitate entry into cells.

Immunohistochemistry

For whole mount immunohistochemistry, embryos were dissected from the filter paper after fixation and washed in TBS containing 0.1% Triton and 1% DMSO (TBTD). Embryos were blocked for 2 hours in TBTD supplemented with 10% donkey serum and incubated in primary antibody diluted in blocking solution, overnight at 4°C. The following primary antibodies were used: anti-Lin28a mouse monoclonal (DSHB, 1:4), anti-Sox10, goat polyclonal (R&D Systems, AF2864, 1:50), anti-Foxd3, rabbit polyclonal (1:200, gift from Patricia Labosky), anti-Pax7, mouse (DSHB AB528428, 1:4), anti-Cas9, rabbit polyclonal (Takara 632607, 1:200) anti-WNT1, rabbit polyclonal (Abcam, ab15251, 1:200), anti-Ctnnb1 (β -catenin) mouse monoclonal (BD Transduction Laboratories, 610154 1:100), anti-Tuj1 (BioLegend, 801202, 1:200), anti-pH3(S10) (Abcam, ab47297, 1:200), anti-Caspase3 (R&D Systems, AF835, 1:100) and anti-mCherry, rabbit polyclonal (Abcam, ab167453, 1:200). Secondary antibodies used included donkey anti-mouse/goat/rabbit IgG conjugated with Alexa Fluor 350/488/568/647 or goat anti-mouse Alexa 633 (Molecular Probes, 1:3000). Quantification of fluorescence for phenotype quantification in gain- and loss-of-function studies was performed with ImageJ.

Cryosectioning

Fixed embryos were washed in 5% sucrose for 3 hours at R.T. and in 15% sucrose solution overnight at 4°C. Next, they were incubated in 7.5% porcine gelatin for 3 hours at 37°C, embedded in silicone molds, snap frozen in liquid nitrogen, and stored at -80°C. 5-10 μ m sections were obtained using the CryoStar NX50 (Thermo Fisher). For imaging, the slides were immersed in PBST at 42°C for 15 mins for gelatin removal, washed in PBS, and mounted with Fluoromount-G (Southern Biotech, 0100-01).

Expression vectors

The *Lin28a* expression construct was assembled by insertion of the full-length coding sequence of avian *Lin28a* in a pCI-H2B-RFP backbone. The coding sequence of *Lin28a* was PCR amplified from an HH8 cDNA library. To generate the mCCHC mutant version of the *Lin28a* expression construct, we introduced two mutations: H147A and H169A in the CCHC domain of the protein, which has been previously reported to abolish *Lin28a* binding to the stem-loop region of *pre-let-7* miRNAs (Heo et al., 2008). The pRNA-U6-*let-7* sponge construct was a gift from Philip Zamore (Addgene plasmid # 35664). The *let-7* sensor was constructed by cloning the *let-7* sponge sequence (amplified from the pRNA-U6-*let-7* sponge plasmid) downstream of a destabilized mCherry coding sequence (mCherry-PEST) in a pCAGGS backbone. The specificity of the sensor was assayed by electroporation of a *let-7a* mimic molecular, which resulted in a strong loss of mCherry expression. All expression vectors were sequenced to ensure that no additional mutations were present.

Nanostring Analysis

To identify the genes regulated by the *Lin28a/let-7* circuit, we performed Nanostring analysis in two experimental conditions: *Lin28a* morpholino-mediated knockdown and *let-7* mimic treatment. Stage HH4 chick embryos were electroporated with control morpholino on the left side and with *Lin28a* morpholino or *let-7* mimic on the right side. Both morpholinos were injected at a concentration of 1.5 mM, and the *let-7* mimic was diluted to a final concentration of 100 $\mu\text{M}/\mu\text{l}$. Post-electroporation, embryos were incubated at 37°C for ~12 hours until they reached stage HH9. The control and targeted dorsal neural folds of embryos were microdissected and lysed in RNAqueous lysis buffer (RNAqueous-Micro Kit, AM1931). RNA lysates were hybridized at 65°C for 12 hours to a Nanostring probe set

containing ~100 probes for neural crest, placodal, and neural genes (Simoes-Costa et al., 2015).

Analysis of Nanostring data was performed with the nSolver software.

3'UTR reporter assay

3'UTR reporter constructs for *FoxD3*, *Pax7*, *Myc*, *Sox10*, *Zic1*, and *Sox8*, were built by amplifying the 3'UTR regions of these genes (as annotated in UCSC genome browser, Galgal 5.0) from an HH8 cDNA library prepared with oligo dT primers. Each 3'UTR was fused to a destabilized mCherry reporter (mCherry-PEST) in a pCAAGS vector backbone. In gastrulating embryos, 1 $\mu\text{g}/\mu\text{l}$ of a 3'UTR reporter construct was transfected in the control side of the embryo (left) and co-transfected with *let-7a* mimic in the experimental side of the embryo (left). As a transfection control, 1 $\mu\text{g}/\mu\text{l}$ of an eGFP expression construct built with the same vector backbone but lacking the 3'-UTR regions was co-injected with the mCherry reporters. After incubation of embryos at 37°C for 12 hours, control and *let-7* mimic transfected halves of the same embryo were dissected and processed independently for flow cytometry. Fluorescent intensity of mCherry and GFP in dual positive cells were quantified, and the ratio of mCherry/GFP intensity was used as a parameter for measuring reporter activity.

Embryo dissociation and cell sorting

For isolation of neural crest cells, embryos were transfected with 1 $\mu\text{g}/\mu\text{l}$ of an enhancer of the *Tfap2a* gene (Attanasio et al., 2013) (*Tfap2aE1*) cloned into PTK-eGFP (Uchikawa et al., 2003). To obtain neural crest cells from different stages, embryos were cultured until HH6 (8 hours), HH8 (~11 hours), HH10 (13-14 hours), HH12 (~18 hours), and HH14 (~23 hours), and screened for robust GFP expression in neural crest cells. Embryo heads were dissected in Ringers solution, washed with dPBS, and incubated in Accumax (Accutase SCR006) cell dissociation solution for 40 minutes at R.T. under mild agitation. Following this, dissociated cells were passed through a cell strainer (Pluriselect USA,

Mini Cell Strainer II, 45-09840-50) and centrifuged at 400g for 10 minutes. The supernatant was carefully discarded, and cells were resuspended in 200 μ l of HANKS buffer supplemented with 0.5% BSA. At least 1500 GFP+ and GFP- cells from each stage were sorted directly into 50 μ l of lysis buffer from Power SYBR Green Cells-to-CT Kit (ThermoFisher, 4402953) using BD AriaFusion cell sorter. To assay for the effects of *Lin28a* overexpression or *let-7* sponge over-expression in neural crest cells, embryos were bilaterally electroporated with *Tfap2aE1* on the left and with *Tfap2aE1*+*Lin28a*-RFP / *Tfap2aE1*+ *let-7* sponge expression vector on the right. Following incubation at 37°C for 18-19 hours, each half of the head of individual embryos were dissected and processed separately for FACS sorting. GFP+/RFP+ populations of cells were sorted from the experimental side of the embryo, while GFP+ neural crest cells were collected from the control side. We employed RT-RT-PCR to compare gene expression levels between control and targeted cells obtained from the same embryo.

Quantitative reverse transcription PCR (RT-PCR)

To quantify changes in gene expression caused by perturbation/reprogramming experiments, we microdissected single neural folds from control and targeted side of the embryo, which were subsequently lysed in lysis buffer from Power SYBR Green Cells-to-CT Kit. RNA extraction and cDNA preparation were performed according to the kit's protocol). RT-PCR was performed using Power Sybr Green PCR master mix (Thermo Fisher, 4368577) in an ABI viia7 RT-PCR machine. Ct values of all genes were normalized to reference gene *hprt* and expressed as a fold change compared to the control sample.

Cornish pasty culture of chick embryos

To assess the long term consequences of *Lin28a* knockdown, cornish pasty culture (Nagai et al., 2011) was performed with HH4 chick embryos bilaterally transfected with *Lin28* MO. Following

electroporation, embryos were transferred to Panett Compton media (12 ml of Solution1 +18ml of Solution 2 +270 ml of dH₂O), released from the filter paper and folded along the anterior-posterior axis with the dorsal side out. Embryos were allowed to rest in Panett-Compton solution for about 30 mins. The excess extra-embryonic membrane was cut with fine surgical scissors, and the embryos were transferred to a media composed of 2:1 ratio of Albumin to Panett Compton solution. Embryos were incubated in this media at 37°C for 48-50 hours until they reached HH15.

Quantification of mature let-7

To measure levels of mature *let-7* miRNAs, RNA was extracted from dissected control and *Lin28a*MO-targeted neural folds with RNeasy Plus Micro kit (Qiagen, 74034), following the guidelines for small RNA extraction. Poly(A) tailing and cDNA synthesis were performed using the qScriptTM microRNA cDNA Synthesis Kit (Quanta Biosciences, 95107-025). RT-PCR for individual *let-7*s was done as suggested by the kit, with mature miRNA-specific primers, and a universal primer against the poly-A tail. Ct values were normalized to 18S rRNA and expressed as a fold change compared to the control sample.

CRISPR-Cas9 mediated knockdown of gene expression

To knock down *Lin28a* using CRISPR-Cas9, gRNAs targeting the first exon of *Lin28a* were designed using online resources (crispr.mit.edu). A combination of two gRNAs was cloned downstream of the U6 promoters in the pX333 vector (a gift from Dr. Andrea Ventura, AddGene plasmid #64073)(Maddalo et al., 2014). To assay for the effects of *Lin28a* knockdown, HH4 embryos were bilaterally electroporated with empty *Cas9* vector on the left and with *Cas9*+*Lin28a* gRNA construct on the right. The embryos were incubated at 37°C until they reached HH9+ when they were stained with FoxD3 (Mundell and Labosky, 2011) and Sox10 (R&D Systems, AF2864) antibodies, sectioned and analyzed for phenotype.

Individual FoxD3 and Sox10+ cells were counted from multiple sections obtained from three bilaterally transfected embryos. For disrupting *let-7* binding sites on 3'-UTR of *FoxD3* and *Pax7*, a unique gRNA was designed for each of the sites, as well as for a control region in the 3'-UTR of FoxD3 which did not contain *let-7* binding sites. The gRNAs were individually cloned under an U6-promoter in a modified pX333 vector, which had a GFP sequence cloned downstream of the Cas9 (Cas9-GFP). To assay for the effect of disruption of *let-7* binding sites on FoxD3 and Pax7, HH4 embryos were electroporated with empty Cas9-GFP vector on the left and with Cas9-GFP+UTR gRNA construct on the right. The embryos were allowed to develop until HH12, after which the control and targeted halves of the head of individual embryos were dissected separately, and FACS sorted for GFP+ cells. Finally, we employed RT-PCR to measure the expression of *FoxD3* or *Pax7*, in cells obtained from the control vs. experimental side.

Single-cell clonal analysis

Single-cell clonal analysis to assay for neural crest multipotency, for performed as described previously (Lahav et al., 1998). Briefly, HH4 quail embryos were injected with a control (pCI:H2B-RFP) or a *Lin28a* O/E (pCI:LIN28A-H2B-RFP) construct. The embryos were incubated at 37°C to develop until the 6-somite stage (HH9-) and screened for robust RFP expression. To isolate neural crest cells, we dissected 6-8 neural folds from control and *Lin28a* o/e quail embryos and plated them on collagen coated tissue-culture dish containing 10%FBS-DMEM media. The explants were incubated for ~36 hours (at 37° C and 5% CO₂ conditions), until neural crest cells had migrated out and a halo of cells was visible around the explanted neural folds. The remaining neural fold tissue was removed, and the migratory neural crest cells were dissociated with Accumax. The cells were resuspended in fresh media (10% FBS-DMEM + 2% Chicken Embryo Extract), and sparsely plated on collagen coated 6-well plates. After 3-4hours, once the cells had attached and spread out, the plates were screened to make sure that the > 90% of the plated cells were isolated, with only 1 cell/field visible using a 10X objective.

The cells were allowed to differentiate over a period of 10 days, after which the different cell types were assayed using immunofluorescence. Immunofluorescence of neural crest clones was performed as described previously. Briefly, the cells were fixed in 4% PFA at R.T. for 10 mins. Following fixation, the cells were permeabilized using 0.1% NP-40 solution in PBS at 37C for 30 mins. Next, the cells were blocked in 1%BSA solution at 37C for 30 mins, after which they were incubated with primary antibody cocktail at 1 h at 37 C. The primary antibodies used for detecting different cell types was as follows: anti-SMA for myofibroblast (rabbit, Abcam), anti-Runx2 for cartilage (mouse IgG2a, DSHB), anti-GFAP for glia (rabbit, Abcam), anti-Neurofilament for neurons (mouse IgG2a, Biolegend) and anti-MelEM for melanocytes (mouse IgG1a, DSHB). Following incubation with primary antibody, the cells were washed and incubated with corresponding secondary antibodies for 90 mins at 37°C. Finally, the cells were washed, stained with DAPI and imaged using a Nikon eclipse inverted microscope. A total of 50 colonies per condition were scored for developmental potential (bipotent, tripotent or multipotent); a subset of these (30 per condition) were analyzed for cell composition to identify the progenitor time.

let-7 sensor activity assay

For single-cell Quantification of sensor activity, the *let-7* sensor (mCherry) construct was co-transfected with *Tfap2aE1*(GFP) in HH4 embryos. Embryos were incubated until desired stages, and dissected heads were processed (as described above) for flow cytometry. Fluorescent intensity of mCherry was measured in GFP positive neural crest cells, and mCherry/GFP intensity ratio in each cell was used as a readout of sensor activity.

Wnt loss- and gain-of-function

Wnt signaling was disrupted using two different strategies: combined inhibition of Wnt1 and Wnt4 with morpholinos (Wnt1: 5'-GATGATGCCCCTACGGAGCGGGAAT-3', Wnt 4: 5'-

GCGCAGGAAATACTCCGGGCTCATC-3')(Simoes-Costa et al., 2015), and using a Wnt dominant negative vector(Garcia-Castro et al., 2002). Morpholinos targeted to Wnt ligands were used at a concentration of 1.1 μ M each, and the Wnt dominant negative construct was electroporated at 1 μ g/ μ l. For activation of Wnt signaling, we employed a vector driving expression of Wnt1, which was electroporated at a concentration of 1 μ g/ μ l.

Enhancer Analysis

To identify the enhancer controlling *Lin28a* expression, we cloned 2-3kb DNA fragments from the *Lin28a* locus that contained regions of open chromatin as determined by ATAC-seq. We tested eleven ATAC peaks in the proximity of the *Lin28a* gene locus (chr23:169040-169093; chr23:169493-169873; chr23:173387-173487; chr23:173838-174004; chr23:175014-175120; chr23:176664-176839; chr23:176849-177557; chr23:177927-178209; chr23:179032-179308; chr23:183167-183754; chr23:184717-184992). Peaks in close proximity were grouped and as a result, we cloned 6 regions in ptk-eGFP vector backbone (Fig.5-fig. supplement 1a). These putative enhancer constructs were electroporated (at a concentration of 1 μ g/ μ l) in HH4 embryos and analyzed for eGFP expression at later stages. The region 39.1 (chr23:176849-177557) which is evolutionarily conserved in amniotes and was subsequently identified as the *Lin28a* enhancer, was henceforth referred to as *Lin28E1*.

Computational analysis of the *Lin28E1* sequence with Jaspar(Mathelier et al., 2016)

(jaspar.genereg.net) revealed the presence of four TCF/LEF binding sites. The mutant *Lin28E1* construct was constructed by replacing the 4 TCF/LEF1 binding sites in the *Lin28E1* construct with a repetitive "CTCTCT" sequence of the same length as the binding sites. To assay for *Lin28E1* activity, the transcript levels of the enhancer-driven GFP were measured by RT-PCR, and GFP expression was normalized to *hprt* levels.

Chromatin Immunoprecipitation

For each experiment, chromatin was isolated from 16 cranial neural folds dissected from HH8-9 embryos in Ringer's solution. Immunoprecipitation was performed as described (Simoes-Costa et al., 2014) using the *Lef1* (Millipore, #17-604) and *Ctnnb1* (B.D. Biosciences, #610154) antibodies and normal mouse IgGs (Millipore, #17-604) as controls. In *Wnt* dominant negative assays, HH5 embryos were electroporated with a *Wnt* dominant negative construct (Garcia-Castro et al., 2002); at stages HH8-9, cranial neural folds were dissected. Chromatin isolation and immunoprecipitation were performed as previously described (Garcia-Castro et al., 2002).

Enhancer pull down

For enhancer pull-down experiments, HH4 embryos were electroporated with a Flag-tagged *Lef1* construct cloned in a pCAAGS-H2B-RFP backbone. Embryos were incubated until stage HH9. After electroporation efficiency was confirmed, embryos were dissected in Ringer's solution (n=4 embryos per sample). Nuclear protein extracts were obtained as previously described (Simoes-Costa et al., 2015). *Lin28E1*, and *MUT Lin28E1* were PCR amplified from pTK-eGFP vectors using a biotinylated forward primer (5'-AAAATAGGCTGTCCCCAGTG-3') and an untagged reverse primer (5'-ATATTTCTTCCGGGGACACC-3'). Immobilization of nucleic acids was performed using Dynabeads™ MyOne™ Streptavidin T1 (Invitrogen, #2023-11-30) following manufacturer's protocol. For enhancer pull down, nuclear protein extracts were diluted in 10mM Tris-HCl, 1mM EDTA, 0.5mM EGTA, 10% Glycerol, 0.25% NP-40 supplemented with 10ug Poly(dI-dC) and incubated with biotinylated DNA coated Dynabeads for 90 minutes in rotation at 4 °C. Magnetic beads were washed 4 times in with Washing Buffer (10mM Tris-HCl, 1mM EDTA, 0.5mM EGTA, 100mM NaCl, 10% Glycerol, 0.25% NP-40), and proteins eluted in RIPA modified buffer containing 1x Sample Reducing Agent and 1x LDS Sample Buffer (Invitrogen, #B0009 and #B0007) for 15 minutes at 80°C and 1400rpm. Proteins were separated by electrophoresis on Bolt 4-12% Bis-Tris Plus mini gels (Invitrogen, #NW04120BOX),

followed by immunoblotting on nitrocellulose membranes using Monoclonal anti FLAG antibodies (Sigma, #A8592).

Quantification of Wnt and let-7 activity in single neural crest cells

To measure levels of activation of the canonical Wnt pathway and *let-7* sensor activity in single neural crest cells, we conducted confocal microscopy of 5 μ m embryo midbrain sections immunostained for β -catenin (Ctnnb1) and mCherry (*let-7* sensor). For image analysis, ImageJ software was used to quantify fluorescence intensity of individual neural crest cells within sections. To estimate *let-7* activity in neural crest cells, we measured mCherry intensity in *Tfap2aE1* positive cells. The perimeter of each neural crest cell was defined by expression of the enhancer. To quantify the activity of the canonical Wnt pathway, we measured the presence of β -catenin in nuclei of neural crest cells (DAPI was used to define the nucleus). The distance between each cell examined, and a fixed point on the dorsal neural tube was also measured. The distance from the dorsal neural tube was plotted linearly vs the fluorescence intensity of nuclear β -catenin or vs the *let-7* sensor intensity. A line of best fit was applied to the data to examine the correlation between migration and the levels of canonical Wnt and *let-7* activity.

Quantification of NC2-mCherry-PEST reporter activity and FoxD3 protein levels in migrating neural crest cells

For assessing the correlation between enhancer activity and protein levels of FoxD3, HH4 chick embryos were electroporated with a *FoxD3*- NC2 enhancer driven mCherry-PEST reporter construct, which is active in cranial migratory N.C. cells (Simoës-Costa et al., 2012). These embryos were subsequently incubated until HH12, and whole embryo immunostaining was performed for FoxD3. The embryos were embedded and cryosectioned, and single cell measurement of enhancer reporter

fluorescence and the FoxD3 antibody was performed with 5 μ m embryo midbrain sections. Image analysis and quantification was performed using ImageJ as described above.

Statistical analysis

At least 10 embryos were analyzed in immunohistochemistry and *in-situ* hybridization assays performed downstream of genetic perturbations experiments. For 3'-UTR reporter assay and *let-7* sensor activity assay, 5 embryos were analyzed per reporter construct or per developmental stage, respectively. CHIP experiments were repeated 3 times, and the results of a representative experiment are shown in the figures. To assess the long term consequence of *Lin28a* knockdown, Cornish pasty culture and subsequent analysis of trigeminal ganglia formation was performed for atleast 6 embryos. Single-cell quantification of canonical Wnt and *let-7* activity was performed in ~100 cells in three embryos, with consistent results. The Nanostring experiments were performed with 3 replicates of the *let-7* gain-of-function (3 experimental and 3 controls) and 2 replicates of *Lin28a* loss-of-function experimental conditions (2 experimental and 2 controls). Student's t-test (one-tailed) was performed to calculate p-values and $p < 0.05$ were considered to be significant. Mann-Whitney test was used to calculate p-values for 3'-UTR reporter and *let-7* sensor activity assay, given the non-parametric and non-Gaussian distribution of intensity values of cells analyzed in the assays.

Figure S1

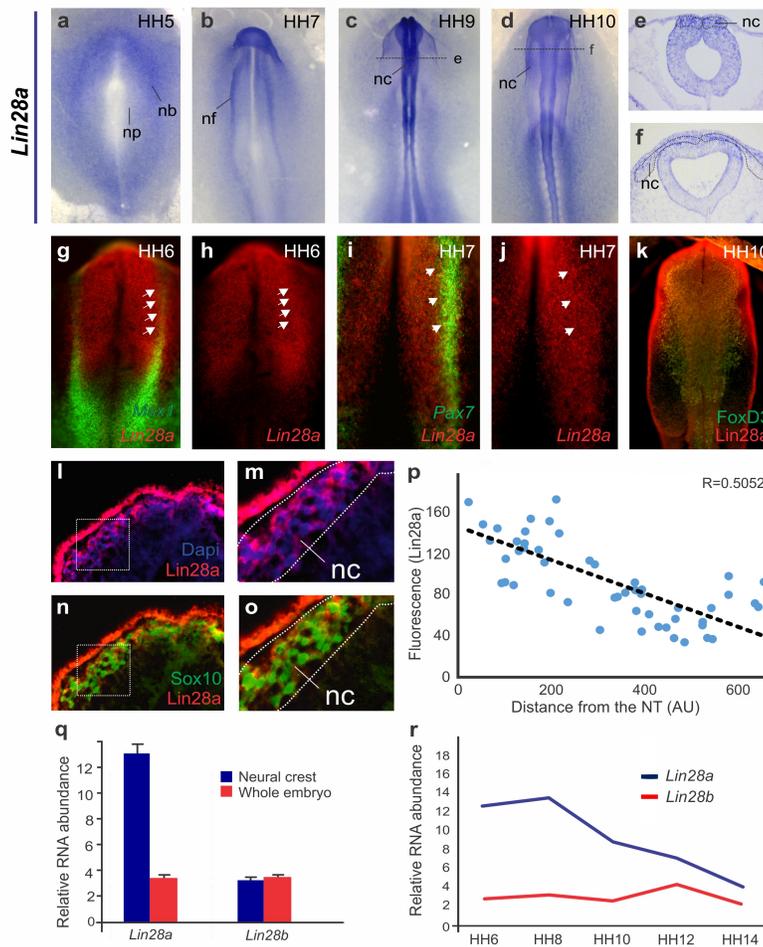


Figure S1. Expression patterns of *Lin28a* and *Lin28b* mRNA and *Lin28a* protein during early chick development.

(a-f) Colorimetric *in situ* hybridization for *Lin28a* in chick embryos of different developmental stages. *Lin28a* mRNA is enriched in the neural plate border at HH5 (a), in the dorsal neural folds at stage HH7-9 (b-c) and in migrating neural crest at stage HH10 (d). Transverse sections showing *Lin28a* expression in pre-migratory and migratory neural crest cells (e-f). (g-j) Fluorescent *in situ* hybridization for *Lin28a* and early neural crest genes *Msx1* and *Pax7*. At HH6, *Lin28a* expression overlaps with *Msx1* (g-h) and at HH7 with *Pax7* in the neural plate border (arrowheads) (i-j). (k) Immunohistochemistry for *Lin28a* protein, and neural crest markers *FoxD3*. In HH10 embryos, *Lin28a* protein (red) is expressed in *FoxD3*⁺ (green) neural crest cells (l-o). Transverse sections showing the localization of the *Lin28a* protein in the cytoplasm (l-m) of *Sox10*⁺ migratory neural crest cells (n-o). (p) Quantification of *Lin28a* fluorescence in migratory neural crest cells, showing that levels of *Lin28a* protein decrease as cells migrate away from the neural tube. (q) RT-PCR for *Lin28a* and *Lin28b* in FACS sorted neural crest (N.C.) cells and in whole embryo (WE) at HH8, showed that *Lin28a*, but not *Lin28b*, is significantly enriched in neural crest cells. (r) Relative expression levels of *Lin28b* paralog (red line) in FACS sorted neural crest cells at different developmental stages highlight that *Lin28b* is lowly expressed in neural crest cells and does not recapitulate the expression dynamics of *Lin28a*. The expression level of *Lin28a* (blue line) at the same developmental timepoints, shown in Fig1, has been included here for comparison. A.U.: arbitrary units. np: Neural plate, nb: neural plate border, nf: neural fold, nc: neural crest, nt: neural tube.

Figure S2

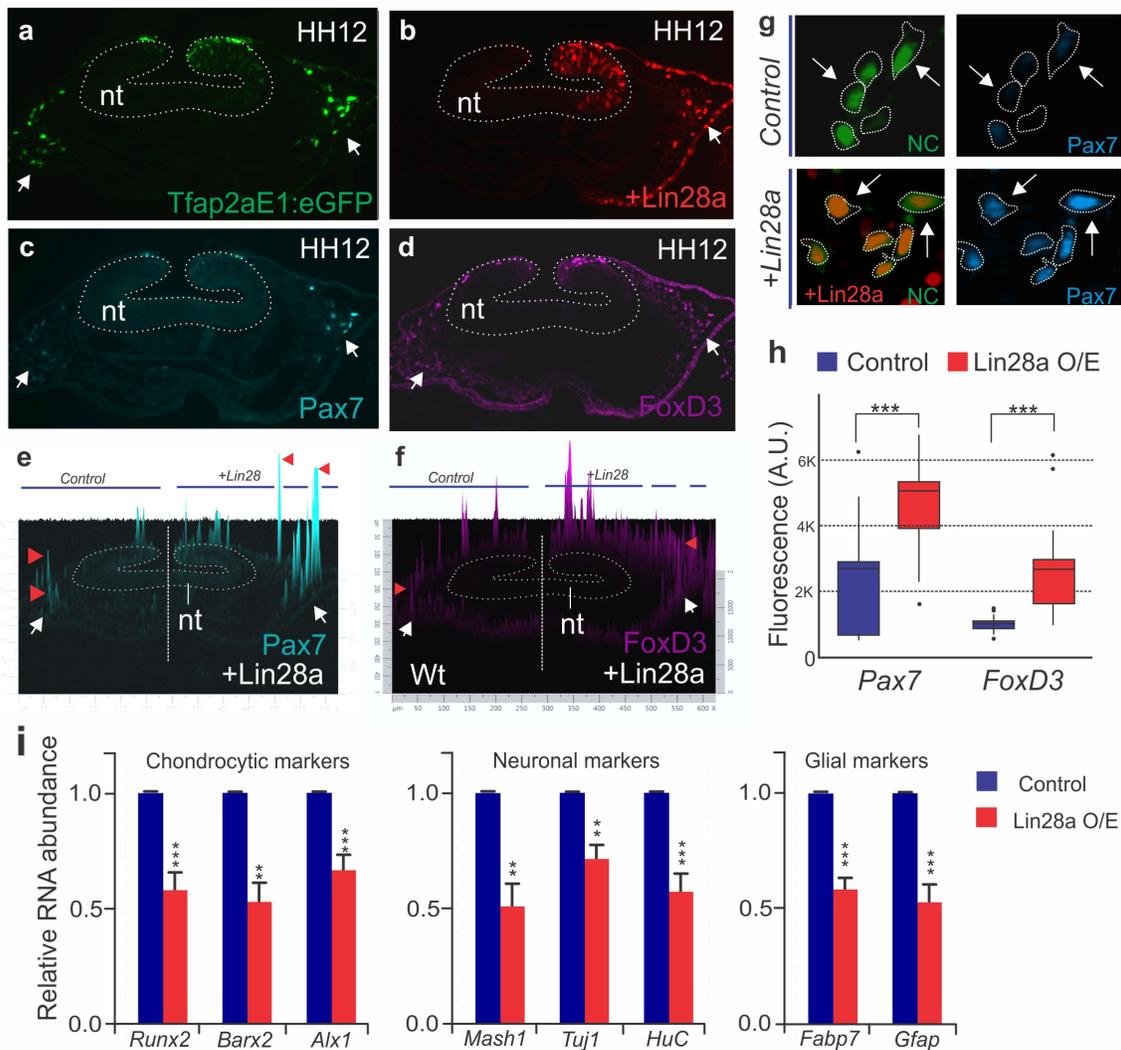


Figure S2 : Ectopic expression of Lin28a prevents silencing of early neural crest genes and delays differentiation.

Representative transverse section of an HH12 embryo, bilaterally injected with *Tfp2aE1:eGFP* on the left, and with *Tfp2aE1:eGFP* and *pCl:Lin28a-H2B-RFP* on the right. GFP expressing neural crest cells are shown in green (arrows) (a) and cells expressing Lin28a-H2B-RFP are shown in red (b). Immunostaining for Pax7(c) and FoxD3 (d) revealed a sustained expression of these proteins in the migratory neural crest on cells transfected with the Lin28a expression vector (arrows, e-f). 2.5D fluorescence intensity plots show that Pax7(e) and FoxD3 (f) intensity is higher in neural crest cells on the right side of the embryo, which express Lin28a in high levels. The height of the peaks (red arrowheads) is a measure of the fluorescence intensity of the proteins in the corresponding cells. (g) Insets showing Pax7 protein expression in control neural crest cells (green) and in cells overexpressing Lin28a (red). Dotted lines mark outlines of neural crest cells. (h) Boxplots showing the fluorescence intensity of FoxD3 and Pax7 in cells from control and Lin28a overexpression side of the embryo. Nt: neural tube, H.H.: Hamburger and Hamilton stages. (i) RT-PCR comparing the expression of differentiation markers in control and LIN28a o/e neural crest cells from HH12 embryos, show that chondrocytic (*Runx2*, *Barx2*, and *Alx1*), neuronal (*Mash1*, *Tuj1* and *HuC*) and glial differentiation (*Fabp7* and *Gfap*) is delayed following sustained Lin28a expression. A.U.: Arbitrary units, nt: Neural tube, o/e: over-expression.

Figure S3

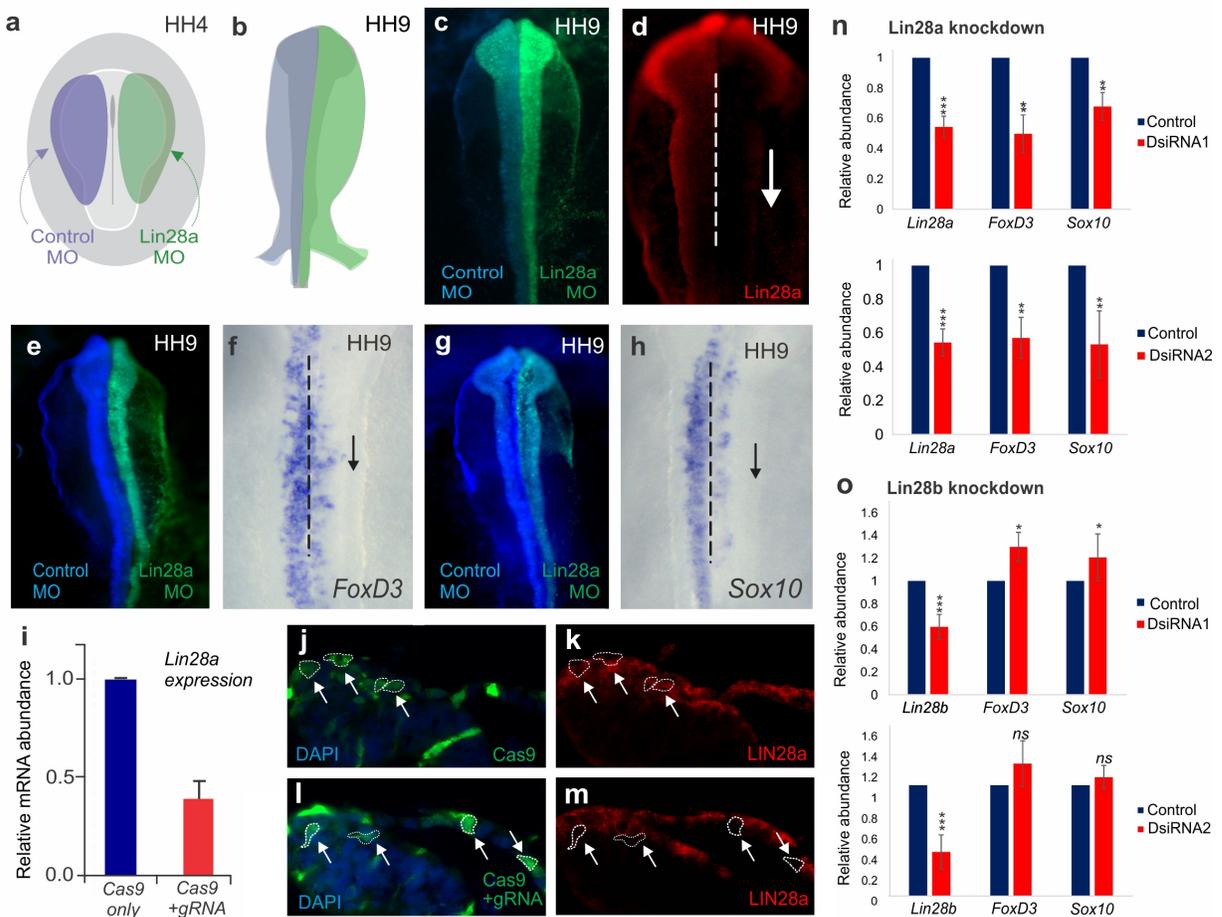


Figure S3. Supporting data for *Lin28a* and *Lin28b* loss-of-function analysis.

(a-b) Transfection strategy for knockdown experiments. HH4 embryos were injected with the control morpholino (blue) on the left and with the *Lin28a* morpholino (green) on the right (a). Following electroporation, embryos were cultured in albumin and incubated until HH9, when they were screened for efficient transfection and analyzed (b).

(c) Dorsal view of a bilaterally electroporated HH9 embryo, showing control MO (left) and *Lin28a* MO (right). (d) Immunohistochemistry for *Lin28a* showed loss of the protein on the experimental side (downward arrow). (e-h) *In situ* hybridization for neural crest markers *FoxD3* and *Sox10* following knockdown of *Lin28a*. (i-m) Disruption of *Lin28a* expression with CRISPR/Cas9. (i) Quantitative analysis of *Lin28a* transcripts following CRISPR-cas9 targeting of the protein through RT-PCR. Immunohistochemistry for Cas9 and *Lin28a* on transverse sections of embryos electroporated with either Cas9 alone (j-k) or with Cas9 vector containing a pair of gRNAs targeting *Lin28a*. (l-m) Dotted lines outline cells expressing Cas9 (arrows).

(n) RT-PCR for *Lin28a*, *FoxD3* and *Sox10*, following knockdown of *Lin28a* using two different DsiRNAs, confirm that loss of the protein results in reduced expression of these early neural crest markers. (o) Downregulation of *Lin28b* using two independent DsiRNAs followed by Quantification of *FoxD3* and *Sox10* transcripts reveal that *Lin28b* is not required for neural crest formation. MO: Morpholino, DsiRNA: Dicer-substrate siRNA, gRNA: guide-RNA, HH: Hamburger and Hamilton stages.

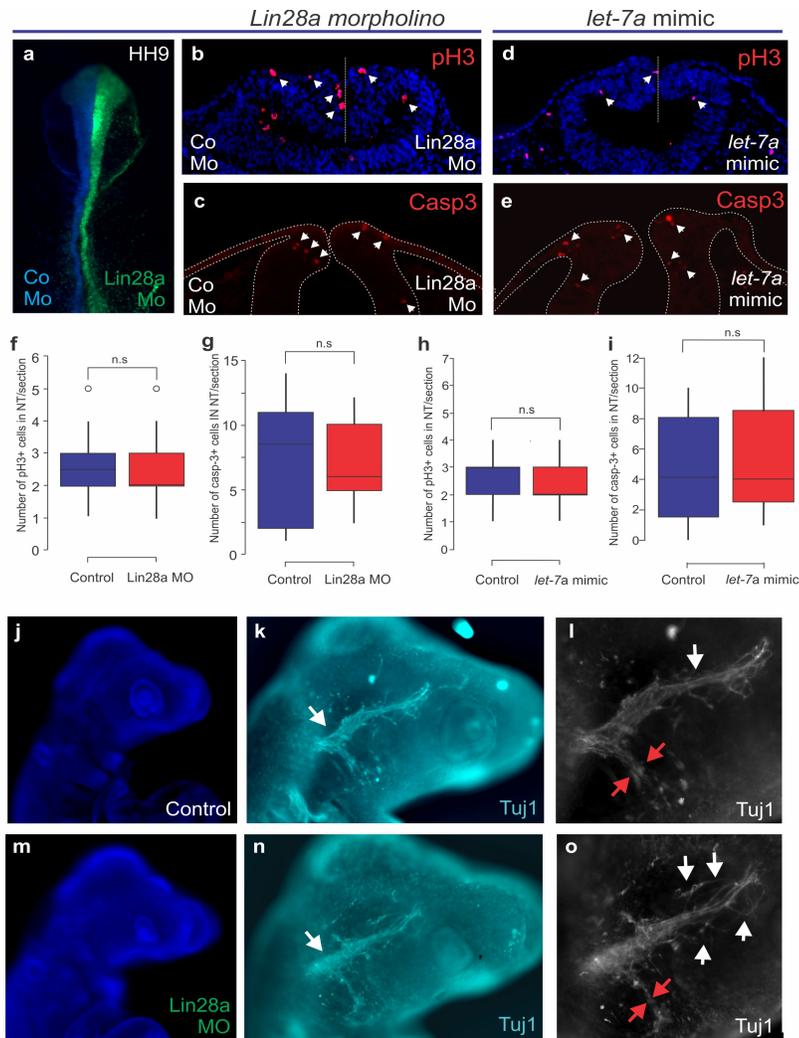


Figure S4. Effects of *Lin28a* loss-of-function on cell death, proliferation and the morphology of cranial ganglia.

(a-i) Disruption of *Lin28a/let-7* axis does not cause cell death or proliferation defects in the dorsal neural tube. (a) Whole mount view of an HH9 embryo bilaterally transfected with control MO (blue) on the left and the *Lin28a* MO (green) on the right. Transverse sections showing the immunostaining for phospho-H3 (S10) (b) and Caspase-3 (c) in these embryos. (d-e) Transverse sections of HH9 embryos bilaterally transfected with *let-7* mimic immunostained for phospho-H3 (d) or Caspase-3 (e). Quantification of average number of phospho-H3 (f) and Caspase-3 (g) positive cells on the neural tube/section in embryos transfected with *Lin28a* MO. Quantification of average number of phospho-H3 (h) and Caspase-3 positive (i) cells on the neural tube/section in embryos transfected with *let-7* mimic. (j-o) Loss of *Lin28a* during early development disrupts the formation of neural crest derived trigeminal ganglia. Dorsal view of Dapi-stained HH15 embryos showing control (j) and *Lin28a* MO (m) transfected side of the same embryos (control images were flipped horizontally to facilitate comparison). Immunohistochemistry for neuronal marker Tuj1 in control (k) and morpholino treated (n) side of the embryo shows the morphology of the trigeminal ganglion (white arrows point to the base of the ganglia). Higher magnification images of the trigeminal ganglia on the control (l) and morpholino-transfected (o) side of the embryo show disorganization of the trigeminal ganglia (n=6/8). The maxillomandibular branch (red arrows) is significantly thinner in the MO transfected side. The white arrows point to the neurons of the ophthalmic branch which are disorganised on the morphants. MO: Morpholino, ns: not significant.

Figure S5

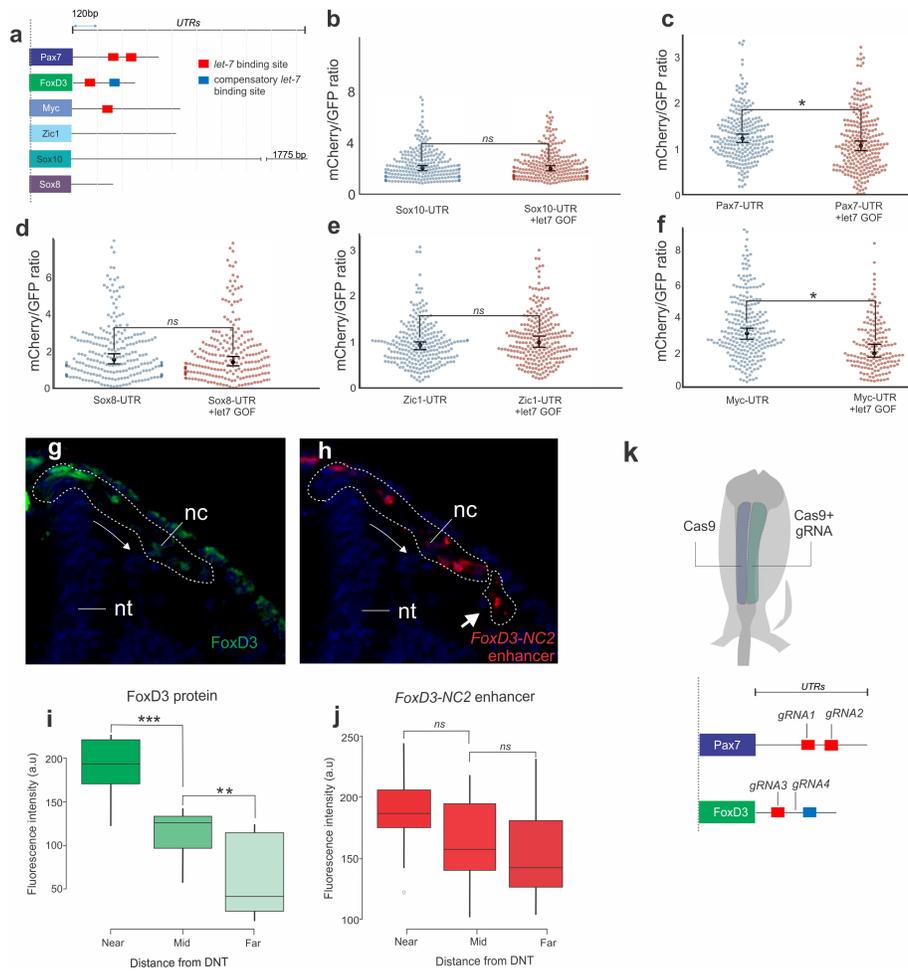


Figure S5. *let-7* miRNAs regulate 3'-UTRs of neural crest genes.

(a) Position of *let-7* binding sites in the 3'-UTRs of neural crest genes assessed in the reporter assay. Red boxes correspond to regions in the UTRs that are complementary to the *let-7* seed sequence, while the blue boxes correspond to *let-7* compensatory sites in the UTR, which are complementary to a region of the miRNA other than the seed sequence. Direct *let-7* targets (*Pax7*, *FoxD3*, *Myc*) have seed sequence complementarity and, in the case of *FoxD3*, a compensatory binding site. These are absent in the UTRs of *Sox10*, *Zic1*, and *Sox8*. (b-f) Representative scatter plots of *Sox10*, *Pax7*, *Sox8*, *Zic1* and *cMyc* UTR-reporter assay, showing the mCherry/GFP intensity ratio of each cell analyzed from the control (gene-UTR) and *let-7a* mimic transfected (gene-UTR+*let7* GOF) halves of the same embryo. Each dot represents a single cell, and the medians and the 99% confidence intervals are overlaid on the scatter plots. (g-j) Single cell measurement of *FoxD3* protein and *NC2* mCherry-PEST reporter construct fluorescence in migrating N.C. cells. Transverse section of HH12 embryos, showing *FoxD3* protein immunostaining (g) and *FoxD3-NC2* enhancer reporter construct (h) in N.C. cells. Dotted lines show the migrating N.C. cells. The farthest migrated cells expressing *FoxD3-NC2* enhancer (lower dotted region on (h)) are not positively stained for *FoxD3* immunostaining, indicating decreased protein in these cells (g) Boxplots quantifying the fluorescent intensity of *FoxD3* protein (i) *NC2*-mCherry-PEST reporter construct (j) in single neural crest cells as a function of the distance of the cells from the DNT. "NEAR", "MID" and "FAR" corresponds to cells within 0-200 a.u, 201-350 a.u and 350-600 a.u from the DNT. (k) Bilateral electroporations of control vs. targeted Cas9 expression vectors were used to disrupt individual *let-7* sites in the UTRs of the neural crest genes. UTR- Un-Translated Region, DNT- Dorsal Neural Tube, nt: neural tube, nc: neural crest, a.u- Arbitrary Units (as measured using ImageJ).

Figure S6

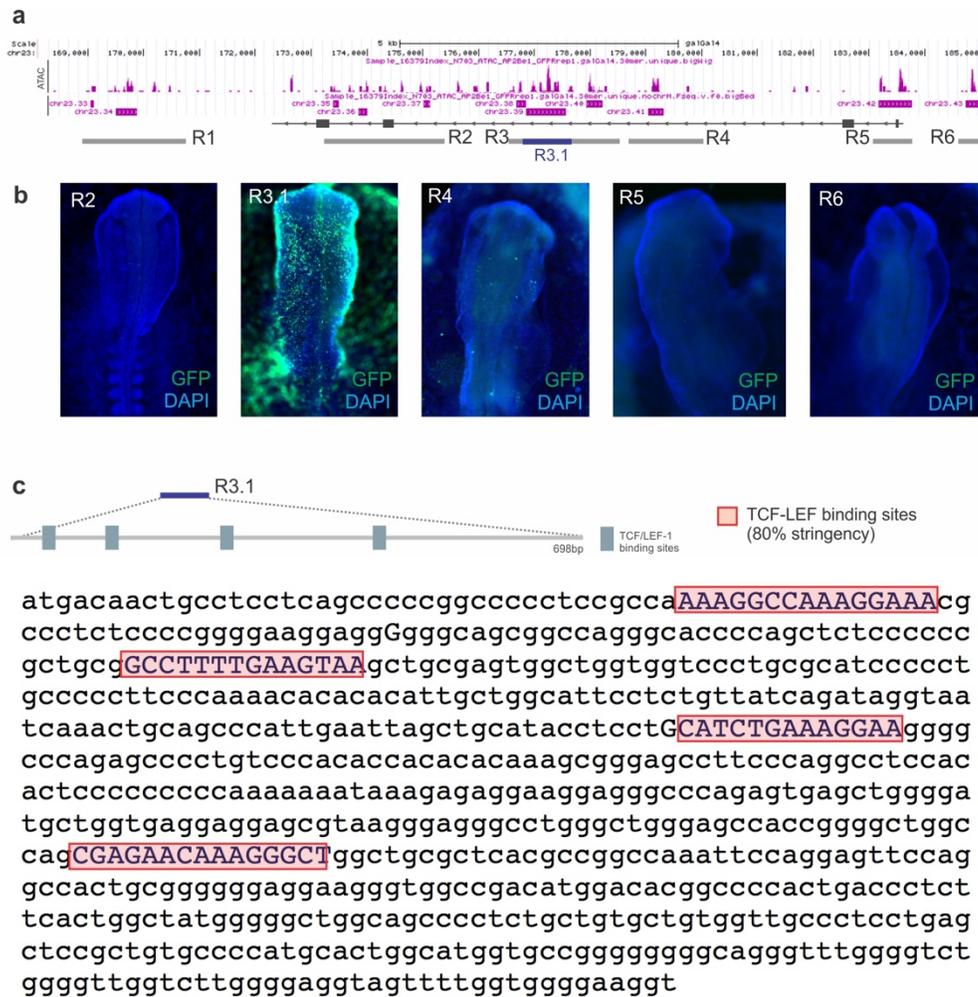


Figure S6. Identification of a *Lin28a* intronic enhancer.

(a) Genome browser snapshot of the ATAC seq peaks on chromosome 23 (*Galgal* 4.0) around and within the *Lin28a* gene locus. R1-R6 corresponds to the 6 regions (consisting of multiple ATAC peaks) which were cloned and assayed for enhancer activity (b). Representative images demonstrate that only R3/R3.1 were able to drive GFP expression in the embryo, in a pattern similar to that of *Lin28a* expression. Further analysis showed that R3.1, a 698 bp region located in the second intron of *Lin28a*, was sufficient to drive reporter expression. R3.1 (which was then renamed *Lin28E1*) has 4 TCF/LEF1 binding sites, as predicted by bioinformatics analysis. (c) Sequence of R3.1 (*Lin28E1*) showing the four TCF/LEF1 binding sites demarcated with red boxes.

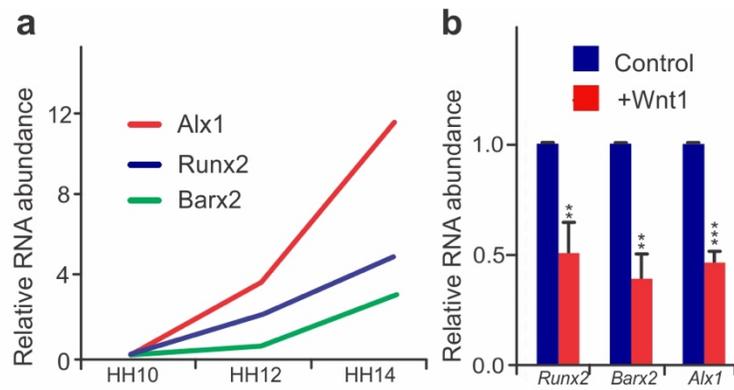


Figure S7. Prolonged Wnt signaling affects neural crest differentiation

(a) Normalized expression levels of differentiation genes *Runx2*, *Alx1* and *Barx2* in migratory neural crest cells in different stages of development. (b) RT-PCR comparing the levels of *Runx2*, *Alx1* and *Barx2* in control vs. Wnt1 overexpressing neural crest cells sorted from HH12 embryo. Error bars represent standard error.

References

- Attanasio, C., Nord, A.S., Zhu, Y., Blow, M.J., Li, Z., Liberton, D.K., Morrison, H., Plajzer-Frick, I., Holt, A., Hosseini, R., *et al.* (2013). Fine tuning of craniofacial morphology by distant-acting enhancers. *Science* 342, 1241006.
- Balzer, E., Heine, C., Jiang, Q., Lee, V.M., and Moss, E.G. (2010). LIN28 alters cell fate succession and acts independently of the let-7 microRNA during neurogliogenesis in vitro. *Development* 137, 891-900.
- Basch, M.L., Bronner-Fraser, M., and Garcia-Castro, M.I. (2006). Specification of the neural crest occurs during gastrulation and requires Pax7. *Nature* 441, 218-222.
- Betancur, P., Bronner-Fraser, M., and Sauka-Spengler, T. (2010). Assembling neural crest regulatory circuits into a gene regulatory network. *Annual review of cell and developmental biology* 26, 581-603.
- Buitrago-Delgado, E., Nordin, K., Rao, A., Geary, L., and LaBonne, C. (2015). NEURODEVELOPMENT. Shared regulatory programs suggest retention of blastula-stage potential in neural crest cells. *Science* 348, 1332-1335.
- Chapman, S.C., Collignon, J., Schoenwolf, G.C., and Lumsden, A. (2001). Improved method for chick whole-embryo culture using a filter paper carrier. *Developmental dynamics : an official publication of the American Association of Anatomists* 220, 284-289.
- Davidson, E. (2009). *The Regulatory Genome: Gene Regulatory Networks in Development and Evolution* (Burlington, MA.: Elsevier).
- Davidson, E.H. (2010). Emerging properties of animal gene regulatory networks. *Nature* 468, 911-920.
- Davidson, E.H., and Levine, M.S. (2008). Properties of developmental gene regulatory networks. *Proceedings of the National Academy of Sciences of the United States of America* 105, 20063-20066.
- Denkers, N., Garcia-Villalba, P., Rodesch, C.K., Nielson, K.R., and Mauch, T.J. (2004). FISHing for chick genes: Triple-label whole-mount fluorescence in situ hybridization detects simultaneous and overlapping gene expression in avian embryos. *Developmental dynamics : an official publication of the American Association of Anatomists* 229, 651-657.
- Dorsky, R.I., Moon, R.T., and Raible, D.W. (1998). Control of neural crest cell fate by the Wnt signalling pathway. *Nature* 396, 370-373.
- Dottori, M., Gross, M.K., Labosky, P., and Goulding, M. (2001). The winged-helix transcription factor Foxd3 suppresses interneuron differentiation and promotes neural crest cell fate. *Development* 128, 4127-4138.
- Erickson, C.A., and Goins, T.L. (1995). Avian neural crest cells can migrate in the dorsolateral path only if they are specified as melanocytes. *Development* 121, 915-924.
- Garcia-Castro, M.I., Marcelle, C., and Bronner-Fraser, M. (2002). Ectodermal Wnt Function As a Neural Crest Inducer. *Science* 13, 13.
- Hamburger, V. (1961). Experimental analysis of the dual origin of the trigeminal ganglion in the chick embryo. *J Exp Zool* 148, 91-123.
- Hamburger, V., and Hamilton, H.L. (1951). A series of normal stages in the development of the chick embryo. *Journal of morphology* 88, 49-92.
- Hari, L., Miescher, I., Shakhova, O., Suter, U., Chin, L., Taketo, M., Richardson, W.D., Kessar, N., and Sommer, L. (2012). Temporal control of neural crest lineage generation by Wnt/beta-catenin signaling. *Development* 139, 2107-2117.
- Heo, I., Joo, C., Cho, J., Ha, M., Han, J., and Kim, V.N. (2008). Lin28 mediates the terminal uridylation of let-7 precursor MicroRNA. *Molecular cell* 32, 276-284.
- Hollyday, M., McMahon, J.A., and McMahon, A.P. (1995). Wnt expression patterns in chick embryo nervous system. *Mech Dev* 52, 9-25.
- Honore, S.M., Aybar, M.J., and Mayor, R. (2003). Sox10 is required for the early development of the prospective neural crest in *Xenopus* embryos. *Developmental biology* 260, 79-96.

- Jho, E.H., Zhang, T., Domon, C., Joo, C.K., Freund, J.N., and Costantini, F. (2002). Wnt/beta-catenin/Tcf signaling induces the transcription of Axin2, a negative regulator of the signaling pathway. *Molecular and cellular biology* 22, 1172-1183.
- Kalkan, T., and Smith, A. (2014). Mapping the route from naive pluripotency to lineage specification. *Philosophical transactions of the Royal Society of London Series B, Biological sciences* 369.
- Kaufman, C.K., Mosimann, C., Fan, Z.P., Yang, S., Thomas, A.J., Ablain, J., Tan, J.L., Fogley, R.D., van Rooijen, E., Hagedorn, E.J., *et al.* (2016). A zebrafish melanoma model reveals emergence of neural crest identity during melanoma initiation. *Science* 351, aad2197.
- Kerosuo, L., and Bronner, M.E. (2016). cMyc Regulates the Size of the Premigratory Neural Crest Stem Cell Pool. *Cell reports* 17, 2648-2659.
- Kim, D.H., Behlke, M.A., Rose, S.D., Chang, M.S., Choi, S., and Rossi, J.J. (2005). Synthetic dsRNA Dicer substrates enhance RNAi potency and efficacy. *Nature biotechnology* 23, 222-226.
- Krishnakumar, R., Chen, A.F., Pantovich, M.G., Danial, M., Parchem, R.J., Labosky, P.A., and Blelloch, R. (2016). FOXD3 Regulates Pluripotent Stem Cell Potential by Simultaneously Initiating and Repressing Enhancer Activity. *Cell stem cell* 18, 104-117.
- Kumar, M.S., Erkeland, S.J., Pester, R.E., Chen, C.Y., Ebert, M.S., Sharp, P.A., and Jacks, T. (2008). Suppression of non-small cell lung tumor development by the let-7 microRNA family. *Proceedings of the National Academy of Sciences of the United States of America* 105, 3903-3908.
- Lahav, R., Dupin, E., Lecoin, L., Glavieux, C., Champeval, D., Ziller, C., and Le Douarin, N.M. (1998). Endothelin 3 selectively promotes survival and proliferation of neural crest-derived glial and melanocytic precursors in vitro. *Proceedings of the National Academy of Sciences of the United States of America* 95, 14214-14219.
- Le Douarin, N. (1982). *The neural crest* (Cambridge Cambridgeshire ; New York: Cambridge University Press).
- Le Douarin, N., and Kalcheim, C. (1999). *The neural crest*, 2nd edn (Cambridge: Cambridge University Press).
- Le Douarin, N.M., and Dupin, E. (2016). The Pluripotency of Neural Crest Cells and Their Role in Brain Development. *Curr Top Dev Biol* 116, 659-678.
- Lee, H.Y., Kleber, M., Hari, L., Brault, V., Suter, U., Taketo, M.M., Kemler, R., and Sommer, L. (2004). Instructive role of Wnt/beta-catenin in sensory fate specification in neural crest stem cells. *Science* 303, 1020-1023.
- Lignell, A., Kerosuo, L., Streichan, S.J., Cai, L., and Bronner, M.E. (2017). Identification of a neural crest stem cell niche by Spatial Genomic Analysis. *Nature communications* 8, 1830.
- Loh, K.M., van Amerongen, R., and Nusse, R. (2016). Generating Cellular Diversity and Spatial Form: Wnt Signaling and the Evolution of Multicellular Animals. *Developmental cell* 38, 643-655.
- Luo, T., Lee, Y.H., Saint-Jeannet, J.P., and Sargent, T.D. (2003). Induction of neural crest in *Xenopus* by transcription factor AP2alpha. *Proceedings of the National Academy of Sciences of the United States of America* 100, 532-537.
- Maddalo, D., Manchado, E., Concepcion, C.P., Bonetti, C., Vidigal, J.A., Han, Y.C., Ogradowski, P., Crippa, A., Rekhman, N., de Stanchina, E., *et al.* (2014). In vivo engineering of oncogenic chromosomal rearrangements with the CRISPR/Cas9 system. *Nature* 516, 423-427.
- Mathelier, A., Fornes, O., Arenillas, D.J., Chen, C.Y., Denay, G., Lee, J., Shi, W., Shyr, C., Tan, G., Worsley-Hunt, R., *et al.* (2016). JASPAR 2016: a major expansion and update of the open-access database of transcription factor binding profiles. *Nucleic acids research* 44, D110-115.
- Mead, T.J., Wang, Q., Bhattaram, P., Dy, P., Afelik, S., Jensen, J., and Lefebvre, V. (2013). A far-upstream (-70 kb) enhancer mediates Sox9 auto-regulation in somatic tissues during development and adult regeneration. *Nucleic acids research* 41, 4459-4469.
- Meulemans, D., and Bronner-Fraser, M. (2004). Gene-regulatory interactions in neural crest evolution and development. *Developmental cell* 7, 291-299.
- Moris, N., Pina, C., and Arias, A.M. (2016). Transition states and cell fate decisions in epigenetic landscapes. *Nature reviews Genetics* 17, 693-703.

- Morrison, S.J., and Spradling, A.C. (2008). Stem cells and niches: mechanisms that promote stem cell maintenance throughout life. *Cell* 132, 598-611.
- Mundell, N.A., and Labosky, P.A. (2011). Neural crest stem cell multipotency requires Foxd3 to maintain neural potential and repress mesenchymal fates. *Development* 138, 641-652.
- Nagai, H., Lin, M.C., and Sheng, G. (2011). A modified cornish pasty method for ex ovo culture of the chick embryo. *Genesis* 49, 46-52.
- Newman, M.A., Thomson, J.M., and Hammond, S.M. (2008). Lin-28 interaction with the Let-7 precursor loop mediates regulated microRNA processing. *RNA* 14, 1539-1549.
- Nieto, M.A. (2013). Epithelial plasticity: a common theme in embryonic and cancer cells. *Science* 342, 1234850.
- O'Donnell, M., Hong, C.S., Huang, X., Delnicki, R.J., and Saint-Jeannet, J.P. (2006). Functional analysis of Sox8 during neural crest development in *Xenopus*. *Development* 133, 3817-3826.
- Plouhinec, J.L., Roche, D.D., Pegoraro, C., Figueiredo, A.L., Maczkowiak, F., Brunet, L.J., Milet, C., Vert, J.P., Pollet, N., Harland, R.M., *et al.* (2014). Pax3 and Zic1 trigger the early neural crest gene regulatory network by the direct activation of multiple key neural crest specifiers. *Developmental biology* 386, 461-472.
- Raible, D.W., and Ragland, J.W. (2005). Reiterated Wnt and BMP signals in neural crest development. *Seminars in cell & developmental biology* 16, 673-682.
- Rybak, A., Fuchs, H., Smirnova, L., Brandt, C., Pohl, E.E., Nitsch, R., and Wulczyn, F.G. (2008). A feedback loop comprising lin-28 and let-7 controls pre-let-7 maturation during neural stem-cell commitment. *Nat Cell Biol* 10, 987-993.
- Sauka-Spengler, T., and Barembaum, M. (2008). Gain- and loss-of-function approaches in the chick embryo. *Methods in cell biology* 87, 237-256.
- Shiau, C.E., Lwigale, P.Y., Das, R.M., Wilson, S.A., and Bronner-Fraser, M. (2008). Robo2-Slit1 dependent cell-cell interactions mediate assembly of the trigeminal ganglion. *Nat Neurosci* 11, 269-276.
- Shyh-Chang, N., and Daley, G.Q. (2013). Lin28: primal regulator of growth and metabolism in stem cells. *Cell stem cell* 12, 395-406.
- Sieber-Blum, M., and Cohen, A.M. (1980). Clonal analysis of quail neural crest cells: they are pluripotent and differentiate in vitro in the absence of noncrest cells. *Developmental biology* 80, 96-106.
- Simoës-Costa, M., and Bronner, M.E. (2015). Establishing neural crest identity: a gene regulatory recipe. *Development* 142, 242-257.
- Simoës-Costa, M., and Bronner, M.E. (2016). Reprogramming of avian neural crest axial identity and cell fate. *Science* 352, 1570-1573.
- Simoës-Costa, M., Stone, M., and Bronner, M.E. (2015). Axud1 Integrates Wnt Signaling and Transcriptional Inputs to Drive Neural Crest Formation. *Developmental cell* 34, 544-554.
- Simoës-Costa, M., Tan-Cabugao, J., Antoshechkin, I., Sauka-Spengler, T., and Bronner, M.E. (2014). Transcriptome analysis reveals novel players in the cranial neural crest gene regulatory network. *Genome research* 24, 281-290.
- Simoës-Costa, M.S., McKeown, S.J., Tan-Cabugao, J., Sauka-Spengler, T., and Bronner, M.E. (2012). Dynamic and differential regulation of stem cell factor FoxD3 in the neural crest is Encrypted in the genome. *PLoS Genet* 8, e1003142.
- Stark, M.R., Sechrist, J., Bronner-Fraser, M., and Marcelle, C. (1997). Neural tube-ectoderm interactions are required for trigeminal placode formation. *Development* 124, 4287-4295.
- Trentin, A., Glavieux-Pardanaud, C., Le Douarin, N.M., and Dupin, E. (2004). Self-renewal capacity is a widespread property of various types of neural crest precursor cells. *Proceedings of the National Academy of Sciences of the United States of America* 101, 4495-4500.
- Tsialikas, J., and Romer-Seibert, J. (2015). LIN28: roles and regulation in development and beyond. *Development* 142, 2397-2404.
- Uchikawa, M., Ishida, Y., Takemoto, T., Kamachi, Y., and Kondoh, H. (2003). Functional analysis of chicken Sox2 enhancers highlights an array of diverse regulatory elements that are conserved in mammals. *Developmental cell* 4, 509-519.

Wilkinson, D.G. (1992). *In situ hybridization : a practical approach* (Oxford ; New York: IRL Press at Oxford University Press).

Yokoyama, S., Hashimoto, M., Shimizu, H., Ueno-Kudoh, H., Uchibe, K., Kimura, I., and Asahara, H. (2008). Dynamic gene expression of Lin-28 during embryonic development in mouse and chicken. *Gene expression patterns : GEP* 8, 155-160.

Yu, J., Vodyanik, M.A., Smuga-Otto, K., Antosiewicz-Bourget, J., Frane, J.L., Tian, S., Nie, J., Jonsdottir, G.A., Ruotti, V., Stewart, R., *et al.* (2007). Induced pluripotent stem cell lines derived from human somatic cells. *Science* 318, 1917-1920.

Zhang, J., Ratanasirintrao, S., Chandrasekaran, S., Wu, Z., Ficarro, S.B., Yu, C., Ross, C.A., Cacchiarelli, D., Xia, Q., Seligson, M., *et al.* (2016). LIN28 Regulates Stem Cell Metabolism and Conversion to Primed Pluripotency. *Cell stem cell* 19, 66-80.

Chapter III: Pluripotency factors are repurposed to shape the epigenomic landscape in neural crest stem cells

Alternative citation²:

Hovland AS⁺, Bhattacharya D⁺, Rothstein M, Simoes-Costa M (2021). Pluripotency factors are repurposed to shape the epigenomic landscape in neural crest stem cells. In Revision, *Developmental Cell*

⁺ Denotes equal contribution

Abstract

Yamanaka factors play an essential role in the establishment and maintenance of pluripotency in embryonic stem cells (ESCs), but their function in specialized stem cell populations is still poorly understood. Here we show that SOX2 and OCT4 cooperate with tissue-specific regulators to promote a multipotent epigenomic state in neural crest stem cells. By isolating primary neural crest cells at distinct stages of development, we characterized the transcriptomic and chromatin changes that take place during specification, migration, and early differentiation. Analysis of these datasets revealed that the OCT4-SOX2 dimer controls thousands of genomic elements specifically active in multipotent neural crest cells. We found that the emergence of this epigenomic state requires the translocation of OCT4-SOX2 to tissue-specific *cis*-regulatory regions. By examining genome reorganization during the induction of human ESCs into neural crest cells, we observed that the patterns of genomic occupancy of the dimer are modified during specification. Dimer translocation is guided by neural crest-specific pioneer transcription factors, which physically interact with the OCT4-SOX2 heterodimer to modify their genomic targets. Our results demonstrate how the ESC pluripotency program is repurposed in specialized stem cells to control chromatin organization and define the developmental potential of embryonic progenitors.

This work is a manuscript with the above citation, which is currently under revision. The study is a collaborative effort between Austin Hovland and Debadrita Bhattacharya. Austin Hovland performed the timecourse RNA-seq and ATAC-seq experiments conducted with chick embryos and performed the bioinformatic analysis for these datasets (Fig 1 & 2). Debadrita Bhattacharya carried out the CUT&RUN studies, functional experiments and performed genomic analysis of human neural crest cells (Fig 3-6). Megan Rothstein performed the proximity ligation experiments described in Fig 7.

Introduction

Understanding the molecular properties that allow multipotent progenitors to generate an array of cell types is a central question in biology, with broad implications for development and regenerative medicine. Stem cell populations that exist in the embryo and the adult body display varied degrees of plasticity and can be harnessed for tissue engineering applications. Research performed in the last decade has brought substantial progress in understanding the biology of embryonic stem cells (ESCs). In particular, the identification of the Yamanaka factors (*Oct3/4*, *Sox2*, *Klf4*, *c-Myc*) as central regulators of ESC pluripotency has transformed our understanding of the molecular control of developmental potential (Takahashi and Yamanaka, 2006). Nonetheless, the mechanisms underlying multipotency of adult and specialized stem cell populations are less understood. Defining the differences and similarities between these genetic programs will allow us to establish new strategies for cell and tissue engineering and inform upon the molecular basis of cellular plasticity.

A specialized stem cell population of particular interest for regenerative medicine is the neural crest. These cells develop adjacent to the central nervous system, from where they delaminate to populate the embryonic body. Neural crest derivatives include the craniofacial skeleton, the peripheral nervous system, the pigmentation of the skin, and cardiac tissue (Le Douarin and Kalcheim, 1999). Multiple lines of evidence have demonstrated that neural crest cells are multipotent and contribute to many tissues and organs. Classic experiments involving the transplantation of quail cells to chick embryos identified the extensive array of derivatives formed by the neural crest (Le Douarin, 1973). *In vivo* cell labeling with intracellular injection of vital dyes demonstrated that individual neural crest cells give rise to distinct cell types including neurons, glia, and pigmented cells (Bronner-Fraser and Fraser, 1988). More recently, genetic fate mapping of trunk neural crest cells with Confetti mice confirmed that both premigratory and migratory neural crest are multipotent (Baggiolini et al., 2015). This study also

showed that the number of multipotent cells does not change during early migration, indicating that fate restriction takes place at later stages of development.

The formation and differentiation of neural crest cells is orchestrated by a complex gene regulatory network (GRN), which has been substantially expanded in the last years with genomic analysis and functional studies (Martik and Bronner, 2017; Simoes-Costa and Bronner, 2015; Williams et al., 2019). Yet, we still have a superficial understanding of the regulation of developmental plasticity in this cell population. Currently, there are two opposing views on the molecular basis of neural crest multipotency. While some studies liken this genetic program to ESC pluripotency, others propose that it represents a new regulatory state that emerges during neural crest specification. For instance, Buitrago-Delgado and colleagues proposed that neural crest cells retain blastula-stage genes that are important for pluripotency (Buitrago-Delgado et al., 2015). Moreover, Zalc and colleagues examined the expression and function of OCT4 and suggested that the ESC pluripotency program is reactivated in the cranial neural crest lineage (Zalc et al., 2021). Both studies provide support for the idea that neural crest multipotency is linked to the ESC pluripotency network. This contrasts with the classical view that neural crest multipotency is controlled by a new regulatory circuit, distinct from that of ESCs, which emerges during neural crest specification (Dupin et al., 2018). This view is supported by transcriptome analyses indicating that neural crest cells express only a subset of the Yamanaka factors (Prasad et al., 2020; Williams et al., 2019). Additionally, disruption of known neural crest-specific transcription factors like *PAX7*, *FOXD3*, and *SOX10* leads to an imbalance in multipotency, suggesting an independent mechanism from that of ESCs (Dottori et al., 2001; Honoré et al., 2003; Kondoh et al., 2012; Lukoseviciute et al., 2018). Furthermore, developmental trajectories from single-cell transcriptome analysis has revealed that neural crest cells diverge from naïve progenitors during specification (Soldatov et al., 2019).

To define the molecular basis of neural crest multipotency, we surveyed the transcriptional and epigenomic dynamics of neural crest development, focusing on the transition from multipotency to differentiation. By conducting genomic analysis of primary cells isolated at a range of developmental

stages, we characterized how the cranial neural crest regulatory landscape changes during fate restriction. This allowed for the identification of the pioneer transcription factors that promote the epigenomic signature of naive neural crest cells, including pluripotency factors OCT4 and SOX2 (Soufi et al., 2012). Notably, we found that the OCT4-SOX2 dimer plays a central role in the neural crest multipotency program and that its genomic targets in neural crest are distinct from those in pluripotent embryonic stem cells. Translocation of the dimer from ESC to neural crest *cis*-regulatory regions requires the activation of pioneer factor TFAP2A, which promotes accessibility of neural crest enhancers. Thus, we propose that a subset of ESC pluripotency transcription factors is repurposed in neural crest cells to promote a tissue-specific multipotency signature. This model conciliates two seemingly contradictory views on the establishment neural crest multipotency and suggests how components of the pluripotency network may be co-opted in specialized stem cell populations.

Results

Transcriptomic and epigenomic dynamics of neural crest development

Genomic studies have identified new players in neural crest development that allowed for substantial expansion of GRN circuitry (Simoes-Costa and Bronner, 2013). Yet, these datasets lack the temporal resolution to uncover the transcriptional changes that drive the early steps of neural crest induction and the subsequent cell state transitions. To address this, we assembled an RNA-Seq time-course of primary neural crest cells at stages corresponding to induction, specification, migration and early differentiation. We labeled neural crest cells from Hamburger and Hamilton (HH) stage 6 to HH16 using a neural-crest specific enhancer (*Tfap2aE1*) driving the expression of GFP (Hamburger and Hamilton, 1951; Rothstein and Simoes-Costa, 2020) (Fig 1A). *Tfap2aE1* is a neural crest-specific enhancer conserved in amniotes that resides 584kb upstream of the TFAP2A promoter. Unlike other neural crest enhancers (Barenbaum and Bronner, 2013; Simoes-Costa et al., 2012), *Tfap2E1* is active throughout the formation, migration, and differentiation of neural crest cells (Fig S1A). *Tfap2E1* drives robust GFP expression at the neural

plate border (HH6), in pre-migratory and migratory neural crest (HH8-12), and in the cranial mesenchyme, trigeminal ganglia, and branchial arches (HH14-16) (Fig 1A). Histological analysis confirmed that *Tfap2E1* is specific to the neural crest lineage, as reporter expression was not detected in other ectodermal cell populations (Fig S1B).

Utilizing this enhancer-reporter construct, we obtained GFP⁺ cranial neural crest cells and GFP⁻ embryonic head cells by fluorescence-activated cell sorting (FACS) at six developmental time points (HH6, HH8, HH10, HH12, HH14, and HH16) and performed bulk RNA-seq (Fig 1B). Principal Component Analysis (PCA) of our RNA-seq datasets showed that GFP⁺ neural crest samples align chronologically along the first principal component (Fig 1C), while the GFP⁺ neural crest and GFP⁻ embryonic head samples separate along the second principal component (Fig S1C). In particular, the GFP⁺ samples formed a stereotypical arc-like pattern that demonstrated a clear trajectory of neural crest differentiation (Fig 1C). Next, we compared the transcriptomes of neural crest (GFP⁺) and whole embryo cells (GFP⁻) using a likelihood ratio test (LRT), to identify a neural crest-specific gene signature along any timepoint (Love et al., 2014). The resulting gene sets included bona fide markers such as *SOX10*, *FOXD3*, and *TFAP2B* as well as hundreds of factors not previously implicated in neural crest development (Fig 1D). Particularly, our analysis revealed stage-specific enrichment of several transcription factors including *SKOR2*, *MAFA*, *KAT6B*, *NFAT5* etc., which we further validated by *in situ* hybridization (Fig 1H, Fig S1D).

Next, to characterize the transcriptional changes associated with loss of neural crest multipotency, we performed pairwise differential gene expression analysis between neural plate border (HH6) and differentiating neural crest cells (HH16) using a negative binomial Wald test (Fig 1E). This analysis revealed that neural plate border cells have high expression of several canonical pluripotency genes like *OCT4* (*POU5F3* in chick), *SOX2*, *LIN28A*, and *KLF5*, which are lost in later developmental stages.

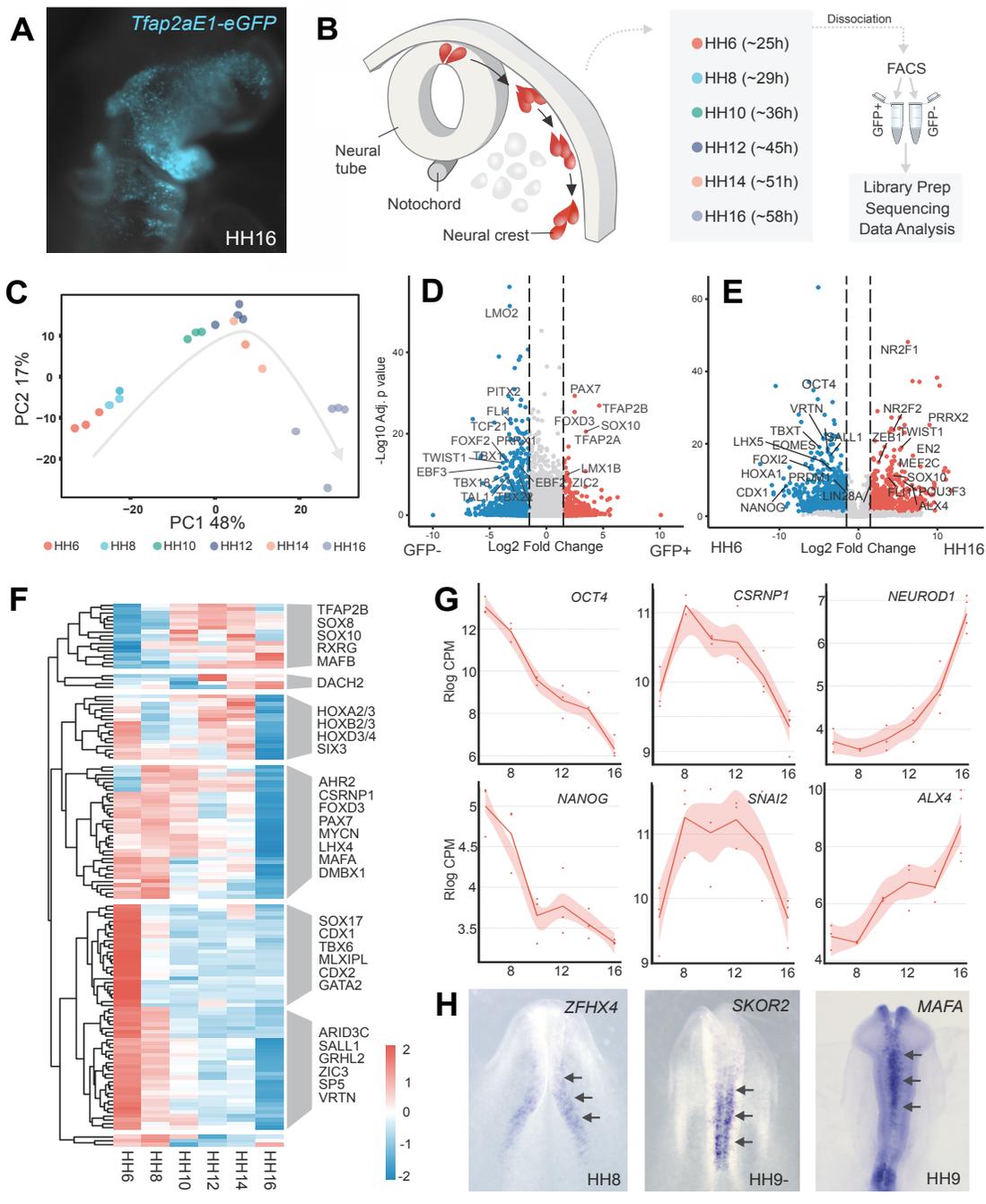


Figure1 – Transcriptional dynamics of cranial neural crest development

(A) *Ex ovo* electroporation of the neural crest-specific *Tfp2aE1* enhancer driving GFP faithfully labels neural crest derivatives in HH16 embryos. Expression is seen in the facial mesenchyme, the trigeminal ganglia, and branchial arches. **(B)** A schematic of the cross-section of the developing avian embryo head. Neural crest and whole embryo cells (highlighted in red) were isolated at six hamburger and Hamilton (HH) developmental stages by FACS sorting for GFP⁺ neural crest or GFP⁻ cells from the embryonic head. RNA-Seq libraries were prepared, sequenced, and analyzed in triplicate per condition/stage. **(C)** Principal Component Analysis (PCA) of GFP⁺ neural crest RNA-Seq

libraries. The majority (48%) of variance captured in the first principal component aligns with developmental time. **(D)** A volcano plot of the likelihood ratio test (LRT) comparing neural crest (GFP⁺) and embryonic head RNA-Seq samples. This test identifies genes enriched or depleted at any time during neural crest development. Significantly ($p\text{-adj}^* < 0.05$ & $\log_2\text{FC} > 1.5$) enriched (red) or depleted (blue) genes are colored, and most transcription factors with a $p\text{-adj}^*$ of $1e-8$ or lower are labeled. **(E)** A volcano plot of a pairwise negative binomial Wald test between the earliest (HH6) and latest (HH16) neural crest samples identifies genes significantly ($p\text{-adj}^* < 0.05$ & $\log_2\text{FC} > 1.5$) enriched (red) or depleted (blue) in differentiating neural crest. Most transcription factors with a $p\text{-adj}^*$ of $1e-8$ or lower are labeled. **(F)** Significantly enriched ($p\text{-adj}^* < 0.05$) transcription factors identified from ImpulseDE2 are hierarchically clustered into seven groups based on Rlog-normalized counts from DESeq2. Modules of the neural crest GRN (induction, specification, and migration) display sequential activation, and various transcription factors of interest are highlighted. **(G)** Stem cell (*Oct4*, *Nanog*), neural crest (*Csrnp1*, *Snai2*), and differentiation (*NeuroD1*, *Alx4*) markers expression visualization from our RNA-Seq Shiny app. Expression (y-axis) is measured by Rlog-normalized counts from DESeq2 across developmental stages (HH, x-axis). **(H)** Whole-mount *in-situ* hybridization for significantly enriched transcription factors identified through RNA-Seq analysis. Arrows indicate enrichment within developing neural crest. *Bonferroni-corrected p-value

We also observed increased transcription of lineage-specific genes such as *NEUROD1*, *ALX1*, *PRRX2*, and *NR2F1* in HH16 cells, corresponding to early stages of neural crest differentiation (Soldatov et al., 2019) (Fig 1E, G). We next examined the dynamics of these gene expression changes by performing clustering analysis with neural crest-enriched transcription factors identified by the LRT test (Fig 1F). This allowed us to group components of the GRN into distinct gene modules according to their temporal patterns of expression. The largest module contains 59 genes, including general pluripotency factors that are downregulated upon differentiation. This analysis also confirmed the sequential activation of early (*PAX7*, *CSRNP1*, and *TFAP2A*) and late (*TFAP2B*, *SOX10*, and *RXRG*) specifier genes (Fig 1F, G). Consistent with previous observations (Bhattacharya et al., 2018), we also observed that both pluripotency and neural crest specification genes undergo a concerted downregulation in late migrating and post-migratory neural crest cells (Fig 1F-G). Taken together, these analyses allowed for an unbiased delineation of temporal modules of the GRN and revealed the transcriptional shifts that characterize the cell state changes observed during neural crest development. These datasets have been compiled into an R Shiny app, which can be accessed at ash274.shinyapps.io/RNA-Seq_App/.

Figure 2

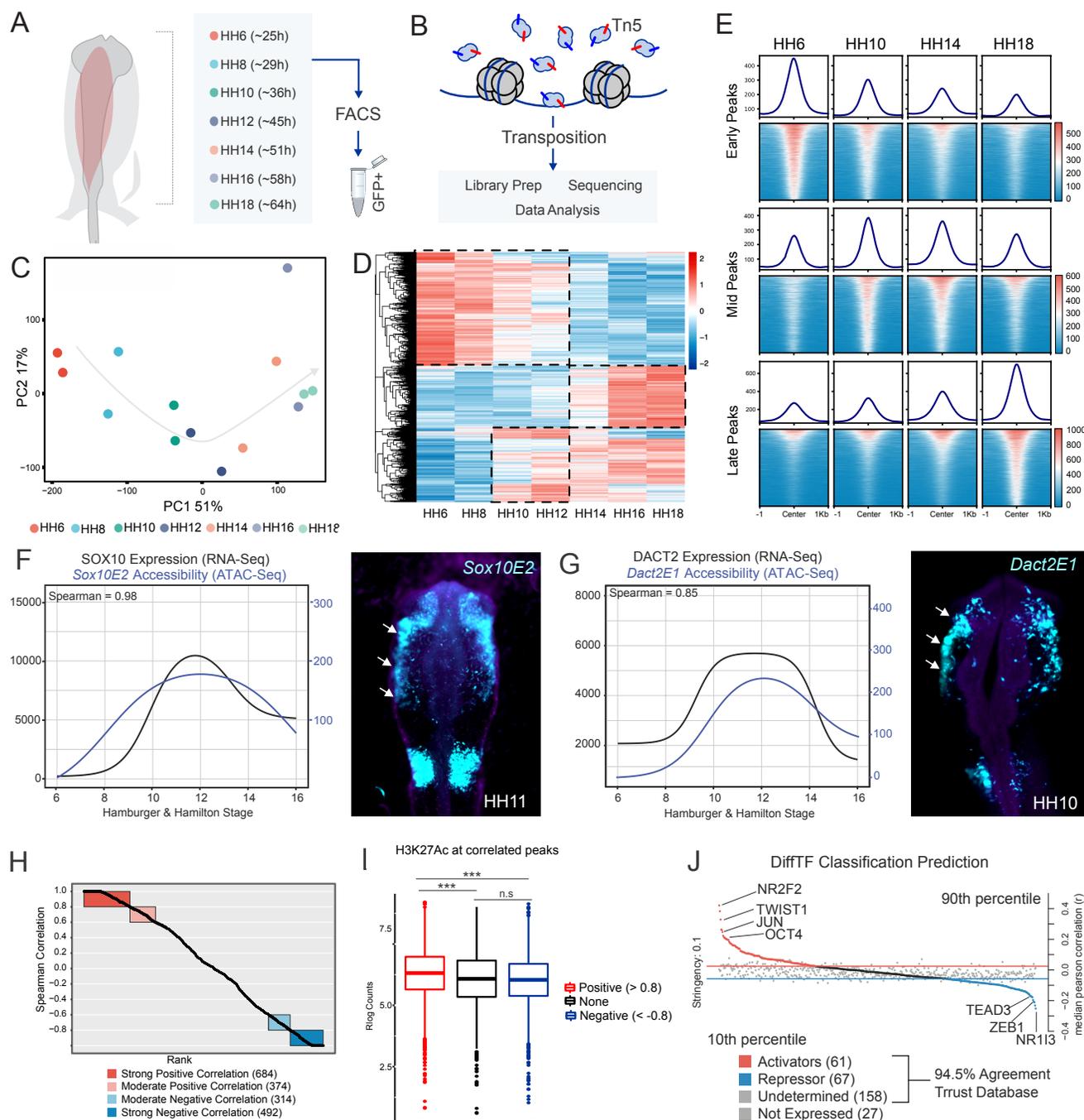


Figure 2 – Profiling the chromatin landscape of cranial neural crest cells with Omni ATAC-seq

(A) *Ex ovo* electroporation of *Tfap2aE1-GFP* was used to label developing neural crest cells. GFP⁺ neural crest cells were isolated at seven developmental stages (HH6-18) via FACS. (B) FACS isolated nuclei from each stage were subjected to Tn5-mediated transposition with Omni-ATAC-Seq and were processed for short-read sequencing. (C) Principal Component Analysis (PCA) of neural crest ATAC-Seq libraries. The first principal component, representing 51% of sample variation, separates samples according to developmental time. (D) Averaged, Rlog normalized counts for each timepoint at differential peaks (FDR < 0.05) from pairwise Wald tests comparing early (HH6 and HH8), mid (HH10-HH14) and late (HH16 and HH18) groups. Three large clusters of peaks (dashed lines) were grouped together into early, mid, and late peaksets. (E) Tornado plots showing the

genome-wide accessibility profile of each peakset (from D) at four different stages of neural crest development. **(F)** ImpulseDE2 models of the *Sox10E2* enhancer and the *Sox10* gene show a highly positive (0.98) Spearman correlation. Electroporation of this element driving a GFP reporter *in vivo* shows remarkable similarity to SOX10 expression. **(G)** ImpulseDE2 models of the *Dact2E1* enhancer and the *Dact2* gene also share a highly positive (0.85) Spearman correlation and similar *in vivo* reporter activity compared to DACT2 expression. **(H)** ImpulseDE2 models for both expression (RNA-Seq) and accessibility (ATAC-Seq) of 146 neural crest genes and their nearby genomic elements reveal 684 elements with a strong (>0.8) positive correlation. These putative regulatory regions may be used to expand the neural crest gene regulatory network. **(I)** H3K27Ac, an activating chromatin mark, is enriched in highly positively correlated elements compared to highly negative or a size and GC-adjusted random background. **(J)** DiffTF was used to investigate ATAC-Seq peaks for specific transcription factor motifs and correlate the changes in the accessibility levels of these peaks with the expression of transcription factors themselves. This classification revealed a number of strongly activating (motif presence associated with increased accessibility) factors including: NR2F2, TWIST1, and OCT4 along with repressing factors including TEAD3, ZEB1, and NR1H3. ***P≤0.001

To understand how neural crest developmental transitions are regulated at the chromatin level, we assayed the epigenomic landscape of neural crest cells at distinct developmental stages. We employed the Omni-ATAC protocol (Corces et al., 2017) to map chromatin accessibility in *Tfap2aE1* labelled neural crest cells isolated at seven embryonic timepoints (HH6-HH18) (Fig 2A-B). PCA analysis of ATAC-seq datasets showed that the samples segregated according to developmental stage in the first principal component, representing 51% of captured variance (Fig 2C). Pairwise differential peak enrichment analysis with DiffBind (Rory and Gord, 2011; Ross-Innes et al., 2012) between early (HH6-8), mid (HH10-14), and late (HH16-18) samples identified ~18,000 highly variable peaks, whose accessibility differed significantly through neural crest development ($p\text{-adj} < 0.05$, Fig. 2D). Clustering of these chromatin regions further allowed us to classify these peaks according to the temporal patterns of their ATAC signal (Fig 2D,E). Consistent with our RNA-seq analysis (Fig. 1F) we found a large number of ‘early’ peaks, which displayed high accessibility in multipotent neural crest (HH6-HH8) but were closed upon differentiation (Fig 2 D,E). Alternatively, ‘mid’ and ‘late’ peaks displayed increased accessibility at stages HH10-H12 and HH14-HH18, respectively (Fig 2D,E). These data highlight the extensive epigenomic changes that occur in the loci of many neural crest genes during early development.

To examine how these epigenomic changes were coupled with shifts in neural crest gene expression, we integrated our ATAC-seq and RNA-seq datasets using ImpulseDE2 (Fischer et al., 2018). By

considering nearby ATAC peaks for every gene, we correlated changes in chromatin accessibility to temporal patterns of gene expression. We observed that the accessibility of known neural crest enhancers, like *Sox10E2* and *Dact2E1*, displayed a strong temporal correlation with the transcriptional dynamics of their cognate genes (Fig 2F-G) (Betancur et al., 2010; Rothstein and Simoes-Costa, 2020). This prompted us to expand our analysis to identify putative enhancers of all significantly enriched neural crest genes as identified by RNA-Seq. We found 684 out of 3526 peaks tested were strongly positively correlated (Spearman > 0.8) with expression levels of 146 neural crest genes (Fig 2H). Testing four of these regions (previously characterized in (Azambuja, 2021)) in transgenic avian embryos confirmed that they display neural crest-specific activity similar to the associated gene (*Msx1E4.4*, *FoxD3E8.5*, *Sox8E14.1*, and *Sox9E18.1* average spearman = 0.895, Figure S2). Notably, we observed that ATAC-seq peaks with strong correlation scores displayed higher levels of H3K27Ac than regions with weak or negative correlations (Fig. 2I). Finally, we employed DiffTF to classify transcription factors as activating or repressing during neural crest differentiation (Berest et al., 2019). This tool incorporates the expression levels of transcription factors across time and correlates it with the accessibility of peaks that contain their binding motifs. This analysis revealed novel transcription factors that act as activators (OCT4, NR2F2, TWIST1 and JUN) or repressors (TEAD3, ZEB1 and NR1I3) in neural crest cells, underscoring how distinct sets of regulators mediate cell state transitions during development. Taken together, these datasets represent the most comprehensive transcriptomic and epigenetic profiling of early neural crest development within the same model system.

The OCT4-SOX2 dimer regulates the early neural crest epigenome

Our genomic analyses revealed that the transition of neural crest cells from a multipotent to differentiated state involves striking changes in the regulatory and transcriptomic landscape of the cells.

Figure 3

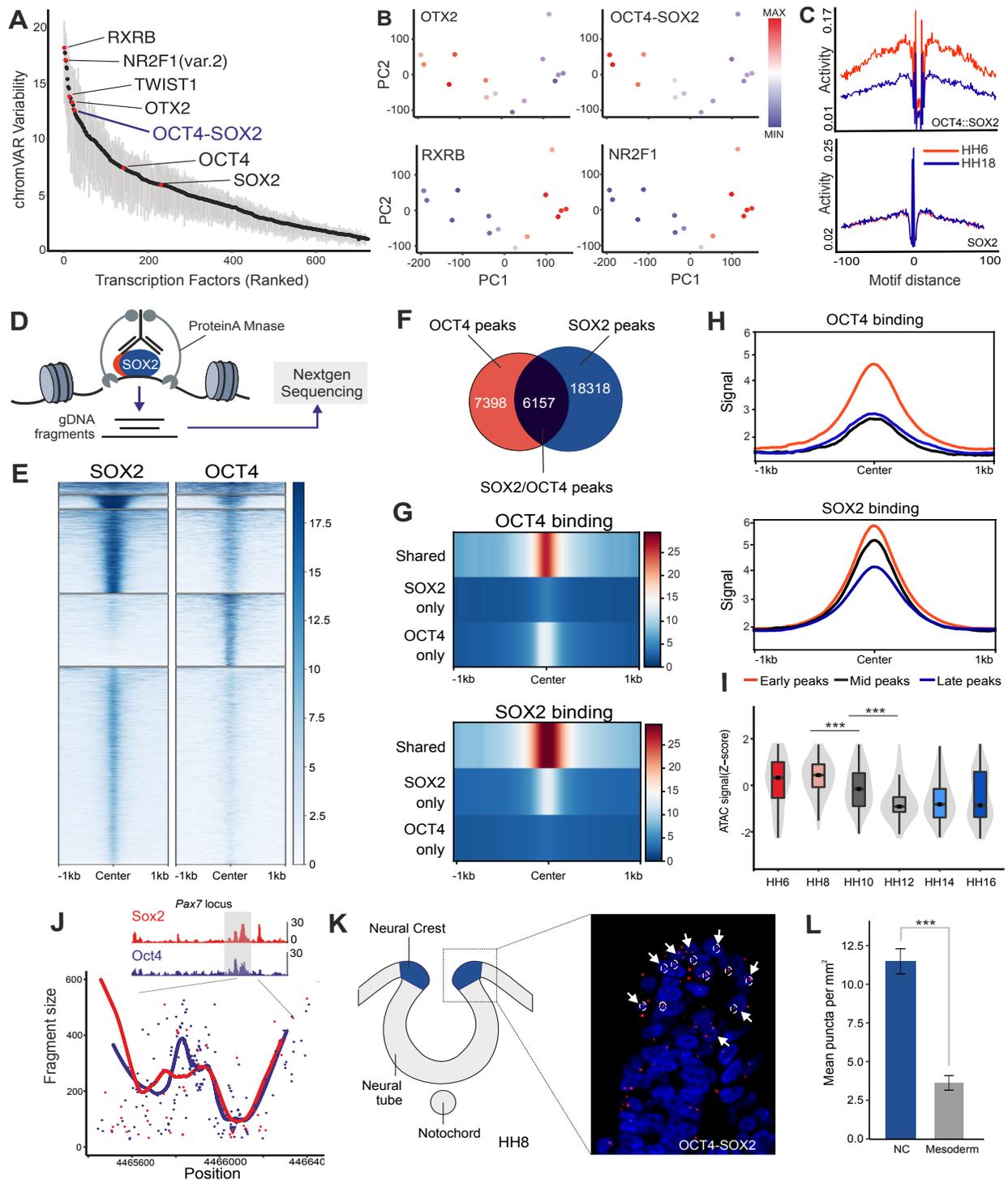


Figure 3 – OCT4 and SOX2 regulate the epigenetic landscape of early neural crest cells.

(A) A ranked-motif plot of neural crest ATAC-Seq motif enrichment via chromVAR. Some highly variable motifs across neural crest development include: RXRB, OTX2, and OCT4-SOX2. **(B)** Motif Z-scores from chromVAR for OTX2, OCT4-SOX2, RXRB, and NR2F1 projected onto the ATAC-Seq PCA of neural crest samples. **(C)** Individual footprinting plots for the OCT4-SOX2 and SOX2 motifs at the earliest (HH6) and latest (HH18) developmental

timepoints. These plots display a specific enrichment for the OCT4-SOX2 motif at HH6. **(D)** CUT&RUN profiles transcription factor binding genome-wide. An antibody for the transcription factor of interest is bound by a proteinA-MNase complex. Upon the addition of Ca_2^{++} , MNase cuts transcription factor-bound genomic DNA, which is then captured and processed for short-read sequencing. **(E)** Tornado plots depicting the genomic occupancy of SOX2 and OCT4 in neural crest cells at HH8. K-means clustering reveal 5 major groups of peaks that differ in relative OCT4 and SOX2 binding. **(F)** Pie chart showing the overlap in SOX2 and OCT4 CUT&RUN peaks in HH8 neural folds samples. **(G)** Compressed genomic occupancy profiles show that the strongest binding events occur at shared OCT4 and SOX2 peaks. **(H)** CUT&RUN signal enrichment centered at the early, mid and late ATAC-seq peaksets show that both OCT4 and SOX2 have highest binding at early ATAC peaks. **(I)** Normalized ATAC-Seq accessibility (Rlog counts Z-score) of shared OCT4 and SOX2 peaks over our developmental time course. **(J)** EChO analysis of SOX2 and OCT4 CUT&RUN show shared binding events of the pluripotency factors around the *Pax7* locus. **(K)** Schematic of the cross-section of HH8 avian embryo head with the position of neural crest cells depicted in blue. Proximity ligation between OCT4 and SOX2 appear as red dots and are highlighted with white arrows. Interactions located within the neural fold are marked in dotted white circles. **(L)** Quantification of PLA puncta show significantly more interactions/ mm^2 between OCT4 and SOX2 in neural crest compared to the surrounding mesenchyme. Significance was assessed via students t-test. * $P \leq 0.05$, ** $P \leq 0.01$, *** $P \leq 0.001$

To identify the key players that control this developmental transition, we utilized chromVAR to find transcription factor motifs that display high variability across developmental stages (Schep et al., 2017). The top ranked motifs included *RXR*B, *NR2F1*, *OTX2* and the motif for the OCT4-SOX2 dimer (Fig 3A). Consistent with their expression dynamics, the variability scores of *OTX2* and the OCT4-SOX2 dimer motifs were highest in early multipotent neural crest cells, while the scores for the *RXR*B and *NR2F1* motifs increased upon differentiation (Fig 3B). To validate the physical association of these transcription factors with open regions in neural crest, we next assessed their chromatin footprint using HINT-ATAC (Li et al., 2019). An unbiased comparison between our ATAC-seq data from the earliest (HH6) and latest (HH18) stages confirmed that one of the highest enriched footprints in early neural crest cells was that of the OCT4-SOX2 dimer (Fig S3A-C, Fig 3C). Further, while the footprint of OCT4 alone was also enriched in early neural crest, the SOX2 motif did not display a clear footprint in either early or late neural crest cells (Fig 3C). These results suggest that a dimer complex of OCT4 and SOX2 binds to DNA to promote neural crest multipotency.

Next, we wanted to determine how general pluripotency factors like OCT4 and SOX2 directly modulate the neural crest GRN. To address this, we mapped the genomic occupancy of both SOX2 and

OCT4 in neural crest cells dissected from HH8 embryos using Cleavage Under Targets & Release Using Nuclease (CUT&RUN) (Skene and Henikoff, 2017) (Fig 3D). Peak calling identified ~13,000 and ~24,000 peaks for OCT4 and SOX2 respectively, in neural crest cells, of which 6,157 were shared between the two factors (Fig 3E,F). Consistent with our previous analysis indicating the importance of OCT4-SOX2 dimerization, we found that the binding of both OCT4 and SOX2 was significantly higher at co-occupied regions than at those bound individually, suggesting a more stable chromatin association of these factors when bound together (Fig 3G). We also observed that OCT4 and SOX2 preferentially occupied the early neural crest peaks identified by our ATAC-seq analysis and displayed little binding in chromatin regions associated with differentiating cells (Fig3H). The shared OCT4-SOX2 CUT&RUN peaks were specifically open in the multipotent cells and lost accessibility over neural crest differentiation (Fig 3H-I). These results support the cooperative action of the two factors in early neural crest cells.

To confirm that SOX2 and OCT4 are indeed functioning as heterodimers in neural crest cells, we performed Enhanced Chromatin Occupancy (EChO) with our CUT&RUN datasets to determine transcription factor binding events at a base-pair resolution (Meers et al., 2019). This analysis identified over 12,000 “foci” within the CUT&RUN peaks that were simultaneously bound by OCT4 and SOX2 in neural crest cells. We observed co-binding events of these factors around neural plate border genes such as *Pax7*, *Sp5* and *Zic1*, consistent with the role of the heterodimer as a direct regulator of the early neural crest GRN (Fig 3J, Fig S3D-E). Lastly, we performed Proximity Ligation Assays (PLAs) to identify molecular interactions between endogenous OCT4 and SOX2 proteins (Alam, 2018). Comparison of HH8 neural folds and surrounding mesenchyme showed a significant enrichment of OCT4 and SOX2 interactions in early neural crest cells (Fig 3K, L). Taken together, these observations reveal that the OCT4-SOX2 heterodimer regulates the epigenomic signature of early neural crest cells which we hypothesize is consequential for multipotency.

Figure 4

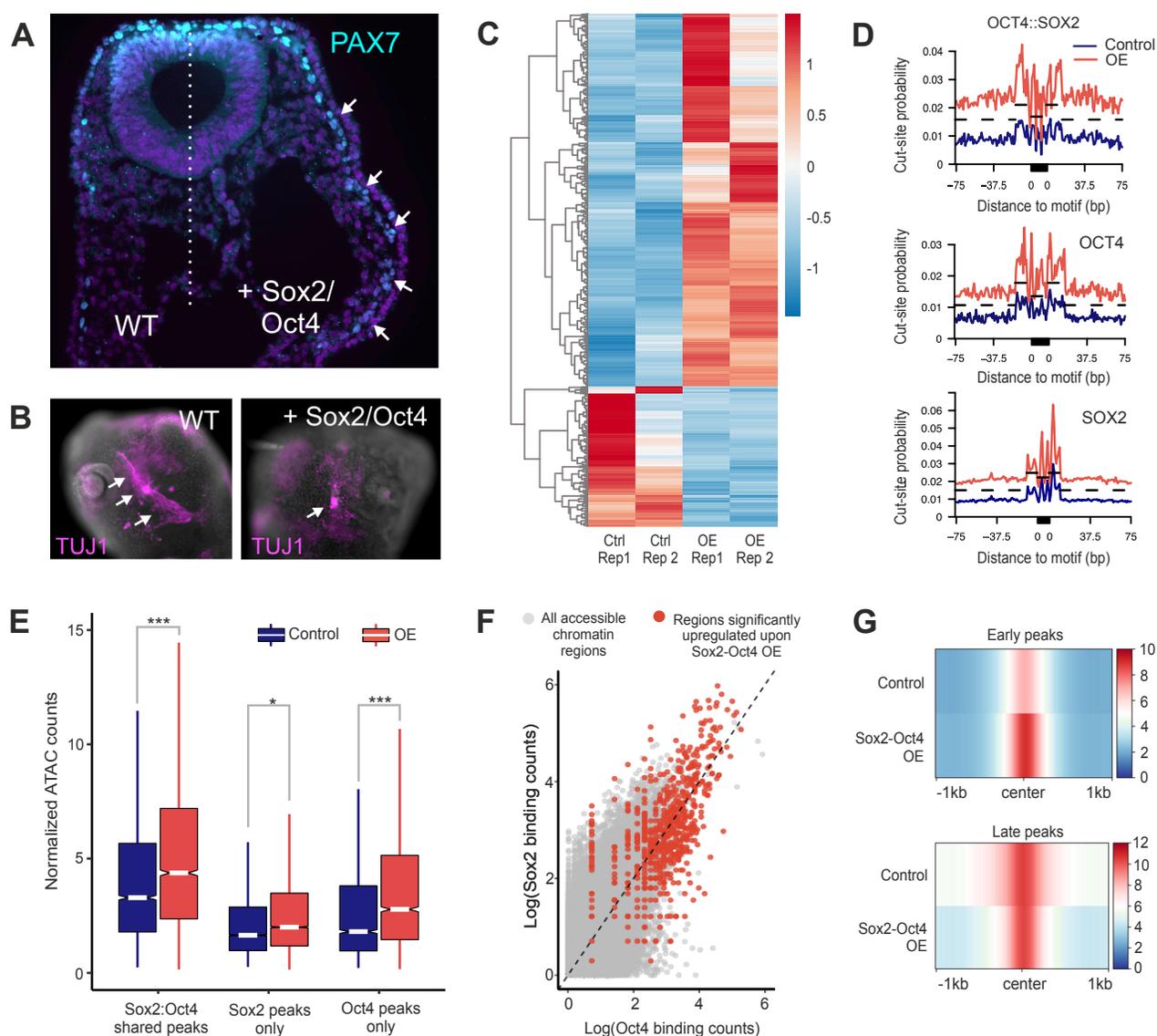


Figure 4 - Overexpression of SOX2 and OCT4 sustains multipotency in differentiating neural crest cells. (A) Chick embryos were bilaterally electroporated with control (left side) and OCT4-SOX2 overexpression constructs (right side) at HH4 and developed *ex ovo* until HH13. Staining of the embryo section with PAX7, a marker of multipotent neural crest, shows sustained expression on the protein in late migratory neural crest cells on the overexpression side of the embryo. (B) Immunostaining for TUJ1, a marker of neural-crest derived trigeminal ganglia, shows improper differentiation of sensory neurons on the SOX2-OCT4 overexpression side of the embryo head. (C) Heatmap showing ATAC signal at genomic regions whose accessibility was significantly altered in HH13 embryos (> 2 fold) upon OCT4-SOX2 overexpression. Row-normalized heatmap values are scaled from -1.5 to 1.5. (D) Differential footprinting analysis of OCT4-SOX2, OCT4, and SOX2 in control versus OE samples. The footprint of the dual motif is most elevated upon overexpression. (E) Boxplot showing change in accessibility (Rlog normalized counts) of different CUT&RUN peaks upon overexpression. (F) Density plot showing OCT4 and SOX2 binding at all ATAC peaks, with red dots being peaks that are significantly altered upon OE. (G) A compressed genome plot showing increased accessibility at the early ATAC at HH13 in overexpression samples. In contrast, chromatin accessibility at late ATAC peaks was not significantly different between control and OE samples. *P<0.05, **P<0.01, ***P<0.001

To establish the relevance of OCT4 and SOX2 in regulation of neural crest multipotency, we next performed a series of functional experiments manipulating the temporal expression pattern of these pluripotency factors in neural crest cells. To this end, we bilaterally electroporated gastrula-stage chicken embryos (HH4) with an expression vector encoding both SOX2 and OCT4 (*pCI-OCT4-T2A-SOX2-H2B-RFP*) to drive sustained expression of these proteins in late neural crest cells. We found that prolonged expression of SOX2 and OCT4 resulted in maintenance of high levels of early neural crest gene PAX7 in the late migratory neural crest (HH13), at stages when these genes would normally be downregulated (Fig 4A). This manipulation also prevented differentiation of neural crest cells as evidenced by gross morphological defects in the trigeminal ganglia in HH16 embryos. TUJ1 staining to label the cranial ganglia revealed morphological defects and showed a reduced number of sensory neurons on the right (overexpression) side of the embryo (Fig. 4B) compared to the control side which had a wild type ganglia (n = 6). These experiments indicate that OCT4 and SOX2 promote expression of early GRN components, and that their downregulation is required for proper differentiation of neural crest derivatives. We thus postulated that the dimer helps to initiate and maintain chromatin accessibility that is characteristic to the multipotent neural crest, which is lost as these cells become fate restricted.

To test this hypothesis, we next assessed whether sustained expression of OCT4 and SOX2 is sufficient to maintain the epigenomic landscape of multipotent cells in differentiating neural crest. Similar to the experimental design described above, we electroporated HH4 chick embryos with control and OCT4-SOX2 overexpression vectors (RFP) alongside the neural crest-specific *Tfap2aE1-GFP* reporter construct. We FACS sorted GFP⁺/RFP⁺ cells from each side of the embryonic head at HH13 to perform ATAC-Seq. Differential peak analysis revealed a large number of peaks (~700) whose accessibility significantly increased (> 2-fold) upon sustained expression of OCT4-SOX2 (Fig 4C). As expected, overexpression of OCT4 and SOX2 augmented chromatin footprinting scores for the OCT4-SOX2, OCT4 and SOX2 motifs, with the strongest effect seen for the dimer motif (Fig 4D). Assessing chromatin accessibility changes at the OCT4/SOX2 CUT&RUN peaks further confirmed that shared peaks co-occupied by the two factors at HH8 maintained high ATAC signal even at HH13, in the overexpression

samples (Fig 4E). Consistent with this, we observed that the largest change in accessibility (in overexpression samples) was observed for peaks that had the highest occupancy of OCT4 and SOX2 in multipotent neural crest cells (Fig 4F). This analysis also revealed that sustained expression of OCT4-SOX2 specifically increases chromatin accessibility at early neural crest peaks without significantly altering that of the late peaks (Fig 4G). Taken together, the data from these functional experiments show that maintenance of OCT4-SOX2 activity can prolong the multipotent state of neural crest cells both at the epigenomic and phenotypic level, thus establishing an essential role of these factors in regulation of stem cell identity of neural crest.

TFAP2A mediates translocation of OCT4-SOX2 during neural crest induction

Our genomic and functional analysis has demonstrated a central role for OCT4-SOX2 in the regulation of neural crest multipotency. Our data shows that the dimer promotes an epigenomic signature that is present in multipotent neural crest cells, but that is subsequently lost when SOX2 and OCT4 are downregulated during migration and differentiation. To define whether neural crest formation involves (i) the retention of a subset of the ESC pluripotency program or (ii) the establishment of a new, tissue-specific epigenomic signature, we employed an *in vitro* model of neural crest formation. We performed a five day induction protocol based on acute treatment of hESCs with the GSK3-inhibitor CHIR99021 (Gomez et al., 2019) (Fig 5A). We also used this protocol to differentiate these cells into smooth muscle following incubation in neutral MEM media supplemented with 10% FBS. Human induced neural crest cells (hiNCCs) display loss of NANOG, but expression of SOX10, OCT4, and SOX2 at D5 (Fig 5B). To examine accessibility dynamics along *in-vitro* neural crest induction, we employed ATAC-seq to examine chromatin accessibility in cells from day (D) zero (hESCs), D3 (early hiNCCs) D5 (hiNCCs), and D14 (Myo) of the induction protocol (Fig 5A). These datasets revealed that neural crest induction involves an extensive remodeling of the epigenomic landscape.

Figure 5

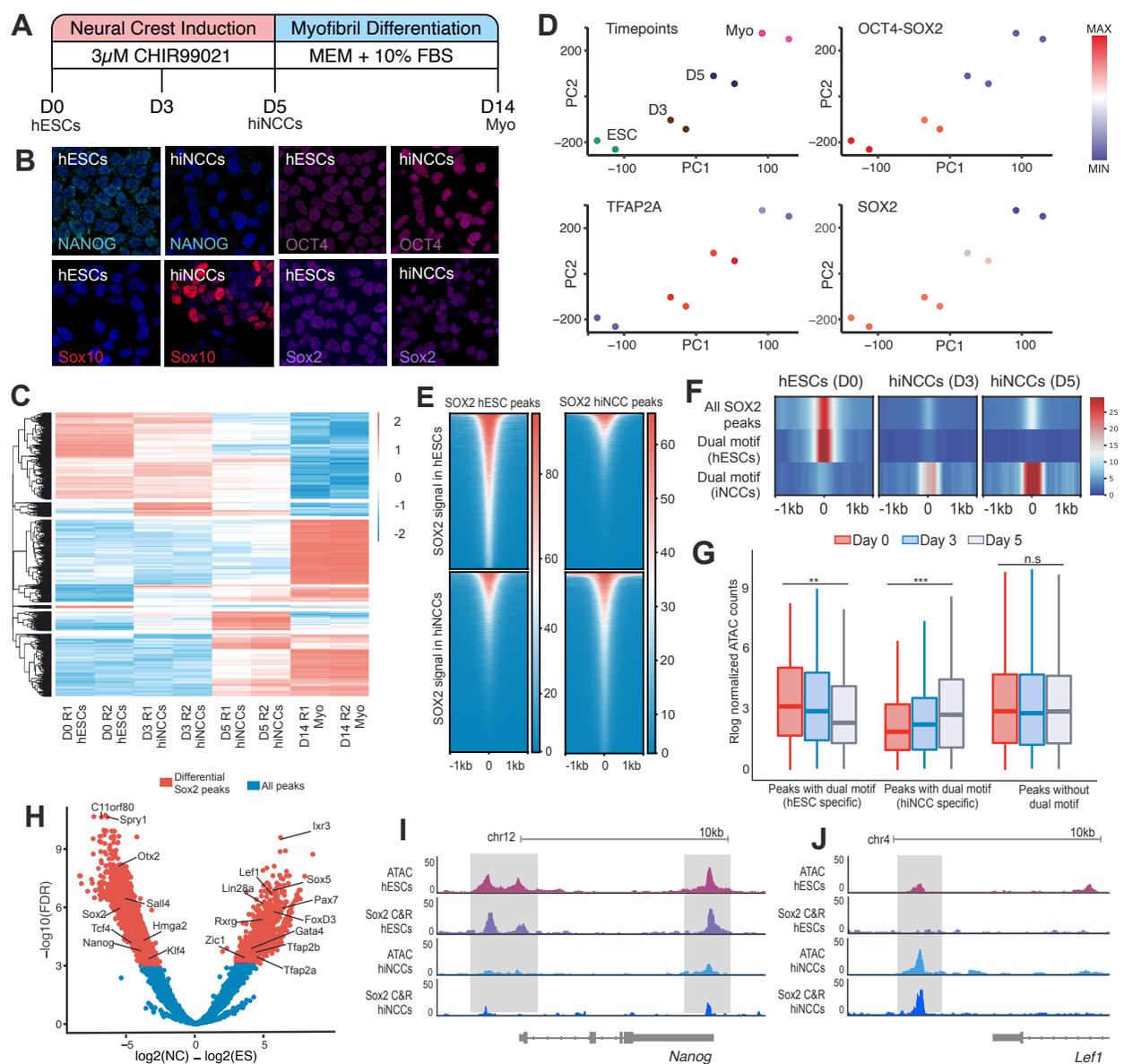


Figure 5 – SOX2 and OCT4 control distinct sets of regulatory regions in pluripotent ES cells and multipotent neural crest cells

(A) A schematic of the 5-day protocol used for obtaining induced human neural crest cells (hiNCCs) from human embryonic stem cells (hESCs). WA01 cells were treated for 3-day with 3µM CHIR99021 followed by two-day incubation in basal media to obtain neural crest cells. The subsequent differentiation of hiNCCs into myofibrils (Myo) was accomplished by growing the cells in MEM + 10% FBS media for a total of 9 days. **(B)** Immunohistochemistry for NANOG, OCT4, SOX10, and SOX2 in hESCs (D0) and hiNCCs (D5) show a loss in NANOG expression and gain in SOX10 upon successful neural crest induction. In contrast OCT4 and SOX2 levels were found to be comparable between D0 and D5 of induction. **(C)** A heatmap showing the differentially (FDR < 0.05, Pairwise Wald Test) accessible ATAC-seq peaks between D0, D3, D5, and D14 induced neural crest cells. **(D)** Principal component analysis dimensionality reduction shows PC1 represents progression from hESC to fully differentiated Myofibrils after fourteen days. Motif enrichment analysis via chromVAR shows

enrichment of OCT4-SOX2, TFAP2A, and SOX2 in the different ATAC samples. **(E)** A tornado plot showing CUT&RUN signal of SOX2 at genomic peaks bound by the factor at D0 hESCs and D5 hiNCCs. These plots show that SOX2 binds to distinct set of genomic regions in hESCs and hiNCCs. **(F)** A compressed genomic occupancy plot of SOX2 CUT&RUN signal at all SOX2 peaks, SOX2 peaks that are enriched in hESCs with an OCT4-SOX2 dual motif, and SOX2 peaks that are enriched in hiNCCs with an OCT4-SOX2 dual motif. These plots demonstrate a distinct and unique set of binding events in D5 hiNCCs. **(G)** Rlog normalized ATAC-Seq signal at sites containing hESC-specific dual motifs, iNCC-specific dual motifs, and peaks without a dual motif at D0, D3, and D5. **(H)** Differential analysis for SOX2 CUT&RUN peaks in D0 vs D5. Peaks displaying significantly altered SOX2 binding are colored in red. Nearby genes are annotated for some differential peaks, highlighting the distinct targets of SOX2 in hESCs and hiNCCs. **(I)** A genome browser view of ATAC-Seq and SOX2 CUT&RUN at the *Nanog* locus. Highlighted areas in grey show a loss of ATAC-Seq and SOX2 signal in hiNCCs. **(J)** A genome browser view of ATAC-Seq and SOX2 CUT&RUN at the *Lef1* locus. A highlighted area in grey shows a gain of ATAC-Seq and SOX2 signal in hiNCCs.

*P≤0.05, **P≤0.01, ***P≤0.001

We performed pairwise comparisons using DiffBind for each stage, revealing over 42,900 unique dynamic loci (adj-pval < 1e-6, Fig 5C). We were able to identify clusters of peaks that lose or gain accessibility during NC induction and differentiation. We also found a cluster transiently gains accessibility during neural crest induction that is lost upon differentiation. This dataset allowed us to investigate if the motifs of OCT4 and SOX2 display similar dynamics within human neural crest cells *in vitro* as we had observed in chick embryos *in vivo*. We next examined the function of OCT4-SOX2 in the remodeling of chromatin that takes place during induction. Analyzing motif enrichment via chromVAR, we found high scores for the OCT4-SOX2 motif in both ESCs and early hiNCCs (Fig 5D). The motif score for the neural crest-specific pioneer factor TFAP2A was drastically increased as early as D3 during hiNCC induction. We next wanted to determine whether OCT4-SOX2 (i) promotes maintenance of the embryonic stem cell signature, or (ii) promotes accessibility of neural crest-specific regulatory regions. We thus examined how the occupancy of OCT4-SOX2 changes as cells transition from hESC to hiNCCs during the induction process. To this end, we first performed CUT&RUN for SOX2 at D0 and D5 of neural crest induction. The results show a drastic shift in the patterns of SOX2 occupancy in hESCs and hiNCCs. The genomic regions bound by SOX2 in hESCs rapidly lost transcription factor binding during induction, while a distinct set of genomic regions gained SOX2 binding in Day5 hiNCCs (Fig 5E). Next, we used FIMO to scan for occurrences of the OCT4-SOX2 dual motif within all SOX2 peaks, SOX2 peaks enriched in

hESCs, and SOX2 peaks enriched in hiNCCs. When we investigated the SOX2 CUT&RUN signal at these locations, we observed this factor relocates to unique dual-motif locations in D5 hiNCCs (Fig 5F). Consistent with the well-documented role of SOX2 as a driver of chromatin accessibility, we also found that ATAC signal of dimer-associated regions changed according to SOX2 occupancy (Fig 5G). These data support the hypothesis that OCT4-SOX2 dimers translocate and bind to a new set of neural crest *cis*-regulatory regions and promote neural crest multipotency. We next used Diffbind to quantify SOX2 CUT&RUN data in hESCs and D5 hiNCCs and identify differentially bound regions (Fig 5H). Labeling of some of the most enriched and depleted regions demonstrated enrichment of SOX2 binding near many neural crest genes (*LEF1*, *LIN28A*, *FOXD3*, and *TFAP2A*) and loss of SOX2 binding near canonical pluripotency targets (*NANOG*, *SOX2*, and *KLF4*). Inspection of ATAC-Seq and SOX2 binding at the *NANOG* and *LEF1* loci demonstrated individual regulatory elements nearby with a loss or gain of signal across both assays, respectively. These data strongly implicate that both OCT4 and SOX2 play an important role in activation of the neural crest GRN.

Our results show that OCT4-SOX2 dimer promotes distinct epigenomic states in ESCs vs. neural crest cells. We next sought to identify molecular players that allow for the translocation of the dimer to a new set

Figure 6

Figure 6. A pioneer transcription factor that plays a central role in the formation of neural crest cells is TFAP2A (Li and Cornell, 2007; Mullins et al., 2019; Pla and Monsoro-Burq, 2018; Rothstein and Simoes-Costa, 2020). Analysis of previously reported TFAP2A CUT&RUN (Rothstein et al., 2018) performed in stage-matched neural crest cells revealed robust association of the pioneer factor in loci co-occupied by OCT4-SOX2 (Fig. 6A), and a strong correlation of binding levels in regulatory regions throughout the genome (Fig 6B). SOX2 and OCT4 binding was higher at peaks co-occupied by all three factors (Fig 6C). Furthermore, loci co-occupied by the three factors also displayed higher accessibility and association with super enhancer marker BRD4 (Fig 6D). This evidence for cooperativity led us to postulate that TFAP2A, which becomes active as OCT4-SOX2 changes genomic targets (Fig 5D), could be involved in the translocation of the dimer.

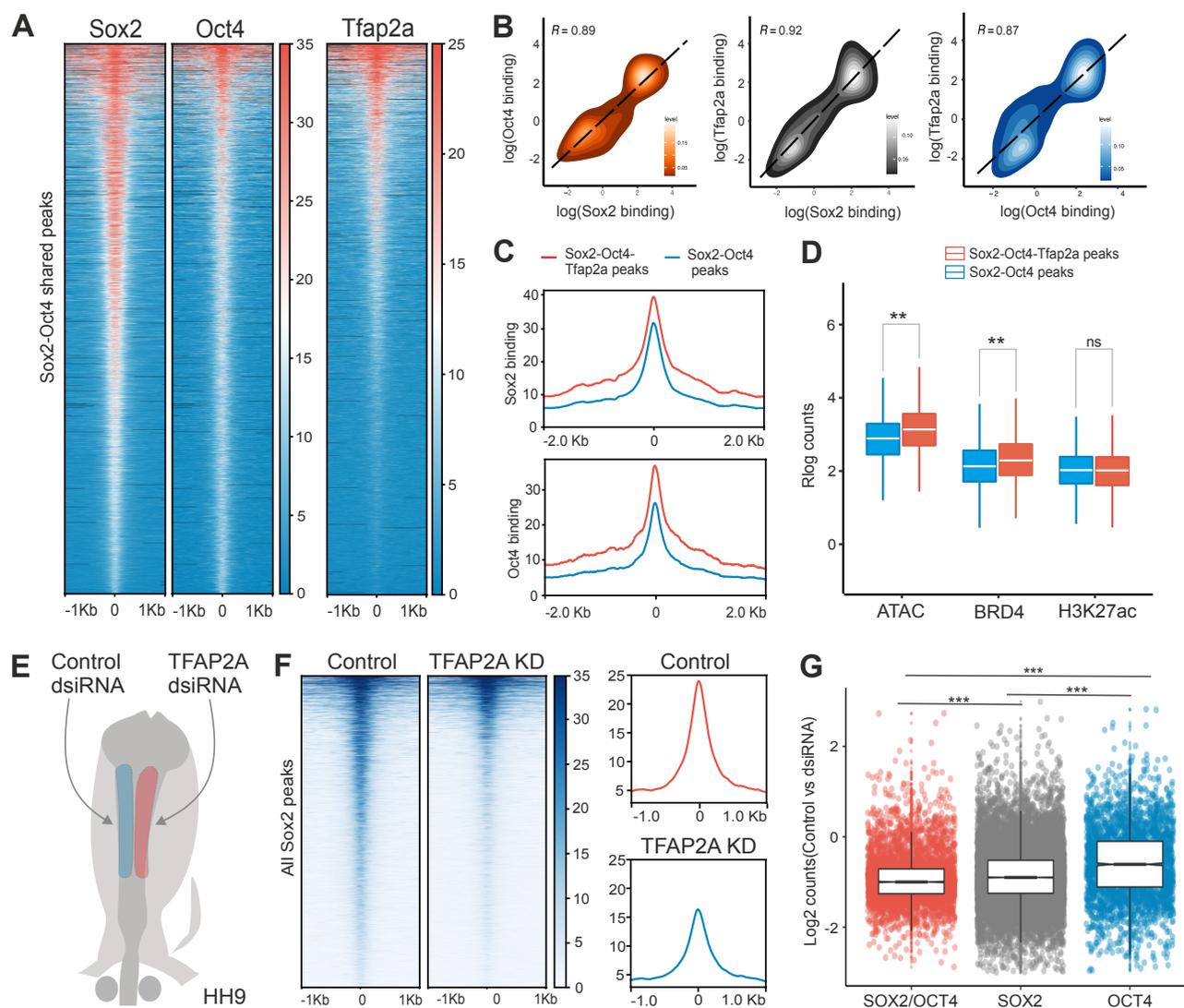


Figure 6 – The pioneer factor TFAP2A regulates tissue-specific function of SOX2 and OCT4 in neural crest cells.

(A) A tornado plot of SOX2-OCT4 shared CUT&RUN peaks and the binding of TFAP2A at these locations. The genomic regions that display strong OCT4-SOX2 signal are also occupied by TFAP2A. **(B)** Density plots showing pairwise correlations between all profiled transcription factors at shared peaks. **(C)** A genome pile-up plot showing peaks that are bound by all three factors have the highest levels of SOX2 and OCT4 binding. **(D)** Boxplot showing that ATAC signal and BRD4 binding is higher in peaks bound by all 3 factors than those only bound by Sox2 and Oct4. H3K27ac signal at these regions was not significantly different. **(E)** An electroporation scheme for a HH9 embryo in which control dsRNA reagents are added on the left side of the embryo and TFAP2A targeting dsRNA are added to the right. **(F)** Genomic tornado plots comparing SOX2 signal at all SOX2 bound peaks in control vs TFAP2A dsRNA treated NC cells. Summary plots (to the right) show a global decrease in SOX2 binding upon TFAP2A knockdown. **(G)** Boxplots showing the fold change in SOX2 binding in Control vs TFAP2A knockdown for Shared, SOX2 only, and OCT4 only peaks. SOX2 binding is most strongly reduced at shared peaks.

To test this, we examined how loss-of-function (LOF) of TFAP2A affected the genome-wide occupancy of SOX2 in neural crest cells. We transfected the right side of avian embryos with a dsiRNA targeting TFAP2A (Fig S4) and performed CUT&RUN in both control and TFAP2A-depleted neural folds (Fig 6E). We observed that loss of TFAP2A resulted in a significant reduction of SOX2 binding (Fig 6F). Unchanging binding events were broad and nonspecific background signal. Notably, peaks normally occupied by both OCT4-SOX2 were most affected by the loss of function (Fig 6G). This suggests that TFAP2A is required for OCT4-SOX2 to interact with and promote accessibility of the genomic regions that comprise the neural crest epigenomic signature.

Finally, we employed proximity ligation assays (PLAs) to examine the potential interactions between OCT4/SOX2 and their putative co-factors *in vivo*. We hypothesized that during neural crest induction, the dimer switches from interacting with its pluripotency partners (NANOG, KLF4, etc.) (Tapia et al., 2015) to cooperate with TFAP2A to promote a cell type specific epigenomic signature. Accordingly, we performed PLA for OCT4/SOX2 and either NANOG or TFAP2A in hESCs (D0) and hiNCCs (D5) (Fig 7A-B). Consistent with our hypothesis, we observed a striking decrease of interactions of both SOX2 and OCT4 with NANOG, indicating the dissolution of the pluripotency protein complex during neural crest induction (Fig 7C). Concomitantly, we found an increase in the interaction frequency between TFAP2A and both OCT4 and SOX2 in hiNCCs at D5 of induction, indicating these factors may act together to shape the neural crest epigenome. Taken together, these results support a model where TFAP2A directs the OCT4-SOX2 dimer to enhancers of many GRN components, promoting epigenomic landscape that is distinct from ESCs and underlies neural crest multipotency (Fig 7D).

Figure 7

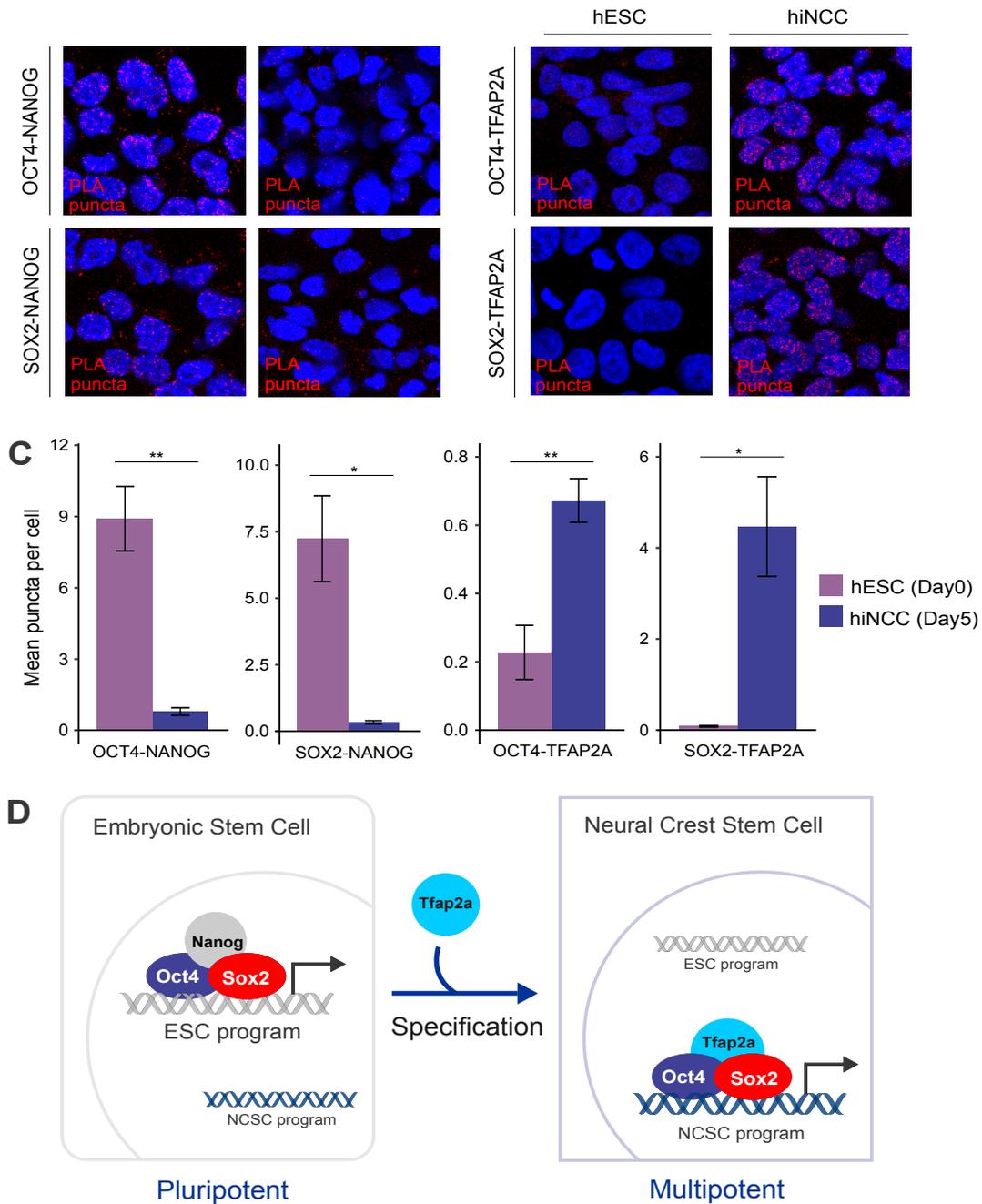


Figure 7 – OCT4-SOX2 have distinct interaction partners in pluripotent ES cells and multipotent neural crest cells

(A) Proximity ligation assays (PLAs) for OCT4-NANOG and SOX2-NANOG in hESCs and hiNCCs. **(B)** PLAs for OCT4-TFAP2A and SOX2-TFAP2A in hESC and hiNCCs. **(C)** Quantification of PLAs shows a decrease in OCT4-NANOG and SOX2-NANOG interactions in upon NCC induction and an increase in OCT4-TFAP2A and SOX2-TFAP2A interactions in hiNCCs. Significance was assessed via student's t-test. **(D)** A model for epigenomic repurposing of the pluripotency factors OCT4 and SOX2. In pluripotent ESCs, these factors interact with NANOG and promote an ESC program. Upon activation of the pioneer factor TFAP2A, these factors lose interactions with NANOG and are reprogrammed to a multipotent neural crest stem cell (NCSC) program. * $P \leq 0.05$, ** $P \leq 0.01$, *** $P \leq 0.001$

Discussion

In this study we present a time-resolved transcriptomic and epigenomic characterization of neural crest development. To examine cell state changes during neural crest development, we generated more than a hundred RNA-seq, ATAC-seq and CUT&RUN datasets from avian and human cells at multiple developmental stages. This analysis revealed an extensive reorganization of the epigenome during neural crest formation and differentiation, resulting in the identification of a chromatin state unique to multipotent cells. This epigenomic signature allowed us to expand our understanding of the molecular basis of neural crest multipotency. First, we found that the regulatory state of neural crest cells is distinct from that of ESCs, suggesting divergent mechanisms for developmental plasticity. Second, chromatin footprinting analysis and mapping of protein-DNA interactions revealed that the neural crest epigenome is regulated by the OCT4-SOX2 heterodimer, a central player in ESC pluripotency. Finally, we identified a mechanism whereby the neural crest pioneer factor TFAP2A modulates the target specificity of the dimer. These results point to a model for the control of neural crest plasticity where components of the ESC pluripotency machinery cooperate with tissue-specific factors to create a specialized multipotent state.

A defining property of developmental gene regulatory networks is that they are not fixed but instead change their topology as cells progress through the steps of fate commitment (Peter and Davidson, 2015). Our time-course experiments allowed us to delineate the temporal modules within the GRN and characterize the epigenomic changes that occur as cells transition through transcriptional states. Furthermore, by integrating epigenomic and transcriptomic datasets, we were able to identify new *cis* and *trans* components of the network. The generation of impulse models (Fischer et al., 2018) of gene expression and enhancer accessibility allowed us to curate a collection of active (H3K27Ac+) enhancers that likely play a role in the temporal regulation of cognate genes. This strategy may be a useful for enhancer-gene assignment from time-resolved datasets and can be used in the absence of physical

linkage (Lieberman-Aiden et al., 2009) or single-cell accessibility data (Pliner et al., 2018). Finally, we were also able to take advantage of both accessibility and transcriptomic data to assign an activating or repressing role for transcription factors expressed during neural crest development (Berest et al., 2019). These results are largely consistent with those of the Trrust database (Han et al., 2018), but transcription factors that are discordant may be of particular interest for a unique role within neural crest development.

This study also revealed an important role for two activating pioneer factors, *OCT4* and *SOX2*, in *neural crest multipotency*. Our results show that these pluripotency factors are expressed in early neural crest cells and that they exhibit context-specific functions. In the ectoderm, *SOX2* has traditionally been described as a pro-neural factor, which promotes neural plate identity via a mutual cross-repression circuit involving neural crest genes (Hovland et al., 2020; Wakamatsu et al., 2004). Yet, recent studies have demonstrated that neural crest cells express *SOX2* (Roellig et al., 2017), albeit at lower levels than neural plate progenitors. Here we show that, despite relatively low levels of expression, *SOX2* is still critical for neural crest multipotency. We propose that different functions of this pioneer factor observed in neural vs. neural crest progenitors is due to its dimerization to *OCT4*. Consistent with this, *OCT4* is required proper development of the cranial neural crest (Zalc et al., 2021). While previous studies have proposed that the pluripotency program is either retained (Buitrago-Delgado et al., 2015) or regained (Zalc et al., 2021) during neural crest development, our results indicate that Yamanaka factors also perform specialized functions in the neural crest that are distinct from their roles in ESCs.

Our model for the epigenomic control of neural crest multipotency (Fig 7D) explains how pluripotency factors are repurposed to generate a distinct chromatin landscape in specialized cells. We found that the tissue-specific pioneer transcription factors (e.g. *TFAP2A*) cooperate with the *OCT4-SOX2* heterodimer to generate a new epigenomic signature. This is yet another example of how almost every transcription factor acts in multiple developmental, disease, or tissue-specific contexts. From a mechanistic standpoint, the cooption of *OCT4-SOX2* from the pluripotency circuit makes sense, as these

factors have very strong pioneering activity (Michael et al., 2020; Whyte et al., 2013) and thus can promote overarching changes in the chromatin landscape. Another key difference between the neural crest multipotency circuit compared to the stem cell pluripotency circuit is the loss of NANOG. Without this factor, the self-promoting feedback loop of canonical embryonic pluripotency is broken (Akberdin et al., 2018; Niwa et al., 2000; Thomson et al., 2011). This rapid feedback may help explain why such a complex epigenetic rewiring is able to occur within the short time period between gastrulation and neural crest formation. In conclusion, our results demonstrate how modifications in the ESC program, both by the loss of bona fide regulators (NANOG) and by the gain of new, tissue-specific partners (TFAP2A), can generate specialized cellular states with different degrees of developmental plasticity.

Methods

Chick embryo collection and electroporation

Fertilized chicken embryos (Leghorn White) were obtained from the Department of Animal Science at the University of Connecticut. Eggs were incubated at 37°C until they reached the desired Hamburger and Hamilton (HH) (Hamburger and Hamilton, 1951) developmental stage. Embryos were collected and cultured *ex-ovo* following the EC method (Chapman et al., 2001) for HH6-16 and a Cornish Pastry (MC) method for HH16+ (Nagai et al., 2011). Enhancer, overexpression, and morpholino constructs were electroporated into HH4 embryos as previously described (Bhattacharya et al., 2018). Briefly, these constructs were injected between the epiblast and vitelline membrane of HH4 chicken embryos and electroporated using platinum electrodes (five 50ms pulses at 5.1V, 100ms resting interval).

Cloning and expression constructs

To label neural crest cells, we electroporated an enhancer of the TFAP2A gene (TFAP2a-E1) driving GFP (Rothstein and Simoes-Costa, 2020) and allowed embryos to develop *ex ovo* until the desired stage. To create a OCT4-SOX2 expression vector, we amplified the msSox2-P2A-msOct4 fragment from the

FUW-SOKM plasmid (Addgene #20325) with Gibson assembly overhangs to a H2B-RFP expression vector. We assembled the vector, leading to a CMV driven Sox2-P2A-Oct4-IRES-H2B-RFP ubiquitous expression vector with a fluorescent readout of electroporation efficiency.

Embryo dissociation and cell sorting

Neural crest cells were dissociated as previously described¹². Briefly, robustly GFP or RFP expressing embryos were dissected in Ringer's solution, dissociated using Accumax (Accutase SCR006), filtered, and resuspended in Hank's solution. Single-cell solutions were then sorted using a Sony MA900. Forward and side scatter were used to exclude debris and doublets, and another gate of either GFP (*TFAP2aE1-GFP*) or RFP (*OCT4-T2A-SOX2-H2B-RFP*) was used to isolate cells of interest.

RNA extraction, library preparation, and sequencing

For chicken neural crest RNA-Sequencing, we FACS isolated at least 5,000 cranial neural crest cells into the lysis buffer from the RNA Aqueous micro kit (ThermoFisher #AM1931). RNA was extracted according to the manufacturer's protocol, quantified using a Qubit RNA HS Assay (ThermoFisher #Q32852), and analyzed for quality on an ABI 3730xl DNA Analyzer. RNA-Sequencing libraries were prepared using the NEBNext® Ultra™ II Directional RNA Library Prep Kit (NEB #E7765) according to the manufacturer's protocol. Depending on the input RNA amount, libraries were PCR amplified 13-16 cycles. Libraries were quantified using a Qubit DNA HS Assay (ThermoFisher #Q33230) and checked for fragment size distribution and quality on an ABI 3730xl DNA Analyzer. Individual samples were pooled at an equimolar ratio calculated using the KAPA Library Quantification Kit (Roche #07960336001) and sequenced in a single-end configuration on an Illumina NextSeq500 using the High Output 75bp kit at the Cornell Institute of Biotechnology.

RNA-Seq Analysis

Raw sequencing reads were demultiplexed by the Cornell Institute of Biotechnology and trimmed using CutAdapt (v2.10) (Martin, 2011) with the TruSeq sequencing adapter (AGATCGGAAGAGCACACGTCTGAACTCCAGTCAC) and a minimum read length of at least 25. Reads were then aligned to the ENSEMBL galGal6 genome using HiSat2 (Kim et al., 2019) using reverse strandedness and discarding unaligned reads. Counts were assigned to genes using featureCounts (Liao et al., 2014) at ENSEMBL gene annotation (v93). Downstream differential expression analysis, PCA dimensionality reduction, clustering, and GO term enrichment analysis were performed in R. Details can be found in the scripts available under Data Availability and Code.

ATAC-Seq library preparation and sequencing

For chicken neural crest ATAC-Seq, at least 5,000 cells were isolated via FACS, gently centrifuged at 500rcf, and resuspended in CryoStor freezing media (Stem Cell Technologies #CS10), and frozen at -80°C until library preparation. For human ESC and iNCC ATAC-Seq, cells were dissociated using Accutase and then counted. One hundred thousand cells were resuspended in CryoStor media and frozen at -80°C until library preparation. Cells were processed using the OMNI-ATAC-Seq protocol (Corces et al., 2017). Frozen cells were quickly thawed at 37°C and ATAC-RSB buffer was added to a total volume of 1.5mL. Samples were centrifuged at 500rcf for 5 minutes at 4°C and the supernatant was removed. Next, cells were resuspended in 100µl of ATAC-RSB-LYSIS and kept on ice for 3-5 minutes, depending on input cell type. To stop lysis, 1mL of ATAC-RSB-WASH was added to each sample and they were again centrifuged for 5 minutes at 500rcf. The supernatant was removed, and cells were resuspended in 50µl of OMNI-ATAC Mix (~100nM concentration of Illumina TDE1 enzyme). Cells were then tagmented on a mixing (500rpm) thermoblock at 37°C for one hour. Tagmented DNA was recovered using a Qiagen MinElute Kit (#28204), with 21µl of elution buffer warmed to 55°C. Library amplification PCR was performed with the NEBNext Ultra II Q5 2X Master Mix (NEB #M0544S) using Nextera primers for 13-15 cycles. DNA concentration was measured using a Qubit DNA HS Assay (ThermoFisher

#Q33230) and stereotypical nucleosomal banding was observed using an ABI 3730xl DNA Analyzer. Individual samples were pooled at an equimolar ratio calculated using the KAPA Library Quantification Kit (Roche #07960336001) and sequenced on an Illumina NextSeq500 at the Cornell Institute of Biotechnology using the 75bp kit in a paired end configuration (32x32).

ATAC-Seq sequence processing

Raw sequencing reads were demultiplexed by the Cornell Institute of Biotechnology and trimmed using CutAdapt (v2.10) in paired-end mode with the forward and reverse Nextera sequencing adapters (Fwd – CTGTCTCTTATACACATCT, Rev- AGATGTGTATAAGAGACAG) and a minimum read length of at least 25. Next, we used Bowtie2 (Langmead and Salzberg, 2012) to align paired-end reads to the ENSEMBL galGal6 or UCSC hg38 genomes for chicken and human samples, respectively. We used the options “--local --very-sensitive-local --no-unal --no-mixed --no-discordant” and excluded alignments > 850bp. PicardTools (<https://github.com/broadinstitute/picard>) was used to mark duplicates, which were then filtered from the BAM files using samtools (Li et al., 2009). MACS2 was used to call peaks genome-wide, with a q value of 0.05 and the arguments “-f BAMPE -g \$GENOME_SIZE --nomodel --shift 37 --extsize 73” (Zhang et al., 2008). All downstream analysis including peak set generation, PCA, DiffBind, HOMER, chromVAR, and HINT-ATAC are available as scripts under Data Availability and Code.

CUT&RUN library preparation and sequencing

CUT&RUN was performed as previously described (Rothstein and Simoes-Costa, 2020). For chicken neural crest samples, 20 neural folds were dissected per sample. For human hESC and hiNCC samples, cultures were briefly dissociated using Accumex, counted, and 500,000 cells were used for each sample.

CUT&RUN sequence processing

Raw sequencing reads were demultiplexed by the Cornell Institute of Biotechnology and trimmed using CutAdapt (v2.10) in paired-end mode with the forward and reverse TruSeq sequencing adapters and a

minimum read length of at least 25. Next, chicken and human samples were aligned to ENSEMBL galGal6 or UCSC hg38, respectively, using Bowtie2. We used the following options “--local --very-sensitive-local --no-unal --no-mixed --no-discordant” and excluded alignments > 1000bp. PicardTools (<https://github.com/broadinstitute/picard>) was used to mark duplicates, which were then filtered from the BAM files using samtools. MACS2 was used to call peaks genome-wide, with a q value of 0.05 and the arguments “-f BAMPE -g \$GENOME_SIZE -q 0.05 --call-summits.” Subsequent analysis including peak filtering, motif enrichment, clustering, and comparison to ATAC-Seq can be found as R scripts within Data Availability and Code.

Immunohistochemistry

Whole-mount chicken embryos were isolated and fixed in PFA-PB for 10-20 minutes at room temperature. After dissection, embryos were washed in TBS containing 0.1X Triton X-100 and 1% DMSO (TBTD). Primary antibodies, diluted in 10% donkey serum blocking solution, were incubated at 4°C overnight, and secondary antibodies were incubated for 2 hours at room temperature.

Generation of induced neural crest (hiNCC)

The induction of human neural crest cells from WA01 (H1) embryonic stem cells was performed as previously described (Gomez et al., 2019). Briefly, H1 cells cultured in mTESR were dissociated with Accumex, and plated on matrigel coated dishes at a density of 20K cells/cm² in DMEM/F12 media supplemented with 1% B27, 0.5%BSA and 3uM CHIR (induction media). For the first 24h, 10uM ROCK inhibitor (Y-27632) was added to the induction media to improve cell survival. After 72 hours (Day 3), induction media was substituted for basal media constituting of DMEM/F12, 1% B27 and 0.5%BSA. The cells were grown in basal media for an additional 48 hours (Day 5), after which they were processed for downstream analysis.

Proximity Ligation Assays (PLA)

To identify close proximity between various combinations of transcription factors, we performed proximity ligation assays as previously described (Alam, 2018) using the DuoLink Fluorescence approach (Sigma-Aldrich, DUO92101) with minor modifications. For PLA in chick embryo sections, embryos were fixed with phosphate buffer (PB) containing 4% PFA for 20 minutes and cryosectioned in OCT compound. Primary antibody pairs used are as follows: goat anti-Sox2 (R&D Systems AF2018), rabbit anti-Oct4 (Invitrogen 701756). For PLA in human cells, we fixed cells at day 0 and day 5 of neural crest induction in 4% PFA for 10 minutes at room temperature. We then permeabilized the cells with 0.1% NP-40 in PBS for 30 minutes and 37°C, followed by blocking and primary antibody incubation in 1% BSA in PBS overnight at 4°C. Primary antibody pairs used are as follows: rabbit anti-Sox2 (Abcam ab97959), mouse anti-Nanog (Santa Cruz Biotechnology sc-293121); rabbit anti-Oct4 (Abcam ab109183), mouse anti-Nanog (Santa Cruz Biotechnology sc-293121); rabbit anti-Sox2 (Abcam ab97959), mouse anti-Tfap2a (DSHB 3B5); rabbit anti-Oct4 (Abcam ab109183), mouse anti-Tfap2a (DSHB 3B5)

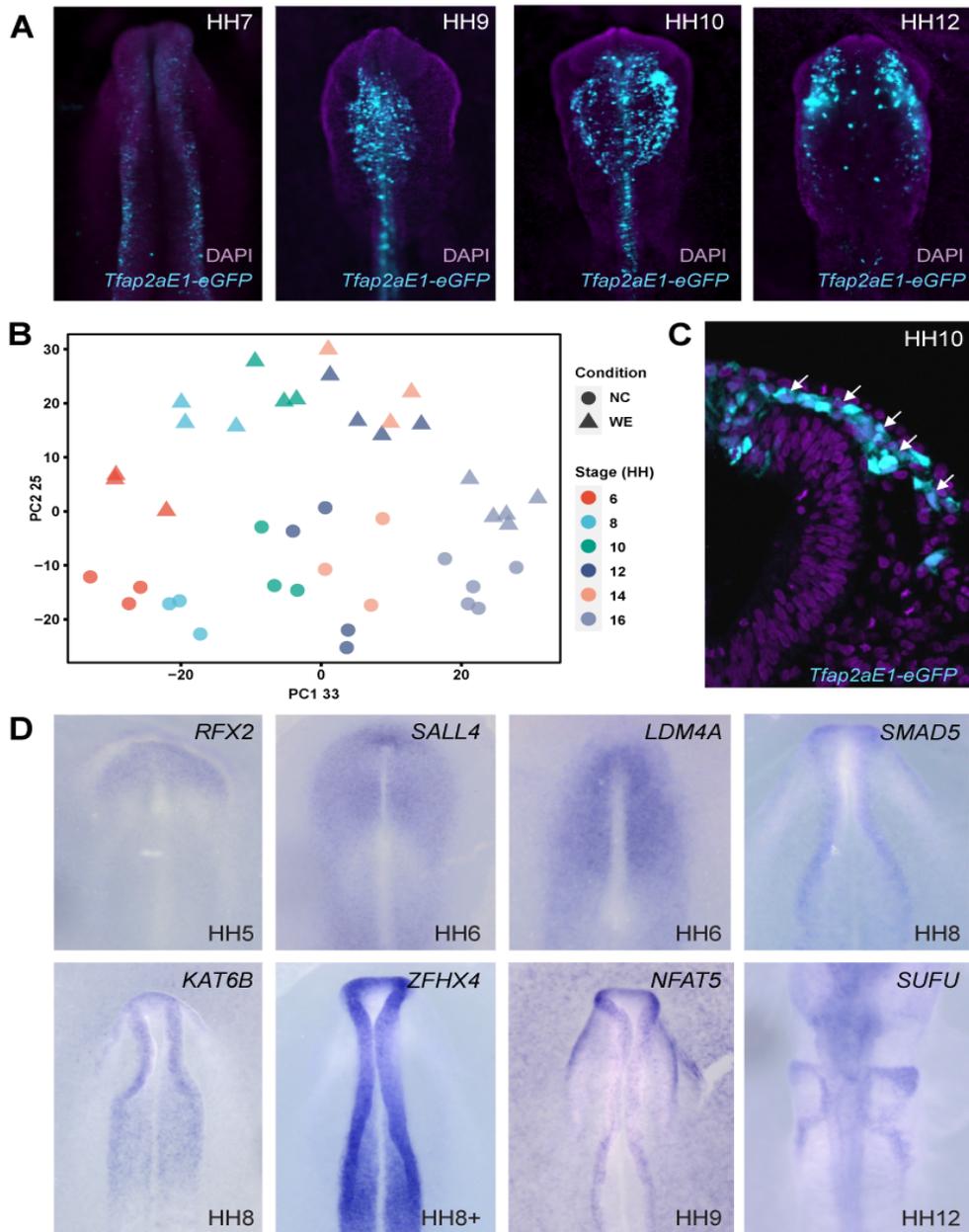
Data Availability and Code

Genomic data for this work has been submitted to the NCBI Gene Expression Omnibus (GEO) under the accession number GSE163961. Reviewers may ask for access if required.

All analysis scripts can be found at https://github.com/Simoes-Costa-Lab/NC_Timecourse.

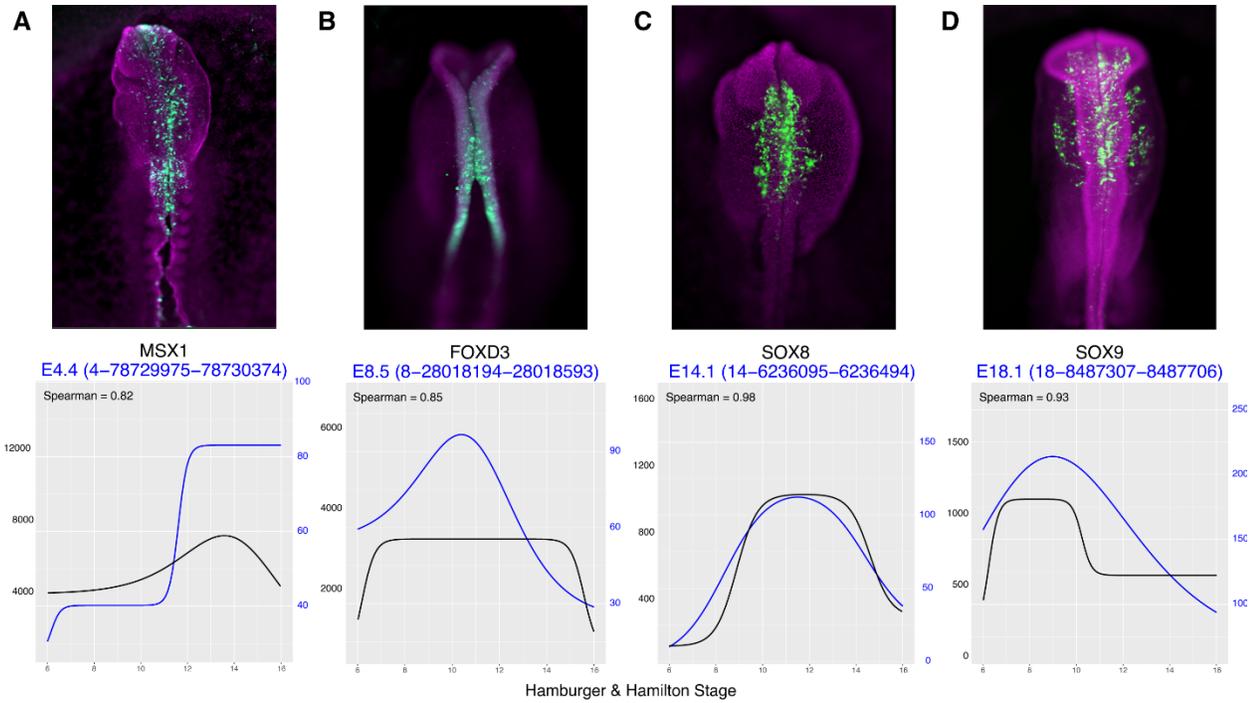
Publicly available samples were downloaded as FASTQs and processed identically to our other CUT&RUN datasets.

Figure S1



Supplemental Figure 1 (Supporting Figure 1).

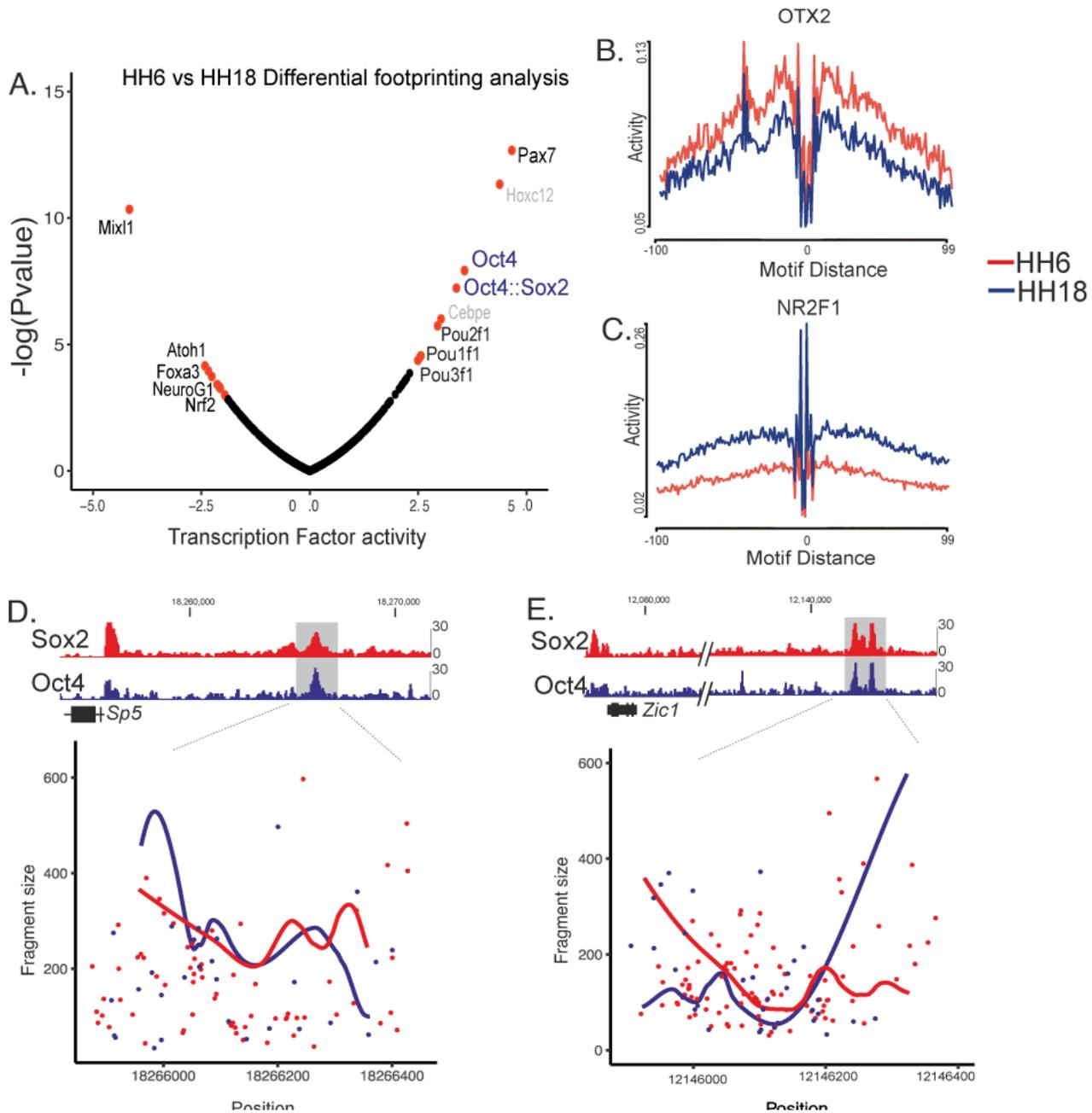
(A) Ex-ovo electroporation of *Tfp2aE1-eGFP* shows expression in the neural plate border at HH7, in the dorsal neural tube and migrating neural crest at HH9 and HH10 and in migratory crest at HH12. **(B)** A principal component analysis of both neural crest (NC) and whole embryo (WE) RNA-Seq samples. Principal component 1 (33% of variance) separates both conditions by developmental time, while principal component 2 (25% of variance) separates neural crest and whole embryo samples. **(C)** A cross-section of a HH10 chicken embryo, showing that the *Tfp2aE1-eGFP* enhancer is specific to neural crest and is not expressed in the ectoderm. **(D)** Whole mount *in-situ* hybridizations for newly discovered neural crest-enriched factors.



Supplemental Figure 2 (Supporting Figure 2)

(A-D) Transgenic embryos transfected with enhancer reporter constructs (Msx1E4.4-eGFP, FoxD3E8.5-eGFP, Sox8E14.1-eGFP and Sox9E18.1-eGFP) containing cis-regulatory elements that are active in neural rest cells. Below, ImpulseDE expression models show changes in accessibility of each element (black) correlate with mRNA expression of the cognate gene (black).

Figure S3



Supplemental Figure 3 (corresponds to Figure 3)

(A) Differential footprinting analysis between HH6 and HH18 ATAC-Seq samples shows Oct4 and Oct4::Sox2 as two of the most highly enriched motifs in HH6.

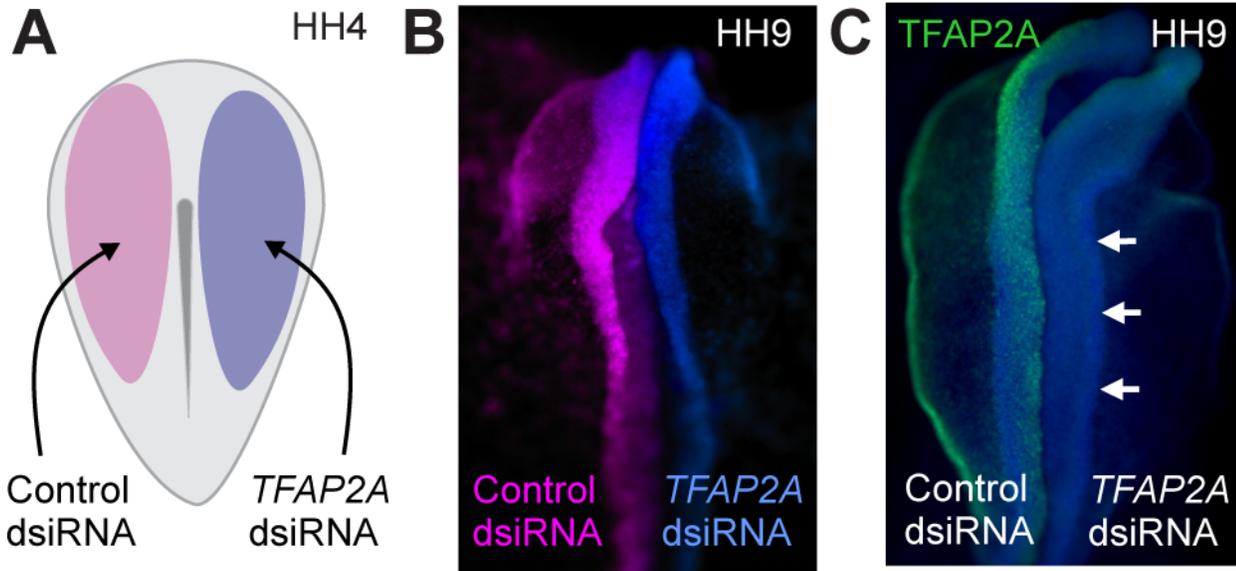
(B) A motif footprinting plot for Otx2, showing enrichment in HH6 ATAC-Seq samples.

(C) A motif footprinting plot for Nr2F1, showing enrichment in HH18 ATAC-Seq samples.

(D) A genome browser plot of a putative regulatory region for *Sp5*. EChO analysis of OCT4 and SOX2 CUT&RUN data shows co-binding events (local minima).

(E) A genome browser plot of a putative regulatory region for *Zic1*. EChO analysis of OCT4 and SOX2 CUT&RUN data shows co-binding events in one of the peaks.

Figure S4



Supplemental Figure 4 (Corresponding to Figure 6)

(A) A bilateral electroporation scheme in which a control dsiRNA reagent is introduced to the left side of the embryo and a *Tfap2a* dsiRNA reagent is introduced into the right side.

(B) A bilaterally electroporated embryo with the *Tfap2a* dsiRNA reagent showing right-side restriction of knockdown.

(C) A representative image of a *Tfap2a* dsiRNA electroporation. Immunohistochemistry for TFAP2A shows a striking decrease in protein expression.

References

- Akberdin, I.R., Omelyanchuk, N.A., Fadeev, S.I., Leskova, N.E., Oschepkova, E.A., Kazantsev, F.V., Matushkin, Y.G., Afonnikov, D.A., and Kolchanov, N.A. (2018). Pluripotency gene network dynamics: System views from parametric analysis. *PLOS ONE* *13*, e0194464.
- Alam, M.S. (2018). Proximity Ligation Assay (PLA). *Current Protocols in Immunology* *123*, e58.
- Azambuja, A.P.S.-C.M. (2021). The connectome of neural crest enhancers reveals regulatory features of signaling systems. *Developmental Cell*, in press.
- Baggiolini, A., Varum, S., Mateos, José M., Bettosini, D., John, N., Bonalli, M., Ziegler, U., Dimou, L., Clevers, H., Furrer, R., *et al.* (2015). Premigratory and Migratory Neural Crest Cells Are Multipotent In Vivo. *Cell Stem Cell* *16*, 314-322.
- Barembaum, M., and Bronner, M.E. (2013). Identification and dissection of a key enhancer mediating cranial neural crest specific expression of transcription factor, Ets-1. *Developmental biology*.
- Berest, I., Arnold, C., Reyes-Palomares, A., Palla, G., Rasmussen, K.D., Giles, H., Bruch, P.M., Huber, W., Dietrich, S., Helin, K., *et al.* (2019). Quantification of Differential Transcription Factor Activity and Multiomics-Based Classification into Activators and Repressors: diffTF. *Cell Rep* *29*, 3147-3159 e3112.
- Betancur, P., Bronner-Fraser, M., and Sauka-Spengler, T. (2010). Genomic code for Sox10 activation reveals a key regulatory enhancer for cranial neural crest. *Proc Natl Acad Sci U S A* *107*, 3570-3575.
- Bhattacharya, D., Rothstein, M., Azambuja, A.P., and Simoes-Costa, M. (2018). Control of neural crest multipotency by Wnt signaling and the Lin28/let-7 axis. *Elife* *7*.
- Bronner-Fraser, M., and Fraser, S.E. (1988). Cell lineage analysis reveals multipotency of some avian neural crest cells. *Nature* *335*, 161-164.
- Buitrago-Delgado, E., Nordin, K., Rao, A., Geary, L., and LaBonne, C. (2015). NEURODEVELOPMENT. Shared regulatory programs suggest retention of blastula-stage potential in neural crest cells. *Science* *348*, 1332-1335.
- Chapman, S.C., Collignon, J.r.m., Schoenwolf, G.C., and Lumsden, A. (2001). Improved method for chick whole-embryo culture using a filter paper carrier. *Developmental Dynamics* *220*, 284-289.
- Corces, M.R., Trevino, A.E., Hamilton, E.G., Greenside, P.G., Sinnott-Armstrong, N.A., Vesuna, S., Satpathy, A.T., Rubin, A.J., Montine, K.S., Wu, B., *et al.* (2017). An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. *Nat Methods* *14*, 959-962.
- Dottori, M., Gross, M.K., Labosky, P., and Goulding, M. (2001). The winged-helix transcription factor Foxd3 suppresses interneuron differentiation and promotes neural crest cell fate. *Development* *128*, 4127-4138.
- Dupin, E., Calloni, G.W., Coelho-Aguiar, J.M., and Le Douarin, N.M. (2018). The issue of the multipotency of the neural crest cells. *Dev Biol*.
- Fischer, D.S., Theis, F.J., and Yosef, N. (2018). Impulse model-based differential expression analysis of time course sequencing data. *Nucleic Acids Research*.
- Gomez, G.A., Prasad, M.S., Wong, M., Charney, R.M., Shelar, P.B., Sandhu, N., Hackland, J.O.S., Hernandez, J.C., Leung, A.W., and García-Castro, M.I. (2019). WNT/ β -catenin modulates the axial identity of embryonic stem cell-derived human neural crest. *Development* *146*, dev175604.
- Hamburger, V., and Hamilton, H.L. (1951). A series of normal stages in the development of the chick embryo. *J Morphol* *88*, 49-92.
- Honoré, S.M., Aybar, M.J., and Mayor, R. (2003). Sox10 is required for the early development of the prospective neural crest in *Xenopus* embryos. *Developmental Biology* *260*, 79-96.
- Hovland, A.S., Rothstein, M., and Simoes-Costa, M. (2020). Network architecture and regulatory logic in neural crest development. *Wiley Interdiscip Rev Syst Biol Med* *12*, e1468.
- Kim, D., Paggi, J.M., Park, C., Bennett, C., and Salzberg, S.L. (2019). Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat Biotechnol* *37*, 907-915.

- Kondoh, H., Simões-Costa, M.S., McKeown, S.J., Tan-Cabugao, J., Sauka-Spengler, T., and Bronner, M.E. (2012). Dynamic and Differential Regulation of Stem Cell Factor FoxD3 in the Neural Crest Is Encrypted in the Genome. *PLoS Genetics* 8, e1003142.
- Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9, 357-359.
- Le Douarin, N. (1973). A biological cell labeling technique and its use in experimental embryology. *Developmental Biology* 30, 217-222.
- Le Douarin, N.M., and Kalcheim, C. (1999). *The Neural Crest*, 2nd edn (The Pitt Building, Trumpington Street, Cambridge CB2 1RP, United Kingdom: Cambridge University Press).
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and Genome Project Data Processing, S. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25, 2078-2079.
- Li, W., and Cornell, R.A. (2007). Redundant activities of Tfap2a and Tfap2c are required for neural crest induction and development of other non-neural ectoderm derivatives in zebrafish embryos. *Dev Biol* 304, 338-354.
- Li, Z., Schulz, M.H., Look, T., Begemann, M., Zenke, M., and Costa, I.G. (2019). Identification of transcription factor binding sites using ATAC-seq. *Genome Biology* 20.
- Liao, Y., Smyth, G.K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* 30, 923-930.
- Lieberman-Aiden, E., Van Berkum, N.L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B.R., Sabo, P.J., Dorschner, M.O., *et al.* (2009). Comprehensive Mapping of Long-Range Interactions Reveals Folding Principles of the Human Genome. *Science* 326, 289-293.
- Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 15, 550.
- Lukoseviciute, M., Gavriouchkina, D., Williams, R.M., Hochgreb-Hagele, T., Senanayake, U., Chong-Morrison, V., Thongjuea, S., Repapi, E., Mead, A., and Sauka-Spengler, T. (2018). From Pioneer to Repressor: Bimodal foxd3 Activity Dynamically Remodels Neural Crest Regulatory Landscape In Vivo. *Developmental Cell* 47, 608-628.e606.
- Martik, M.L., and Bronner, M.E. (2017). Regulatory Logic Underlying Diversification of the Neural Crest. *Trends in Genetics* 33, 715-727.
- Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnetjournal* 17, 10.
- Meers, M.P., Janssens, D.H., and Henikoff, S. (2019). Pioneer Factor-Nucleosome Binding Events during Differentiation Are Motif Encoded. *Mol Cell* 75, 562-575 e565.
- Michael, A.K., Grand, R.S., Isbel, L., Cavadini, S., Kozicka, Z., Kempf, G., Bunker, R.D., Schenk, A.D., Graff-Meyer, A., Pathare, G.R., *et al.* (2020). Mechanisms of OCT4-SOX2 motif readout on nucleosomes. *Science* 368, 1460-1465.
- Mullins, M.C., Dooley, C.M., Wali, N., Sealy, I.M., White, R.J., Stemple, D.L., Collins, J.E., and Busch-Nentwich, E.M. (2019). The gene regulatory basis of genetic compensation during neural crest induction. *PLOS Genetics* 15, e1008213.
- Nagai, H., Lin, M.C., and Sheng, G. (2011). A modified cornish pasty method for ex ovo culture of the chick embryo. *Genesis* 49, 46-52.
- Niwa, H., Miyazaki, J.-I., and Smith, A.G. (2000). Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. *Nature Genetics* 24, 372-376.
- Peter, I.S., and Davidson, E.H. (2015). *Genomic control process: development and evolution* (32 Jamestown Road, London NW1 7BY, UK: Academic Press).
- Pla, P., and Monsoro-Burq, A.H. (2018). The neural border: Induction, specification and maturation of the territory that generates neural crest cells. *Dev Biol* 444 Suppl 1, S36-S46.
- Pliner, H.A., Packer, J.S., McFaline-Figueroa, J.L., Cusanovich, D.A., Daza, R.M., Aghamirzaie, D., Srivatsan, S., Qiu, X., Jackson, D., Minkina, A., *et al.* (2018). Cicero Predicts cis-Regulatory DNA Interactions from Single-Cell Chromatin Accessibility Data. *Mol Cell* 71, 858-871 e858.

- Prasad, M.S., Charney, R.M., Patel, L.J., and Garcia-Castro, M.I. (2020). Distinct molecular profile and restricted stem cell potential defines the prospective human cranial neural crest from embryonic stem cell state. *Stem Cell Res* 49, 102086.
- Roellig, D., Tan-Cabugao, J., Esaian, S., and Bronner, M.E. (2017). Dynamic transcriptional signature and cell fate analysis reveals plasticity of individual neural plate border cells. *Elife* 6.
- Rory, S., and Gord, B. (2011). DiffBind: differential binding analysis of ChIP-Seq peak data.
- Ross-Innes, C.S., Stark, R., Teschendorff, A.E., Holmes, K.A., Ali, H.R., Dunning, M.J., Brown, G.D., Gojis, O., Ellis, I.O., Green, A.R., *et al.* (2012). Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. *Nature* 481, 389-393.
- Rothstein, M., Bhattacharya, D., and Simoes-Costa, M. (2018). The molecular basis of neural crest axial identity. *Developmental biology* 444 Suppl 1, S170-S180.
- Rothstein, M., and Simoes-Costa, M. (2020). Heterodimerization of TFAP2 pioneer factors drives epigenomic remodeling during neural crest specification. *Genome Res* 30, 35-48.
- Schep, A.N., Wu, B., Buenrostro, J.D., and Greenleaf, W.J. (2017). chromVAR: inferring transcription-factor-associated accessibility from single-cell epigenomic data. *Nat Methods* 14, 975-978.
- Simoes-Costa, M., and Bronner, M.E. (2013). Insights into neural crest development and evolution from genomic analysis. *Genome Res* 23, 1069-1080.
- Simoes-Costa, M., and Bronner, M.E. (2015). Establishing neural crest identity: a gene regulatory recipe. *Development* 142, 242-257.
- Simoes-Costa, M.S., McKeown, S.J., Tan-Cabugao, J., Sauka-Spengler, T., and Bronner, M.E. (2012). Dynamic and differential regulation of stem cell factor FoxD3 in the neural crest is Encrypted in the genome. *PLoS Genet* 8, e1003142.
- Skene, P.J., and Henikoff, S. (2017). An efficient targeted nuclease strategy for high-resolution mapping of DNA binding sites. *eLife* 6, e21856.
- Soldatov, R., Kaucka, M., Kastriti, M.E., Petersen, J., Chontorotzea, T., Englmaier, L., Akkuratova, N., Yang, Y., Haring, M., Dyachuk, V., *et al.* (2019). Spatiotemporal structure of cell fate decisions in murine neural crest. *Science* 364, 1-67.
- Soufi, A., Donahue, G., and Kenneth (2012). Facilitators and Impediments of the Pluripotency Reprogramming Factors' Initial Engagement with the Genome. *Cell* 151, 994-1004.
- Takahashi, K., and Yamanaka, S. (2006). Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. *Cell* 126, 663-676.
- Tapia, N., MacCarthy, C., Esch, D., Gabriele Marthaler, A., Tiemann, U., Arauzo-Bravo, M.J., Jauch, R., Cojocar, V., and Scholer, H.R. (2015). Dissecting the role of distinct OCT4-SOX2 heterodimer configurations in pluripotency. *Sci Rep* 5, 13533.
- Thomson, M., Siyuan, Zou, L.-N., Smith, Z., Meissner, A., and Ramanathan, S. (2011). Pluripotency Factors in Embryonic Stem Cells Regulate Differentiation into Germ Layers. *Cell* 145, 875-889.
- Wakamatsu, Y., Endo, Y., Osumi, N., and Weston, J.A. (2004). Multiple roles of Sox2, an HMG-box transcription factor in avian neural crest development. *Dev Dyn* 229, 74-86.
- Whyte, W.A., Orlando, D.A., Hnisz, D., Abraham, B.J., Lin, C.Y., Kagey, M.H., Rahl, P.B., Lee, T.I., and Young, R.A. (2013). Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell* 153, 307-319.
- Williams, R.M., Candido-Ferreira, I., Repapi, E., Gavriouchkina, D., Senanayake, U., Ling, I.T.C., Telenius, J., Taylor, S., Hughes, J., and Sauka-Spengler, T. (2019). Reconstruction of the Global Neural Crest Gene Regulatory Network In Vivo. *Dev Cell* 51, 255-276 e257.
- Zalc, A., Sinha, R., Gulati, G.S., Wesche, D.J., Daszczuk, P., Swigut, T., Weissman, I.L., and Wysocka, J. (2021). Reactivation of the pluripotency program precedes formation of the cranial neural crest. *Science* 371.
- Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nussbaum, C., Myers, R.M., Brown, M., Li, W., *et al.* (2008). Model-based Analysis of ChIP-Seq (MACS). *Genome Biology* 9, R137.

Chapter IV: Metabolic reprogramming promotes neural crest migration via Yap/Tead signaling

Alternate Citation³:

D Bhattacharya, AP Azambuja and M Simoes-Costa (2020) Metabolic Reprogramming promotes neural crest migration via Yap/Tead signaling. *Developmental cell*, 53(2) 199-211. e6
[doi:10.1016/j.devcel.2020.03.005](https://doi.org/10.1016/j.devcel.2020.03.005)

Abstract

The Warburg effect is one of the metabolic hallmarks of cancer cells, characterized by enhanced glycolysis even under aerobic conditions. While this physiological adaptation is associated with metastatic behavior, we still have a superficial understanding of how it affects cell behavior during embryonic development. Here we report that the neural crest, a migratory stem cell population in vertebrate embryos, undergoes an extensive metabolic remodeling to engage in aerobic glycolysis prior to delamination. This increase in glycolytic flux promotes Yap/Tead signaling, which directly activates the expression of a set of transcription factors to drive epithelial-to-mesenchymal transition. Our results demonstrate how shifts in carbon metabolism can trigger the gene regulatory circuits that control complex cell behaviors. These findings support the hypothesis that the Warburg effect is a precisely regulated developmental mechanism that is anomalously reactivated during tumorigenesis and metastasis.

Introduction

Shifts in carbon metabolism play essential roles in the regulation of cellular properties in eukaryotes. Under well-defined conditions, cells employ distinct strategies for energy production, with consequences that extend beyond changes in ATP levels. In stem cells, heightened levels of glycolysis stimulate histone acetylation to promote cell proliferation and inhibit differentiation (Cha et al., 2017; Moussaieff et al.,

This work is a manuscript that has been published with the above citation. Debadrita Bhattacharya performed most of the experiments described in this paper including all the metabolic analysis and functional studies. Ana Paula Azambuja carried out the CUT&RUN experiments, the subsequent bioinformatic analysis and the enhancer experiments described in Fig 6.

2015). Cancer cells also display high glycolytic activity, which is linked to key features of oncogenesis such as rapid cell division and tissue growth (Hsu and Sabatini, 2008; Vander Heiden et al., 2009). This metabolic adaptation, known as the Warburg effect, is characterized by preferential activation of glycolysis even in environments that are rich in oxygen. While the Warburg effect was initially described as an anomalous mode of metabolism (Warburg, 1925), growing evidence points to glycolysis as a crucial regulator of cell identity and behavior under physiological conditions. Recent studies indicate that aerobic glycolysis is deployed in a tissue-specific manner to influence the activity of signaling pathways in vertebrate embryos (Bulusu et al., 2017; Oginuma et al., 2017). These findings reinforce the notion that metabolism has an active role in the regulation of cell identity and behavior during embryonic development.

A process that may be especially prone to metabolic regulation is cell motility. The association between the Warburg effect and metastatic behavior is well documented (Lu, 2019; Lu et al., 2015), and motile cells from the immune system display enhanced glycolytic flux (Kaushik et al., 2019; Kishore et al., 2017). Despite this, we still have a limited understanding of the importance of metabolism for cell migration in embryonic development. To explore a possible link between energy production and cellular movement, we examined the metabolic shifts that take place during the development of neural crest cells. The neural crest is a migratory stem cell population that contributes to a variety of derivatives including peripheral neurons, glia, facial cartilage, and skin pigments. During early vertebrate development, neural crest cells delaminate from the dorsal neural tube, undergo epithelial-to-mesenchymal transition (EMT) and migrate extensively within the developing embryo (Simoes-Costa and Bronner, 2013). This striking migratory ability has made it an important *in vivo* model for investigating processes such as EMT, cell motility and migration (Theveneau and Mayor, 2012). Furthermore, numerous studies have drawn parallels between neural crest migration and cancer metastasis to show that the two processes are physiologically similar (Acloque et al., 2009). Neural crest-derived cancers like melanoma and neuroblastoma de-differentiate into a neural crest-like state before metastasis (Tsoi et al., 2018). Similar

to cancer cells, neural crest cells are also highly sensitive to oxidative stress during migration (Chen et al., 2013; Laforgia et al., 2018), suggesting a role of energy metabolism in the regulation of this process.

Here we show that carbon metabolism is a major regulator of cell delamination and migration during embryonic development. Upon specification, neural crest cells display an increase in the expression of glycolytic enzymes and engage in aerobic glycolysis. This metabolic transition is necessary for proper cell movement since the chemical inhibition of the Warburg effect prevented EMT and migration. We found that an increase in glycolytic flux promotes the interaction between the YAP1 and TEAD1 proteins, resulting in the activation of the Yap/Tead signaling pathway. YAP/TEAD directly binds to tissue-specific enhancers and promotes the expression of EMT factors in neural crest cells. While enhanced glycolysis may be important for faster ATP production during migration (Epstein et al., 2017), our results indicate that the Warburg effect precedes delamination and activates the gene regulatory circuit that controls EMT. Taken together, these findings demonstrate how intracellular shifts in bioenergetics can modulate gene regulatory networks to ultimately control cell state transitions during embryonic development.

Results

Neural crest cells display the Warburg effect at the onset of migration

To identify shifts in gene expression that are important for neural crest EMT and migration, we performed an extensive transcriptome analysis of FACS sorted cells at six developmental time-points (Hamburger Hamilton stages 6-16, Fig. 1a) (Hamburger and Hamilton, 1951). The RNA-seq data recapitulated the expected transcriptional changes, such as a marked increase in the expression of EMT factors such as *SNAI1*, *SNAI2*, *TWIST*, and *ZEB1* during migration (Fig. 1b). We also observed an unexpected increase in the expression of rate-limiting glycolytic enzymes such as *PFKP*, *GAPDH*, *PFKFB3*, and *LDHA*, at the onset of migration (Fig. 1c). Hierarchical clustering analysis with a set of carbon metabolism genes indicated that glycolytic enzymes were upregulated during EMT and migration,

whereas most components of the Krebs Cycle displayed a distinct pattern of gene expression (Fig. 1d-e).

Figure 1

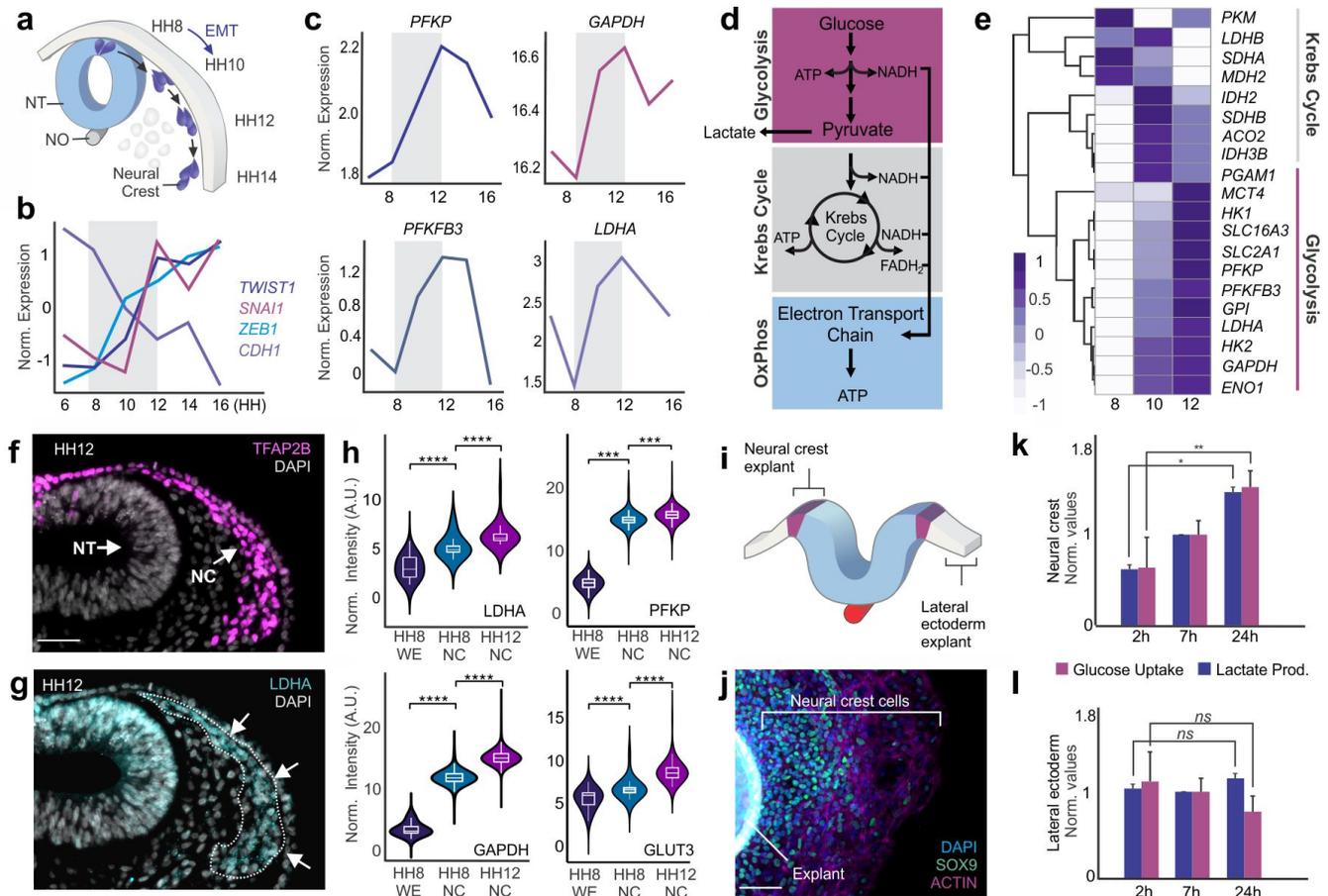


Figure 1. The expression of glycolytic enzymes is upregulated in migratory neural crest cells

(a) Diagram of avian neural tube transverse section, showing the position of cranial neural crest cells at distinct developmental stages. **(b)** Transcriptional profiles of EMT regulators in developing neural crest cells. Expression of *TWIST1*, *SNAI1*, *ZEB1* increases during migration while there is a reduction in *CDH1* (E-cadherin) transcript levels. **(c)** Expression levels of rate-limiting glycolytic enzymes increase during neural crest EMT. **(d)** Simplified schematic of carbohydrate metabolism in eukaryotic cells. **(e)** Clustering analysis showing an increase in the expression of glycolytic enzymes in pre-migratory (HH8), migratory (HH10) and late migratory (HH12) neural crest cells. The heatmap shows z-score normalized expression levels of each metabolic genes. **(f-g)** Immunostaining for glycolytic enzyme LDHA reveals high protein expression in TFAP2B+ migratory neural crest cells. **(h)** The protein levels of glycolytic enzymes LDHA, PFKP, GAPDH, and GLUT3 are enriched in pre-migratory neural crest cells (HH8) and increases further upon neural crest migration. **(i)** Diagram indicating the regions of the ectoderm that were dissected for explant culture. **(j)** Neural fold explant with a halo of migratory SOX9+ neural crest cells. **(k-l)** Cultured neural crest cells increase in glucose uptake and lactate production during migration, while these parameters remain constant in lateral ectoderm explants. Ect: Ectoderm; EMT: epithelial-mesenchymal transition;

HH: Hamburger and Hamilton stage; NC: Neural Crest; NT: neural tube; NO: notochord. Error bars represent \pm S.E.M. Scale bar : 50 μ m (f-g), 200 μ m (j).

Immunohistochemistry for glycolytic enzymes (Fig. 1f-g) and flow cytometry analysis (Fig. 1h) confirmed that protein expression levels reflected these transcriptomic changes. LDHA, PFKP, GAPDH, and GLUT3 were enriched in neural crest cells, and we observed a significant rise in the levels of these glycolytic enzymes between stages HH8 and HH12 (Fig. 1h). Thus, neural crest migration involves a transient increase in the expression of glycolytic enzymes. Since chick embryos develop in a normoxic environment (Philips, 1941; Romanoff and Romanoff, 1949), these changes in gene expression were reminiscent of the Warburg effect observed in cancer cells (Ngo et al., 2015; Warburg et al., 1924).

To investigate if higher levels of glycolytic enzymes were linked to aerobic glycolysis, we measured the glycolytic flux at different stages of neural crest migration. First, we employed colorimetric and luminescent assays to quantify lactate production and glucose consumption. In these experiments, we performed explant cultures to isolate populations of migratory neural crest cells (Coles et al., 2007). Dorsal neural folds of HH9 avian embryos were dissected and cultured on fibronectin-coated plates. Within 2h, the neural folds attach to the substrate; by 7h of incubation most neural crest cells have delaminated, and by 24h migratory cells form a halo around the explant (Fig. 1i-j). Consistent with high glycolytic flux, we observed a gradual increase in both glucose uptake and lactate levels as neural crest cells transition from the pre-migratory to the migratory stages in the explant system (Fig. 1k). This metabolic shift was specific to neural crest cells since it was not observed in lateral ectoderm explants maintained under identical culture conditions (Fig. 1l).

Next, we used the Seahorse XFp to measure the levels of glycolysis vs. oxidative phosphorylation in neural crest cells. We compared the basal Extracellular Acidification Rate (ECAR) and the Oxygen Consumption Rate (OCR) of neural crest progenitors at the onset of specification (HH7), and

delamination (HH9). Quantification of the ECAR and OCR rates revealed that, during EMT, neural crest cells transition from quiescence to a state of enhanced glycolysis (Fig. 2a-c).

Figure 2

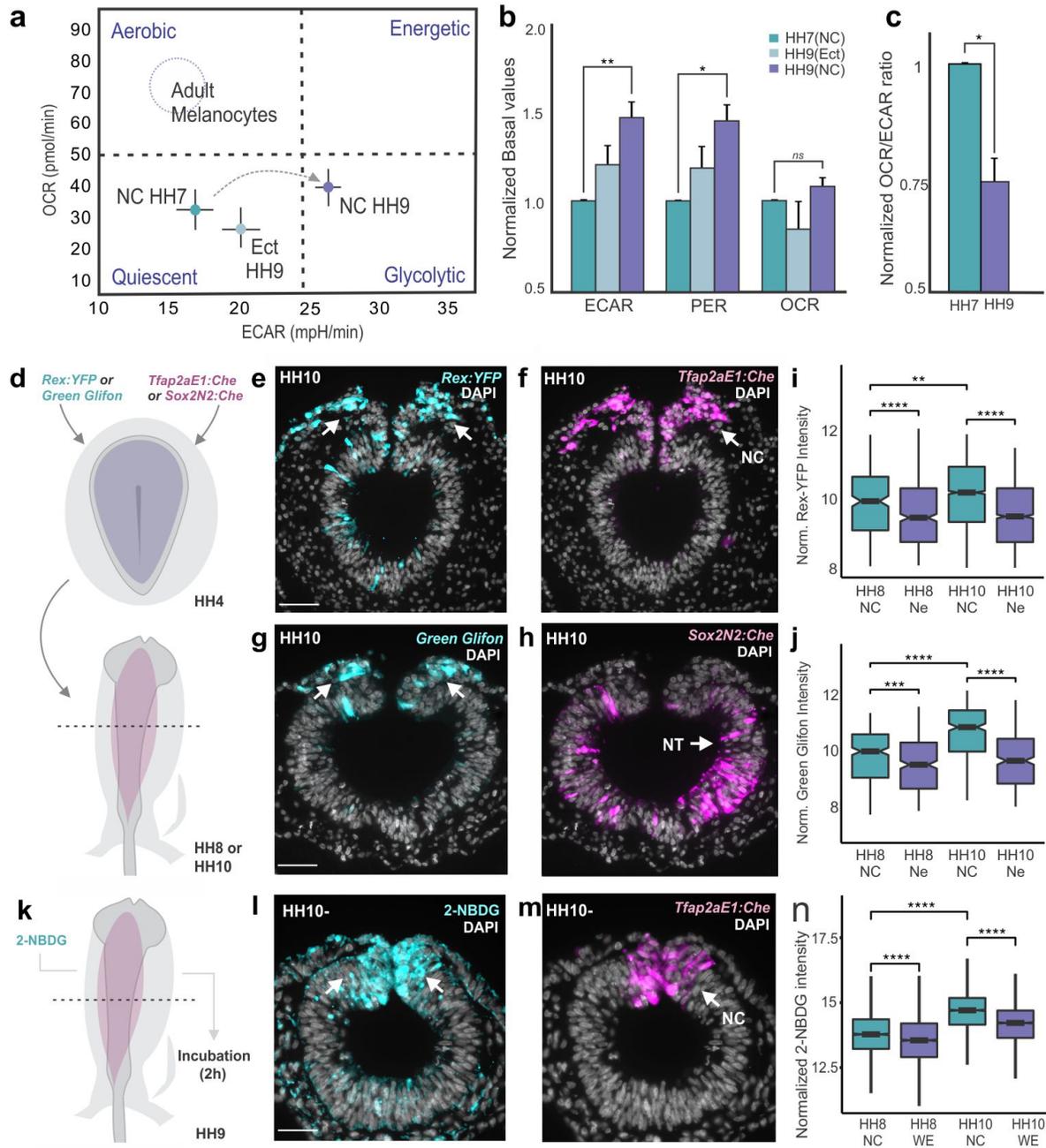


Figure 2. Cranial neural crest cells display the Warburg effect at the onset of migration

(a) Basal oxygen consumption and extracellular acidification rates measured using Seahorse XF indicate that metabolically quiescent neural crest cells (NC HH7) become highly glycolytic during delamination (NC HH9).

Lateral ectoderm cells remain quiescent (Ect HH9), whereas neural crest-derived melanocytes have been reported to be highly aerobic. **(b-c)** Despite the increased rates of extracellular acidification (ECAR) and proton efflux (PER) in HH9 neural crest, the oxygen consumption rates (OCR) in these cells is not significantly different from other tissues (b). This manifests in reduced OCR/ECAR ratio in HH9 cells compared to HH7 neural crest (c). **(d)** Diagram showing the electroporation scheme for assessing the activity of metabolic reporters in avian neural crest cells and neural progenitors at different developmental stages. **(e-h)** Transverse sections of HH10 transgenic embryos transfected with reporters of intracellular NAD^+/NADH ratio (*Rex-YFP*) and *Tfap2aE1:Cherry* (e-f) and cytoplasmic glucose levels (*Green Glifon50*) and *Sox2N2:Cherry* (g-h). **(i-j)** Flow cytometry quantification of *Rex:YFP* (i) and *Green Glifon50* (j) activity in isolated neural crest (NC) and neural cells at two developmental stages. Both metabolic reporters have high activity in *Tfap2E1+* neural crest cells and display low overlap with *Sox2N2+* neural progenitor cells. **(k)** Diagram depicting *Tfap2aE1:Cherry* transgenic embryos treatment with fluorescent glucose analog 2-NBDG. **(l-m)** Transverse section of an HH10 embryo showing colocalization of 2-NBDG and *Tfap2aE1+* neural crest cells. **(n)** Boxplots showing the intensity of 2-NBDG in *Tfap2aE1+* neural crest cells and *Tfap2aE1-* whole embryo cells at stages HH8 and HH10. che: Cherry ECAR: Extracellular Acidification Rate; Ect: Ectoderm; EMT: epithelial-mesenchymal transition; HH: Hamburger and Hamilton stage; PER: Proton Efflux Rate, NC: Neural Crest; Ne: Neural, OCR: Oxygen Consumption Rate. Error bars represent \pm S.E.M. Also see Fig. S1. Scale bar : 50 μm .

This shift was again specific to neural crest cells, as ectodermal cells obtained from HH9 embryos were metabolically quiescent (Fig. 2a,b). Since neural crest-derived melanocytes engage in aerobic respiration and display a high OCR to ECAR ratio (Hall et al., 2013), these results indicate that the metabolic state of neural crest cells is highly dynamic through development.

To confirm that these metabolic changes occur during neural crest migration *in vivo*, we employed a combination of molecular sensors to monitor the glycolytic flux in developing avian embryos (Fig 2d). An increased glycolytic flux raises the cellular NAD^+/NADH ratio, which is up to five to ten-fold higher in cancer cells compared to the normal cells (Ngo et al., 2015). Thus, we first examined the NAD^+/NADH ratio of ectodermal cells using the *Rex:YFP* reporter (Bilan et al., 2014). We co-transfected gastrula stage (HH5) avian embryos with pCI-*Rex-YFP* (Fig. 2d) and either a neural crest (*Tfap2aE1:mCherry*) (Fig. 2f) or a neural (*Sox2N2:mCherry*) (Fig. 2h) specific enhancer. Consistent with our previous experiments, transfection of the reporter revealed that neural crest cells have a higher NAD^+/NADH ratio than adjacent tissues like the neural tube, as evidenced by microscopy (Fig. 2e-f) and flow cytometry (Fig. 2i). Notably, migratory neural crest cells, which have the highest expression levels of glycolytic genes (Fig. 1e), displayed higher *Rex:YFP* activity than pre-migratory cells (Fig. 2i). Next, we utilized a *Green Glifon50*

(green glucose indicating fluorescent protein sensor) to monitor the cytoplasmic glucose levels in neural crest cells. Transfection of avian embryos with a vector encoding *Green Glifon50* (Mita et al., 2019) revealed that neural crest cells contain higher levels of intracellular glucose than neural progenitors, and that levels of glucose increase in migratory cells (Fig. 2g-h, 2j). Finally, incubation of embryos with the fluorescent glucose analog 2-NBDG (Fig. 2k) (Oginuma et al., 2017) confirmed that neural crest cells display higher glucose uptake than neighboring cells (Fig. 2l-m) and that this is also exacerbated during migration (Fig. 2n). Taken together, these experiments show that neural crest cells exhibit the Warburg effect, which is characterized by a pronounced increase in lactate production, glucose uptake, high levels of intracellular glucose and low OCR/ECAR but high NAD^+/NADH ratios. These observations led us to hypothesize that enhanced glycolysis plays a role in the migration of neural crest cells.

Enhanced glycolysis is necessary for neural crest EMT

To establish the importance of the Warburg effect for neural crest development, we treated these cells with 2-deoxy-D-glucose (2-DG), a competitive inhibitor of glucose transporters and the hexokinase enzyme (Barban and Schulze, 1961). Treatment of neural crest explants with 2-DG drove cells towards aerobic respiration, as evidenced by an increase in the OCR/ECAR ratio (Fig. 3a, Fig. S2a-b), and decreased Rex-YFP fluorescence (Fig. S2c). This also resulted in a decrease in lactate production (Fig. 3b) but did not alter total cellular ATP levels (Fig. 3c), cell proliferation or survival (Fig. S2d-e). Strikingly, we found that inhibition of glycolysis had a strong effect on cell migration. Treatment with 2-DG suppressed neural crest migration, resulting in explants with smaller diameters than controls (Fig. 3d-e). This phenotype was reversible, as removal of the 2-DG prompted cells to resume migration (Fig. S2 f-h). Targeting glycolysis with 3-bromopyruvate (3-BP) and 6-aminonicotinamide (6-AN) yielded similar results (Fig. S3a-e), but inhibition of the electron transport chain with either oligomycin or sodium azide did not affect cell migration or explant size (Fig. S3f-h). Thus, high glycolytic flux is specifically required for proper neural crest migration. To further characterize this phenotype, we employed time-lapse imaging and

quantified cell movement following inhibition of glycolysis (Supplemental Movie 1 and 2). By tracking the nuclei of individual migratory cells (Fig. 3f), we observed that 2-DG treatment significantly decreased the maximum linear distance traveled, the total displacement of the cells, and their mean speed (Fig. 3g-i). These changes in cell migration suggested that 2-DG treated neural crest cells were unable to complete EMT.

Figure 3

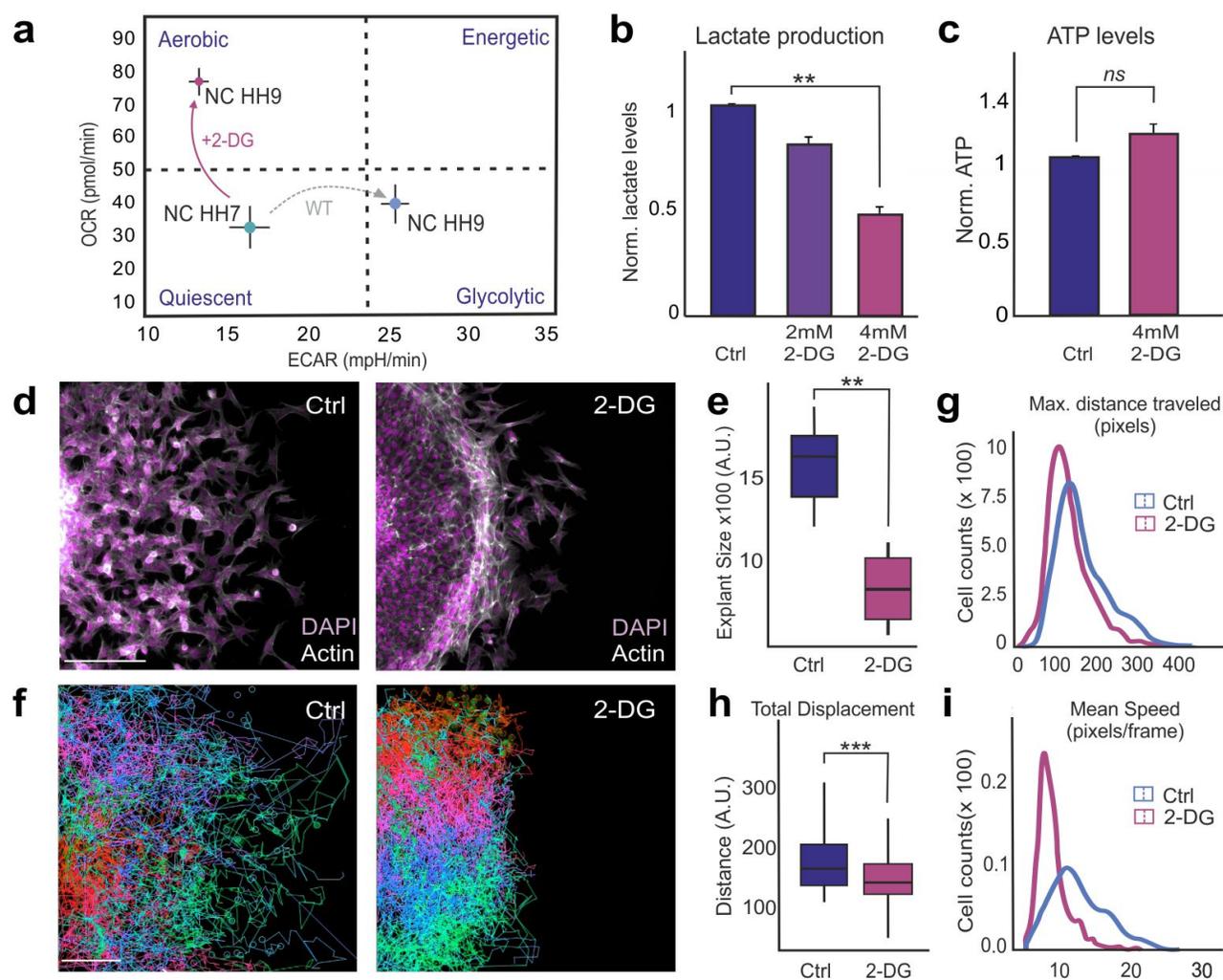


Figure 3. Inhibition of glycolysis disrupts neural crest migration

(a) Measurement of basal OCR and ECAR show that inhibition of glycolysis with 2-DG drives neural crest cells to adopt aerobic, rather than glycolytic, metabolism. **(b)** Lactate production by neural crest cells is inhibited by 2-DG treatment in a dose-dependent manner. **(c)** Despite reducing glycolytic flux, 2-DG treatment does not affect ATP levels in neural crest cells. **(d)** Images of control and 2-DG-treated neural crest explants after 24h of incubation.

(e) Inhibition of glycolysis results in smaller neural crest explants. **(f)** Overlay of tracks of individual cells from a 12h time-lapse movie of a neural fold explant. Neural crest cells were transfected with vectors expressing H2B-RFP and Actin-GFP for live imaging. Cell trajectories in control and 2-DG treated explants show decreased neural crest migration following inhibition of aerobic glycolysis. **(g-i)** Treatment with 2-DG results in reduction of the maximum distance traveled, total displacement and mean speed of individual neural crest cells. ECAR: Extracellular Acidification Rate; HH: Hamburger and Hamilton stage; NC: neural crest; OCR: Oxygen Consumption Rate. 2-DG : 2- Deoxy-Glucose, Ctrl: Control, HH: Hamburger and Hamilton stage .Error bar represents \pm S.E.M. Scale bar: 200um. Also see Fig. S2 and Fig. S3.

Immunofluorescence for different EMT markers confirmed this hypothesis. Under control conditions, cells at the migratory front of the explants exhibited mesenchymal characteristics, such as low levels of membrane E-cadherin, Paxillin localized to focal adhesions and elongated actin stress fibers (Fig. 4a). However, 2-DG treatment caused these cells to retain epithelial features, such as high levels of membrane-localized CDH1 (E-cadherin), cytoplasmic Paxillin, and a dense Actin mesh (Fig. 4a). To test if aerobic glycolysis was also important for EMT *in vivo*, we treated stage HH8 chick embryos with 2-DG in EC culture (Chapman et al., 2001) until they reached HH12. While these embryos had normal gross morphology, staining for neural crest marker TFAP2A revealed an EMT phenotype, with the majority of cells trapped adjacent to the dorsal neural tube (Fig. 4b). Notably, these cells also retained high levels of CDH1, when compared to the neural crest cells from control embryos (Fig. S4).

Next, we sought to identify the molecular mechanism linking the Warburg effect to neural crest EMT. To survey the impact of glycolytic inhibition on neural crest gene expression, we performed a Nanostring analysis with a set of 200 probes targeting genes involved in neural crest, neural and placodal development (Bhattacharya et al., 2018). This analysis showed that neural crest cells treated with 2-DG displayed a robust reduction in the expression of EMT transcription factors such as *SNAI2*, *FOXD3*, and *ETS1* (Fig. 4c). Expression of many neural crest genes (e.g. *TFA2B*, *CRABP1*, *LMO4*) remained unchanged, indicating that glycolysis was preferentially required for the expression of factors related to EMT (Fig. 4d). This was confirmed by qPCR analysis, which revealed a robust downregulation of the bona fide EMT regulators following treatment with 2-DG (Fig. 4e). These results suggested the existence of a mechanism coupling cell metabolism to the neural crest gene regulatory network (Simoes-Costa and

Bronner, 2015). To identify a possible molecular pathway that mediates transcriptional changes in response to glycolytic flux, we examined the regulatory regions of the genes most affected by 2-DG treatment in our Nanostring analysis. We compiled the ten active chromatin regions (H3K27Ac+) closest to glycolysis-responsive genes (Fig. 4f) and performed motif enrichment analysis with TRAP (Bailey et al., 2009). This analysis identified an abundance of TEAD1 binding motifs in these sequences (p-value = 5.42e-07), suggesting the involvement of the Yap/Tead signaling pathway in the generation of a transcriptional response to changes in the metabolism of neural crest cells.

Figure 4

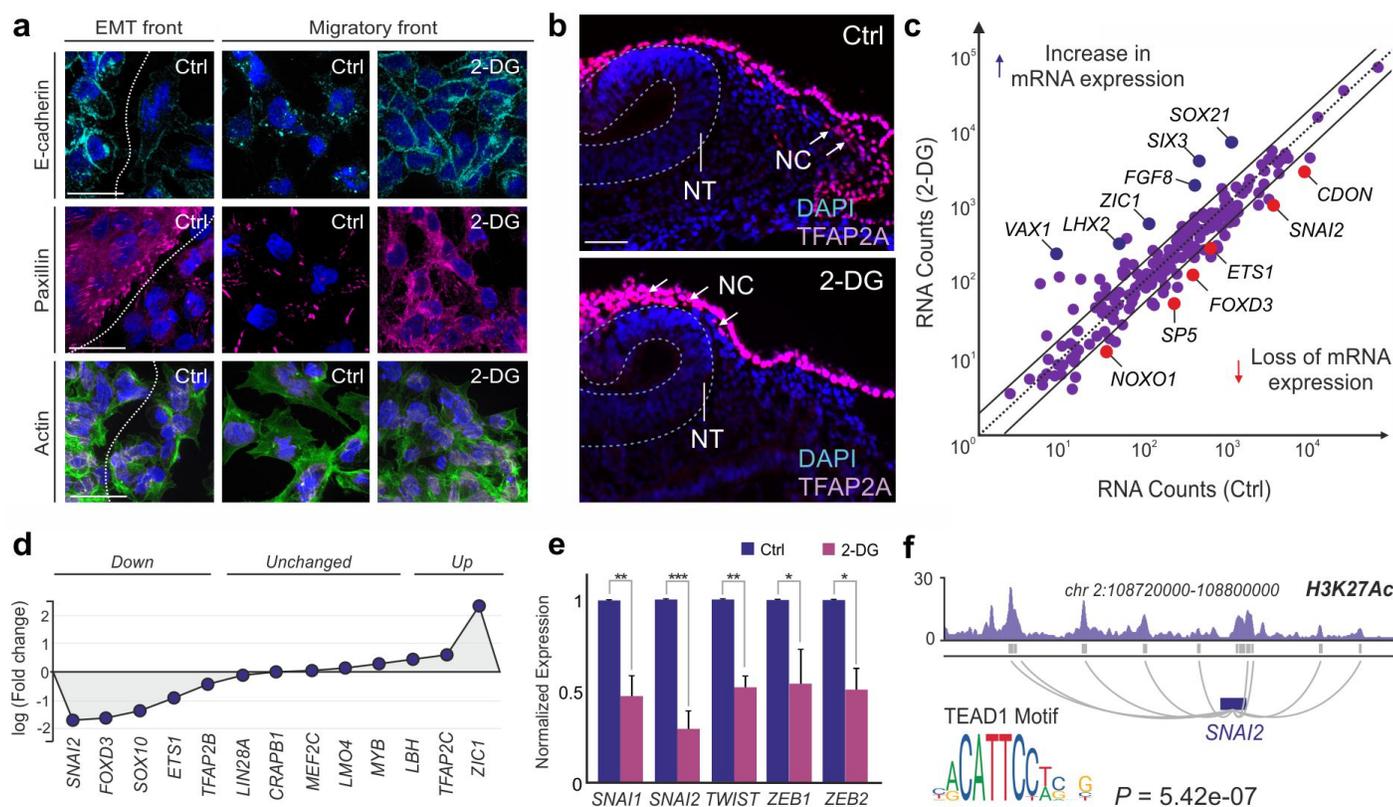


Figure 4. Aerobic glycolysis is required for neural crest EMT

(a) Immunofluorescence for different EMT markers reveals that cells in 2-DG treated explants retain epithelial features. **(b)** Transverse sections of HH12 embryos show that 2-DG treatment inhibits neural crest EMT and migration *in vivo*, resulting in the majority of cells being trapped adjacent to the dorsal neural tube. **(c)** Nanostring assay comparing gene expression profiles of control and 2-DG treated explants. Genes below the diagonal lines are significantly downregulated following inhibition of aerobic glycolysis. **(d)** Neural crest genes related to EMT are

more strongly affected by 2-DG treatment, while many markers remain unchanged. **(e)** qPCR analysis confirms the loss of expression of bona fide EMT markers following inhibition of aerobic glycolysis. **(f)** Motif enrichment analysis performed with H3K27ac peaks in the *loci* of glycolysis-responsive genes identified the TEAD1 as one of the motifs to be statistically overrepresented. 2-DG : 2- Deoxy-Glucose, Ctrl: Control, NC: neural crest; NT: neural tube. Error bar represents \pm S.E.M. Scale bar : 35 μ m (a), 50 μ m (b).

High glycolytic flux promotes activation of the Yap/Tead pathway

YAP and TEAD are components of the Hippo signaling system that can respond to a variety of stimuli to regulate transcription (Hansen et al., 2015). YAP lacks a DNA binding domain, but it is able to translocate to the nucleus and interact with TEAD (TEA domain family members 1-4) transcription factors to induce gene expression (Zhao et al., 2008). Yap/Tead signaling has been shown to promote invasiveness and metastasis in several cancer types (Zanconato et al., 2016). Recent studies in zebrafish, mouse and neural crest cell culture models have also implicated this pathway in the control of neural crest migration and differentiation (Dooley et al., 2019; Hindley et al., 2016; Wang et al., 2016). Consistent with this, enrichment of active, non-phosphorylated YAP1 in the nuclei of migratory neural crest cells supported that this pathway was important in chick neural crest development (Fig. 5a-c). To test that YAP1/TEAD1 responds to glycolytic flux in neural crest cells, we employed the HOP-GFP and HOP-FLASH reporter transgenes, which consist of multiple TEAD1 sites upstream of a minimal promoter and a reporter gene (Kim and Gumbiner, 2015). Transfection of HOP-GFP revealed that migratory neural crest cells from explants displayed high YAP/TEAD activity, which was repressed upon inhibition of glycolysis with 2-DG (Fig. 5d). Luciferase assays confirmed that the activity of HOP-FLASH was significantly reduced upon treatment with 2-DG (Fig. 5e), while the HIP-FLASH control construct (which contains mutated TEAD1 binding sites) was not affected by the drug (Fig. S5a). Next, we examined how YAP/TEAD is regulated by glycolysis. There are two critical steps for the regulation of the pathway: (i) translocation of YAP to the nucleus or (ii) interaction of YAP1 with TEAD1 (Enzo et al., 2015; Meng et al., 2016). We employed immunofluorescence and Proximity Ligation Assays (PLA) (Greenwood et al., 2015) to examine how inhibition of glycolysis affected each of these processes. While immunostaining for active YAP1 showed no significant changes in nuclear translocation upon 2-DG treatment of explants

(Fig. 5f-g) or *in vivo* (Fig. S5b-c), PLA revealed that the inhibitor greatly reduced the number of YAP-TEAD interactions in the nuclei of neural crest cells (Fig. 5h-i). These results demonstrate that the Warburg effect activates of YAP1/TEAD1 signaling by promoting the assembly of its nuclear effector protein complex.

Figure 5

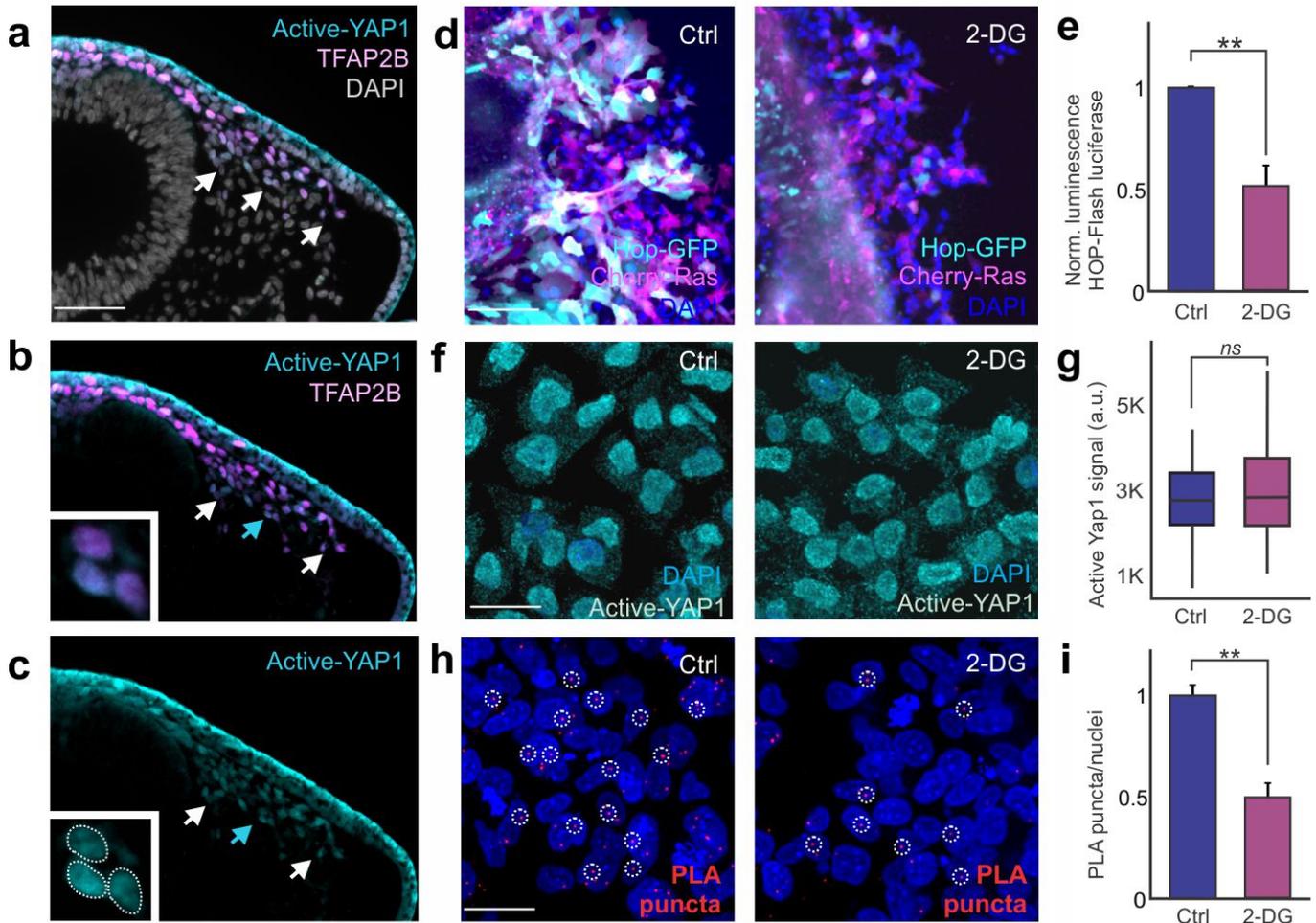


Figure 5. Glycolysis activates Yap/Tead signaling in neural crest cells by promoting YAP1-TEAD1 interaction

(a-c) Transverse section showing immunofluorescence for the non-phosphorylated form of YAP1 (Active YAP) and neural crest marker TFAP2B in an HH12 embryo. Arrows show enrichment of Active YAP in migratory neural crest cells. Inserts (b-c) depict nuclei marked by blue arrows. **(d)** Effect of inhibition of aerobic glycolysis on Yap signaling reporter. Neural crest cells electroporated with HOP-GFP reporter and the transfection control plasmid pCI-Cherry-Ras in control and in the presence of 2-DG. **(e)** Luminescence assay with luciferase version of the reporter (HOP-Flash) shows a reduction in Yap/Tead signaling activity. **(f)** Immunofluorescence for Active YAP1 in control and 2-DG explants. **(g)** Quantification of the nuclear intensity of active YAP1 staining shows a non-significant difference between control and 2-DG treated cells. **(h)** Proximity ligation assay (PLA) performed to

detect YAP1 and TEAD1 interaction frequency in control vs 2-DG treated explants. **(i)** Quantification of PLA puncta revealed a significant decrease in YAP1-TEAD1 interaction frequency upon glycolytic inhibition. 2-DG: 2-Deoxy-Glucose, HH: Hamburger and Hamilton, PLA: Proximity Ligation Assay. Error bar represents \pm S.E.M. Scale bar: 50 μ m (a-c,d), 35 μ m (f,h). Also see Fig. S5.

Next, we utilized orthogonal approaches to test if disruption of YAP1-TEAD1 interaction was sufficient to affect neural crest EMT. Treatment of explants with Verteporfin (Feng et al., 2016), a small molecule that inhibits the pathway by preventing YAP1-TEAD1 association (Fig. S6), disrupted neural crest EMT and resulted in a phenotype indistinguishable from 2-DG treatment (Fig. 6a-b,6c-d). Consistent with this, transfection of neural crest cells with a dominant-negative form of YAP1 unable to interact with TEAD (YAP-5SA/S94A), phenocopied 2-DG and Verteporfin treatments (Fig. 6c-d). Both pharmacological and genetic inhibition of YAP1-TEAD1 interaction also resulted in decreased expression of EMT transcription factors in neural crest cells (Fig. 6e). To place Yap/Tead signaling as the metabolism-sensing pathway that promotes EMT, we tested if constitutive activation of the pathway could rescue neural crest cell migration in explants treated with 2-DG. We designed a construct expressing the DNA binding domain of TEAD1 (TEA domain) fused to the activating VPR domain (TEA-VPR). Consistent with our hypothesis, explants expressing TEA-VPR were able to undergo complete EMT and migration, even in the presence of 2-DG (Fig. 6f). TEA-VPR transfected neural crest cells also displayed mesenchymal features like the absence of membrane CDH1 (Fig. 6g). These results show that the Yap/Tead pathway is an essential regulator of neural crest EMT, acting to modulate gene expression and cell behavior in response to glycolytic flux.

YAP/TEAD interact with tissue-specific enhancers to drive EMT

The outcome from our Nanostring experiment and motif enrichment analysis (Fig. 4c, f) suggests that Yap/Tead signaling directly activates many glycolysis-responsive genes. To define how this signaling pathway promotes neural crest EMT, we employed CUT&RUN (Skene and Henikoff, 2017) (Fig. 7a, Fig. S7a-b) to map the genomic occupancy of YAP1. YAP1 was preferentially bound to intergenic regions

(Fig. 7b, Fig. S7c) of open chromatin that are positive for the H3K27Ac mark (Fig. 7c). Motif enrichment analysis of these regions identified the TEAD-binding motif as the top hit (p -value <0.00001) (Fig. 7d), indicating that YAP1 binds to these genomic regions with its canonical partner.

Figure 6

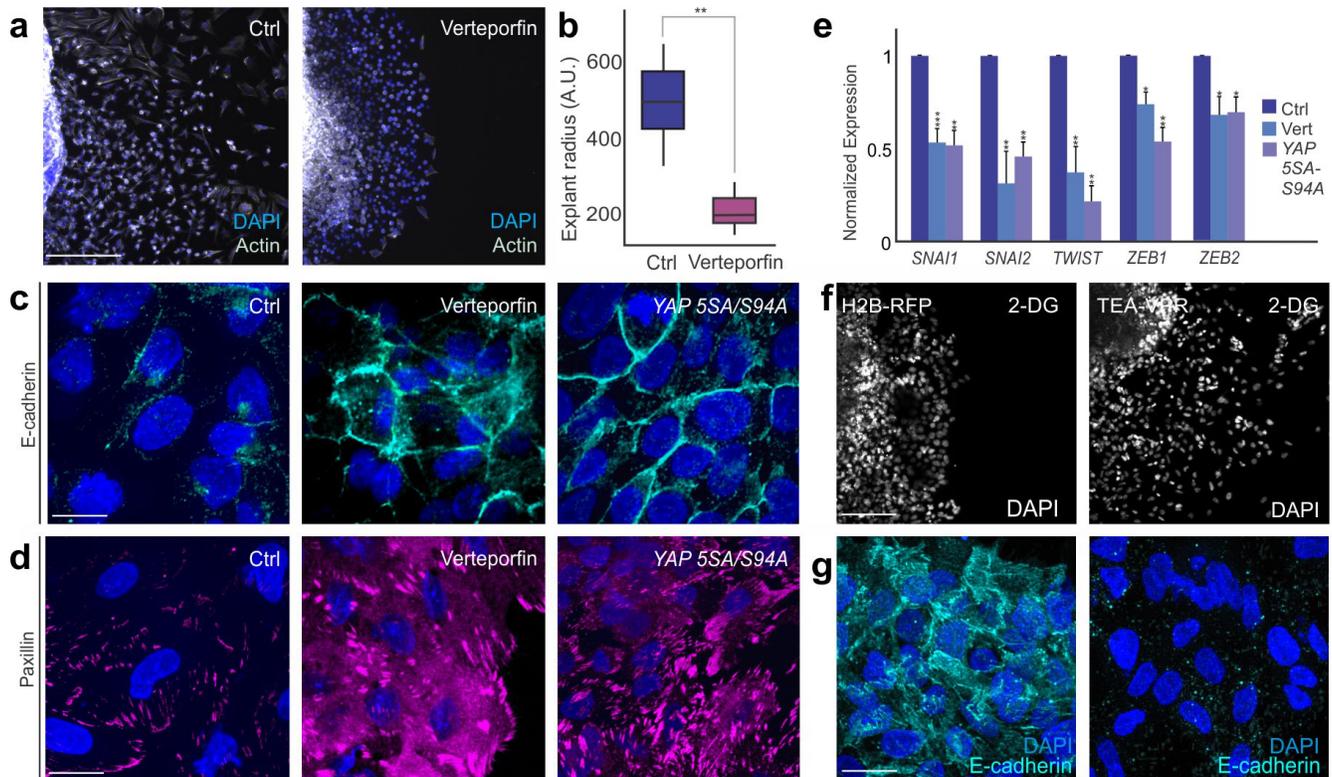


Figure 6. Pharmacological inhibition of YAP1-TEAD1 interaction prevents neural crest EMT.

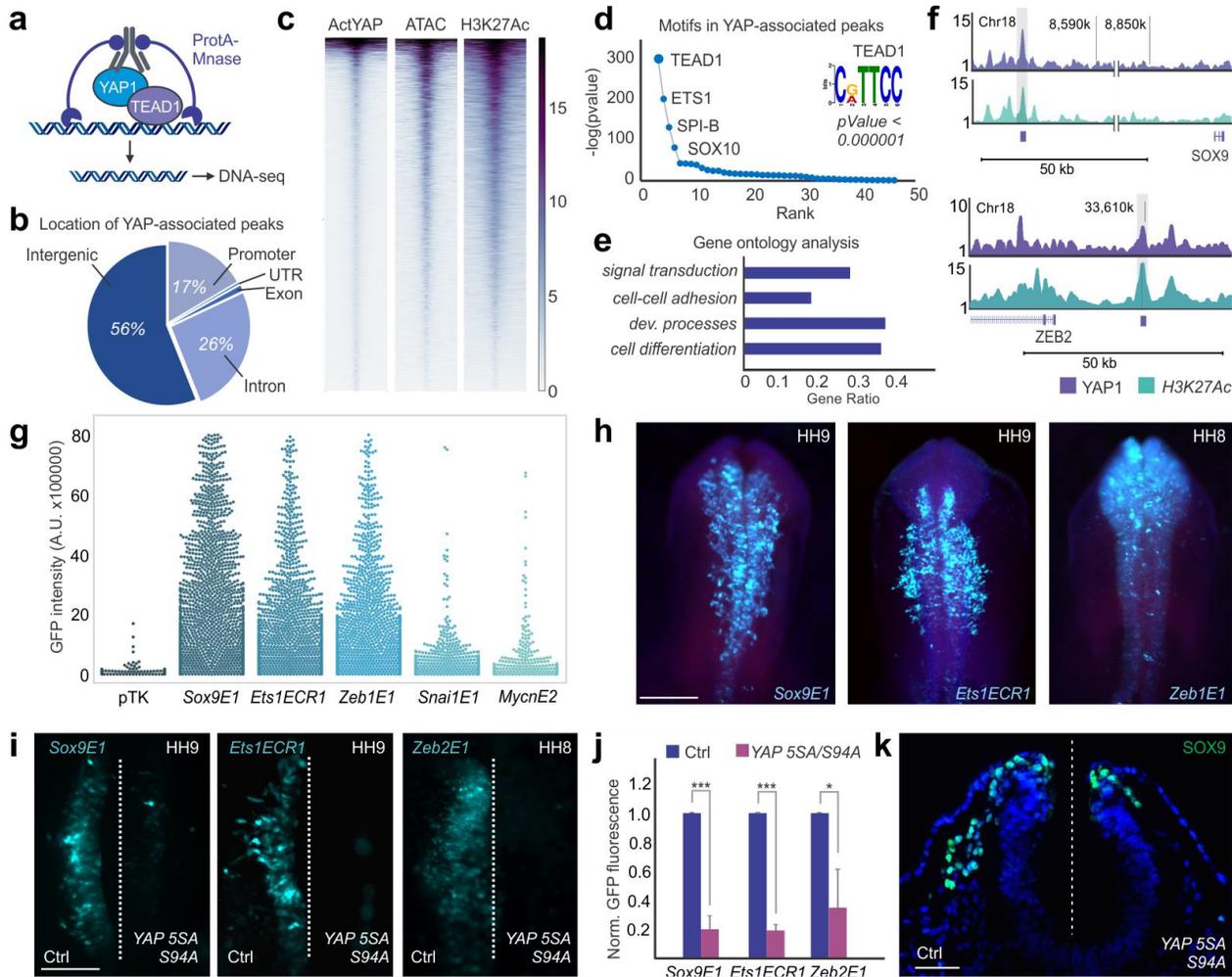
(a-b) Inhibition of YAP1-TEAD1 interaction with Verteporfin inhibits neural crest migration and results in smaller neural crest explants. **(c-d)** Images of CDH1 and Paxillin staining in control, Verteporfin treated and dominant negative YAP 5SA-S94A transfected neural crest explants, show retention of epithelial features following inhibition of Yap/Tead signaling. **(e)** qPCR for EMT transcription factors following pharmacological (Vert) and competitive inhibition (mutant YAP5SA-S94A construct) of YAP1-TEAD1 interaction in neural crest cells. **(f)** Expression of TEA-VPR construct that mimics constitutively active TEAD1, rescues phenotypes of glycolytic inhibition. DAPI staining of control and TEA-VPR overexpressing explants treated with 2-DG. **(g)** Immunofluorescence for CDH1 shows successful EMT in TEA-VPR expressing explants, even in the presence of 2-DG. Error bars represents \pm S.E.M. Scale bar : 200um (a,f), 35um (c-d,g). Also see Fig. S6.

GO-term analysis of the genes closest to YAP1-bound elements confirmed that the pathway regulates a large number of genes involved in cell-cell adhesion (Fig. 7e, Fig. S7d), including both structural proteins

and EMT transcription factors. Accordingly, our CUT&RUN analysis identified several YAP1 peaks in the vicinity of EMT genes such as *SOX9*, *ZEB2*, *ETS1*, *SNAI2* and *MYCN* (Fig. 7f). Thus, the genomic profiling of YAP1 indicates that the Yap/Tead1 signaling pathway directly regulates the expression of factors that drive neural crest delamination and migration.

To test whether the YAP1 bound peaks are active enhancers in neural crest cells, we cloned five of these regions in pTK-GFP (Uchikawa et al., 2003) and performed transient transgenesis assays in avian embryos. One of these putative enhancers (*Ets1ECR1*) had previously been shown to regulate *Ets1* expression during early neural crest development (Barembaum and Bronner, 2013). To quantify enhancer activity, we co-transfected the GFP constructs with a neural crest reporter (*Tfap2aE1:mCherry*) and used flow cytometry to measure GFP levels in HH10 neural crest cells. All five regions tested were able to induce reporter expression above background levels (Fig. 7g), with the enhancers from the *SOX9*, *ETS1*, and *ZEB2 loci* driving the strongest GFP fluorescence in the neural crest cells (Fig. 7h). To further confirm the requirement of YAP1-TEAD1 interaction for the activity of these regulatory elements, we employed the dominant negative YAP-5SA/S94A construct. Embryos were bilaterally electroporated with each enhancer, whereas only the left side of the embryo was co-transfected with YAP-5SA/S94A expression vector. As expected, we observed a significant loss in enhancer activity of *Sox9E1*, *Zeb2E1*, and *Ets1ECR1* in the presence of the dominant-negative YAP1 construct (Fig. 7i-j), which also prevented proper neural crest migration (Fig. 7k). Notably, the three enhancers were also strongly affected by treatment with 2-DG (Fig. S7e-g), indicating their activity is dependent on glycolysis. Taken together, these results indicate that Yap/Tead signaling is a direct regulator of the EMT program and that it promotes neural crest delamination by activating tissue-specific enhancers.

Figure 7



(a) Diagram of CUT&RUN experiments to map genome occupancy of ActYAP1. **(b)** YAP-associated regions are mostly intergenic. **(c)** Heatmaps displaying ATAC-seq and H3K27ac signal at ActYAP1 bound regions show that YAP1 binds to open, active chromatin regions. **(d)** Transcription factor binding sites identified in YAP1-associated peaks. Tead1 was the enriched motif with the highest confidence score. **(e)** Gene ontology analysis of the genes closest to the YAP1-associated peaks suggests that Yap/Tea signaling is a major regulator of cell signaling and cell-cell adhesion. The plot shows the four most significant biological processes identified by the GO analysis. **(f)** Examples of regulatory regions associated with YAP1 in the vicinity of EMT regulators Sox9 and Zeb2. **(g)** Swarm plots showing GFP intensity driven by YAP-associated regions in neural crest cells. Neural crest cells were identified by a specific reporter (*Tfap2E1:Che*), and GFP expression was measured by flow cytometry. The regions tested were in the *loci* of the *SOX9*(*Sox9E1*), *ETS1*(*Ets1ECR1*), *ZEB2*(*Zeb2E1*), *SNAI2*(*Snai2E1*) and *MYCN*(*MycNE1*) genes. GFP intensity analysis was performed in 3500 *Tfap2aE1:Che*⁺ cells. **(h)** Transgenic chick embryos showing tissue-specific activity of *Sox9E1*, *Ets1ECR1*, and *Zeb2E1*. **(i-j)** Images of embryos transfected with YAP5SA-S94A mutant construct and *Sox9E1*, *Ets1ECR1* and *Zeb2E1* respectively. Embryos were injected with the enhancer and a control vector on the left side and the same enhancer and the Yap5SA-S94A on the right side. The expression of enhancer driven GFP was lost on the embryo side transfected with mutant YAP1 construct. **(k)** Inhibition of Yap/Tea signaling with YAP5SA-S94A prevents neural crest EMT and migration. HH: Hamburger and Hamilton. Scale bar: 200um. Error bars represents \pm S.E.M. Also see Fig. S7.

Discussion

Here we report that the Warburg effect plays an essential function in the control of cell migration during embryonic development. Our results are consistent with previous studies that describe a large degree of metabolic plasticity of embryonic cells. In *Drosophila* embryos, rapidly proliferating cells upregulate the expression of enzymes and glucose transporters to increase glycolytic flux (Tennesen et al., 2011). During implantation, mouse embryos preferentially produce ATP via glycolysis (Krisher and Prather, 2012), but subsequently undergo a shift to oxidative phosphorylation during organogenesis (Houghton et al., 1996). Furthermore, recent studies in mouse and chick embryos have shown that progenitor cells in the pre-somitic mesoderm display Warburg-like metabolism. In these cells, increased glycolysis mediates the crosstalk between FGF and Wnt signaling to promote a paraxial mesoderm fate (Bulusu et al., 2017; Oginuma et al., 2017). These studies underscore that cellular metabolism has to be tightly regulated during development. Our results support this observation, while also indicating that aerobic glycolysis itself can regulate cellular behavior by impinging on signaling systems, independent of its role in ATP production.

We also describe how metabolism is coupled with genetic programs that control cell identity and behavior. Neural crest development is orchestrated by a complex gene regulatory network, composed by transcription factors, signaling molecules and epigenetic regulators (Simoës-Costa and Bronner, 2015). This system needs to operate in coordination with extrinsic and intrinsic conditions for the accurate progression of morphogenetic events. Extracellular signaling systems, like Wnts and BMPs (Kleber et al., 2005), relay environmental signals to neural crest cells by directly activating nodes of the network. Our results identify Yap/Tea signaling as an intracellular sensor of metabolism, which modulates gene expression in response to glycolytic flux. Importantly, Yap/Tea signaling targets a specific subset of neural crest genes that are part of a regulatory sub-circuit that promotes EMT. Thus, EMT is robustly

linked to the energetic status of the cell. This mechanism may ensure that only cells that have been metabolically primed are able to delaminate to engage in long-range migration.

The identification of a metabolic requirement for neural crest migration also carries important clinical implications. Defects in neural crest development result in neurocristopathies, which are amongst the most common congenital disabilities. These conditions are also associated with metabolic abnormalities. Babies born to insulin-dependent diabetic mothers often have oral clefts and other facial defects (Spilson et al., 2001), which suggests excess glucose disrupts neural crest development. Furthermore, babies suffering from Fetal Alcohol Syndrome (FAS) who have reduced fetal glucose levels (Shibley and Pennington, 1997), also develop craniofacial malformations (Johnson et al., 1996). Our findings shed light on the etiology for these congenital diseases, by showing that improper fluctuations in glucose metabolism can adversely affect neural crest EMT and migration. Fetuses of diabetic mothers are exposed to high levels of insulin that promote excess aerobic respiration and oxidative stress in the embryo (Gabbay-Benziv et al., 2015; Morgan et al., 2008). As indicated by our observations, oxidative phosphorylation inhibits neural crest migration. Conversely, in FAS patients, reduced glucose uptake by neural crest cells can prevent aerobic glycolysis in delaminating neural crest cells, thus delaying EMT. While it remains to be established if disruption of neural crest metabolism is the main cause of these conditions, our study opens new avenues for research into the role of metabolism in neurocristopathies.

Finally, this study also highlights an additional parallel between neural crest development and metastasis. The idea of a link between cancer and embryonic development has gained traction in the last decades, as genomic studies have systematically uncovered shared genetic circuits between embryonic progenitor and neoplastic cells. This has been reported for diverse cancer types such as medulloblastoma, glioblastoma (Azzarelli et al., 2018), and more recently, in neural crest-derived cancers (Kaufman et al., 2016; Maguire et al., 2015). Our results show that cancer cells and neural crest cells share metabolic states and that enhanced glycolytic flux promotes cell invasion by regulating the EMT

program. This suggests that the Warburg effect is a developmental mechanism with essential functions in the regulation of cell adhesion and migration. Thus, the cooption of aerobic glycolysis by adult cells may underlie the emergence of cellular behaviors that characterize tumorigenesis and malignancy.

Acknowledgments

We are indebted to Adam Wojno for cell-sorting assistance at the BRC Flow Cytometry Cell Sorting Facility in Cornell University. We thank Dr. Rebecca M. Williams and Carol Bayles for support in image acquisition through the BRC Imaging Facility. We are also thankful to Austin Hovland for support with RNA-seq data analysis. This work was supported by NIH grants R00DE024232 and R01DE028576 to M.S.-C, a Basil O'Connor Starter Scholar Award from March of Dimes to M.S.-C, and a Cornell CVG Scholars Award to D.B.

Methods

Chick embryo collection and electroporation

Fertilized Leghorn White chicken eggs were obtained from the Department of Animal Science, University of Connecticut. Eggs were incubated at 37°C until embryos reached the desired developmental stage. Embryos were collected and cultured according to the EC protocol (Chapman et al., 2001), and staged using the Hamburger and Hamilton staging system (Hamburger and Hamilton, 1951). Metabolic reporters, over-expression constructs and enhancer plasmids were transfected in chick embryos at HH5 by *ex ovo* electroporation, as previously described (Simoes-Costa et al., 2015). Briefly, expression vectors were injected between the epiblast and vitelline membrane of embryos at a concentration of 1-2 µg/ul and electroporated with platinum electrodes (five 50ms pulses of 5.1V, with an interval of 100ms between pulses). Whole embryo injections were performed for neural crest sorting experiments, HOP-GFP transfections, and enhancer reporter assays, while mutant YAP 5SA/S94A and TEA-VPR constructs were transfected in only one side of the embryo. Following electroporation, embryos were cultured in

albumin at 37°C until they reached appropriate developmental stages. Embryo survival was >90% and all embryos were screened to ensure uniform electroporation and proper embryo morphology prior to further downstream analysis. For drug treatment experiments, HH8 embryos (4-6 somites) were dissected and incubated in albumin plates containing 2mM of 2-DG, 40 uM of 3-BP and 10uM of Verteporfin. Control embryos were incubated in albumin containing appropriate volumes of the drug solvent (PBS/DMSO). The control and treated embryos were then developed until HH12 and further processed for downstream analysis.

Explant cultures

Neural crest and lateral ectoderm explant cultures were performed as previously described (Coles et al., 2007). Neural crest explants were derived via microdissection of neural folds of HH9 embryos (8ss), and lateral ectoderm was obtained from dissection of ectodermal tissue adjacent to the neural tube. The dissected tissues were then transferred to 6-well cell culture plates or chamber slides coated with fibronectin and containing appropriate volumes of 10%FBS DMEM (10mM glucose + Glutamax) culture media. Unless specified, the explants were incubated for up to 18 hours at 37°C, under normoxic conditions in a CO₂ incubator. For drug treatments, the dorsal neural folds were directly plated in DMEM media having a final concentration of 4mM of 2-DG, 40uM of 3-BP, 50uM of 6-NA, 10uM of Verteporfin, 1mM of NaN₃ or 1uM of oligomycin (the same volume of vehicle was used for control explants). The explants were also incubated for 18 hours at 37°C in 5% CO₂ and normoxic conditions.

Lactate assays

To obtain samples for this assay, neural crest and ectodermal explants were cultured in 35mm plates for 2h, 7h and 24h. A total of 10 explants were cultured on a plate corresponding to each time point. Following incubation, the media was removed and explants were washed with cold PBS. 100ul of PBS containing 0.5% Tween and protease inhibitors were added to each plate, and the cells were detached by scraping. The lysate was collected and vortexed to ensure complete cell lysis. The samples were immediately

frozen in liquid N₂ and stored at -80°C, or processed for downstream assay. Cellular lactate levels were measured using the Lactate Assay Kit (Biovision, #K607) in accordance with the manufacturer's protocol. Values obtained from this assay were normalized to the total protein amount of each sample as measured using the Pierce BCA protein assay kit (Thermo Fisher).

Glucose uptake assays

For performing glucose uptake assay, single dorsal neural folds or ectodermal explants were cultured in individual wells of a clear bottom 96-well plate (Corning). Following appropriate incubation, glucose uptake was measured using Glucose Uptake-Glo™ Assay (Promega, #J1341) according to the manufacturer's instructions. Briefly, glucose-containing media was removed, and explants were incubated in 1mM of 2-DG in PBS for 20 mins at RT. Following cell lysis and neutralization, samples were incubated in 2DG6P detection reagent for 1h and analyzed with a luminometer.

Seahorse assay for measuring basal OCR and ECAR levels in live cells

Chick embryos were incubated until HH7+ (2ss) and HH9+ (8ss), and their dorsal neural folds were micro-dissected in Ringers. 20 neural folds or 15 neural folds were used per HH7+ and HH9+ samples, respectively. For the HH9+ ectodermal samples, adjacent ectodermal regions were dissected in parallel with the neural folds. After collection, the samples were dissociated in 100ul of Accumax for 5 mins. Following dissociation, the samples were centrifuged at 450g for 10 mins, resuspended in 200ul of Seahorse XF complete DMEM media (10mM glucose, 1mM pyruvate, 2mM L-glutamine), and 100ul of the sample was added to an individual well of an 8-well Seahorse cell microplate coated with fibronectin (two technical replicates/sample). The total volume of each well was brought up to 180ul with Seahorse XF DMEM, and the microplate was incubated in a non-CO₂ 37°C incubator for 1h. Following incubation, the seahorse assay was run in Seahorse XFp analyzer according to the manufacturer's instructions. Post basal ECAR and OCR measurements, the cells from each well were fixed with 4% PFA, stained with DAPI and quantified. The raw OCR and ECAR values were normalized for differences in cell number

between individual wells. For ECAR/OCR measurements following 2-DG treatment, explant culture was performed with neural tubes from HH9 embryos in control and 4mM 2-DG conditions. After 18 hours of culture, the explants were dissociated with Accumax, resuspended in 200ul of Seahorse media and plated as described above. Measurements were performed after 1h incubation in a non-CO₂ 37°C incubator.

Immunohistochemistry

For whole-mount immunohistochemistry, embryos were collected at appropriate developmental stages and fixed in 4% PFA-PB for 20 mins at RT. Post fixation, embryos were dissected from the filter paper and washed in TBS containing 0.1% Triton and 1% DMSO (TBTD). Embryos were blocked at RT for 2 hours in TBTD supplemented with 10% donkey serum and incubated in primary antibody diluted in blocking solution, overnight at 4°C. The following primary antibodies were used: rabbit anti-PFKP (Abcam, 1:200), rabbit anti-LDHA (Abcam, 1:200), mouse anti-GAPDH (Millipore, 1:500), mouse anti-TFAP2B (Santa Cruz Biotechnology, 1:500), rabbit anti-Act/YAP1 (Abcam, 1:500), rabbit anti-TFAP2B (Abcam, 1:250), rabbit anti-Sox9 (EMD Millipore, 1:500), mouse anti-E-Cadherin (BD Biosciences, 1:200). Following the primary antibody incubation, embryos were washed, blocked for 30 mins at RT, and stained with appropriate secondary antibodies for 2h at RT. Secondary antibodies used included donkey anti-mouse/rabbit IgG conjugated with Alexa Fluor 488/568/647 or goat anti-mouse Alexa 480/568 (Molecular Probes, 1:3000). Following the secondary antibody step, the embryos were washed, stained with Dapi and post-fixed with 4% PFA for 1h, prior to imaging. Whole-mount images were collected using an upright Zeiss Axio Imager fluorescent microscope and processed as described below.

Cryosectioning

To obtain embryo sections following immunohistochemistry, fixed embryos were washed in 5% sucrose (in PBS) for 3 hours at RT, and in 15% sucrose solution overnight at 4°C. Next, they were incubated in

7.5% porcine gelatin (dissolved in 15% sucrose solution) for 3 hours at 37°C, embedded in silicone molds, snap-frozen in liquid nitrogen and stored at -80°C. 8µM sections were obtained using the CryoStar NX50 (Thermo Fisher). For imaging, the slides were immersed in PBST at 42°C for 15 mins for gelatin removal, washed in PBS and mounted with Vectashield medium containing DAPI (Vector Labs).

Immunohistochemistry of cell suspensions

Immunohistochemistry of dissociated embryos was performed as previously described (Wang et al., 2014). Briefly, the heads of HH8 and HH12 wild type embryos were micro-dissected in Ringers, and dissociated in Accumax Cell dissociation solution (Accutase, SCR006) for 15-20 mins at RT. At least eight HH8 embryo heads and six HH12 heads were used for each antibody staining experiment. Following dissociation, cells were washed in PBS and fixed in 0.4% PFA solution for 15 mins at RT. Post-fixation, cells were permeabilized with PBS+0.3% Triton solution and blocked with 1% BSA solution (in PBS+0.1% Tween 20) for 1h at RT. The cells were subsequently incubated overnight with appropriate dilutions of primary antibody in blocking solution. The antibodies rabbit anti-PFKP, rabbit anti-LDHA, rabbit anti-GLUT3 and rabbit anti-TFAP2B were used at a dilution of 1:200, while mouse anti-GAPDH and mouse anti-TFAP2B were used at a dilution of 1:500. Following primary antibody staining, the cells were washed in PBS+0.1% Tween 20 solution and incubated with appropriate secondary antibodies diluted to a concentration of 1:1000 in blocking solution for 1h at RT. The samples were then washed twice in PBS+0.1% Tween 20 and the antibody staining intensity was measured by FACS analysis using the Attune Nxt flow cytometer at the Cornell Flow facility. The cytometry data was analyzed using the FCS Express 6 software as described below.

Metabolic reporter fluorescence analysis

To measure the ratio of NAD⁺/NADH in the neural crest and neural tube, HH5 embryos were co-electroporated with 1 µg/ul of pCI-Rex-YFP, which was sub-cloned from the original plasmid provided by Vsevolod Belousov (Addgene plasmid #48247) (Bilan et al., 2014). pCI-Rex-YFP was co-transfected with

1.5ug/ul of either neural crest specific enhancer *Tfap2aE1*-mCherry (Bhattacharya et al., 2018) or neural tube specific enhancer *Sox2N2*-mCherry enhancer (Uchikawa et al., 2003). To measure intracellular glucose levels, embryos were electroporated with the glucose reporter Green Glifon50, a gift from Tetsuya Kitaguchi (Addgene plasmid #126206) (Mita et al., 2019), along with the *Tfap2aE1* or *Sox2N2* reporter constructs. Post electroporation, the embryos were incubated at 37°C *ex ovo* until they reached HH8 (~8 hours) or HH10 (~10 hours). Embryo heads were dissected in Ringers and rapidly dissociated in Accumax for 15 mins at RT. Following dissociation, the fluorescence intensity of the Rex:YFP and Green Glifon50 in *Tfap2aE1*-mCherry+ neural crest cells and *Sox2N2*-mCherry+ neural progenitor cells was measured using the Attune NxT Flow Cytometer (Thermo Fisher). Additionally, HH10 embryos electroporated with each metabolic reporter and enhancers construct combination were also fixed in 4% PFA for 2h at RT and further processed for sectioning and imaging. Both the metabolic reporter constructs were subcloned into the pCI-H2B-RFP (Betancur et al., 2010) vector backbone to ensure robust expression in the chick embryo. To confirm that this subcloning does not bias the expression of the constructs, we measured the fluorescence activity of the pCI:EGFP vector in neural crest and neural tube cells of HH8 embryos by flow cytometry analysis. Upon performing this experiment, we observed no significant difference in GFP intensity between the tissues (Fig. S1a).

Validation of metabolic reporters

To validate that the Rex:YFP reporter responds to changes in cellular NAD⁺/NADH ratio, HH8 embryos transfected with pTK-*Tfap2E1*-mCherry and the pCI-Rex-YFP plasmids were incubated on albumin plates containing 10μM pyruvate for 20 mins. The embryos were then quickly dissociated in Accumax and the intensity of Rex-YFP in mCherry+ neural crest cells was measured in control vs pyruvate treated samples by flow cytometry. To confirm the activity of Green Glifon50, HH8 embryos co-electroporated with pTK-*Tfap2E1*-mCherry and the glucose sensor were incubated on albumin plates containing 50μM glucose for 40 mins. Following incubation, the embryos were processed as described above. The

intensity of Green Glifon50 in control and glucose treated *Tfap2aE1+* neural crest cells was assayed by FACS analysis in the Attune NxT cytometer (Fig. S1b-c).

2-NBDG Uptake assay

20ul of 1mM solution of the fluorescent glucose analog 2-NBDG was applied to the dorsal and ventral side of the HH8 and HH10 chick embryos electroporated with the pTK-*Tfap2E1*-mCherry enhancer construct. The embryos were incubated in albumin plates for 2h at 37°C. Following incubation, embryos were dissected and processed for flow cytometry analysis as described above. A subset of the HH10 embryos incubated with the 2-NBDG reporter was also fixed in 4% PFA for 2h at RT and further processed for sectioning and imaging.

Luciferase Assays

To quantitatively measure Yap/Tead signaling activity *in vivo*, HH5 avian embryos were co-electroporated with HOP-flash plasmid or its mutated version HIP flash (Addgene plasmid # 83467) (Kim and Gumbiner, 2015) and the pRL-TK control reporter vector (Promega, #E2231). The embryos were then cultured *ex ovo* until HH9. The two dorsal neural folds of each embryo were micro-dissected, and one explant was cultured in control media while the other was cultured in 2-DG containing media for 18h, in white, clear-bottom 96 well plates. Following explant culture, a luciferase assay was performed using the Dual-Luciferase Assay kit (Promega, #E1910) according to the manufacturer's protocol. Post luminometer reading, firefly luciferase values were normalized to Renilla luciferase measurements to account for differences in transfection efficiency and number of cells. Normalized luciferase values were compared between control and 2-DG treated explants obtained from the same embryo.

Live imaging of neural crest explants

HH5 embryos were co-electroporated with pCI-H2B-RFP and pCAG-mGFP-Actin (Addgene plasmid # 21948)(Murakoshi et al., 2008) and cultured *ex ovo* until HH9-. The dorsal neural fold of each embryo

was micro-dissected and plated on fibronectin-coated 4-well chamber slides, with each well containing either control DMEM or DMEM+2-DG. Before imaging, the explants were incubated for 3h until they were completely attached to the plates. Live imaging was performed using the inverted Andor/Olympus Spinning Disk Confocal microscope at the BRC Imaging Facility, Cornell University. The chamber slides were maintained at 37°C at a 5% CO₂ condition and imaged for a total of 12h (1 frame/12mins) using the 10X objective. For all live imaging experiments, control and 2-DG treated explants were imaged simultaneously. Post imaging, nuclear tracking analysis was performed using the TrackMate plug-in in Fiji (Tinevez et al., 2017).

Immunohistochemistry of explant cultures and proximity ligations assay (PLA)

For immunohistochemistry experiments, explants were fixed with 4% PFA for 10 mins at RT, followed by permeabilization with 0.1% NP40 for 30 mins at 37°C. Next, the cells were blocked with 1% BSA in PBS for 30 mins at 37°C, prior to incubation with the primary antibody for 1h at 37°C. Primary antibodies were prepared in blocking solution, and were used in the following dilutions: mouse anti-E-cadherin antibody (BD Biosciences) (1:200), mouse anti-Paxillin antibody (BD Biosciences) (1:200), and rabbit anti-Active Yap1 (Abcam)(1:200), anti-pH3(S10) (Abcam, 1:200), anti-Caspase3 (R&D Systems, 1:200). Following primary antibody incubation, the explants were washed 4X times in PBS for 10 mins and incubated in a secondary antibody cocktail for 30 mins at 37°C. Finally, cells were washed 3X times in PBS and stained with DAPI or Phalloidin for 20 mins at RT. Post antibody staining, imaging was performed using Andor/Olympus Spinning Disk Confocal microscope at BRC facility, Cornell University.

For Proximity Ligation Assay (PLA), explants were fixed and permeabilized as described above. The primary antibodies used were: Mouse Anti-YAP1 (DSHB) (1:5) and Rabbit Anti-TEAD1 (Abcam) (1:200). Following primary antibody incubation, the remaining steps of PLA were performed using reagents from the Duolink PLA detection kit (Sigma Aldrich, DUO92101) according to the manufacturer's protocol.

Quantitative reverse transcription PCR (RT-PCR)

To quantify changes in gene expression following 2-DG/verteporfin treatments, we compared neural crest explants (obtained from the two sides of the same embryo), cultured in control vs. drug treatment media. The explants were lysed in the lysis buffer from Power SYBR Green Cells-to-CT Kit (Thermo Fisher). RNA extraction and cDNA preparation were performed according to the suggested protocol. RT-PCR was performed using the Power Sybr Green PCR master mix (Thermo Fisher, 4368577) in an ABI viera7 RT-PCR machine. Ct values of all genes were normalized to reference gene *HPRT* and presented as a fold-change of the control sample (ddCT).

Nanostring analysis

To access the global effect of glycolytic inhibition on the neural crest gene regulatory network, Nanostring analysis was performed for control and 2-DG treated neural crest explants. As described above, pairs of dorsal neural folds were dissected from individual HH9 embryos, and one neural tube was cultured in control DMEM, while the other explant was cultured in DMEM+2-DG for 18 hrs. Following culture, the cells were lysed in 5ul of Cell-to-CT lysis buffer (Cell-to-CT kit, Thermo Fisher). The cell lysate was hybridized to a Nanostring Probe Set containing 200 probes for neural crest, placode and neural genes (Bhattacharya et al., 2018), at 65°C for 16 hrs. Nanostring data was analyzed using the nSolver software package.

Cloning of expression vectors

The TEA-VPR expression vector is a fusion construct of the sequence of the TEA domain of the avian TEAD1 protein and the sequence of VPR activating domains. The sequence of the TEA domain was amplified from a Gallus gallus HH9 cDNA library, and the VPR sequence was amplified from PB-TRE-dcas9-VPR vector, a gift from George Church (Addgene plasmid # 63800) (Chavez et al., 2015). The two sequences were cloned in the pCI-H2B-RFP vector (Betancur et al., 2010), upstream of the H2B-RFP coding sequence. The YAP-5SA/S94A was a gift from Kunliang Guan (Addgene plasmid # 33103) (Zhao

et al., 2008). The HOP-GFP vector was derived from the HOP-Flash vector by swapping the Firefly luciferase sequence with the EGFP sequence amplified from pTK-EGFP vector backbone.

CUT&RUN

Neural folds were dissected from HH9+ embryos (n=20 per CUT&RUN experiment). Cells were dissociated in Accumax (Accutase, SCR006) for 20 minutes at RT under mild agitation. CUT&RUN experiments were carried out as described previously described (Rothstein and Simoes-Costa, 2019). Briefly, cells were immobilized on BioMag Plus Concanavalin A magnetic beads (Bangs Laboratories, BP531) and incubated with rabbit anti-Active-Yap1 (Abcam, ab205270) (1:50) or anti-Histone H3 (acetyl K27) (Abcam, ab177178) (1:50) antibody overnight at 4°C. After washing away unbound antibody, protein A-MNase was added to a final concentration of 700ng/mL and incubated for 1 h at 4°C. Cells were cooled to 0°C and CaCl₂ was added to a final concentration of 2 mM to activate the MNase enzyme. MNase digestion was performed for 45 minutes and terminated by the addition of 2XSTOP buffer containing heterologous *Saccharomyces cerevisiae* spike-in DNA at a concentration of 2pg/mL. The protein-DNA complexes were released by centrifugation and digested with proteinase K for 10 minutes at 70°C. DNA fragments were isolated via phenol-chloroform extraction and ethanol precipitation. Protein A-MNase and spike-in DNA were kindly provided by Dr. Steven Henikoff (Skene and Henikoff, 2017).

CUT&RUN library preparation

CUT&RUN libraries were prepared using the NEBNext Ultra II DNA Library Prep Kit (New England Biolabs, E7645) using the suggested protocol. Fragment analysis was performed with ABI 3730xl DNA Analyzer to perform quality control for the libraries. Equimolar concentrations of the libraries were pooled using the KAPA Library Quantification Kit - ROX Low (Roche, 07960336001) and sequenced with paired-end 37bp reads on an Illumina NextSeq500 instrument.

Enhancer analysis

Yap-bound putative enhancers in the *loci* of to neural crest genes were amplified from HH10 chicken gDNA and cloned in the pTK-EGFP (Uchikawa et al., 2003). To assess enhancer activity, HH5 embryos were co-electroporated with *Tfap2aE1-mCherry* and the pTK-EGFP constructs, and incubated until HH9. To measure the activity of each enhancer in mCherry+ neural crest cells, embryos were dissociated and analyzed in the Attune NxT cytometer. These values were plotted to determine the specificity and strength of each enhancer in delaminating neural crest cells.

Functional analysis of enhancer activity

To assess the effect of 2-DG treatment on the activity of neural crest specific enhancers, HH5 avian embryos were electroporated with *Ets1ECR1*, *Sox9E1* or *Zeb2E1* along with control plasmid pCI-Cherry-Ras. The embryos were cultured *ex ovo* until HH9. The two dorsal neural folds of each embryo were micro-dissected, and one explant was cultured in control media while the other was cultured in 2-DG containing media in chamber slides. Following 18h of incubation, the explants were fixed in 4% PFA for 10 mins at RT, stained with DAPI and imaged using an inverted Olympus Spinning Disk Confocal microscope.

Flow cytometry data analysis

Flow cytometry data obtained from the metabolic reporter assays and the antibody-staining experiments were analyzed using the FCS Express 6 software package. After assigning appropriate forward scatter and side scatter gates, the neural crest and/or neural populations were identified as cells having enhancer reporter fluorescence or antibody staining intensity above the threshold that was set based on the negative and single-color controls. Next, we obtained the intensity values of individual cells in the population of interest (neural crest, neural or whole embryo). To compare samples and replicates across different experiments, the intensity values were log-transformed and normalized by the variance of each sample.

Live imaging analysis

Analysis of live-cell imaging experiments was performed using the Trackmate plugin in ImageJ. To track the movement of individual cell nuclei, the program was provided with 60 frames/movie of the H2B-RFP channel (frames were captured every twelve minutes). Minimum threshold values of track duration and track displacement were set in order to restrict the analysis to (i) nuclei that could be tracked for the entire duration of the movie and (ii) cells with an average displacement that was above background levels. From this analysis, we obtained values for total displacement (in pixels) and the mean speed of each nucleus (in pixels/frame) for the two conditions. Furthermore, we measured the maximum distance each nucleus travels in a straight line before changing direction (maximum distance traveled). The values for these parameters were obtained for all tracked nuclei from a total of 5 movies/condition. We used these values to plot the graphs depicted in Fig. 3 and to perform the corresponding statistical analysis.

CUT&RUN data analysis

Paired-end sequencing reads from the CUT&RUN libraries were trimmed using Cutadapt (Martin, 2011). Reads were filtered for those with a minimum length of 25bp or longer and aligned to the reference chicken Galgal5 assembly using Bowtie2 (Langmead and Salzberg, 2012). Picard MarkDuplicates tool was used to mark duplicate reads and BAM files were filtered with SAMtools to discard unmapped reads (those that were not the primary alignment, reads failing platform/vendor quality checks, and PCR/optical duplicates (-f 2 -F 1804)). Peak calling was performed using MACS version 2.1 with a p-value cutoff of 0.01. Representative heatmaps showing the H3K27ac and ATAC-seq signal at YAP1 bound peaks were generated using the deepTools2 package (Ramirez et al., 2016). We also utilized this program to determine the correlation between the CUT&RUN replicates for ActYAP1. To identify transcription factor motifs in the genomic regions occupied by Yap1, motif enrichment analysis was performed using the TRAP web tools (Thomas-Chollier et al., 2011) on 250bp sequences flanking the summit of each peak on either side. We also used the Bioconductor package CHIPseeker to annotate the Yap1 bound peaks and determine their occupancy genome-wide compared to a random peak-set (Yu et al., 2015). Lastly,

to identify putative genes regulated by Yap/Tead signaling, the closest gene was assigned to each peak using the Bedtools closest function. The GO-category analysis was performed on these putative Yap targets by utilizing the ClusterProfiler package (Yu et al., 2012) to assay for over-represented Biological Processes (BP) having a p-value cut off of 0.01 post Bonferonni correction.

Additional Statistical Analyses

The metabolic assays including lactate production and Seahorse analysis were performed with 3 independent biological replicates (each iteration having two technical replicates). The glucose uptake assay was performed with eight individual neural crest and lateral ectoderm replicates. The flow cytometry analysis for metabolic reporter activity included at least six embryos per stage per reporter construct/enhancer, and the data from multiple embryos were pooled to plot the graphs. The immunohistochemistry experiments in the embryo were performed with at least 5 embryos, and the functional experiments following 2-DG and 3-BP drug treatment included six control and six treated embryos. For all functional experiments utilizing neural crest explants at least 10 explants were analyzed per condition in each experiment. The Nanostring experiment was performed with three replicates of control and 2-DG treated neural fold explants. The CUT&RUN experiment for ActYAP1 was repeated three times, while CUT&RUN for H3K27ac was replicated twice. The n and p values of all quantitative experiments are listed in Table S1. Student's t-test (one-tailed) was performed to calculate p-values for functional experiments and $p < 0.05$ were considered to be significant. Wilcoxon test was used to calculate p-values for all flow cytometry data analysis and for comparing the distributions of mean speed and total distance traveled (as obtained from the live imaging experiments), to account for the non-parametric distribution of these values.

Data and Software availability: The CUT&RUN datasets for ActYAP1 and H3K27ac in chick embryos have been deposited to the Gene Expression Omnibus GSE142101.

Supplementary Figures

Figure S1

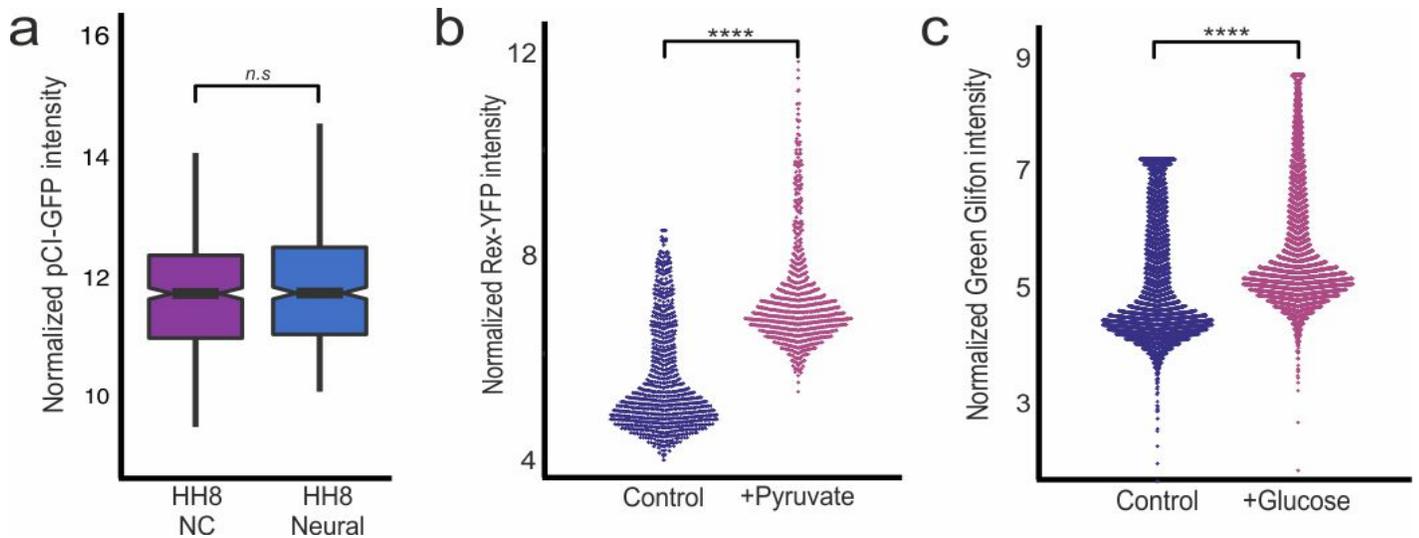


Figure S1. (Related to Figure 2) Additional controls for the Rex-YFP and the *Green Glifon 50* metabolic reporters in the avian embryo

(a) Quantification of ubiquitous pCI-GFP plasmid expression in *Tfap2E1*⁺ neural crest cells and *Sox2N2*⁺ neural cells of HH8 embryos. (b-c) Validation of metabolic reporter activity. Swarm plot depicting the change in the fluorescence output of the metabolic reporters Rex-YFP (b) and Green Glifon50 (c) in *Tfap2aE1*⁺ neural crest cells upon treatment of the embryos with exogenous pyruvate and glucose respectively. In all plots, intensity values were log-transformed and normalized by the variance of each sample. Mann-Whitney test was performed to assess statistical significance. NC: Neural crest.

Figure S2

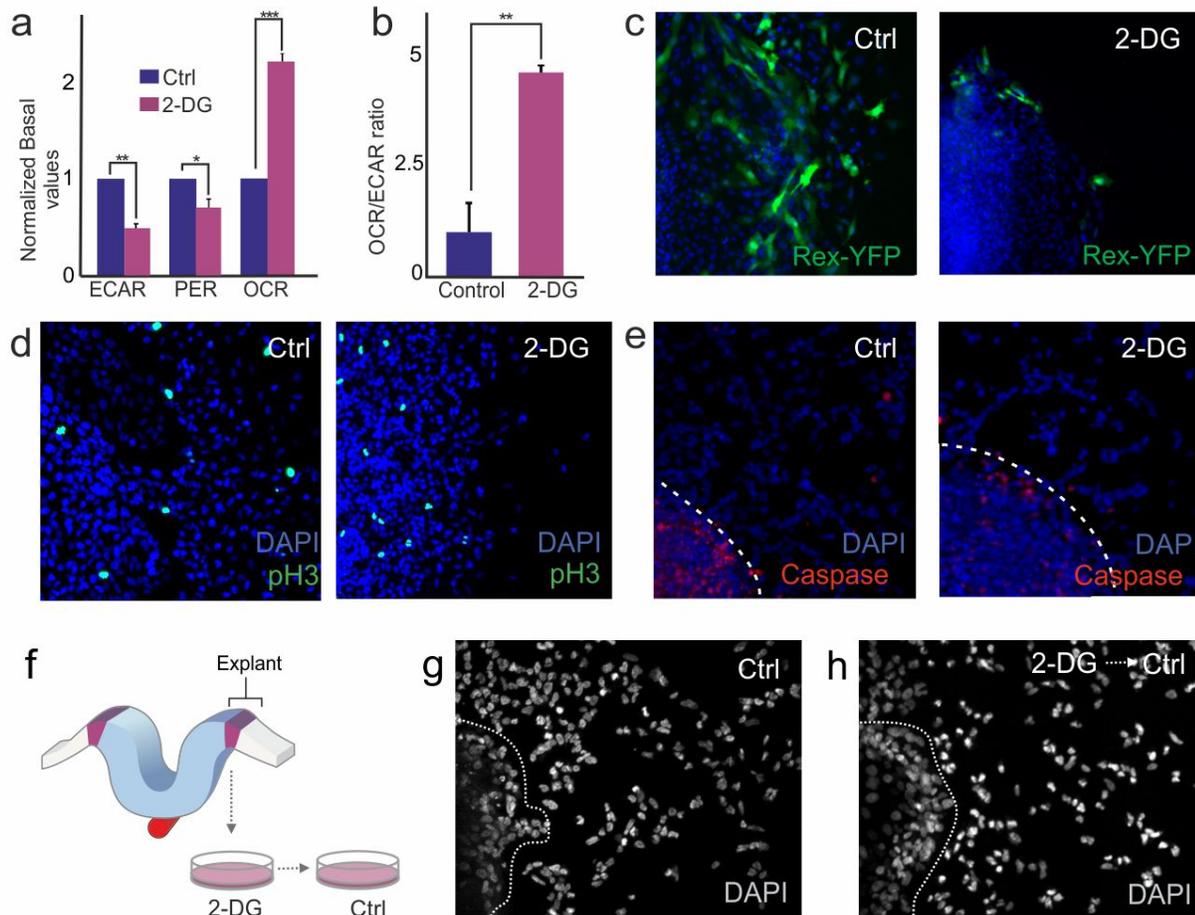


Figure S2. (Related to Figure 3) Glycolytic inhibition alters neural crest metabolism and reduces migration

(a-c) Inhibition of glycolysis with 2-DG causes neural crest cells to engage in aerobic respiration. Analysis of basal Extracellular Acidification Rate (ECAR), Proton Efflux Rate (PER) and Oxygen Consumption Rate (OCR) in HH9 neural crest explants treated with 2-DG (a). Glycolytic inhibition significantly alters neural crest metabolism and results in increased OCR/ECAR ratio in neural crest cells following 2-DG treatment (b). Fluorescence of Rex-YFP reporter construct is decreased in 2-DG treated neural crest explants, indicating a lower intracellular NAD⁺/NADH ratio upon glycolytic inhibition(c). **(d-e)** Immunofluorescence for proliferation marker pH3 (d) and cell death marker Caspase-3 (e) reveal no significant difference between control and 2-DG treated explants. **(f-g)** The short-term effect of 2-DG treatment on neural crest migration is reversible. Neural crest explants treated with 2-DG for 5 hours, followed by incubation in control media, were able to overcome 2-DG inhibition, as neural crest cells engaged in migratory behavior comparable to control explants (g). Dotted lines show the position of the neural fold explant. 2-DG: 2-deoxy-glucose.

Figure S3

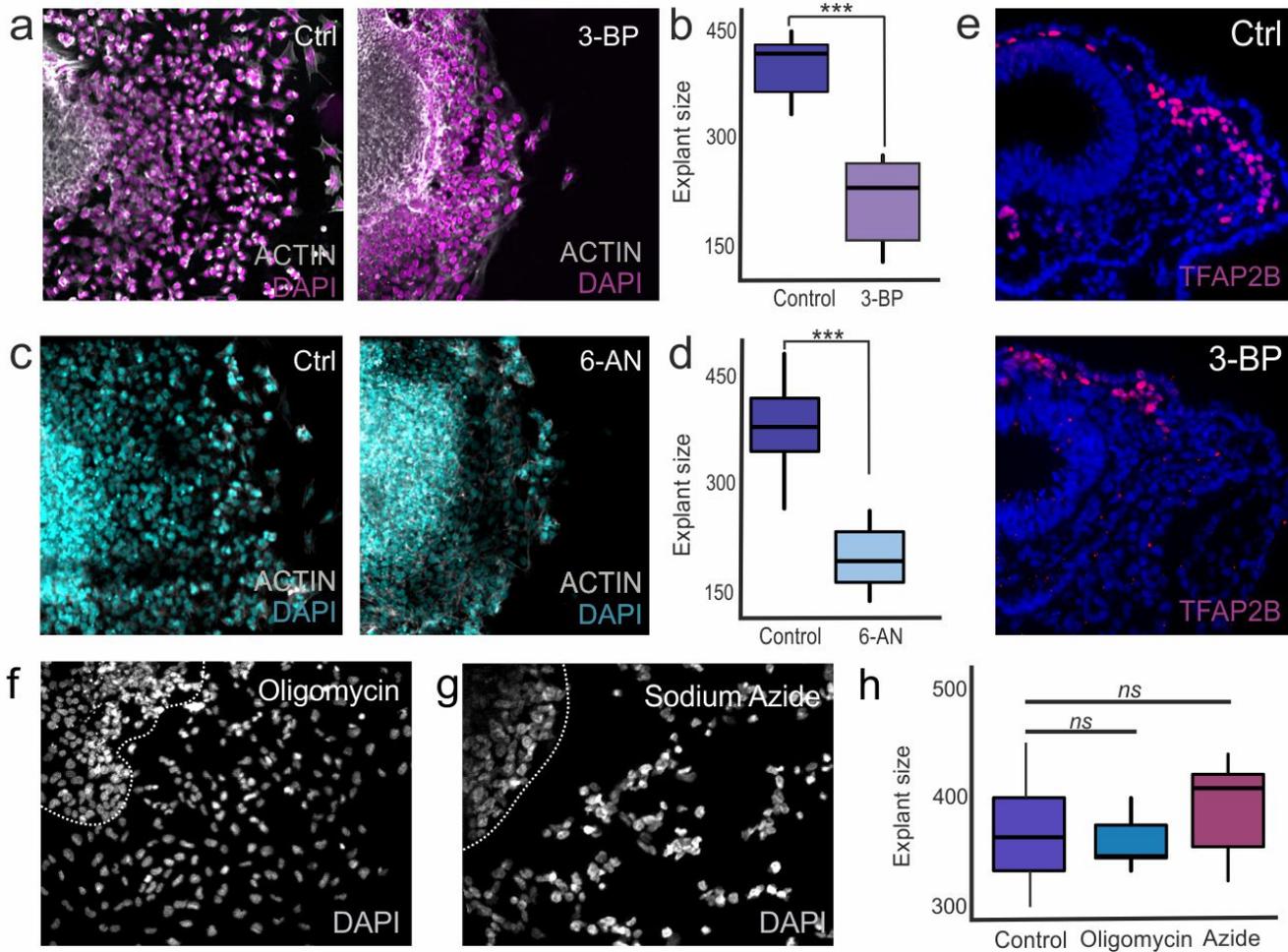


Figure S3. (Related to Figure 3) Neural crest migration is not dependent upon aerobic respiration

(a-d) Representative images of control and neural crest explants treated with glycolytic inhibitors 3-bromopyruvate (3-BP) (a,b) and 6-aminonicotinamide (6-AN) (c,d). Similar to the effect of 2-DG, treatment with both these inhibitors reduce neural crest migration, resulting in smaller explant size (b,d). (e) Transverse sections of immunohistochemistry for neural crest marker TFAP2B show reduced migration of neural crest cells in 3-BP treated embryos. (f-h) The inhibition of the Electron Transport Chain (ETC) does not affect neural crest migration in the explant system. Treatment of explants with ETC inhibitors oligomycin (f) or sodium azide (g) does not significantly affect neural crest explant size (h). 2-DG: 2-deoxy-glucose, 3-BP: 3-bromopyruvate, 6-NA: 6-aminonicotinamide.

Figure S4

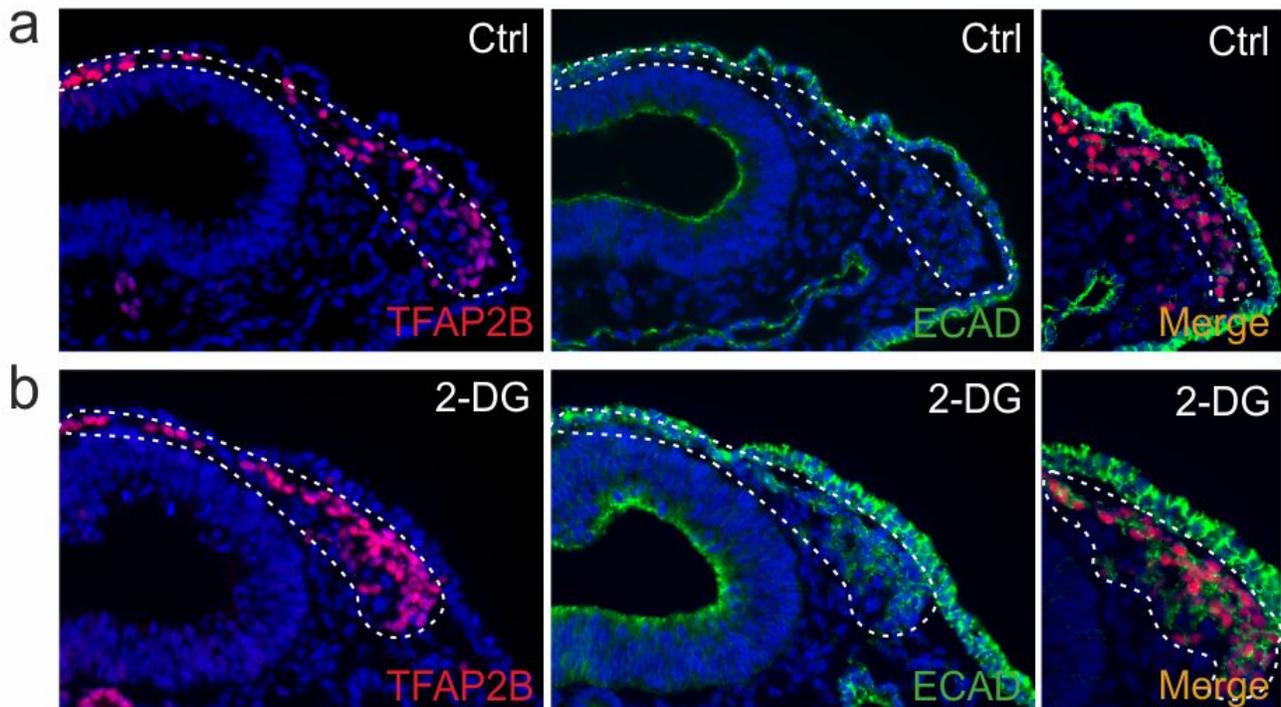


Figure S4. (Related to Figure 4) Enhanced glycolysis is required for downregulation of E-cadherin in neural crest cells

(a-b) Transverse sections showing immunostaining for E-cadherin (ECAD) and TFAP2B in control and 2-DG treated HH12 embryos. Inhibition of glycolysis *in vivo* prevents complete neural crest EMT, as evidenced by the persistence of the epithelial marker E-cadherin in TFAP2B+ migratory cells (dotted lines).

Figure S5

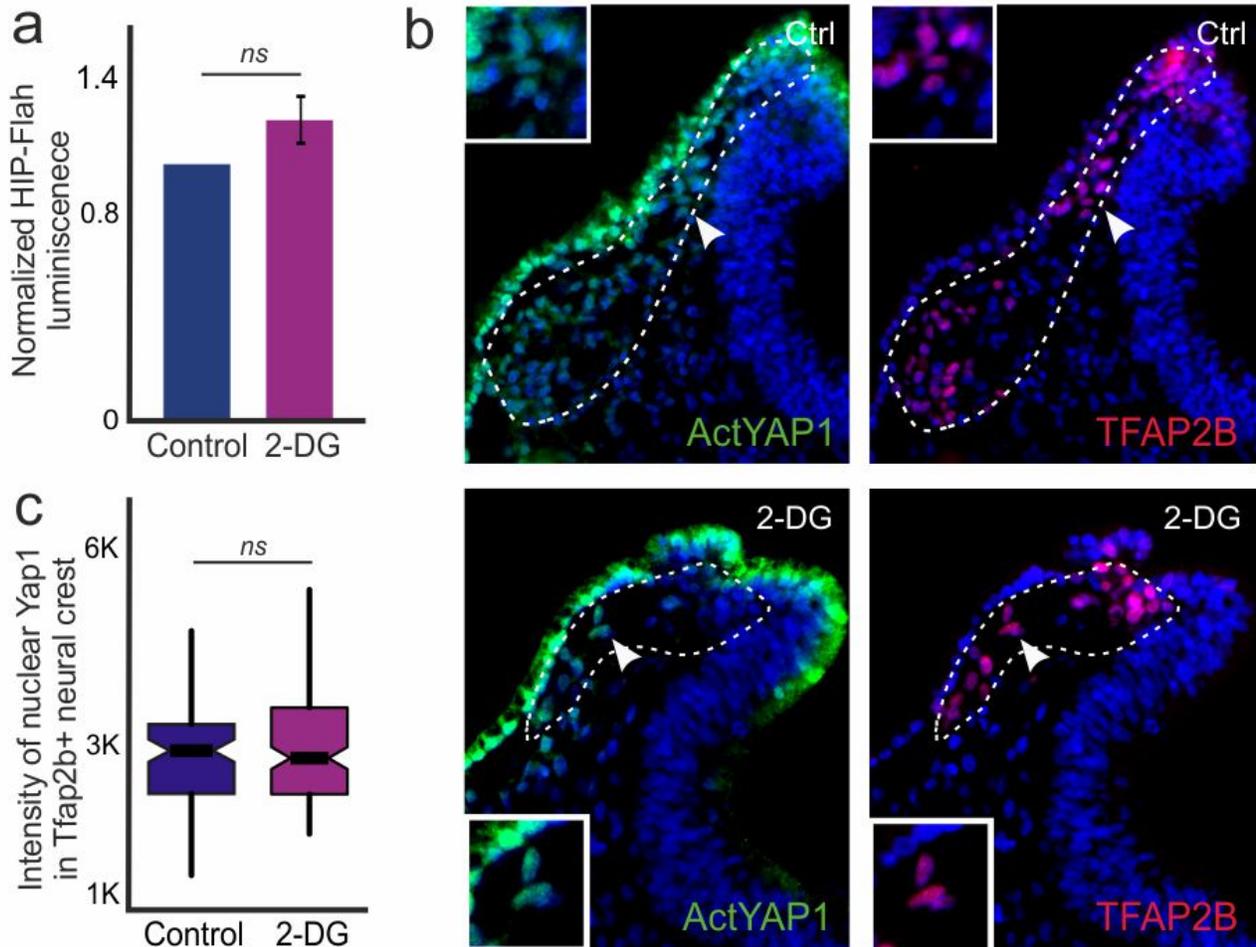


Figure S5. (Related to Figure 5) Treatment with 2-DG does not prevent translocation of YAP to the nuclei of neural crest cells

(a) Luminescence assay with the mutant Yap/Tead signaling reporter HIP-Flash reveals that it is not affected by 2-DG treatment in neural crest explants, unlike the HOP-Flash construct (Fig 5e). **(b-c)** Immunohistochemistry for active YAP1 (actYAP1) and neural crest marker TFAP2B show that YAP1 translocates to the nucleus even in 2-DG treated embryos (insets show cells indicated by white arrowheads) **(b)**. Quantitation of nuclear fluorescence of YAP1 in TFAP2B+ cells shows no significant difference between control and 2-DG-treated embryos **(c)**.

Figure S6

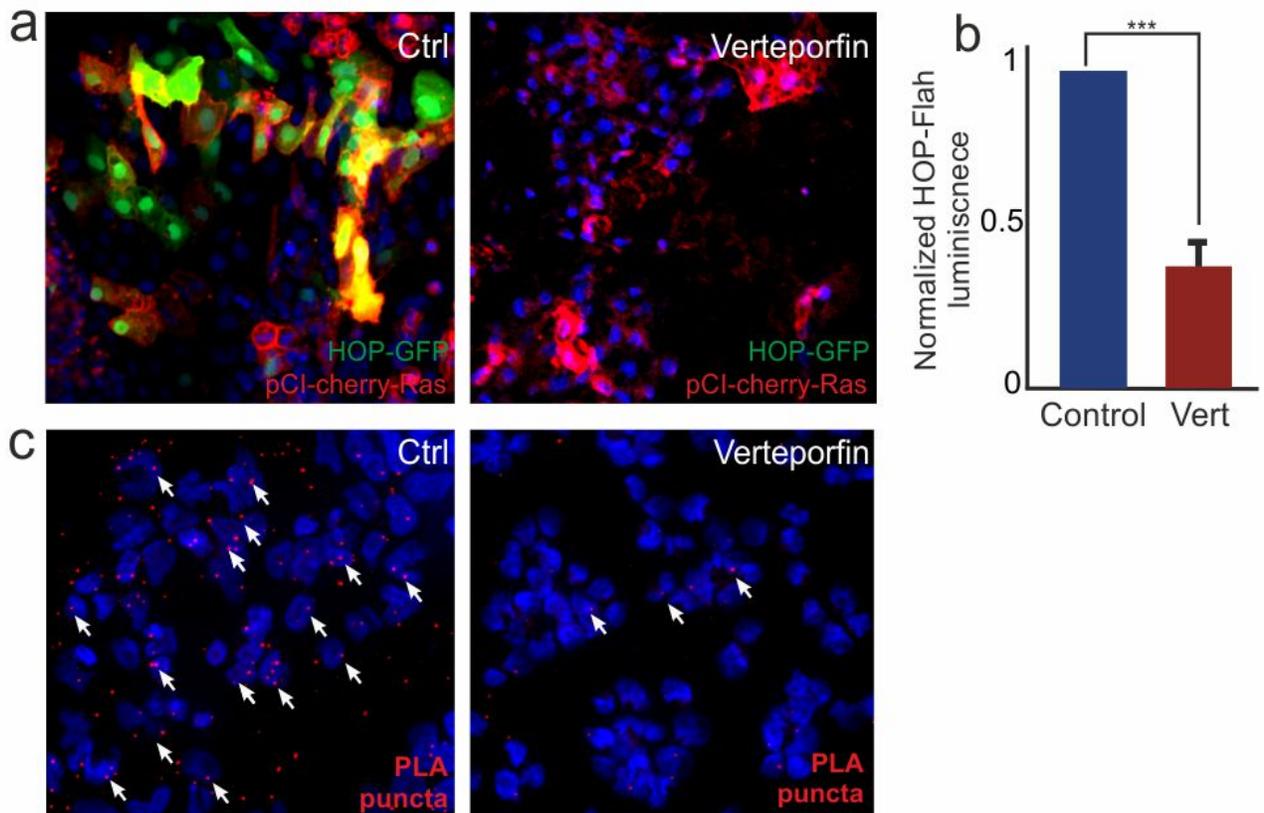


Figure S6. (Related to Figure 6) Verteporfin inhibits Yap/Tead signaling in neural crest explants

(a-b) Neural crest cells transfected with Yap/Tead signaling reporter HOP-GFP have decreased reporter expression following Verteporfin treatment **(a)**. This observation is confirmed by luciferase assay measuring the activity of HOP-Flash (the luciferase version of the reporter) in control and verteporfin treated explants **(b)**. **(c)** Images of Proximity Ligation Assay (PLA) for YAP1 and TEAD1 in control and Verteporfin treated explants. PLA puncta were found to be significantly reduced on Verteporfin treatment, indicating disruption of YAP/TEAD interaction. Vert: Verteporfin

Figure S7

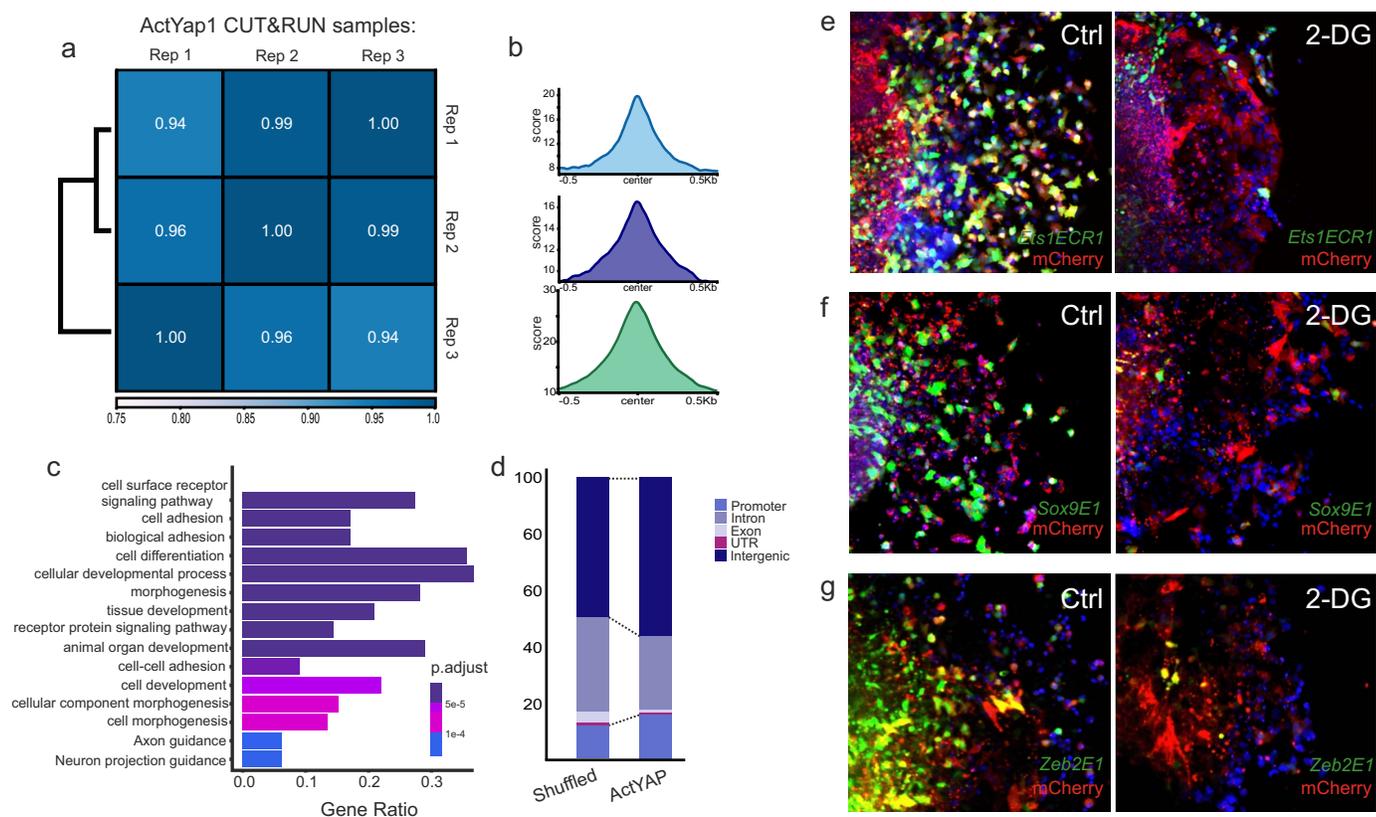


Figure S7. (Related to Figure 7) Correlation analysis of YAP1 CUT&RUN experiments and additional enhancer assays

Pairwise Pearson analysis (**a**) and summit plots (**b**) from the active-YAP1 CUT&RUN experiments show a high correlation between the biological replicates. (**c**) Top 15 biological processes identified by Gene ontology analysis of genes closest to the YAP1-associated peaks. (**d**) Feature annotation of YAP1 bound peaks reveals an increased binding at intergenic and promoter regions in the chicken genome when compared to a random set of genomic regions. (**e-g**) Effect of glycolytic inhibition on activity of Yap/Tead regulated neural crest enhancers. Images of control and 2-DG treated explants expressing the neural crest enhancers *Ets1ECR1* (**e**), *Sox9E1* (**f**) and *Zeb2E1* (**g**) and control plasmid pCI-Cherry-Ras. The activity of all three Yap/Tead regulatory regions is significantly reduced upon 2-DG treatment.

References

- Acloque, H., Adams, M.S., Fishwick, K., Bronner-Fraser, M., and Nieto, M.A. (2009). Epithelial-mesenchymal transitions: the importance of changing cell state in development and disease. *The Journal of clinical investigation* *119*, 1438-1449.
- Azzarelli, R., Simons, B.D., and Philpott, A. (2018). The developmental origin of brain tumours: a cellular and molecular framework. *Development* *145*.
- Bailey, T.L., Boden, M., Buske, F.A., Frith, M., Grant, C.E., Clementi, L., Ren, J., Li, W.W., and Noble, W.S. (2009). MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res* *37*, W202-208.
- Barban, S., and Schulze, H.O. (1961). The effects of 2-deoxyglucose on the growth and metabolism of cultured human cells. *J Biol Chem* *236*, 1887-1890.
- Barembaum, M., and Bronner, M.E. (2013). Identification and dissection of a key enhancer mediating cranial neural crest specific expression of transcription factor, *Ets-1*. *Dev Biol* *382*, 567-575.
- Betancur, P., Bronner-Fraser, M., and Sauka-Spengler, T. (2010). Genomic code for Sox10 activation reveals a key regulatory enhancer for cranial neural crest. *Proceedings of the National Academy of Sciences of the United States of America* *107*, 3570-3575.
- Bhattacharya, D., Rothstein, M., Azambuja, A.P., and Simoes-Costa, M. (2018). Control of neural crest multipotency by Wnt signaling and the Lin28/let-7 axis. *Elife* *7*.
- Bilan, D.S., Matlashov, M.E., Gorokhovatsky, A.Y., Schultz, C., Enikolopov, G., and Belousov, V.V. (2014). Genetically encoded fluorescent indicator for imaging NAD(+)/NADH ratio changes in different cellular compartments. *Biochim Biophys Acta* *1840*, 951-957.
- Bulusu, V., Prior, N., Snaebjornsson, M.T., Kuehne, A., Sonnen, K.F., Kress, J., Stein, F., Schultz, C., Sauer, U., and Aulehla, A. (2017). Spatiotemporal Analysis of a Glycolytic Activity Gradient Linked to Mouse Embryo Mesoderm Development. *Developmental cell* *40*, 331-341 e334.
- Cha, Y., Han, M.J., Cha, H.J., Zoldan, J., Burkart, A., Jung, J.H., Jang, Y., Kim, C.H., Jeong, H.C., Kim, B.G., *et al.* (2017). Metabolic control of primed human pluripotent stem cell fate and function by the miR-200c-SIRT2 axis. *Nat Cell Biol* *19*, 445-456.
- Chapman, S.C., Collignon, J., Schoenwolf, G.C., and Lumsden, A. (2001). Improved method for chick whole-embryo culture using a filter paper carrier. *Developmental dynamics : an official publication of the American Association of Anatomists* *220*, 284-289.
- Chavez, A., Scheiman, J., Vora, S., Pruitt, B.W., Tuttle, M., E, P.R.I., Lin, S., Kiani, S., Guzman, C.D., Wiegand, D.J., *et al.* (2015). Highly efficient Cas9-mediated transcriptional programming. *Nat Methods* *12*, 326-328.
- Chen, X., Liu, J., and Chen, S.Y. (2013). Over-expression of Nrf2 diminishes ethanol-induced oxidative stress and apoptosis in neural crest cells by inducing an antioxidant response. *Reprod Toxicol* *42*, 102-109.
- Coles, E.G., Taneyhill, L.A., and Bronner-Fraser, M. (2007). A critical role for Cadherin6B in regulating avian neural crest emigration. *Developmental biology* *312*, 533-544.
- Dooley, C.M., Wali, N., Sealy, I.M., White, R.J., Stemple, D.L., Collins, J.E., and Busch-Nentwich, E.M. (2019). The gene regulatory basis of genetic compensation during neural crest induction. *PLoS Genet* *15*, e1008213.
- Enzo, E., Santinon, G., Pocaterra, A., Aragona, M., Bresolin, S., Forcato, M., Grifoni, D., Pession, A., Zanconato, F., Guzzo, G., *et al.* (2015). Aerobic glycolysis tunes YAP/TAZ transcriptional activity. *EMBO J* *34*, 1349-1370.
- Epstein, T., Gatenby, R.A., and Brown, J.S. (2017). The Warburg effect as an adaptation of cancer cells to rapid fluctuations in energy demand. *PLoS One* *12*, e0185085.
- Feng, J., Gou, J., Jia, J., Yi, T., Cui, T., and Li, Z. (2016). Verteporfin, a suppressor of YAP-TEAD complex, presents promising antitumor properties on ovarian cancer. *Onco Targets Ther* *9*, 5371-5381.

- Gabbay-Benziv, R., Reece, E.A., Wang, F., and Yang, P. (2015). Birth defects in pregestational diabetes: Defect range, glycemic threshold and pathogenesis. *World J Diabetes* 6, 481-488.
- Greenwood, C., Ruff, D., Kirvell, S., Johnson, G., Dhillon, H.S., and Bustin, S.A. (2015). Proximity assays for sensitive quantification of proteins. *Biomol Detect Quantif* 4, 10-16.
- Hall, A., Meyle, K.D., Lange, M.K., Klima, M., Sanderhoff, M., Dahl, C., Abildgaard, C., Thorup, K., Moghimi, S.M., Jensen, P.B., *et al.* (2013). Dysfunctional oxidative phosphorylation makes malignant melanoma cells addicted to glycolysis driven by the (V600E)BRAF oncogene. *Oncotarget* 4, 584-599.
- Hamburger, V., and Hamilton, H.L. (1951). A series of normal stages in the development of the chick embryo. *Journal of morphology* 88, 49-92.
- Hansen, C.G., Moroishi, T., and Guan, K.L. (2015). YAP and TAZ: a nexus for Hippo signaling and beyond. *Trends Cell Biol* 25, 499-513.
- Hindley, C.J., Condurat, A.L., Menon, V., Thomas, R., Azmitia, L.M., Davis, J.A., and Pruszak, J. (2016). The Hippo pathway member YAP enhances human neural crest cell fate and migration. *Sci Rep* 6, 23208.
- Houghton, F.D., Thompson, J.G., Kennedy, C.J., and Leese, H.J. (1996). Oxygen consumption and energy metabolism of the early mouse embryo. *Mol Reprod Dev* 44, 476-485.
- Hsu, P.P., and Sabatini, D.M. (2008). Cancer cell metabolism: Warburg and beyond. *Cell* 134, 703-707.
- Johnson, V.P., Swayze, V.W., II, Sato, Y., and Andreasen, N.C. (1996). Fetal alcohol syndrome: craniofacial and central nervous system manifestations. *Am J Med Genet* 61, 329-339.
- Kaufman, C.K., Mosimann, C., Fan, Z.P., Yang, S., Thomas, A.J., Ablain, J., Tan, J.L., Fogley, R.D., van Rooijen, E., Hagedorn, E.J., *et al.* (2016). A zebrafish melanoma model reveals emergence of neural crest identity during melanoma initiation. *Science* 351, aad2197.
- Kaushik, D.K., Bhattacharya, A., Mirzaei, R., Rawji, K.S., Ahn, Y., Rho, J.M., and Yong, V.W. (2019). Enhanced glycolytic metabolism supports transmigration of brain-infiltrating macrophages in multiple sclerosis. *The Journal of clinical investigation* 129, 3277-3292.
- Kim, N.G., and Gumbiner, B.M. (2015). Adhesion to fibronectin regulates Hippo signaling via the FAK-Src-PI3K pathway. *J Cell Biol* 210, 503-515.
- Kishore, M., Cheung, K.C.P., Fu, H., Bonacina, F., Wang, G., Coe, D., Ward, E.J., Colamatteo, A., Jangani, M., Baragetti, A., *et al.* (2017). Regulatory T Cell Migration Is Dependent on Glucokinase-Mediated Glycolysis. *Immunity* 47, 875-889 e810.
- Kleber, M., Lee, H.Y., Wurdak, H., Buchstaller, J., Riccomagno, M.M., Ittner, L.M., Suter, U., Epstein, D.J., and Sommer, L. (2005). Neural crest stem cell maintenance by combinatorial Wnt and BMP signaling. *J Cell Biol* 169, 309-320.
- Krisher, R.L., and Prather, R.S. (2012). A role for the Warburg effect in preimplantation embryo development: metabolic modification to support rapid cell proliferation. *Mol Reprod Dev* 79, 311-320.
- Laforgia, N., Di Mauro, A., Favia Guarnieri, G., Varvara, D., De Cosmo, L., Panza, R., Capozza, M., Baldassarre, M.E., and Resta, N. (2018). The Role of Oxidative Stress in the Pathomechanism of Congenital Malformations. *Oxid Med Cell Longev* 2018, 7404082.
- Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9, 357-359.
- Lu, J. (2019). The Warburg metabolism fuels tumor metastasis. *Cancer Metastasis Rev* 38, 157-164.
- Lu, J., Tan, M., and Cai, Q. (2015). The Warburg effect in tumor progression: mitochondrial oxidative metabolism as an anti-metastasis mechanism. *Cancer Lett* 356, 156-164.
- Maguire, L.H., Thomas, A.R., and Goldstein, A.M. (2015). Tumors of the neural crest: Common themes in development and cancer. *Dev Dyn* 244, 311-322.
- Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnetjournal* 17, 10-12.
- Meng, Z., Moroishi, T., and Guan, K.L. (2016). Mechanisms of Hippo pathway regulation. *Genes Dev* 30, 1-17.

- Mita, M., Ito, M., Harada, K., Sugawara, I., Ueda, H., Tsuboi, T., and Kitaguchi, T. (2019). Green Fluorescent Protein-Based Glucose Indicators Report Glucose Dynamics in Living Cells. *Anal Chem* 91, 4821-4830.
- Morgan, S.C., Relaix, F., Sandell, L.L., and Loeken, M.R. (2008). Oxidative stress during diabetic pregnancy disrupts cardiac neural crest migration and causes outflow tract defects. *Birth Defects Res A Clin Mol Teratol* 82, 453-463.
- Moussaieff, A., Rouleau, M., Kitsberg, D., Cohen, M., Levy, G., Barasch, D., Nemirovski, A., Shen-Orr, S., Laevsky, I., Amit, M., *et al.* (2015). Glycolysis-mediated changes in acetyl-CoA and histone acetylation control the early differentiation of embryonic stem cells. *Cell Metab* 21, 392-402.
- Murakoshi, H., Lee, S.J., and Yasuda, R. (2008). Highly sensitive and quantitative FRET-FLIM imaging in single dendritic spines using improved non-radiative YFP. *Brain Cell Biol* 36, 31-42.
- Ngo, H., Tortorella, S.M., Ververis, K., and Karagiannis, T.C. (2015). The Warburg effect: molecular aspects and therapeutic possibilities. *Mol Biol Rep* 42, 825-834.
- Oginuma, M., Moncuquet, P., Xiong, F., Karoly, E., Chal, J., Guevorkian, K., and Pourquie, O. (2017). A Gradient of Glycolytic Activity Coordinates FGF and Wnt Signaling during Elongation of the Body Axis in Amniote Embryos. *Dev Cell* 40, 342-353 e310.
- Philips, F.S. (1941). The oxygen consumption of the early chick embryo at various stages of development. *J Exp Zool* 86, 257-289.
- Ramirez, F., Ryan, D.P., Gruning, B., Bhardwaj, V., Kilpert, F., Richter, A.S., Heyne, S., Dundar, F., and Manke, T. (2016). deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res* 44, W160-165.
- Romanoff, A.L., and Romanoff, A.J. (1949). *The avian egg* (New York,: J. Wiley).
- Rothstein, M., and Simoes-Costa, M. (2019). Heterodimerization of TFAP2 pioneer factors drives epigenomic remodeling during neural crest specification. *Genome research*.
- Shibley, I.A., Jr., and Pennington, S.N. (1997). Metabolic and mitotic changes associated with the fetal alcohol syndrome. *Alcohol Alcohol* 32, 423-434.
- Simoes-Costa, M., and Bronner, M.E. (2013). Insights into neural crest development and evolution from genomic analysis. *Genome research* 23, 1069-1080.
- Simoes-Costa, M., and Bronner, M.E. (2015). Establishing neural crest identity: a gene regulatory recipe. *Development* 142, 242-257.
- Simoes-Costa, M., Stone, M., and Bronner, M.E. (2015). Axud1 Integrates Wnt Signaling and Transcriptional Inputs to Drive Neural Crest Formation. *Developmental cell* 34, 544-554.
- Skene, P.J., and Henikoff, S. (2017). An efficient targeted nuclease strategy for high-resolution mapping of DNA binding sites. *Elife* 6.
- Spilson, S.V., Kim, H.J., and Chung, K.C. (2001). Association between maternal diabetes mellitus and newborn oral cleft. *Ann Plast Surg* 47, 477-481.
- Tennessen, J.M., Baker, K.D., Lam, G., Evans, J., and Thummel, C.S. (2011). The *Drosophila* estrogen-related receptor directs a metabolic switch that supports developmental growth. *Cell Metab* 13, 139-148.
- Theveneau, E., and Mayor, R. (2012). Neural crest delamination and migration: from epithelium-to-mesenchyme transition to collective cell migration. *Developmental biology* 366, 34-54.
- Thomas-Chollier, M., Hufon, A., Heinig, M., O'Keeffe, S., Masri, N.E., Roider, H.G., Manke, T., and Vingron, M. (2011). Transcription factor binding predictions using TRAP for the analysis of ChIP-seq data and regulatory SNPs. *Nat Protoc* 6, 1860-1869.
- Tinevez, J.Y., Perry, N., Schindelin, J., Hoopes, G.M., Reynolds, G.D., Laplantine, E., Bednarek, S.Y., Shorte, S.L., and Eliceiri, K.W. (2017). TrackMate: An open and extensible platform for single-particle tracking. *Methods* 115, 80-90.
- Tsoi, J., Robert, L., Paraiso, K., Galvan, C., Sheu, K.M., Lay, J., Wong, D.J.L., Atefi, M., Shirazi, R., Wang, X., *et al.* (2018). Multi-stage Differentiation Defines Melanoma Subtypes with Differential Vulnerability to Drug-Induced Iron-Dependent Oxidative Stress. *Cancer Cell* 33, 890-904 e895.

- Uchikawa, M., Ishida, Y., Takemoto, T., Kamachi, Y., and Kondoh, H. (2003). Functional analysis of chicken Sox2 enhancers highlights an array of diverse regulatory elements that are conserved in mammals. *Developmental cell* 4, 509-519.
- Vander Heiden, M.G., Cantley, L.C., and Thompson, C.B. (2009). Understanding the Warburg effect: the metabolic requirements of cell proliferation. *Science* 324, 1029-1033.
- Wang, C.C., Bajikar, S.S., Jamal, L., Atkins, K.A., and Janes, K.A. (2014). A time- and matrix-dependent TGFBR3-JUND-KRT5 regulatory circuit in single breast epithelial cells and basal-like premalignancies. *Nat Cell Biol* 16, 345-356.
- Wang, J., Xiao, Y., Hsu, C.W., Martinez-Traverso, I.M., Zhang, M., Bai, Y., Ishii, M., Maxson, R.E., Olson, E.N., Dickinson, M.E., *et al.* (2016). Yap and Taz play a crucial role in neural crest-derived craniofacial development. *Development* 143, 504-515.
- Warburg, O. (1925). The metabolism of carcinoma cells. *The Journal of Cancer Research* 9, 148–163.
- Warburg, O., Posener, K., and Negelein, E. (1924). On the metabolism of carcinoma cells. *Biochem Z* 152, 309-344.
- Yu, G., Wang, L.G., Han, Y., and He, Q.Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* 16, 284-287.
- Yu, G., Wang, L.G., and He, Q.Y. (2015). ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. *Bioinformatics* 31, 2382-2383.
- Zanconato, F., Cordenonsi, M., and Piccolo, S. (2016). YAP/TAZ at the Roots of Cancer. *Cancer Cell* 29, 783-803.
- Zhao, B., Ye, X., Yu, J., Li, L., Li, W., Li, S., Yu, J., Lin, J.D., Wang, C.Y., Chinnaiyan, A.M., *et al.* (2008). TEAD mediates YAP-dependent gene induction and growth control. *Genes Dev* 22, 1962-1971.

Chapter V: Neural crest metabolism: at the crossroads of development and disease

Alternative Citation:

D Bhattacharya, BM Khan and M Simoes-Costa (2021) Neural crest metabolism: at the crossroads of development and disease, *Developmental Biology*, <https://doi.org/10.1016/j.ydbio.2021.01.018>

Abstract

The neural crest is a migratory stem cell population that contributes to various tissues and organs during vertebrate embryonic development. These cells possess remarkable developmental plasticity and give rise to many different cell types, including chondrocytes, osteocytes, peripheral neurons, glia, melanocytes, and smooth muscle cells. Although the genetic mechanisms underlying neural crest development have been extensively studied, many facets of this process remain unexplored. One key aspect of cellular physiology that has gained prominence in the context of embryonic development is metabolic regulation. Recent discoveries in neural crest biology suggest that metabolic regulation may play a central role in the formation, migration, and differentiation of these cells. This possibility is further supported by clinical studies that have demonstrated a high prevalence of neural crest anomalies in babies with congenital metabolic disorders. Here, we examine why neural crest development is prone to metabolic disruption and discuss how carbon metabolism regulates developmental processes like epithelial-to-mesenchymal transition (EMT) and cell migration. Finally, we explore how understanding neural crest metabolism may inform upon the etiology of several congenital birth defects.

Introduction

The neural crest is a migratory stem cell population that contributes to multiple components of the vertebrate body plan, including the craniofacial skeleton, the peripheral and enteric nervous systems, and the pigmentation of the skin (Le Douarin and Kalcheim, 1999). During neurulation, neural crest cells

delaminate from the dorsal neural tube and migrate along stereotypical pathways (Fig 1). Upon reaching the appropriate destinations in the embryo, these cells differentiate into more than thirty distinct cell types such as chondrocytes, melanocytes, neurons, etc. (Bronner and Simoes-Costa, 2016). The neural crest has served as model for the study of multipotency and migration for many decades, and interest in this cell population has also been fueled by its medical importance. Many congenital disabilities, broadly termed as neurocristopathies, result from defects in the formation of the neural crest and its derivatives (Vega-Lopez et al., 2018). While some of these conditions are linked to genetic mutations, they can also be caused by environmental stimuli. The high prevalence of neurocristopathies underscores the complexity of neural crest development and its reliance on the coordination of multiple cellular processes (Vega-Lopez et al., 2018).

The different steps in neural crest formation, which include induction, specification, migration, and differentiation, are governed by an expansive gene regulatory network (GRN) (Simoes-Costa and Bronner, 2015). The neural crest GRN is composed of a large number of genetic interactions between signaling systems, transcription factors, and epigenomic regulators (Hovland et al., 2020; Hu et al., 2014). In the past years, studies in multiple model organisms have allowed us to expand the network's architecture and gain insight into its logic. Yet, we still have a superficial understanding of how the genetic programs that control cellular behaviors and developmental transitions are coupled with cellular physiology. The complex tasks performed by neural crest cells during embryonic development require the coordination of a number of biochemical processes beyond transcriptional control. One facet of neural crest development that remains relatively unexplored is its metabolic regulation, which recently has taken center stage in many cellular and developmental contexts.

Figure 1

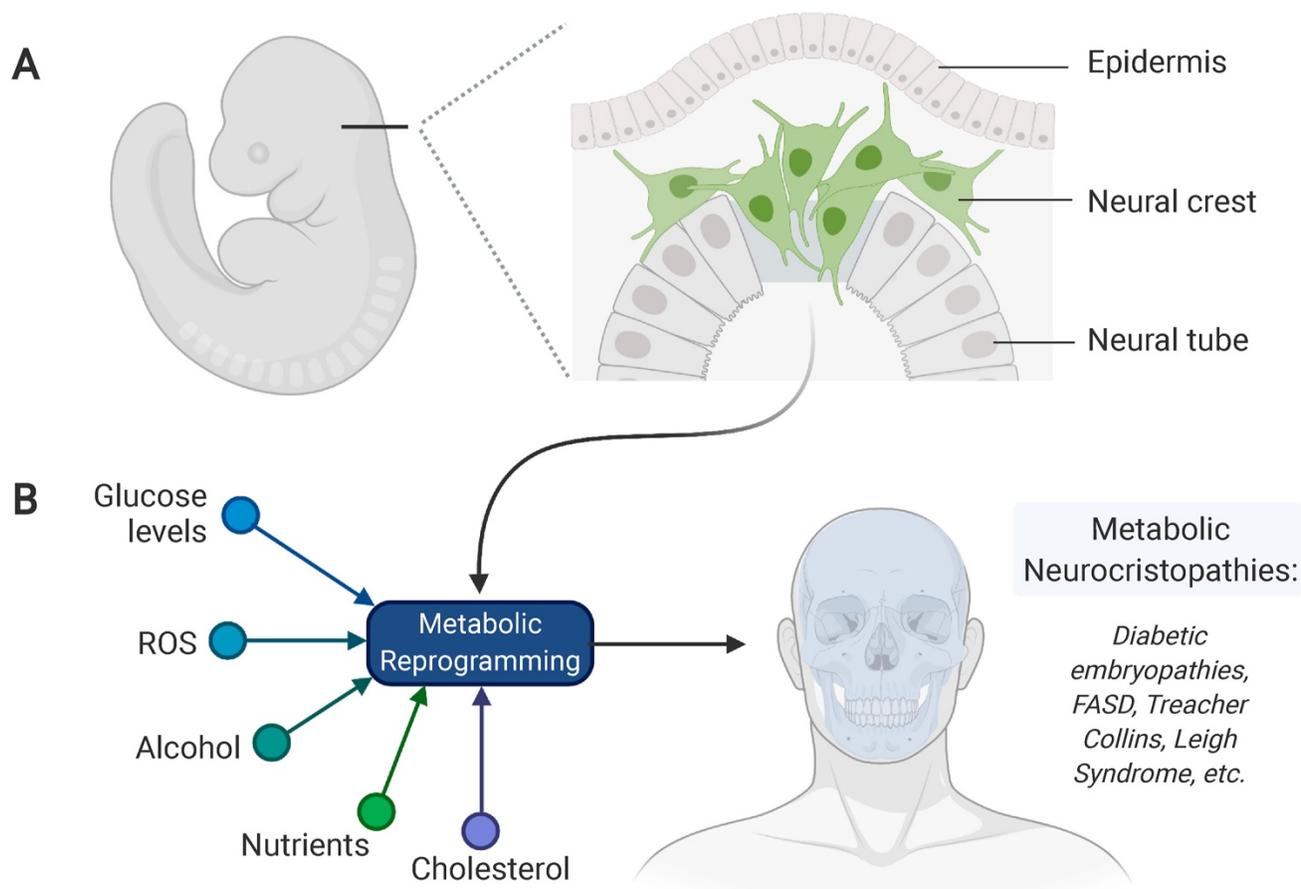


Figure 1. Neural crest development and the genesis of metabolic neurocristopathies.

(A) The neural crest is migratory stem cell population that delaminates from the dorsal neural tube and migrates extensively throughout the embryo. Neural crest development requires drastic changes in cellular metabolism for the proper migration and differentiation of these cells. (B) Environmental or genetic factors may disrupt the metabolic transitions observed in neural crest cells, resulting in congenital disorders known as metabolic neurocristopathies. FASD: Fetal alcohol spectrum disorder, ROS: Reactive oxygen species.

For much of the last century, studies on the links between metabolism and cellular behaviors were bioenergetic in nature, focusing on metabolic changes as a response to cellular needs. An example of this is the Warburg effect (WE), in which cancer cells transition to glycolysis, even in high oxygen environments, to meet the macromolecule demands of persistent proliferation (Hsu and Sabatini, 2008). More recently, numerous studies have suggested that separate from its role in ATP production, metabolism broadly functions as a signaling hub that impinges on other signal transduction pathways as well as on the epigenome and the proteome, to regulate key cellular functions and behaviors (Miyazawa

and Aulehla, 2018). Consistent with this idea, we have recently demonstrated that neural crest cells display WE and that this metabolic adaptation is necessary for EMT and migration (Bhattacharya et al., 2020). The importance of this metabolic adaptation in neural crest development is further underscored by the etiology of many neurocristopathies. Metabolic syndromes or dietary changes that interfere with glucose metabolism often result in defects in neural crest development (Berio, 2011; Chappell et al., 2009; Smith et al., 2014). Furthermore, disruption of other bioenergetic pathways such as lipid metabolism and folate synthesis also has severe consequences on the formation of this stem cell population (See Box 1), and has been implicated in the genesis of congenital malformation. This led us to postulate that the multifactorial effects of cellular metabolism, may be essential for neural crest specification, migration and differentiation by any combination of the following: protecting neural crest cells against oxidative stress, coupling neural crest behavior and physiology, regulating cellular pH, and rapidly producing ATP. In this review, we discuss the divergent roles that aerobic glycolysis plays in stem cells and throughout embryogenesis, with an emphasis on the neural crest. We also examine why neural crest development is particularly prone to disruptions in metabolism and how these disruptions may result in congenital birth defects (Fig. 1).

Metabolic regulation of neural crest development

Central carbon metabolism, which breaks down glucose, is the primary ATP production mechanism in all living cells. It comprises three sequential pathways: Glycolysis, the Krebs Cycle, and the Electron Transport Chain (ETC) through which, in the presence of oxygen, glucose is completely oxidized to produce CO_2 , H_2O , and 38 ATP molecules (Noor et al., 2010). This combined process is referred to as Oxidative Phosphorylation (OXPHOS) and is the default cellular respiration mechanism for somatic cells under aerobic conditions (Gautheron, 1984; Noor et al., 2010) (Fig. 2). Additionally, glucose can also be incompletely oxidized through glycolysis to produce lactate, which does not enter the subsequent pathways. This process is independent of oxygen availability and is known as aerobic

glycolysis when occurring under normoxic conditions (Lunt and Vander Heiden, 2011; Noor et al., 2010) (Fig. 2b). Aerobic glycolysis results in the net gain of only 2 ATP per glucose molecule, and thus increased flux through this pathway necessitates high glucose uptake to meet cellular energy demands. Though not mutually exclusive, aerobic glycolysis and OXPHOS are antagonistic mechanisms of glucose breakdown. The balance between these pathways has been studied extensively in various tissues and distinct biological contexts (Ito and Suda, 2014; Jose et al., 2011; Shyh-Chang and Ng, 2017; Zheng, 2012). In general, most differentiated cells display OXPHOS, while proliferative cells such as stem cells and tumor cells preferentially utilize aerobic glycolysis (Fig 1). The benefits and drawbacks of using one process versus the other are cell type-specific and not only reflect metabolic regulation but also affect cell identity and behavior.

Congenital diseases caused by disrupted glucose metabolism are often associated with severe defects in neural crest-derived tissues, indicating that this metabolic pathway has a lineage-specific

Figure 2 (Appell et al., 2009; Smith et al., 2014). Furthermore, as discussed below, early neural crest cells are susceptible to oxidative stress, a byproduct of excessive OXPHOS, suggesting that a balance between the distinct modes of glucose metabolism may be critical for proper development of the cell type. Despite this, the dynamics of carbon metabolism and its broader functions in neural crest cells have not been resolved. In the 1940s, a study measuring cellular respiration in amphibian neural tube explants reported that, despite aerobic culture conditions, delaminating and migratory neural crest cells display low oxygen uptake rates and that oxygen consumption increases significantly only when the cells begin to differentiate (Flickinger, 1949). This observation was further supported by reports showing that activity of the critical OXPHOS enzyme cytochrome C oxidase is low in undifferentiated migratory quail neural crest cells but is elevated in differentiating cells, particularly in those committed to neuronal lineages (Liu et al., 1990). Additionally, recent studies in murine embryos have revealed that tissue-specific deletion of the mTOR kinase, a critical upstream regulator of OXPHOS, results in apoptosis of only post-migratory neural crest cells (Nie et al., 2018). These observations indicate that OXPHOS plays a more prominent role in later

neural crest development.

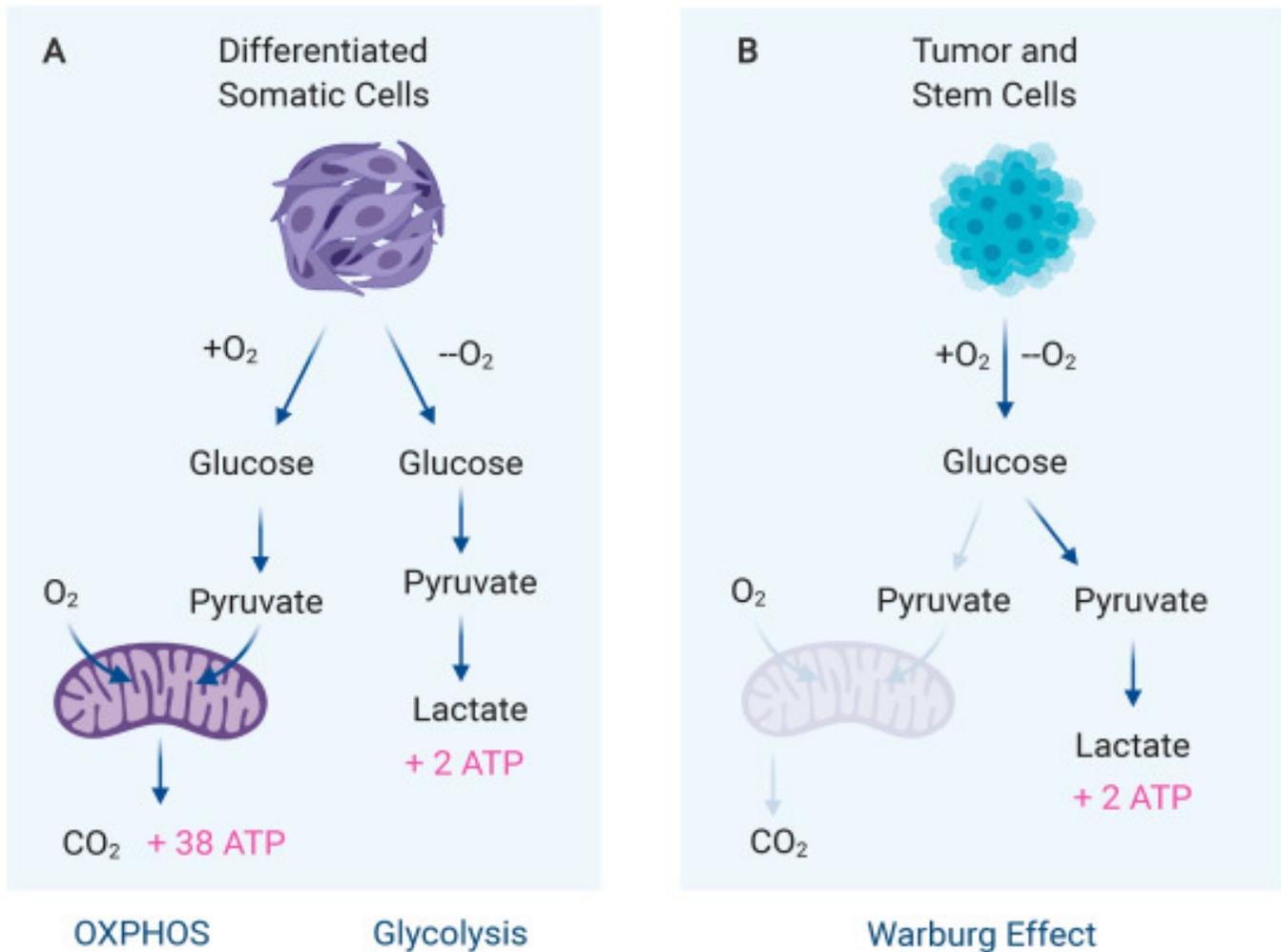


Figure 2. Glucose metabolism in differentiated versus tumor and stem cells. (A)

Oxidative Phosphorylation (OXPHOS) is the default cellular respiration mechanism for somatic cells under aerobic conditions. It results in 38 molecules of ATP per molecule of glucose. In the absence of oxygen, these cells engage in glycolysis to produce only 4 molecules of ATP. (B) Cancer and stem cell populations like the neural crest are highly glycolytic, even in the presence of oxygen. These cells display a metabolic adaptation known as the Warburg Effect and produce approximately 4 molecules of ATP per molecule of glucose.

These findings raised the question of what is the metabolic state of early neural crest cells. Early on, it was hypothesized that undifferentiated neural crest cells maintain a basal state of metabolism (low OXPHOS) and only become energetic (high OXPHOS) during differentiation (Flickinger, 1949, Liu et al., 1990). However, studies from our group and others have recently uncovered that migratory neural

crest cells are energetic and that their primary mode of glucose metabolism is aerobic glycolysis (Bhattacharya et al., 2020; Keuls et al., 2020). This discovery underscores a novel feature of neural crest metabolism that has so far been underappreciated. Below, we have described how this metabolic state is specifically regulated in neural crest cells and its possible impacts on the biology of this progenitor population.

Neural crest and aerobic glycolysis

Elevated flux of a metabolic pathway is usually associated with enhanced expression of its rate-limiting enzymes. We thus utilized a time-course series of RNA-seq data in chick neural crest cells to assess the transcriptional dynamics of enzymes involved in glucose breakdown. This analysis revealed that, before undergoing EMT, cranial neural crest cells specifically upregulate the expression of several glycolytic genes, including *PFKP*, *GAPDH*, and *LDHA*. Consistent with these findings, pre-migratory chick neural crest cells display hallmarks of aerobic glycolysis, including increased glucose uptake, lactate production, and low rates of oxygen consumption. Remarkably, pharmacological inhibition of glycolysis, but not OXPHOS, disrupted EMT and inhibited neural crest migration (Bhattacharya et al., 2020).

By examining the mechanism underlying this phenotype, we uncovered that aerobic glycolysis promotes neural crest EMT by activating the YAP/TEAD pathway, a known regulator of cell migration and invasion (Plouffe et al., 2015) (Fig. 3). In neural crest cells, WE stabilizes the interaction between the transcription factor TEAD1 and its activator YAP1, thus promoting the formation of a functional complex that can activate the transcription of downstream target genes. By profiling the genomic occupancy of the active form of YAP1, we identified several direct targets of YAP/TEAD signaling in neural crest cells, which included EMT factors such as *SOX9*, *ZEB2*, *ETS1*, *TWIST1* and *PRRX2* (Bhattacharya et al., 2020), as well as several genes involved in cell migration and extracellular matrix remodeling (Fig. 3) (Simoes-Costa and Bronner, 2015). Our work directly implicates carbon

metabolism in neural crest development for the first time and highlights how bioenergetic shifts impact GRNs to control cell identity and behavior *in vivo* (Bhattacharya et al., 2020). Notably, another recent study reported that in murine cranial neural crest cells, the switch from the pre-migratory to the migratory state is also accompanied by an increase in glycolytic gene expression and a decrease in the transcription of OXPHOS enzymes (Keuls et al., 2020). These observations support our findings in the chick embryo and suggest that modulation of glucose breakdown is an evolutionarily conserved mechanism of controlling neural crest migration.

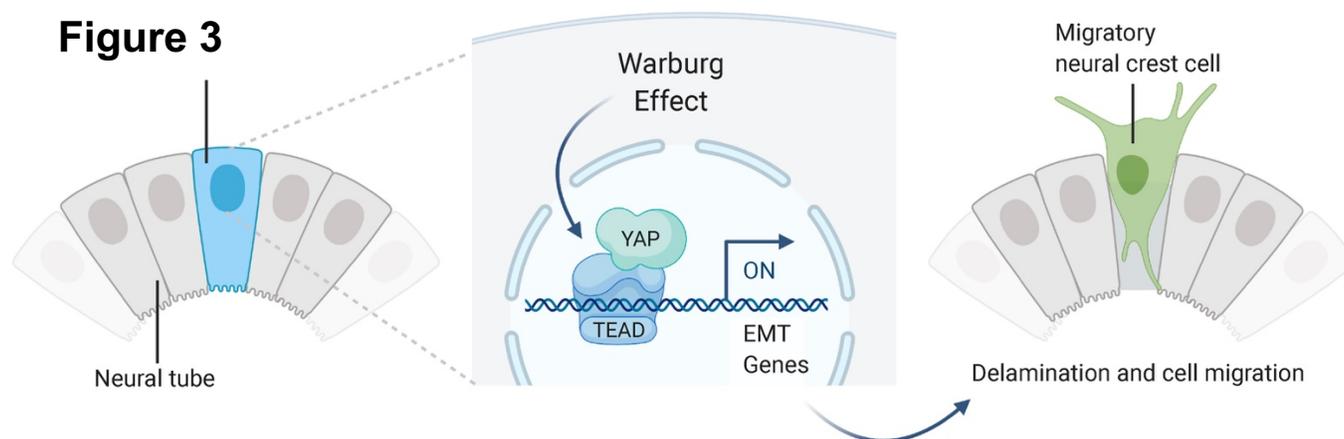


Figure 3. Control of neural crest delamination and migration by the Warburg effect. High glycolytic flux in the pre-migratory neural crest promotes the interaction of YAP and TEAD, leading to the activation of genes that promote epithelial to mesenchymal transition (EMT). YAP/TEAD directly interact with tissue-specific enhancers to promote the EMT regulatory program (Adapted from Bhattacharya et al., 2020).

Potential functions of aerobic glycolysis in migratory neural crest cells

Our study indicates that WE promotes the activation of the YAP/TEAD pathway, promoting EMT and cell migration in neural crest cells. However, given the pleiotropic function of this metabolic adaptation in different contexts, it is likely WE impacts other aspects of neural crest physiology. One such role of WE could be regulating Reactive Oxygen Species (ROS) levels in the migrating cells. Not only ROS can function to mediate cellular signaling (Finkel, 2011), but neural crest cells are especially sensitive to ROS, which can cause oxidative stress and severely affect the development of this cell type (Laforgia et al.,

2018). In respiring cells, most ROS are generated as byproducts of mitochondrial electron transport (Murphy, 2009). However, WE can significantly reduce cellular ROS production by decreasing reliance on mitochondrial respiration (Brand and Hermfisse, 1997; Mullarky and Cantley, 2015). Additionally, pathways promoting redox homeostasis, such as the pentose phosphate pathway (PPP), are often upregulated in cells undergoing aerobic glycolysis. Increased flux through PPP produces the reducing intermediate NADPH, which alongside glutathione, function to quench cellular ROS (Cho et al., 2018; Patra and Hay, 2014). Downstream of WE, these mechanisms may also be essential for preventing oxidative stress in migratory neural crest cells.

Another function of WE could be modulating extracellular matrix proteins to make the microenvironment conducive to neural crest migration. In metastatic tumors, glycolytic cells secrete excess lactate, which renders the environment acidic (Estrella et al., 2013). This shift in pH activates metalloproteases and cathepsins and promotes the degradation of extracellular matrix proteins to facilitate cell migration (Li et al., 2016). Furthermore, metabolic intermediates of aerobic glycolysis and enzymes of this pathway can function as cytokines that promote cell invasion through autocrine signaling (Kathagen- Buhmann et al., 2018). Neural crest migration and cancer invasion implicate many of the same proteases and chemokine signaling pathways (Gallik et al., 2017). Thus, in principle, WE could function through some or all of the above mechanisms to promote neural crest migration.

Apart from their canonical function in glucose metabolism, glycolytic enzymes can also directly control cellular signaling systems. A recent study by Figueiredo et al. reported that the enzyme PFKFB4 is an important regulator of the AKT pathway in neural crest cells of *Xenopus* embryos. By performing stage-specific knockdown of the enzyme, the authors were able to show that PFKFB4 function is essential both for specification and migration of neural crest cells (Figueiredo et al., 2017). Interestingly, the role of this protein during specification was found to be independent of glycolysis and is mediated solely through the activation of the AKT pathway. However, during neural crest migration PFKFB4 function is dependent

both on glycolysis and AKT signaling (Figueiredo et al., 2017). This hints at yet another potential mechanism by which glucose metabolism co-operates with a signaling system to drive neural crest migration. As discussed below, AKT as well as the broader FGF pathway has been implicated in the regulation of WE in other developmental contexts. Thus it is possible that PFKFB4 mediated activation of AKT pathway during neural crest specification may play a deterministic role in the upregulation of glycolysis, later during neural crest migration. Future studies aimed at testing this possibility could inform on upstream regulators of WE in delaminating neural crest cells.

Lastly, aerobic glycolysis may also affect the epigenetic landscape of neural crest cells. As discussed below, a link between glycolytic metabolism and histone acetylation is well documented (Cluntun et al., 2015; Liu et al., 2015; Lu and Thompson, 2012). Removal of glucose causes a global reduction in chromatin acetylation, decreasing the transcription of several genes, including those of glycolytic enzymes. Conversely, histone deacetylation is also sensitive to cellular glycolytic flux (Wellen and Thompson, 2012). The activity of histone deacetylases, such as sirtuins, is modulated by the cytoplasmic NAD⁺/NADH ratio, an increase in which accentuates these enzymes' function (Cluntun et al., 2015; Wellen and Thompson, 2012). Importantly, delaminating neural crest cells also display a high NAD⁺/NADH ratio (Bhattacharya et al., 2020), indicating that aerobic glycolysis may be essential for maintaining a balance between histone acetylation and deacetylation in this progenitor population. Additionally, a recent study reported the association of WE with another novel epigenetic modification: lactate-derived lactylation of histone lysine residues (Zhang et al., 2019). The authors discovered that lactylation is a glycolysis-regulated widespread histone mark that stimulates gene expression. Histone lactylation is, in fact, more specific than acetylation and is only observed in cells displaying high glycolytic flux (Zhang et al., 2019). Thus, the identification of this epigenetic mark provides another clue regarding the non-metabolic functions of WE, which could be relevant to neural crest biology. In summary, a number of developmental processes intrinsic to neural crest cells may be facilitated by WE. We hypothesize that, in coordination with its neural crest specific role as a regulator of the YAP/TEAD pathway, these

downstream functions of WE ensure that the neural crest EMT program is initiated only in metabolically primed cells that are equipped to modify the microenvironment.

Warburg Effect in other developmental contexts

While the physiological role of WE remains controversial, its proposed functions in cancer cells include rapid ATP production, promoting flux through biosynthetic pathways, modulating the tumor microenvironment, and impacting cellular signaling through ROS and chromatin modulation (Liberti and Locasale, 2016). By regulating these distinct aspects of cellular physiology, WE promotes cell survival, proliferation, and migration, making this phenomenon indispensable for tumor growth and metastasis (Liberti and Locasale, 2016). Extensive cell proliferation and migration also underlie embryonic morphogenesis, which suggested to researchers early on that WE may be similarly relevant during normal development (Gardner and Leese, 1987, 1990). The concept of WE as a developmental phenomenon has gained traction and studies over the past decade have provided elegant examples of how aerobic glycolysis regulates cell identity in developmental systems. These findings summarized below, indicate that WE is a tightly-regulated developmental mechanism rather than an anomaly of cancer cells.

Lessons in carbon metabolism from embryonic stem cells

The most well-studied function of aerobic glycolysis in development is its critical role in regulating pluripotency and proliferation of Embryonic Stem Cells (ESCs) (Ito and Suda, 2014; Shyh-Chang and Ng, 2017). Like cancer cells, ESCs in culture proliferate rapidly and have shorter cell cycle than differentiated cells, prompting researchers to investigate their energetic and biosynthetic demands. Studies in both human and mouse primed ESCs revealed that these cells are almost exclusively glycolytic and display increased activity of the PPP pathway under normoxic conditions (Gu et al., 2016; Sperber et al., 2015; Zhou et al., 2012). Together, these metabolic pathways regulate the stemness of ESCs, protect them from oxidative stress, and produce biomolecules (such as ribose and NADPH) required for nucleotide

synthesis and DNA replication (Shyh-Chang and Ng, 2017). Notably, differentiation of ESCs to adult cell types is associated with a metabolic shift to OXPHOS (Varum et al., 2011). Conversely, nuclear reprogramming of somatic cells to induced Pluripotent Stem Cells (iPSCs) results in a metabolic shift to aerobic glycolysis (Folmes et al., 2011). These observations indicate that WE is intricately associated with pluripotency. Indeed, a seminal study by Moussaief and colleagues reported that increased glycolysis in ESCs increases the production of acetyl-CoA and promotes histone acetylation (Moussaieff et al., 2015). The authors further demonstrated that a shift to OXPHOS results in the loss of these acetylation marks within the first hour of differentiation. This can be prevented by adding exogenous acetate, which sustains ESC pluripotency (Moussaieff et al., 2015). Apart from acetylation in ESCs, WE regulate other epigenetic marks such as histone methylation and O-N-acetylglycolysation, which either enhances the transcription of pluripotency factors or silences the expression of differentiation genes in ESCs (Hanover et al., 2012; Jang et al., 2012; Shi et al., 2013). Moreover, as a feed-forward mechanism, key glycolytic enzymes, including HK2 and PKM2 and glucose transporter GLUT1, are direct transcriptional targets of core pluripotency factors c-MYC, SOX2, and OCT4 (Kim et al., 2015; Yu et al., 2019). Another important ESC factor, the RNA binding protein LIN28A/B also inhibits mitochondrial respiration and directly alters the cellular proteome to promote glycolysis (Zhang et al., 2016). This complex crosstalk illustrates that stem cell identity and cellular metabolism are reciprocally regulated in ESCs and emphasizes the importance of metabolism in the regulation of developmental plasticity.

The above observations obtained from cultured ESCs *in vitro* also hold true for pluripotent stem cells that compose the Inner Cell Mass (ICM) of human and murine blastocysts *in vivo*. Analysis of metabolic flux in early murine embryos revealed a transition from OXPHOS to aerobic glycolysis as embryos progress from a single-cell zygote to morula and blastula stages (Kaneko, 2016; Martin and Leese, 1995). This metabolic shift is driven by an increased expression of glycolytic genes and decreased mitochondrial potential of the ICM (Houghton, 2006; Kaneko, 2016). Post implantation, despite a surge in oxygen availability, the murine ICM continues to elevate its glycolytic flux until the beginning of germ layer differentiation (Houghton, 2006; Kaneko, 2016; reviewed by Intlekofer and Finley, 2019). These findings

reiterate the association between WE and pluripotency and show that complementary metabolic transitions drive ICM formation and germ layer differentiation *in vivo*.

Notably, the role of WE in the regulation of stemness extends beyond pluripotent cells and is also relevant in several tissue-specific multipotent progenitor cells such as Neural Stem/Progenitor Cells (NSPCs) and Haemopoietic Stem Cells (HSCs) (reviewed by Candelario et al., 2013; Ito and Suda, 2014). Both NSPCs and HSCs are quiescent stem cells that preferentially utilize glycolysis. In fact, in their respective tissues, these cells exist within a hypoxic niche that promotes WE downstream of HIF1 α -mediated upregulation of glycolytic genes (De Filippis and Delia, 2011; Mohyeldin et al., 2010; Suda et al., 2011). The dependence of both NSPCs and HSCs on WE suggests that, rather than rapid ATP production, the primary function of glycolysis in stem cells is the regulation of developmental potential. Thus, it is likely that the mechanisms by which WE regulates stemness is cell type-specific and dependent upon the functional and environmental context of stem cell populations.

In this context, the development of the vertebrate retina provides another interesting example of how aerobic glycolysis sustains progenitor identity *in vivo*. This is exemplified in a study by Agathocleous and colleagues, which reported that proliferating progenitor cells within the retina of *Xenopus* and Zebrafish embryos are glycolytic (Agathocleous et al., 2012). Unlike their differentiated counterparts, these cells have reduced sensitivity to OXPHOS inhibition. The authors showed that WE is essential for the survival and proliferation of these progenitor cells and that glycolysis mediates this function by directly regulating the cell cycle. Lastly, as observed in several other contexts, the differentiation of retinal progenitors in the *Xenopus* embryo was accompanied by a transition to OXPHOS and reduced dependence on glycolysis (Agathocleous et al., 2012). Taken together, these findings echo a common theme: the balance between aerobic glycolysis and OXPHOS regulates cell state transitions during embryogenesis. Since the neural crest shares many regulatory features with embryonic stem cells, it is likely that many of the mechanisms described here are shared between these cell populations. Comparative studies in metabolic regulation of cell identity will be vital to obtaining a holistic perspective on the role of WE in stem cell populations in the developing embryo.

Aerobic glycolysis in organogenesis

Our understanding of the role of carbon metabolism in neural crest development can benefit from an examination of the metabolic transitions that take place in other embryonic cell types. As discussed above, germ layer differentiation and organogenesis in vertebrate embryos are usually associated with a shift to oxidative metabolism. However, as in the case of neural crest, specific cell types within the developing embryo can transition back to a glycolytic state to engage in distinct biological processes. In particular, migratory progenitor cells commonly upregulate glycolysis to drive organ morphogenesis. A classic example of this is the "tip" and "stalk" Endothelial Cells (EC) that mediate blood vessel formation in the embryo (De Bock et al., 2013). The tip ECs are migratory and promote vessel sprouting while the proliferative stalk ECs function to elongate the blood vessel (Potente et al., 2011). By measuring the metabolic flux of cultured ECs *in vitro*, De Bock and colleagues uncovered that these cells display very high glycolytic flux comparable to that observed in tumor cells (De Bock et al., 2013). This metabolic alternation in ECs is mediated by the action of sprouting signals such as FGF and VEGF, that increase the expression of glycolysis activator enzyme PFKFB3. Depletion of PFKFB3 in ECs expectedly reduces the glycolytic flux and inhibits sprout formation and EC migration *in vitro* and embryonic angiogenesis *in vivo*. This demonstrates that high glycolytic flux is required for proper tip cell positioning in the developing blood vessel, and that PFKFB3 expression is necessary and sufficient to promote the formation and activity of tip cells *in vitro* and *in vivo*. Interestingly, this "pro-tip" function of aerobic glycolysis counters the "pro-stalk" activity of Notch signaling. This balancing act is essential for maintaining the proper ratio of tip:stalk cells in the growing vessel (De Bock et al., 2013). This elegant study thus shows how aerobic glycolysis can mediate the crosstalk between different signaling pathways (such as FGF and Notch) to regulate cellular identity and function during organogenesis.

A series of recent studies characterizing the metabolic transitions that control somite formation also support the premise that WE functions as a mediator of developmental signals in specific cellular contexts. By studying paraxial mesoderm segmentation during somitogenesis, Oginuma and colleagues,

and Bulusu and colleagues independently observed that a glycolytic gradient in the presomitic mesoderm is required for the posterior elongation of the embryonic axis (Bulusu et al., 2017; Oginuma et al., 2017). Furthermore, Oginuma and colleagues showed that this gradient is established downstream of FGF signaling, which increases the expression of glycolytic enzymes in the developing tail bud. In a follow-up study, they reported that aerobic glycolysis-mediated acidosis elevates the intracellular pH of NMPs, leading to the activation of Wnt signaling through acetylation and stabilization of the effector protein β -catenin (Oginuma et al., 2020). This enhanced Wnt activity establishes a mesodermal identity in multipotent NMPs and may also promote their migration towards the forming somite by regulating cell motility. Additionally, Wnt signaling transcriptionally controls the expression of FGF ligands and receptors in the paraxial mesoderm (Oginuma et al., 2020). As discussed above, elevated FGF signaling activates aerobic glycolysis and consequentially the Wnt pathway, thus feeding into a positive feedback loop that ensures tight coordination between Wnt and FGF signaling during somitogenesis.

Crosstalk between aerobic glycolysis and FGF appears to be a recurring theme during organogenesis. A recent study by Kantarci and colleagues indicates that FGF and aerobic glycolysis can reciprocally regulate each other and that this mechanism is broadly relevant in the formation of multiple tissues within the embryo (Kantarci et al., 2020). By harnessing the power of reverse genetics, the authors of this study disrupted the glycolytic enzyme phosphoglycerate kinase-1 (PGK1) in zebrafish embryos and observed defective neuronal development in the otic vesicle and central/peripheral nervous system, as well as disruption of hair cell formation. Interestingly, both neuron and hair cell development are controlled by FGF signaling, prompting the authors to investigate how aerobic glycolysis impacts this signaling pathway. Through a series of functional experiments, the study found that lactate elevates the basal levels of FGF effectors MAPK and ETV5B in glycolytic cells, which primes them to respond efficiently to changes in this signaling pathway (Kantarci et al., 2020). Taken together, the above studies shed light on how signaling systems work in concert with metabolic regulation to drive changes in cell identity.

Effects of metabolic disorders in neural crest development

Our examination of carbon metabolism in neural crest cells indicates that their metabolic states are precisely regulated during development. Prior to delamination, these cells display basal cellular metabolism (Fig. 4). However, as they become primed for migration, there is a striking upregulation of glycolytic enzymes, and the cells become highly glycolytic, exhibiting the hallmarks of the Warburg effect. Glycolytic flux increases during migration, but once neural crest cells reach their final destinations they begin differentiating. As observed in many other developmental cell types, differentiation is accompanied by an increase in the expression of OXPHOS enzymes and reduction in levels of aerobic glycolysis. Thus, there are two major bioenergetic shifts during neural crest development: the first during delamination and the second at the onset of differentiation (Fig. 4). We propose that disruptions in these metabolic transitions may affect the migration and differentiation of neural crest cells, leading to metabolic neurocristopathies (Fig. 1).

Neurocristopathies are a set of congenital disabilities that arise from defects in the formation of the neural crest and its derivatives (Vega-Lopez et al., 2018). These pathologies are often associated with exposure of the fetus to abnormal levels of different metabolites, including high glucose (Chappell et al., 2009) and alcohol (Smith et al., 2014; Zhang et al., 2018). Furthermore, mutations in genes that affect the cellular response to metabolic stress can cause neurocristopathies such as Treacher Collins syndrome (Sakai et al., 2016). Other congenital diseases caused by mutations in mitochondrial proteins, such as Leigh syndrome, also have severe craniofacial phenotypes indicative of improper neural crest development (Berio, 2011). This is consistent with the view that neural crest cells are specifically sensitive to metabolic abnormalities, underscoring the importance of energy homeostasis in the formation of this cell population. Below, we describe the phenotypes associated with some metabolic neurocristopathies and dis

Figure 4

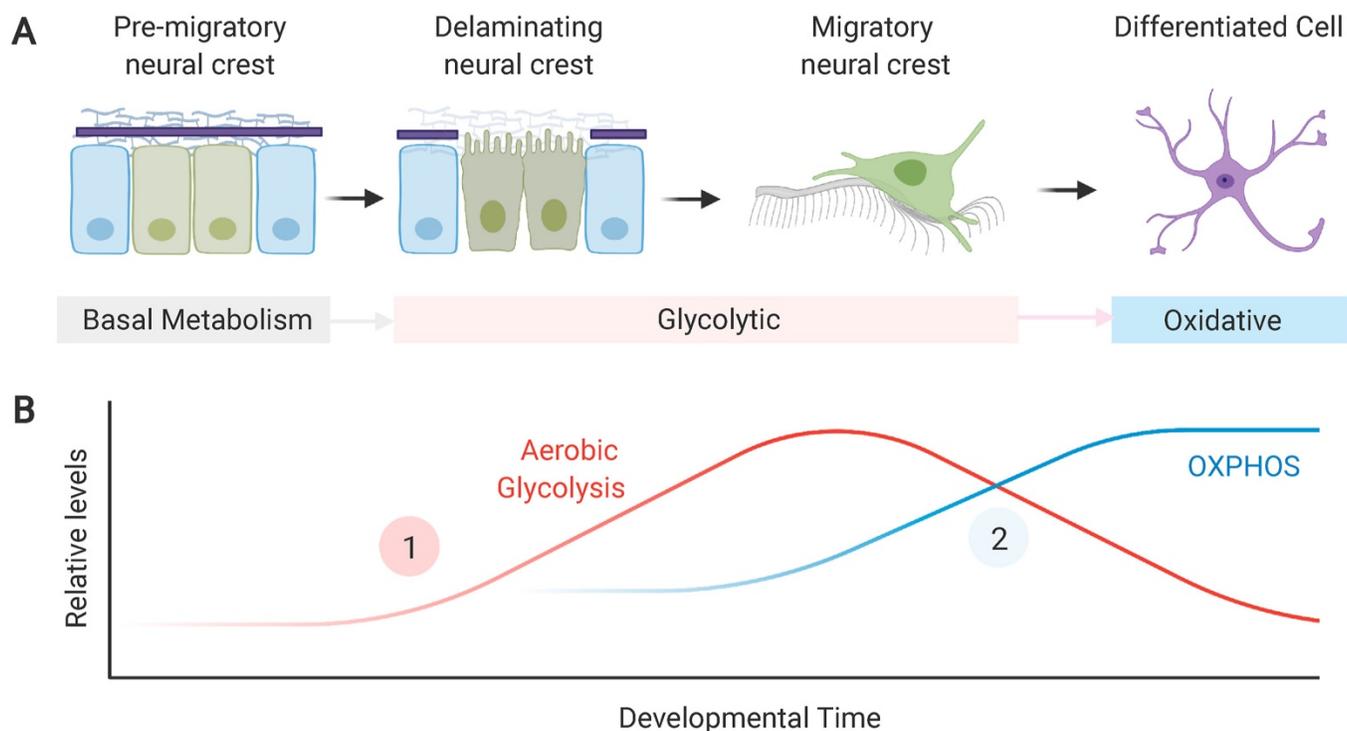


Figure 4. Metabolic transitions during neural crest development. (A) Early pre-migratory neural crest cells initially exhibit low levels of glycolysis and oxidative phosphorylation. As the cells become primed for epithelial to mesenchymal transition, they transition to a highly glycolytic state and undergo delamination and migration. After migration, neural crest cells differentiate and engage in oxidative phosphorylation (OXPHOS). (B) Estimated levels of glycolysis and oxidative phosphorylation during neural crest development. Neural crest cells undergo two main developmental transitions, when they transition from quiescent to glycolytic (1), and when they differentiate and display high levels of mitochondrial respiration (2). OXPHOS: Oxidative phosphorylation.

Diabetic Embryopathies

Maternal diabetes is strongly associated with increased craniofacial malformations such as cleft palate, cardiac outflow tract defects, and holoprosencephaly (Castori, 2013). Though maternal diabetes has been linked to embryopathies since the 1930s, the specific diabetic teratogens causing these malformations remain to be determined. Initial studies in animal models suggested that hyperglycemia is the crucial factor, since exposing mouse and rat embryos to diabetic glucose levels recapitulates the malformations observed in human neonates (Fine et al., 1999). Community-based approaches to monitor and manage maternal blood glucose levels during gestation were also found to be effective in moderating birth defects (Kitzmilller et al., 1991). However, more recent reports implicate maternal hyperketonemia,

disrupted fetal arachidonic acid metabolism, and increased oxidative and cellular stress in diabetic embryopathy (Eriksson and Wentzel, 2016). Given the pleiotropic effect of diabetes on adult tissue (DeFronzo et al., 2015), it is likely that the cause behind maternal diabetes embryopathy is multifactorial. Furthermore, since neural crest cells are significantly affected by high levels of extracellular glucose and oxidative stress, maternal diabetes may impinge on neural crest development at multiple levels and via distinct mechanisms.

Studies in chick, mouse, and rat models have revealed that maternal diabetes induces apoptosis of cranial and cardiac neural crest cells. A study by Wang and colleagues shows that chick embryos developed in the presence of glucose display facial bone defects caused by apoptosis of early neural crest cells (Wang et al., 2015). The authors determined that the hyperglycemic environment activates the PI3K/ERK signaling pathway promoting autophagy, the excess of which results in neural crest apoptosis in a p53-dependent manner (Wang et al., 2015). Consistent with this conclusion, pharmacological inhibition of autophagy rescues glucose-induced cranial neural crest apoptosis, highlighting the therapeutic relevance of this pathway (Wang et al., 2015).

Similar observations were also made in an earlier study by Morgan and colleagues, which demonstrated that hyperglycemia and oxidative stress directly disrupt cardiac neural crest induction (Morgan et al., 2008). In this study, the authors showed that embryos of both diabetic and transiently hyperglycemic mice have increased apoptosis of PAX3⁺ neural crest cells and subsequently display cardiac outflow defects. Interestingly, these phenotypes could be rescued by administering antioxidants to hyperglycemic pregnant mice, leading the authors to conclude that oxidative stress is the primary cause of diabetic embryopathy in cardiac neural crest cells (Morgan et al., 2008). In support of this hypothesis, a transient increase in oxidative stress by subcutaneous injection with actinomycin in pregnant mice recapitulated the cardiac neural crest defects, further indicating that oxidative stress is a downstream effect of hyperglycemia (Morgan et al., 2008).

A separate study by Wentzel and colleagues provides insights into a possible mechanism by which

high glucose levels increase oxidative stress in neural crest cells (Wentzel and Eriksson, 2011). By assessing the gene expression profiles of cranial and trunk neural crest cells isolated from rat embryos, the authors showed that when cultured in the presence of high glucose, cranial neural crest cells specifically downregulate the expression of antioxidant genes such as NRF2, Superoxide dismutase, Catalase, and PARP. (Wentzel and Eriksson, 2011). This suggests that, in addition to altering the cells' metabolic profile, hyperglycemia can impinge on their transcriptional program, reducing the ability of neural crest cells to respond to oxidative stress. The specificity of this transcriptional alteration to the cranial subpopulation further indicates an axial-specific susceptibility of neural crest cells to hyperglycemia and oxidative stress.

Taken together, these studies describe two distinct mechanisms: (1) upregulation of autophagy and (2) increase in oxidative stress, by which maternal diabetes can promote apoptosis of neural crest cells. However, since an important effect of oxidative stress is the upregulation of autophagy (Filomeni et al., 2015; Yun et al., 2020), these two mechanisms likely merge into a single signaling cascade where, downstream of hyperglycemia, cellular oxidative stress upregulates autophagy to cause neural crest apoptosis in a p53-dependent manner. If substantiated, this mechanism emphasizes the potential of dietary antioxidant supplementation to pregnant diabetic mothers as a possible intervention for diabetic embryopathy.

Fetal Alcohol Spectrum Disorder

Fetal Alcohol Spectrum Disorder (FASD) is a neurocristopathy that results from prenatal alcohol exposure (Smith et al., 2014). It can be diagnosed by characteristic craniofacial dysmorphia, including flattened midface and mouth area, thin upper lip, and micrognathia. Infants suffering from FASD may also display sensory and language disorders, neurological abnormalities, and growth defects (Smith et al., 2014; Wozniak et al., 2019). Prenatal alcohol exposure impacts multiple stages of neural crest development. During early gastrulation, high maternal blood alcohol levels disrupt neuroectoderm development and

decrease expression of crucial neural crest markers genes in the embryo (Cartwright and Smith, 1995). Alcohol also suppresses cranial neural crest migration (Cartwright and Smith, 1995; Rovasio and Battiato, 1995). *In vivo* studies in the chick embryo have revealed that upon exposure to ethanol, neural crest cells fail to emerge from the neural tube, indicative of defects in EMT. The few cells that do migrate initially follow appropriate migratory routes but do not arrive at the correct locations in the embryo, further suggesting disruption of signaling cues (Cartwright and Smith, 1995; Rovasio and Battiato, 1995). A recent study in zebrafish embryos demonstrated that upon ethanol challenge, the left-right symmetry of migratory cranial neural crest cells is lost, and the cells display an uncharacteristic retrograde motion that prevents travel over longer distances (Boric et al., 2013). Intriguingly, complementary *in vitro* studies using neural crest explants revealed that the migration defects observed upon ethanol exposure are specific to cranial neural crest cells and are significantly moderate in its trunk counterparts (Czarnobaj et al., 2014; Rovasio and Battiato, 2002), thus affirming the axial differences in the neural crest's ability to respond to metabolic stress.

The above defects in cranial neural crest formation and migration are partly attributed to increased apoptosis of these cells upon alcohol exposure (Debelak-Kragtorp et al., 2003; Flentke et al., 2011; Rovasio and Battiato, 1995; Smith et al., 2014). Even low doses of ethanol induce cell death in up to 50% of delaminating neural crest cells, and inhibition of apoptosis can rescue the facial anomalies observed in FASD (Cartwright et al., 1998; Dunty et al., 2001; Rovasio and Battiato, 2002). At least two distinct mechanisms have been described by which alcohol promotes cell death in the neural crest. The first mechanism involves the interaction of ethanol with a yet unidentified G-protein coupled receptor, which elevates the levels of inositol triphosphate (IP₃) and causes a massive increase in intracellular Ca²⁺ levels (Debelak-Kragtorp et al., 2003; Garic-Stankovic et al., 2005). This transient calcium wave activates the calcium/calmodulin-dependent protein kinase II (CaMKII), specifically in the dorsal neural folds (Garic et al., 2011). CaMKII, in turn phosphorylates and destabilizes β -catenin, thus effectively inhibiting the Wnt signaling pathway (Flentke et al., 2011). This disrupts neural crest specification and migration, leading to apoptosis of this cell population (Flentke et al., 2011).

An alternate mechanism implicates oxidative stress and depletion of intracellular metabolites as the primary cause of ethanol-induced neural crest apoptosis (Chen and Sulik, 1996; Chen et al., 2013; Davis et al., 1990). Alcohol is known to induce oxidative stress by generating excess ROS, which neural crest cells are particularly sensitive to due to their low levels of endogenous antioxidants (Davis et al., 1990). The mechanism through which ethanol metabolism generates ROS in neural crest cells remains unknown. Despite this, several antioxidants, including astaxanthin, vitamin C, vitamin E, and beta-carotene have effectively prevented FASD in animal models (Zhang et al., 2018). Interestingly, neural crest cells increase expression of NRF2, a transcription factor and master regulator of the antioxidant response after ethanol exposure, likely to compensate for the elevated oxidative stress (Chen et al., 2013). While this modest increase in NRF2 transcription is not sufficient to reduce apoptosis, Chen and colleagues observed that overexpression of NRF2 or treatment with its chemical agonists increased the expression of antioxidant factors such as SOD1, catalase, GPX3, and NQO; and decreased ROS generation and cell death in the neural crest (Chen et al., 2013). Ethanol additionally impacts cellular metabolism by decreasing the availability of cholesterol esters and depleting polyamines (Poodeh et al., 2014; Li et al., 2007). The former results in suppression of the Sonic Hedgehog Signaling pathway and has severe consequences on neurogenesis and possibly neural crest development (Li et al., 2007). Reduced availability of polyamines such as putrescine and spermidine negatively affect DNA replication, which may inhibit the proliferation of pre-migratory neural crest cells, resulting in apoptosis (Poodeh et al., 2014).

Prenatal alcohol exposure can also directly impact the migration of neural crest cells. Apart from increased apoptosis, the profound migratory defects described above result from improper EMT of cranial neural crest cells and defects in filopodia formation and substrate adhesion (Hassler and Moran, 1986; Oyedele and Kramer, 2013; Rovasio and Battiato, 2002). Intriguingly, these phenotypes mirror those observed upon inhibition of aerobic glycolysis during the migration of cranial neural crest cells. As discussed above, aerobic glycolysis necessitates increased uptake of glucose by delaminating neural crest cells (Bhattacharya et al., 2020). However, a well-characterized feature of FASD is significantly

decreased fetal blood glucose level and reduced glucose uptake by the embryo. This could prevent aerobic glycolysis in the neural crest due to the unavailability of glucose, which may be an additional mechanism by which prenatal alcohol exposure disrupts delamination and migration of this cell type. Future studies aimed at understanding how ethanol challenge impacts central glucose metabolism could verify this possibility and provide insights on mechanisms by which ethanol impacts OXPHOS to promote ROS production.

Treacher Collins Syndrome

Neurocristopathies can also be caused by an inadequate or insufficient response to metabolic stress. An example of this is Treacher Collins syndrome (TCS), characterized by jawbone and cheekbone hypoplasia, external ear anomalies, cleft palate, and periorbital abnormalities (Algerian and Gilardino, 2019). Of the four causative genes, autosomal dominant mutations in *TCOF1* is the most common and is observed in 90% of TCS patients (Group, 1996). TCS-associated mutations in *TCOF1* result in reduced expression of the nucleolar protein Treacle, which causes impaired ribosome biogenesis (Valdez et al., 2004), increased nucleolar stress (Rubbi and Milner, 2003), and apoptosis of neural crest cells (Jones et al., 2008). Historically, defects in ribosome function were considered to be the primary cause of neural crest apoptosis in TCS patients. However, recent reports indicate that haploinsufficiency of *TCOF1* dampens the oxidative stress response in neuroepithelium, resulting in cell apoptosis in a p53-dependent manner (de Peralta et al., 2016; Sakai et al., 2016).

A study by Sakai and colleagues demonstrated that a key function of Treacle is to interact with the DNA damage sensor MRNM complex (Mre11-Rad51-NBS1-MDC1) and localize at the sites of oxidative stress-induced DNA strand breaks (Sakai et al., 2016). At these damaged loci, Treacle recruits the DNA repair protein BRCA1 to fix the strand breaks and restore cell cycle progression following oxidative stress. The authors further demonstrated that in early embryos, the neuroepithelium exists in a highly oxidative state with elevated ROS levels, which explains why *TCOF1* loss of function has a severe effect on neural

crest development. Importantly, by injecting the antioxidant N-acetyl cysteine into *TCOF1* +/- pregnant mice, the authors could rescue the craniofacial deformations associated with TCS. This suggests that antioxidant treatments can potentially alleviate and even prevent pathogenesis in *TCOF1* +/-embryos (Sakai et al., 2016).

A recent study utilizing a zebrafish model of TCS also confirmed the utility of antioxidants in reducing craniofacial abnormalities in TCS embryos and identified an additional mechanism by which Treacle protects cells from oxidative stress (de Peralta et al., 2016). In this study, Peralta and colleagues showed that Treacle stabilizes ROS protective CNBP protein (Cellular nucleic acid-binding protein), which is crucial for craniofacial development (de Peralta et al., 2016). Though the mechanism by which CNBP promotes antioxidant response is not understood, the authors showed that overexpression of this protein was sufficient to reduce elevated ROS levels and rescue the craniofacial phenotypes of TCS embryos (de Peralta et al., 2016). Altogether, the above studies aimed at delineating the etiology of TCS reiterate the importance of antioxidant response in neural crest development and inform upon therapeutic strategies for managing birth defects arising from the disruption of this protective pathway.

Mitochondrial Diseases

The congenital diseases described above are classical neurocristopathies resulting from defective neural crest metabolism. Intriguingly, mutations in crucial metabolic enzymes also result in birth defects with phenotypes very similar to those observed in FASD and diabetic embryopathy. In this regard, diseases caused by mutations in mitochondrial proteins are particularly relevant as they are often associated with gross craniofacial deformities (Berio, 2011). For example, patients with Leigh syndrome, a life-threatening birth defect caused by mutations in OXPHOS genes, often display craniofacial dysmorphisms, such as maxillary hypoplasia, upturned nose, micrognathia, and smooth philtrum (Berio and Piazzzi, 2007). Other OXPHOS-related diseases, such as Kearns-Sayre syndrome and congenital pyruvate dehydrogenase deficiency,

are also associated with similar facial abnormalities, albeit with varying severity (Berio, 2011; Finsterer et al., 2020). While these diseases are not traditionally categorized as neurocristopathies, likely due to severe comorbidities in other organs, their phenotypes strongly suggest that neural crest development is particularly vulnerable to mitochondrial dysfunction. The molecular basis of this vulnerability, and indeed the role of OXPHOS in neural crest development, remains uncharacterized. However, this is increasingly becoming a relevant area of research in neural crest and developmental biology.

Conclusions

In the last decades, we have learned much about the connections between metabolism and cellular processes, such as signal transduction and gene regulation. Studies on metabolic diseases, such as diabetes, led to the idea that carbon metabolism has a pervasive role in cellular function (DeFronzo et al., 2015). Since then, this framework has become critical to our understanding of cancer, aging, and neurodegeneration (Hsu and Sabatini, 2008; Procaccini et al., 2016; Barzilai et al., 2012). Yet, we are only now gaining insight on how the metabolic state of embryonic cells contributes to their behavior and identity during embryonic development (Miyazawa and Aulehla, 2018). The reciprocal relationships between metabolism and cellular behaviors are evident in the development of the neural crest. A key conclusion emerging from the existing body of research is that the metabolic state of the neural crest is highly dynamic and involves drastic bioenergetic shifts that are consequential to its development.

At least two metabolic reprogramming events define neural crest development: (i) the transition of delaminating cells from basal to a glycolytic state, which is essential for EMT, and (ii) the transition of post migratory neural crest cells to a high OXPHOS state that is required for differentiation (Fig. 3). The importance of these metabolic transitions was initially determined through experiments in model organisms, and their relevance for human health is supported by studies on neurocristopathies linked to metabolic diseases. Notably, these conditions specifically affect either neural crest migration or differentiation and sometimes disrupt both processes. This indicates that these two steps in neural crest

development are the most sensitive to metabolic disruptions. Additional studies on the molecular pathways that work in concert with carbon metabolism are required for the identification of specific mechanisms that are disrupted in metabolic neurocristopathies.

Diseases caused by metabolite imbalance, such as maternal diabetes and FASD, primarily increase cellular ROS levels in delaminating neural crest cells, resulting in their apoptosis and failure to delaminate. We propose that the dysregulation of extracellular glucose levels inherent to these diseases prevents aerobic glycolysis, which can prematurely upregulate OXPHOS and elevate cellular ROS levels. In contrast, congenital diseases caused by mutations in mitochondrial genes disproportionately affect differentiation of neural crest cells, particularly mesenchymal lineages. This is likely because inhibition of mitochondrial respiration prevents the second transition of post migratory neural crest cells to a high OXPHOS state, which appears to be especially critical for craniofacial development.

The phenotypes associated with neurocristopathies also highlight the metabolic heterogeneity of different neural crest populations. As mentioned above, susceptibility of these progenitors to metabolic disruptions appears to be dependent on their location along the embryonic anterior-posterior axis. Neural crest cells derived from the cranial regions appear to be more vulnerable than their trunk counterparts. This observation indicates that the bioenergetic profile and the characteristic metabolic modes of neural crest cells may vary depending on their axial level. As evidenced by recent studies on the metabolic regulation of neuro-mesodermal progenitors, the differentiation of embryonic cells to more mesodermal lineages tend to be significantly reliant on cellular metabolism (Oginuma et al., 2017). This raises the intriguing possibility that the developmental potential of neural crest subpopulations is connected with their bioenergetic state.

Lastly, the above discussion also engenders several outstanding questions, especially regarding the mechanisms controlling the bioenergetic shifts throughout neural crest formation. Our understanding of how metabolic reprogramming is mediated in other developmental contexts could provide essential clues in this regard. In particular, the regulation of aerobic glycolysis in ESCs likely has direct parallels

with how this metabolic adaptation is implemented in migratory neural crest cells. Indeed, key transcription factors like HIF1a and c-MYC, which are master regulators of WE in ESCs (Kim et al., 2015; Varum et al., 2011), are also highly expressed in pre-migratory and migratory neural crest cells in a tissue-specific manner in multiple vertebrate models (Barriga et al., 2013; Kerosuo and Bronner, 2016). The depletion of these factors in delaminating neural crest recapitulate the phenotypes observed upon inhibition of WE, suggesting that these regulators may have a conserved role in pluripotent ESCs and multipotent neural crest cells during embryogenesis (Barriga et al., 2013). Similarly, mechanisms that inhibit WE during ESC differentiation may also be active in post-migratory neural crest cells to mediate their transition from a glycolytic to a high OXPHOS state prior to cell fate commitment. Identifying the key players would provide further clues on how the crosstalk between the neural crest GRN and metabolic pathways guide the formation of this cell type.

Box 1. Other metabolic pathways and metabolites important for neural crest formation

Apart from carbohydrate metabolism, several other accessory metabolic pathways and their derivatives have essential roles in neural crest development. One such process is the folic acid synthesis pathway, which has been extensively studied in the context of early vertebrate embryogenesis. Folate deficiency during embryogenesis severely affects craniofacial development and causes many of the same phenotypes associated with the neurocristopathies described below. However, only recently have we been able to determine the molecular basis of these deformities. To this end, at least three distinct mechanisms have been proposed by which folic acid regulates orofacial development. Firstly, decreased folic acid disrupts methionine metabolism and causes cellular accumulation of the metabolite homocysteine (Boot et al., 2003; Melo et al., 2017). In avian and rat neural tube explants, excess homocysteine promotes neural crest migration but prevents differentiation, particularly into mesenchymal lineages (Boot et al., 2003; Melo et al., 2017). This defect can be fully rescued by adding folic acid to culture media, indicating that the regulation of homocysteine levels is an important function of folate in neural crest cells (Boot et al., 2003). Secondly, during differentiation, folate metabolism is required for the survival of the neural crest-derived mesenchymal cells. A study by Wahl et al. showed that inhibition of the dihydrofolate reductase enzyme (DHFR) required for folate synthesis, caused increased apoptosis and decreased proliferation of cells in the facial prominences and the developing jaw of the *Xenopus* embryo (Wahl et al., 2015). The authors further uncovered a novel collaboration between folate and the Retinoic Acid (RA) signaling pathway in neural crest cells, wherein folic acid can rescue the craniofacial defects observed upon RA deficiency in embryos, and the inhibition of DHFR exacerbates the phenotype of RA signaling disruption (Wahl et al., 2015). Whether folate and RA directly regulate each other or function in parallel remains unresolved. However, this study provides a clinically relevant example of the synergistic

interactions between different metabolites that control organ development. Lastly, being a precursor of S-adenosylmethionine (SAM), folate can control the epigenetic landscape of neural crest cells by impacting histone and DNA methylation. A recent report by Jimenez et al. showed that loss of the neural crest-specific folate transporters, FOLR1 and RFC1, confers a neuronal fate to dorsal neural tube progenitor cells in avian embryos (Jimenez et al., 2018). This is in part due to reduced methylation and de-repression of the neuronal *SOX2* gene locus in the neural crest territory, which causes the expansion of the neural plate fate to the dorsal extremities (Jimenez et al., 2018). Taken together, these mechanisms reveal the requirement of folate at every step of neural crest development: from specification to differentiation, and provide clues to how this metabolic pathway is adapted to perform cell specific functions.

Another metabolite that is crucial to neural crest formation is cholesterol. Cholesterol is the precursor of all steroid hormones and bile salts and also plays a critical role during embryogenesis by moderating the Sonic Hedgehog Signaling (Shh) pathway (Lewis et al., 2001; Riobo, 2012). In humans, mutations in enzymes associated with cholesterol transport and biosynthesis cause embryonic defects in neural tube and limb formation and affect neural crest-derived structures such as the facial mesenchyme and cardiac outflow tract (Porter and Herman, 2011). These defects have been primarily attributed to disrupted Shh signaling, which is indispensable for the development of these affected tissues. In the neural crest, Shh signaling is specifically required for the differentiation of progenitor cells into skeletal and non-skeletal mesenchymal cells that form facial structures and the branchial arches (Quintana et al., 2017). Interestingly, despite its important function in neural tube development, Shh primarily promotes mesenchymal fate in neural crest cells and is not necessary for neuronal differentiation. Additionally, Shh signaling functions to pattern the (neural crest-derived) cardiac outflow tract and is generally required for the survival and proliferation of cranial and cardiac neural crest cells. However, disruption of cholesterol metabolism does not have overt effects on derivatives of trunk neural crest cells, suggesting that the relevance of this pathway maybe axial specific.

Given the ubiquity of cholesterol metabolism, it is important to consider how the tissue/axial-specific function of this pathway is brought about in neural crest cells. A study by Iwata et al. implicates the cranial specific TGF- β signaling system in control of lipid metabolism in neural crest-derived mesenchyme (Iwata et al., 2014). Murine mutants of TGF- β receptors display facial dysmorphisms similar to that observed upon inhibition of cholesterol biosynthesis. By further characterizing palatal mesenchyme cells obtained from these mutants, the authors uncovered that inhibition of TGF- β signaling prevents lipolysis and subsequent lipid turnover in these cells (Iwata et al., 2014). This, in turn compromises Shh signaling, resulting in decreased cell proliferation and apoptosis. However, these cellular defects can be rescued by administration of Shh protein, confirming that this signaling pathway is epistatic to TGF- β in neural crest-derived facial mesenchyme (Iwata et al., 2014). Together with the studies discussed above, these findings provide an important clue to the function of lipid/cholesterol metabolism as a pivotal node that links distinct signaling pathways to mediate tissue-specific functions in neural crest cells.

Acknowledgements

We would like to thank the members of the Simoes-Costa Lab for critical reading of the manuscript and valuable suggestions. The figures in this article were created with Biorender.

References

- Agathocleous, M., Love, N.K., Randlett, O., Harris, J.J., Liu, J., Murray, A.J., Harris, W.A., 2012. Metabolic differentiation in the embryonic retina. *Nat Cell Biol* 14, 859-864.
- Alata Jimenez, N., Torres Perez, S.A., Sanchez-Vasquez, E., Fernandino, J.I., Strobl-Mazzulla, P.H., 2018. Folate deficiency prevents neural crest fate by disturbing the epigenetic Sox2 repression on the dorsal neural tube. *Dev Biol* 444 Suppl 1, S193-S201.
- Algerian, A., Gilardino, M.S., 2019. Treacher Collins Syndrome. *Clin Plast Surg* 46, 197-205.
- Barriga, E.H., Maxwell, P.H., Reyes, A.E., Mayor, R., 2013. The hypoxia factor Hif-1alpha controls neural crest chemotaxis and epithelial to mesenchymal transition. *The Journal of cell biology* 201, 759-776.
- Barzilai, N., Huffman, D.M., Muzumdar, R.H., Bartke, A., 2012. The critical role of metabolic pathways in aging. *Diabetes* 61, 1315-1322.
- Berio, A., 2011. Metabolic syndromes and neural crest development. *Journal of Biological Research*.
- Berio, A., Piazzini, A., 2007. [Leigh syndrome with facial abnormalities: a neurocristopathy]. *Pediatr Med Chir* 29, 50-54.
- Bhattacharya, D., Azambuja, A.P., Simoes-Costa, M., 2020. Metabolic Reprogramming Promotes Neural Crest Migration via Yap/Tead Signaling. *Dev Cell* 53, 199-211 e196.
- Boot, M.J., Steegers-Theunissen, R.P., Poelmann, R.E., Van Iperen, L., Lindemans, J., Gittenberger-de Groot, A.C., 2003. Folic acid and homocysteine affect neural crest and neuroepithelial cell outgrowth and differentiation in vitro. *Dev Dyn* 227, 301-308.
- Boric, K., Orio, P., Vieville, T., Whitlock, K., 2013. Quantitative analysis of cell migration using optical flow. *PLoS One* 8, e69574.
- Brand, K.A., Hermfisse, U., 1997. Aerobic glycolysis by proliferating cells: a protective strategy against reactive oxygen species. *FASEB J* 11, 388-395.
- Bronner, M.E., Simoes-Costa, M., 2016. The Neural Crest Migrating into the Twenty-First Century. *Curr Top Dev Biol* 116, 115-134.
- Bulusu, V., Prior, N., Snaebjornsson, M.T., Kuehne, A., Sonnen, K.F., Kress, J., Stein, F., Schultz, C., Sauer, U., Aulehla, A., 2017. Spatiotemporal Analysis of a Glycolytic Activity Gradient Linked to Mouse Embryo Mesoderm Development. *Developmental cell* 40, 331- 341 e334.
- Candelario, K.M., Shuttleworth, C.W., Cunningham, L.A., 2013. Neural stem/progenitor cells display a low requirement for oxidative metabolism independent of hypoxia-inducible factor- 1alpha expression. *J Neurochem* 125, 420-429.
- Cartwright, M.M., Smith, S.M., 1995. Stage-dependent effects of ethanol on cranial neural crest cell development: partial basis for the phenotypic variations observed in fetal alcohol syndrome. *Alcohol Clin Exp Res* 19, 1454-1462.
- Cartwright, M.M., Tessmer, L.L., Smith, S.M., 1998. Ethanol-induced neural crest apoptosis is coincident with their endogenous death but is mechanistically distinct. *Alcohol Clin Exp Res* 22, 142-149.
- Castori, M., 2013. Diabetic embryopathy: a developmental perspective from fertilization to adulthood. *Mol Syndromol* 4, 74-86.
- Chappell, J.H., Jr., Wang, X.D., Loeken, M.R., 2009. Diabetes and apoptosis: neural crest cells and neural tube. *Apoptosis* 14, 1472-1483.
- Chen, S.Y., Sulik, K.K., 1996. Free radicals and ethanol-induced cytotoxicity in neural crest cells.

Alcohol Clin Exp Res 20, 1071-1076.

- Chen, X., Liu, J., Chen, S.Y., 2013. Over-expression of Nrf2 diminishes ethanol-induced oxidative stress and apoptosis in neural crest cells by inducing an antioxidant response. *Reprod Toxicol* 42, 102-109.
- Cho, E.S., Cha, Y.H., Kim, H.S., Kim, N.H., Yook, J.I., 2018. The Pentose Phosphate Pathway as a Potential Target for Cancer Therapy. *Biomol Ther (Seoul)* 26, 29-38.
- Cluntun, A.A., Huang, H., Dai, L., Liu, X., Zhao, Y., Locasale, J.W., 2015. The rate of glycolysis quantitatively mediates specific histone acetylation sites. *Cancer Metab* 3, 10.
- Czarnobaj, J., Bagnall, K.M., Bamforth, J.S., Milos, N.C., 2014. The different effects on cranial and trunk neural crest cell behaviour following exposure to a low concentration of alcohol in vitro. *Arch Oral Biol* 59, 500-512.
- Davis, W.L., Crawford, L.A., Cooper, O.J., Farmer, G.R., Thomas, D.L., Freeman, B.L., 1990. Ethanol induces the generation of reactive free radicals by neural crest cells in vitro. *J Craniofac Genet Dev Biol* 10, 277-293.
- De Bock, K., Georgiadou, M., Carmeliet, P., 2013. Role of endothelial cell metabolism in vessel sprouting. *Cell Metab* 18, 634-647.
- De Filippis, L., Delia, D., 2011. Hypoxia in the regulation of neural stem cells. *Cell Mol Life Sci* 68, 2831-2844.
- de Peralta, M.S., Mouguelar, V.S., Sdrigotti, M.A., Ishiy, F.A., Fanganiello, R.D., Passos-Bueno, M.R., Coux, G., Calcaterra, N.B., 2016. Cnbp ameliorates Treacher Collins Syndrome craniofacial anomalies through a pathway that involves redox-responsive genes. *Cell Death Dis* 7, e2397.
- Debelak-Kragtorp, K.A., Armant, D.R., Smith, S.M., 2003. Ethanol-induced cephalic apoptosis requires phospholipase C-dependent intracellular calcium signaling. *Alcohol Clin Exp Res* 27, 515-523.
- DeFronzo, R.A., Ferrannini, E., Groop, L., Henry, R.R., Herman, W.H., Holst, J.J., Hu, F.B., Kahn, C.R., Raz, I., Shulman, G.I., Simonson, D.C., Testa, M.A., Weiss, R., 2015. Type 2 diabetes mellitus. *Nat Rev Dis Primers* 1, 15019.
- Dunty, W.C., Jr., Chen, S.Y., Zucker, R.M., Dehart, D.B., Sulik, K.K., 2001. Selective vulnerability of embryonic cell populations to ethanol-induced apoptosis: implications for alcohol-related birth defects and neurodevelopmental disorder. *Alcohol Clin Exp Res* 25, 1523-1535.
- Eriksson, U.J., Wentzel, P., 2016. The status of diabetic embryopathy. *Ups J Med Sci* 121, 96- 112.
- Estrella, V., Chen, T., Lloyd, M., Wojtkowiak, J., Cornnell, H.H., Ibrahim-Hashim, A., Bailey, K., Balagurunathan, Y., Rothberg, J.M., Sloane, B.F., Johnson, J., Gatenby, R.A., Gillies, R.J., 2013. Acidity generated by the tumor microenvironment drives local invasion. *Cancer Res* 73, 1524-1535.
- Filomeni, G., De Zio, D., Cecconi, F., 2015. Oxidative stress and autophagy: the clash between damage and metabolic needs. *Cell Death Differ* 22, 377-388.
- Fine, E.L., Horal, M., Chang, T.I., Fortin, G., Loeken, M.R., 1999. Evidence that elevated glucose causes altered gene expression, apoptosis, and neural tube defects in a mouse model of diabetic pregnancy. *Diabetes* 48, 2454-2462.
- Finsterer, J., Winklehner, M., Stollberger, C., Hummel, T., 2020. Unusual Phenotype and Disease Trajectory in Kearns-Sayre Syndrome. *Case Rep Neurol Med* 2020, 7368527.
- Flentke, G.R., Garic, A., Amberger, E., Hernandez, M., Smith, S.M., 2011. Calcium-mediated repression of beta-catenin and its transcriptional signaling mediates neural crest cell death in an

- avian model of fetal alcohol syndrome. *Birth Defects Res A Clin Mol Teratol* 91, 591- 602.
- Flickinger, R.A., Jr., 1949. A study of the metabolism of amphibian neural crest cells during their migration and pigmentation in vitro. *J Exp Zool* 112, 465-484.
- Folmes, C.D., Nelson, T.J., Martinez-Fernandez, A., Arrell, D.K., Lindor, J.Z., Dzeja, P.P., Ikeda, Y., Perez-Terzic, C., Terzic, A., 2011. Somatic oxidative bioenergetics transitions into pluripotency-dependent glycolysis to facilitate nuclear reprogramming. *Cell Metab* 14, 264- 271.
- Gallik, K.L., Treffy, R.W., Nacke, L.M., Ahsan, K., Rocha, M., Green-Saxena, A., Saxena, A., 2017. Neural crest and cancer: Divergent travelers on similar paths. *Mech Dev* 148, 89-99.
- Garic, A., Flentke, G.R., Amberger, E., Hernandez, M., Smith, S.M., 2011. CaMKII activation is a novel effector of alcohol's neurotoxicity in neural crest stem/progenitor cells. *J Neurochem* 118, 646-657.
- Garic-Stankovic, A., Hernandez, M.R., Chiang, P.J., Debelak-Kragtorp, K.A., Flentke, G.R., Armant, D.R., Smith, S.M., 2005. Ethanol triggers neural crest apoptosis through the selective activation of a pertussis toxin-sensitive G protein and a phospholipase Cbeta-dependent Ca²⁺ transient. *Alcohol Clin Exp Res* 29, 1237-1246.
- Gautheron, D.C., 1984. Mitochondrial oxidative phosphorylation and respiratory chain: review. *J Inherit Metab Dis* 7 Suppl 1, 57-61.
- Group, T.T.C.S.C., 1996. Positional cloning of a gene involved in the pathogenesis of Treacher Collins syndrome. The Treacher Collins Syndrome Collaborative Group. *Nat Genet* 12, 130-136.
- Gu, W., Gaeta, X., Sahakyan, A., Chan, A.B., Hong, C.S., Kim, R., Braas, D., Plath, K., Lowry, W.E., Christofk, H.R., 2016. Glycolytic Metabolism Plays a Functional Role in Regulating Human Pluripotent Stem Cell State. *Cell Stem Cell* 19, 476-490.
- Haghighi Poodeh, S., Alhonen, L., Salonurmi, T., Savolainen, M.J., 2014. Ethanol-induced impairment of polyamine homeostasis--a potential cause of neural tube defect and intrauterine growth restriction in fetal alcohol syndrome. *Biochem Biophys Res Commun* 446, 173-178.
- Hanover, J.A., Krause, M.W., Love, D.C., 2012. Bittersweet memories: linking metabolism to epigenetics through O-GlcNAcylation. *Nat Rev Mol Cell Biol* 13, 312-321.
- Hassler, J.A., Moran, D.J., 1986. Effects of ethanol on the cytoskeleton of migrating and differentiating neural crest cells: possible role in teratogenesis. *Journal of craniofacial genetics and developmental biology. Supplement 2*, 129-136.
- Houghton, F.D., 2006. Energy metabolism of the inner cell mass and trophectoderm of the mouse blastocyst. *Differentiation* 74, 11-18.
- Hovland, A.S., Rothstein, M., Simoes-Costa, M., 2020. Network architecture and regulatory logic in neural crest development. *Wiley Interdiscip Rev Syst Biol Med* 12, e1468.
- Hsu, P.P., Sabatini, D.M., 2008. Cancer cell metabolism: Warburg and beyond. *Cell* 134, 703- 707.
- Hu, N., Strobl-Mazzulla, P.H., Bronner, M.E., 2014. Epigenetic regulation in neural crest development. *Dev Biol* 396, 159-168.
- Intlekofer, A.M., Finley, L.W.S., 2019. Metabolic signatures of cancer cells and stem cells. *Nat Metab* 1, 177-188.
- Ito, K., Suda, T., 2014. Metabolic requirements for the maintenance of self-renewing stem cells. *Nat Rev Mol Cell Biol* 15, 243-256.
- Iwata, J., Suzuki, A., Pelikan, R.C., Ho, T.V., Sanchez-Lara, P.A., Chai, Y., 2014. Modulation of lipid metabolic defects rescues cleft palate in *Tgfr2* mutant mice. *Hum Mol Genet* 23, 182- 193.

- Jang, H., Kim, T.W., Yoon, S., Choi, S.Y., Kang, T.W., Kim, S.Y., Kwon, Y.W., Cho, E.J., Youn, H.D., 2012. O-GlcNAc regulates pluripotency and reprogramming by directly acting on core components of the pluripotency network. *Cell Stem Cell* 11, 62-74.
- Jones, N.C., Lynn, M.L., Gaudenz, K., Sakai, D., Aoto, K., Rey, J.P., Glynn, E.F., Ellington, L., Du, C., Dixon, J., Dixon, M.J., Trainor, P.A., 2008. Prevention of the neurocristopathy Treacher Collins syndrome through inhibition of p53 function. *Nature medicine* 14, 125-133.
- Jose, C., Bellance, N., Rossignol, R., 2011. Choosing between glycolysis and oxidative phosphorylation: a tumor's dilemma? *Biochim Biophys Acta* 1807, 552-561.
- Kaneko, K.J., 2016. Metabolism of Preimplantation Embryo Development: A Bystander or an Active Participant? *Curr Top Dev Biol* 120, 259-310.
- Kantarci, H., Gou, Y., Riley, B.B., 2020. The Warburg Effect and lactate signaling augment Fgf-MAPK to promote sensory-neural development in the otic vesicle. *Elife* 9.
- Kathagen-Buhmann, A., Maire, C.L., Weller, J., Schulte, A., Matschke, J., Holz, M., Ligon, K.L., Glatzel, M., Westphal, M., Lamszus, K., 2018. The secreted glycolytic enzyme GPI/AMF stimulates glioblastoma cell migration and invasion in an autocrine fashion but can have anti-proliferative effects. *Neuro Oncol* 20, 1594-1605.
- Kerosuo, L., Bronner, M.E., 2016. cMyc Regulates the Size of the Premigratory Neural Crest Stem Cell Pool. *Cell reports* 17, 2648-2659.
- Keuls, R.A., Kojima, K., Lozzi, B., Steele, J.W., Chen, Q., Gross, S.S., Finnell, R.H., Parchem, R.J., 2020. MiR-302 Regulates Glycolysis to Control Cell-Cycle during Neural Tube Closure. *Int J Mol Sci* 21.
- Kim, H., Jang, H., Kim, T.W., Kang, B.H., Lee, S.E., Jeon, Y.K., Chung, D.H., Choi, J., Shin, J., Cho, E.J., Youn, H.D., 2015. Core Pluripotency Factors Directly Regulate Metabolism in Embryonic Stem Cell to Maintain Pluripotency. *Stem Cells* 33, 2699-2711.
- Kitzmiller, J.L., Gavin, L.A., Gin, G.D., Jovanovic-Peterson, L., Main, E.K., Zigrang, W.D., 1991. Preconception care of diabetes. Glycemic control prevents congenital anomalies. *JAMA* 265, 731-736.
- Laforgia, N., Di Mauro, A., Favia Guarnieri, G., Varvara, D., De Cosmo, L., Panza, R., Capozza, M., Baldassarre, M.E., Resta, N., 2018. The Role of Oxidative Stress in the Pathomechanism of Congenital Malformations. *Oxid Med Cell Longev* 2018, 7404082.
- Le Douarin, N., Kalcheim, C., 1999. *The neural crest*, 2nd ed. Cambridge University Press, Cambridge, UK ; New York, NY, USA.
- Lewis, P.M., Dunn, M.P., McMahon, J.A., Logan, M., Martin, J.F., St-Jacques, B., McMahon, A.P., 2001. Cholesterol modification of sonic hedgehog is required for long-range signaling activity and effective modulation of signaling by Ptc1. *Cell* 105, 599-612.
- Li, X., Yu, X., Dai, D., Song, X., Xu, W., 2016. The altered glucose metabolism in tumor and a tumor acidic microenvironment associated with extracellular matrix metalloproteinase inducer and monocarboxylate transporters. *Oncotarget* 7, 23141-23155.
- Li, Y.X., Yang, H.T., Zdanowicz, M., Sicklick, J.K., Qi, Y., Camp, T.J., Diehl, A.M., 2007. Fetal alcohol exposure impairs Hedgehog cholesterol modification and signaling. *Lab Invest* 87, 231-240.
- Liberti, M.V., Locasale, J.W., 2016. The Warburg Effect: How Does it Benefit Cancer Cells? *Trends Biochem Sci* 41, 211-218.
- Liu, X.S., Little, J.B., Yuan, Z.M., 2015. Glycolytic metabolism influences global chromatin structure. *Oncotarget* 6, 4214-4225.

- Lu, C., Thompson, C.B., 2012. Metabolic regulation of epigenetics. *Cell Metab* 16, 9-17.
- Lunt, S.Y., Vander Heiden, M.G., 2011. Aerobic glycolysis: meeting the metabolic requirements of cell proliferation. *Annu Rev Cell Dev Biol* 27, 441-464.
- Martin, K.L., Leese, H.J., 1995. Role of glucose in mouse preimplantation embryo development. *Mol Reprod Dev* 40, 436-443.
- Melo, F.R., Bressan, R.B., Costa-Silva, B., Trentin, A.G., 2017. Effects of Folic Acid and Homocysteine on the Morphogenesis of Mouse Cephalic Neural Crest Cells In Vitro. *Cell Mol Neurobiol* 37, 371-376.
- Miyazawa, H., Aulehla, A., 2018. Revisiting the role of metabolism during development. *Development* 145.
- Mohyeldin, A., Garzon-Muvdi, T., Quinones-Hinojosa, A., 2010. Oxygen in stem cell biology: a critical component of the stem cell niche. *Cell Stem Cell* 7, 150-161.
- Morgan, S.C., Relaix, F., Sandell, L.L., Loeken, M.R., 2008. Oxidative stress during diabetic pregnancy disrupts cardiac neural crest migration and causes outflow tract defects. *Birth Defects Res A Clin Mol Teratol* 82, 453-463.
- Moussaieff, A., Rouleau, M., Kitsberg, D., Cohen, M., Levy, G., Barasch, D., Nemirovski, A., Shen-Orr, S., Laevsky, I., Amit, M., Bomze, D., Elena-Herrmann, B., Scherf, T., Nissim-Rafinia, M., Kempa, S., Itskovitz-Eldor, J., Meshorer, E., Aberdam, D., Nahmias, Y., 2015. Glycolysis-mediated changes in acetyl-CoA and histone acetylation control the early differentiation of embryonic stem cells. *Cell Metab* 21, 392-402.
- Mullarky, E., Cantley, L.C., 2015. Diverting Glycolysis to Combat Oxidative Stress, in: Nakao, K., Minato, N., Uemoto, S. (Eds.), *Innovative Medicine: Basic Research and Development*, Tokyo, pp. 3-23.
- Murphy, M.P., 2009. How mitochondria produce reactive oxygen species. *Biochem J* 417, 1-13.
- Nie, X., Zheng, J., Ricupero, C.L., He, L., Jiao, K., Mao, J.J., 2018. mTOR acts as a pivotal signaling hub for neural crest cells during craniofacial development. *PLoS Genet* 14, e1007491.
- Nishimura, M., Uyeda, K., 1995. Purification and Characterization of a Novel Xylulose 5-Phosphate-activated Protein Phosphatase Catalyzing Dephosphorylation of Fructose-6-phosphate, 2-kinase:Fructose-2,6-bisphosphatase. *J Biol Chem* 270, 26341-26346.
- Noor, E., Eden, E., Milo, R., Alon, U., 2010. Central carbon metabolism as a minimal biochemical walk between precursors for biomass and energy. *Mol Cell* 39, 809-820.
- Oginuma, M., Harima, Y., Tarazona, O.A., Diaz-Cuadros, M., Michaut, A., Ishitani, T., Xiong, F., Pourquie, O., 2020. Intracellular pH controls WNT downstream of glycolysis in amniote embryos. *Nature* 584, 98-101.
- Oginuma, M., Moncuquet, P., Xiong, F., Karoly, E., Chal, J., Guevorkian, K., Pourquie, O., 2017. A Gradient of Glycolytic Activity Coordinates FGF and Wnt Signaling during Elongation of the Body Axis in Amniote Embryos. *Dev Cell* 40, 342-353 e310.
- Oyedele, O.O., Kramer, B., 2013. Nuanced but significant: how ethanol perturbs avian cranial neural crest cell actin cytoskeleton, migration and proliferation. *Alcohol* 47, 417-426.
- Patra, K.C., Hay, N., 2014. The pentose phosphate pathway and cancer. *Trends Biochem Sci* 39, 347-354.
- Plouffe, S.W., Hong, A.W., Guan, K., 2015. Disease Implication of the Hippo/YAP pathway. *Trends Mol Med* 21, 212-222.
- Porter, F.D., Herman, G.E., 2011. Malformation syndromes caused by disorders of cholesterol

- synthesis. *J Lipid Res* 52, 6-34.
- Potente, M., Gerhardt, H., Carmeliet, P., 2011. Basic and therapeutic aspects of angiogenesis. *Cell* 146, 873-887.
- Procaccini, C., Santopaolo, M., Faicchia, D., Colamatteo, A., Formisano, L., de Candia, P., Galgani, M., De Rosa, V., Matarese, G., 2016. Role of metabolism in neurodegenerative disorders. *Metabolism* 65, 1376-1390.
- Quintana, A.M., Hernandez, J.A., Gonzalez, C.G., 2017. Functional analysis of the zebrafish ortholog of HMGCS1 reveals independent functions for cholesterol and isoprenoids in craniofacial development. *PLoS One* 12, e0180856.
- Riobo, N.A., 2012. Cholesterol and its derivatives in Sonic Hedgehog signaling and cancer. *Curr Opin Pharmacol* 12, 736-741.
- Rovasio, R.A., Battiato, N.L., 1995. Role of early migratory neural crest cells in developmental anomalies induced by ethanol. *Int J Dev Biol* 39, 421-422.
- Rovasio, R.A., Battiato, N.L., 2002. Ethanol induces morphological and dynamic changes on in vivo and in vitro neural crest cells. *Alcohol Clin Exp Res* 26, 1286-1298.
- Rubbi, C.P., Milner, J., 2003. Disruption of the nucleolus mediates stabilization of p53 in response to DNA damage and other stresses. *The EMBO journal* 22, 6068-6077.
- Sakai, D., Dixon, J., Achilleos, A., Dixon, M., Trainor, P.A., 2016. Prevention of Treacher Collins syndrome craniofacial anomalies in mouse models via maternal antioxidant supplementation. *Nat Commun* 7, 10328.
- Shi, F.T., Kim, H., Lu, W., He, Q., Liu, D., Goodell, M.A., Wan, M., Songyang, Z., 2013. Ten- eleven translocation 1 (Tet1) is regulated by O-linked N-acetylglucosamine transferase (Ogt) for target gene repression in mouse embryonic stem cells. *J Biol Chem* 288, 20776- 20784.
- Shyh-Chang, N., Ng, H.H., 2017. The metabolic programming of stem cells. *Genes Dev* 31, 336-346.
- Simoës-Costa, M., Bronner, M.E., 2015. Establishing neural crest identity: a gene regulatory recipe. *Development* 142, 242-257.
- Smith, S.M., Garic, A., Flentke, G.R., Berres, M.E., 2014. Neural crest development in fetal alcohol syndrome. *Birth Defects Res C Embryo Today* 102, 210-220.
- Sperber, H., Mathieu, J., Wang, Y., Ferreccio, A., Hesson, J., Xu, Z., Fischer, K.A., Devi, A., Detraux, D., Gu, H., Battle, S.L., Showalter, M., Valensisi, C., Bielas, J.H., Ericson, N.G., Margaretha, L., Robitaille, A.M., Margineantu, D., Fiehn, O., Hockenbery, D., Blau, C.A., Raftery, D., Margolin, A.A., Hawkins, R.D., Moon, R.T., Ware, C.B., Ruohola-Baker, H., 2015. The metabolome regulates the epigenetic landscape during naive-to-primed human embryonic stem cell transition. *Nat Cell Biol* 17, 1523-1535.
- Suda, T., Takubo, K., Semenza, G.L., 2011. Metabolic regulation of hematopoietic stem cells in the hypoxic niche. *Cell Stem Cell* 9, 298-310.
- Valdez, B.C., Henning, D., So, R.B., Dixon, J., Dixon, M.J., 2004. The Treacher Collins syndrome (TCOF1) gene product is involved in ribosomal DNA gene transcription by interacting with upstream binding factor. *Proceedings of the National Academy of Sciences of the United States of America* 101, 10709-10714.
- Varum, S., Rodrigues, A.S., Moura, M.B., Momcilovic, O., Easley, C.A.t., Ramalho-Santos, J., Van Houten, B., Schatten, G., 2011. Energy metabolism in human pluripotent stem cells and their differentiated counterparts. *PLoS One* 6, e20914.

- Vega-Lopez, G.A., Cerrizuela, S., Tribulo, C., Aybar, M.J., 2018. Neurocristopathies: New insights 150 years after the neural crest discovery. *Dev Biol* 444 Suppl 1, S110-S143.
- Wahl, S.E., Kennedy, A.E., Wyatt, B.H., Moore, A.D., Pridgen, D.E., Cherry, A.M., Mavila, C.B., Dickinson, A.J., 2015. The role of folate metabolism in orofacial development and clefting. *Dev Biol* 405, 108-122.
- Wang, X.Y., Li, S., Wang, G., Ma, Z.L., Chuai, M., Cao, L., Yang, X., 2015. High glucose environment inhibits cranial neural crest survival by activating excessive autophagy in the chick embryo. *Sci Rep* 5, 18321.
- Wellen, K.E., Thompson, C.B., 2012. A two-way street: reciprocal regulation of metabolism and signalling. *Nat Rev Mol Cell Biol* 13, 270-276.
- Wentzel, P., Eriksson, U.J., 2011. Altered gene expression in rat cranial neural crest cells exposed to a teratogenic glucose concentration in vitro: paradoxical downregulation of antioxidative defense genes. *Birth Defects Res B Dev Reprod Toxicol* 92, 487-497.
- Wozniak, J.R., Riley, E.P., Charness, M.E., 2019. Clinical presentation, diagnosis, and management of fetal alcohol spectrum disorder. *Lancet Neurol* 18, 760-770.
- Yu, L., Ji, K.Y., Zhang, J., Xu, Y., Ying, Y., Mai, T., Xu, S., Zhang, Q.B., Yao, K.T., Xu, Y., 2019. Core pluripotency factors promote glycolysis of human embryonic stem cells by activating GLUT1 enhancer. *Protein Cell* 10, 668-680.
- Yun, H.R., Jo, Y.H., Kim, J., Shin, Y., Kim, S.S., Choi, T.G., 2020. Roles of Autophagy in Oxidative Stress. *Int J Mol Sci* 21.
- Zhang, D., Tang, Z., Huang, H., Zhou, G., Cui, C., Weng, Y., Liu, W., Kim, S., Lee, S., Perez-Neut, M., Ding, J., Czyz, D., Hu, R., Ye, Z., He, M., Zheng, Y.G., Shuman, H.A., Dai, L., Ren, B., Roeder, R.G., Becker, L., Zhao, Y., 2019. Metabolic regulation of gene expression by histone lactylation. *Nature* 574, 575-580.
- Zhang, J., Ratanasirintrao, S., Chandrasekaran, S., Wu, Z., Ficarro, S.B., Yu, C., Ross, C.A., Cacchiarelli, D., Xia, Q., Seligson, M., Shinoda, G., Xie, W., Cahan, P., Wang, L., Ng, S.C., Tintara, S., Trapnell, C., Onder, T., Loh, Y.H., Mikkelsen, T., Sliz, P., Teitell, M.A., Asara, J.M., Marto, J.A., Li, H., Collins, J.J., Daley, G.Q., 2016. LIN28 Regulates Stem Cell Metabolism and Conversion to Primed Pluripotency. *Cell stem cell* 19, 66-80.
- Zhang, Y., Wang, H., Li, Y., Peng, Y., 2018. A review of interventions against fetal alcohol spectrum disorder targeting oxidative stress. *Int J Dev Neurosci* 71, 140-145.
- Zheng, J., 2012. Energy metabolism of cancer: Glycolysis versus oxidative phosphorylation (Review). *Oncol Lett* 4, 1151-1157.
- Zhou, W., Choi, M., Margineantu, D., Margaretha, L., Hesson, J., Cavanaugh, C., Blau, C.A., Horwitz, M.S., Hockenbery, D., Ware, C., Ruohola-Baker, H., 2012. HIF1 α induced switch from bivalent to exclusively glycolytic metabolism during ESC-to-EpiSC/hESC transition. *EMBO J* 31, 2103-2116.

Chapter VI: Co-option of developmental gene circuits in neural crest-derived cancers

Alternative citation⁴:

D Bhattacharya, AP Azambuja, J Copeland, AC White, M Simoes-Costa (2021) Co-option of developmental gene circuits in neural crest-derived cancers. *In prep.*

Abstract:

Re-emergence of stem cell identity in malignant cells involves anomalous activation of developmental pathways that promote proliferation, invasion, and self-renewal. This is exemplified during melanoma progression, where a rare population of de-differentiated cells plays a central role in metastasis and therapy resistance. These progenitor-like melanoma cells display striking similarities to the neural crest, a multipotent stem cell population that gives rise to the melanocytic lineage. While emergence of neural crest identity is crucial to melanoma progression, we still have a superficial understanding of how this stem cell state is established during tumorigenesis. Here we show that the Yap/Tead pathway controls melanoma dedifferentiation by reactivating thousands neural crest-specific regulatory elements in this cancer type. By applying single-cell ATAC-seq to characterize the regulatory programs driving melanoma heterogeneity, we identified a subpopulation of cancer cells that mirror the epigenomic and transcriptional profile of embryonic neural crest cells. These cells were characterized by high activity of Tead factors, which re-deploy developmental gene circuits to sustain this neural crest-like state within melanoma tumors. Crucially, we found that inhibition of Yap-Tead interaction was sufficient to revert melanoma cells to a more differentiated melanoblast state. Our findings demonstrate how developmental regulatory networks are co-opted during cancer progression to promote tumor heterogeneity.

⁴ This work is being prepared as a manuscript for submission. Debadrita Bhattacharya performed all the mouse work, the single-cell ATAC-seq experiments and analysis as well all the functional genomics experiments. Ana Azambuja performed the spatial transcriptomics studies and Jackie Copeland performed the bulk ATAC-seq experiments in human neural crest cells.

Introduction:

Re-activation of developmental pathways underlies key features of cancer cells, including uncontrolled proliferation, cellular plasticity, and invasiveness (Borah et al., 2015; Nwabo Kamdje et al., 2017). Downstream of oncogenic mutations, “cancer-initiating cells” re-acquire progenitor identity, which enables them to form a heterogeneous primary tumor. A few cells within this tumor continue to de-differentiate in response to signals produced by the microenvironment and eventually metastasize to form secondary tumors (Capp, 2019; Friedmann-Morvinski and Verma, 2014). Thus, progressive reversion along the developmental trajectory promotes distinct stages of malignancy (Johnsen et al., 2019; Kulesa et al., 2013; Manoranjan et al., 2012). Despite this, we currently lack an understanding of how embryonic pathways re-emerge during malignant transformation. Given that dedifferentiation to a progenitor state promotes self-renewal, metastasis, and therapy resistance in tumors, defining mechanisms that re-activate developmental processes in cancer cells can provide new strategies for targeting these upstream events during tumorigenesis.

The role of de-differentiated cancer cells in tumor progression is perhaps most relevant in melanoma biology (Kulesa et al., 2013; Shakhova, 2014). Multiple lines of evidence indicate that melanoma progression recapitulates the developmental course of the Neural crest Stem Cells (NCSC), the multipotent embryonic cell population that gives rise to melanocytes (Caramel et al., 2013; Kaufman et al., 2016). These cells are specified in the ectoderm bordering the central nervous system, from which they emerge via epithelial to mesenchymal transition prior to undergoing extensive migration within the developing embryo (Simoës-Costa and Bronner, 2015; Theveneau and Mayor, 2012). This process of neural crest migration has striking similarities to cancer invasion, and transcription factors and signaling systems involved in neural crest development have been implicated in mediating the transition of melanocytes to primary and eventually metastatic melanoma (Kaufman et al., 2016; Rambow et al., 2018; Tsoi et al., 2018).

The re-emergence of NCSC identity also contributes to the phenotypic heterogeneity of melanoma tumors. Bulk and single-cell RNA-seq analysis of patient-derived melanoma biopsies have revealed that these cancer cells transition between an oncogene-addicted “proliferative” state, which sustains tumor growth, and a therapy-resistant “invasive” state, that promotes metastasis (Hoek et al., 2006; Sensi et al., 2011; Tirosh et al., 2016). Crucially, proliferative and invasive melanoma cells are defined by their distinct epigenomic and transcriptional profiles rather than by particular genetic lesions. Thus, without accumulating new mutations, melanoma cells can rapidly switch between the two phenotypes, which essentially involves transitioning from a differentiated melanocytic state (proliferative) to a de-differentiated progenitor state (invasive) (Arozarena and Wellbrock, 2019; Rambow et al., 2019). Recent studies in murine and human melanoma models have evidenced that re-activation of the NC gene regulatory network (GRN) is causal to phenotype switching in these tumors. Indeed, overexpression of NC Transcription Factors (TFs) such as SOX9 and TWIST1 in proliferative melanoma cells is sufficient to mediate their transition to an invasive state (Caramel et al., 2013; Cheng et al., 2015). Additionally, scRNA-seq analysis of patient-derived xenografts has implicated the NC-specific factor RXRG in promoting an invasive cell state that mediates therapy resistance in melanoma (Rambow et al., 2018). These findings indicate that components of the NC GRN are co-opted to drive phenotypic heterogeneity in melanoma tumors and further hint at the existence of common upstream mechanisms that control NCSC fate during embryonic development and tumorigenesis.

To define the molecular mechanisms underlying the re-emergence of NCSC identity, we examined how epigenomic heterogeneity drives cell plasticity in human melanoma tumors. By employing single-cell ATAC-seq, we characterized a continuum of cell states within the tumor, driven by unique regulatory and transcriptional programs. In particular, we identified a subpopulation of melanoma cells whose chromatin accessibility and TF activity profile mirrored that of NCSCs, and which displayed high phenotypic plasticity within the tumor. We uncovered that this NCSC-like cell state is established downstream of the Yap/TeaD

signaling system, which re-deploys the stem cell GRN by activating thousands of NC-specific regulatory elements in melanoma cells. Importantly, silencing of the Yap/Tead pathway using a small molecular inhibitor was sufficient to suppress NCSC fate in melanoma cells, which further underscores the therapeutic relevance of this mechanism. Taken together, our findings reveal how downstream of developmental signaling pathways, embryonic stem cell programs are co-opted to drive intra-tumor heterogeneity during oncogenesis.

Results:

Epigenomic analyses of melanoma xenografts at single-cell resolution:

Recent scRNA-seq analyses of melanoma tumors have identified distinct cellular subpopulations that contribute to the phenotypic heterogeneity of this tumor type (Rambow et al., 2018; Tirosh et al., 2016; Wouters et al., 2020). However, the relevant *cis* and *trans* regulators that drive these cell states within the tumor remain to be determined. To address this, we generated a high-resolution map of the epigenomic landscape of melanoma cells by performing single-cell ATAC-seq (scATAC-seq) on tumor xenografts. We established xenografts from the BRAF^{V600E/K} mutant patient-derived melanoma cell line WM266.4, which was engineered to constitutively express GFP protein (WM266.4-GFP). Immunostaining followed by FACS analysis for the invasive marker AXL and the proliferative marker MITF revealed that ~44% and ~50% WM266.4 cells express these markers, respectively, confirming the inherent phenotypic heterogeneity of this cell line. Subcutaneous injection of 1x10⁶ WM-266.4-GFP cells in age-matched NSG mice (2 males + 2 females) resulted in tumor formation within two weeks and endpoint attainment (1000mm³ tumor volume) within five weeks post-injection. To investigate the diversity of cell states in these xenografts, we FACS-sorted GFP+ cancer cells from individually dissected tumors of each mouse and performed scATAC-seq protocol utilizing the 10x genomics chromium platform.

Altogether we collected 23,665 high-quality cancer cells across eight different xenografts from four individual mice. On average, from our data, we obtained a Fraction of Reads in Peaks (FRiP) score of 77.3% with ~10,000 unique fragments mapping per cell. Next, to identify the distinct cell states, we utilized the Signac package (Satija et al., 2015; Stuart T, 2020) to perform dimensionality reduction and unsupervised clustering of the cancer cells based on variable ATAC peaks. This analysis revealed 13 highly reproducible cell clusters, which were visualized using Uniform Manifold Approximation and Projection (UMAP) (Fig 1B). This single-cell clustering observed in the UMAP space was not driven by sequencing depth, FRiP score, or mouse-specific batch effects.

Previous studies characterizing “proliferative” and “invasive” cells in melanoma tumors have implicated key TFs in the establishment of these phenotypic states. In particular, the transcription factor MITF is the “master regulator” of the proliferative state, while FOSL1 is essential for invasiveness (Verfaillie et al., 2015). Thus, computing TF activity (based on the chromatin accessibility of TF motif) in individual melanoma cells can provide valuable insights into the phenotypic state of the cell. Accordingly, we applied chromVAR analysis on the melanoma scATAC-seq dataset to obtain the “activity score” for all 726 TFs in the JASPAR2020 database. To validate the utility of this approach in defining phenotypic states, we projected the motif scores of MITF and FOSL1 on the UMAP, which revealed complementary activity patterns for the two factors (Fig 1C). Indeed, cell clusters on the right and the middle of the UMAP displayed high MITF activity, while FOSL1 activity was mainly restricted to the far-left clusters of the projection. These visualizations confirmed that our scATAC-seq data recapitulate the phenotypic heterogeneity of melanoma tumors.

To comprehensively characterize the transcriptional programs underlying the distinct epigenomic states, we next defined modules of TFs having synergistic activity across the different clusters. We filtered for the top variable TFs (motif variance > 1.75, n=94) and computed the Pearson correlation score between each motif across all cells in the scATAC-seq dataset. Unsupervised hierarchical clustering

based on correlation coefficients revealed two prominent TF modules, which essentially demarcated the invasive and proliferative regulatory programs, respectively (Fig 1D). Indeed, the first module predominantly consisted of TFs associated with melanoma invasion and cancer metastasis, including FOSL1, JUN, TEADs, TWIST, and RUNX factors. Conversely, the second TF module comprised of MITF, SOX10, IRFs, and YY1, increased activity of which characterize the proliferative state in melanoma tumors (Fig 1D) (Arozarena and Wellbrock, 2019; Verfaillie et al., 2015; Wouters et al., 2020). In addition to these known factors, we also identified previously uncharacterized TF associations. For example, we observed a strong correlation of NeuroD/G, ETV, and PAX motifs with invasive TFs, while GATA factors and SP family motifs were highly correlated with MITF and SOX10. Thus, through this analysis, we unbiasedly defined comprehensive sets of TFs associated with the two phenotypic states of melanoma and compiled our “in-house” Invasive (n=51) and Proliferative (n=43) motif signatures.

We next used these TF motif signatures to establish the phenotypic identity of the different scATAC clusters in our dataset. Interestingly, moving from right to left of the UMAP, we observed a gradual transition from proliferative to invasive phenotype (Fig1E). In agreement with the above observations, we found that Cluster 4 and 11 had the highest enrichment of proliferative motifs in their open chromatin regions, while Clusters 2,9,6 &12 displayed strong activity for invasive motifs (Fig 1E). Clusters located in the middle of the UMAP, including Clusters 0,1 and 5, did not show high enrichment for either signature, suggesting that these could be “intermediate” cell populations transitory between the two extreme phenotypes. In contrast, Clusters 3 and 7 had elevated activity of both proliferative and invasive motifs in their regulatory regions, suggesting that cells of these clusters possess a “dual” phenotypic identity. Taken together, this analysis allowed us to classify the different cell clusters into broad phenotypic groups.

To further validate this phenotypic classification, we computed the “gene score” for individual transcripts by quantifying the ATAC signal at their TSS and associated co-accessible regions. We then used these gene scores to compare the expression of Hoek’s proliferative and invasive gene signatures

across the different clusters (Fig 1F, G) (Hoek et al., 2006). In agreement with the TF motif activity-based classification, we found that Clusters 4,11 and 8 were most “proliferative,” while Clusters 9,12 and 6 had elevated gene scores for the “invasive” signature. Strikingly, Cluster7, but not Cluster3, had uniformly high gene scores for both invasive and proliferative markers, thus confirming that cells in this cluster indeed possess a dual identity (Fig1 F, G). Moreover, Cluster7 also displayed the highest score for several neural crest marker genes, including *Tfap2a*, *Satb2*, and *Ets1*, providing initial clues that these cells may correspond to the “NCSC-like” cells described in melanoma tumors (Fazio et al., 2021; Simoes-Costa and Bronner, 2015). Thus, by combining TF factor activity and gene scores analyses, we characterized a continuum of phenotypic states displayed by melanoma and evidenced the presence of a previously uncharacterized “dual” cell population.

Intrigued by these findings, we next attempted to spatially resolve these different cell states within the melanoma tumor. Accordingly, we performed spatial transcriptomics (10x Genomics Visium platform) on 8um tumor cryosections obtained from four different WM266.4-GFP xenograft tumors. Our spatial transcriptomics data from each tumor section covered over 2800 barcodes and detected ~ 10K unique transcripts in all samples. Standalone analysis of this data identified eight different cell clusters, which broadly corresponded to the phenotypic states identified by our scATAC-seq analysis. However, to ensure that we compare equivalent populations, we utilized Hoek’s gene signatures to determine the proliferative and invasive cell clusters in our spatial transcriptomics data. Using marker genes detected in at least ten barcode we identified ~948 barcodes (33%) that displayed strong expression ($\log_{2}FC \geq 3$) for the proliferative signature and ~ 586 (21%) spots that showed elevated signal ($\log_{2}FC \geq 2$) for the invasive gene set. These percentages agreed with our scATAC-seq data, which had found approximately the same fraction of proliferative (27%) and invasive (20%) cells within our dataset.

Figure 1

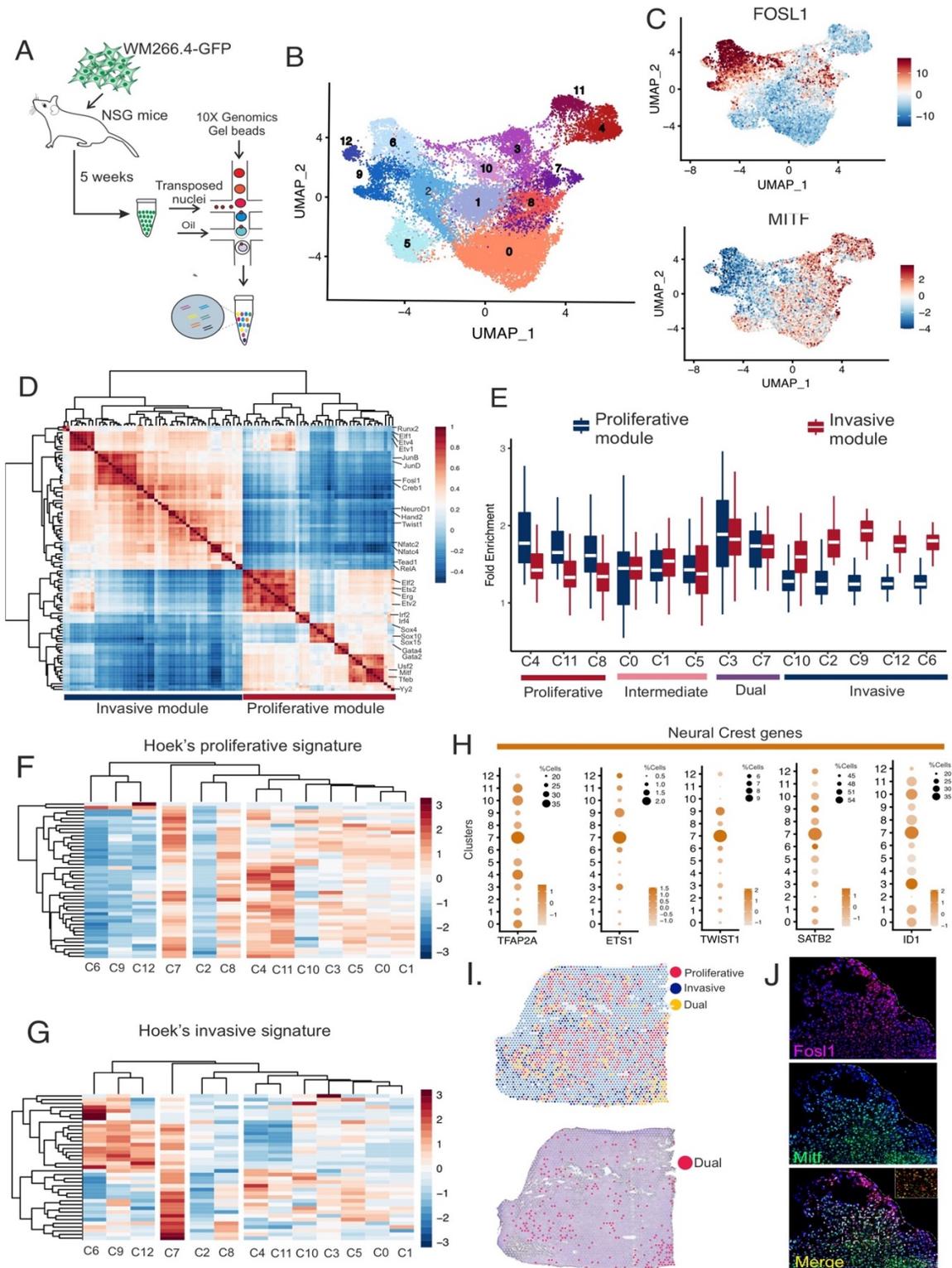


Figure 1: Single-cell chromatin accessibility data defined heterogeneous phenotypic states in melanoma tumor cells

(A) Schematic of the scATAC-seq strategy for profiling GFP⁺ human melanoma xenograft cells using the 10X Chromium platform. **(B)** UMAP visualization of melanoma cells profiled by scATAC-seq. The thirteen cell clusters

identified by nearest neighbor k-means based clustering approach are color-coded and numbered. **(C)** Example TF activity scores are shown on the UMAP for FOSL1 and MITF. **(D)** Hierarchical clustering of top variable TF motifs (variance ≥ 1.75) based on the correlation in their activity scores across individual melanoma cells. The TF motifs associate into two distinct modules that correspond to the invasive and proliferative programs respectively. **(E)** Boxplots comparing the enrichment of invasive and proliferative TF modules in each scATAC-seq cell cluster. Based on differences in combined activity of each module, the clusters can be classified into the four phenotypic states: proliferative, intermediate, dual, and invasive, which are labelled below. **(F-G)** Heatmaps depicting the average gene scores for factors comprising Hoek's proliferative (F) and Invasive (G) signature in the scATAC-seq cell clusters. **(H)** Dot plots showing the gene scores for a set of neural crest markers in the different cell clusters. The size of the dot corresponds to the number of the cells in the cluster that have a finite score for the factor and color represents average gene score. **(I)** *Upper*: Spatial transcriptomics map of a melanoma tumor section showing barcoded spots colored according to their phenotypic identity. Red spots correspond to proliferative state (33% of total), Dark blue spots correspond to Invasive state (21% of total) and the overlap between the two states is depicted in yellow (9%). Light blue spots correspond to background barcodes that do not display high expression for either gene signature. *Lower*: Projection of the "Dual" identity barcoded-spots on the H&E stained image of the tumor section used for spatial RNA-seq. **(J)** Immunostaining for invasive marker FOSL1 and proliferative marker MITF in melanoma tumor section. Individual insets show overlap of Dapi (Blue) and FOSL1 (red), Dapi (Blue) and MITF (green) and the overlaid image of all three channels. Dashed rectangle in inset three corresponds to the tumor region that was magnified to highlight the "Dual" cell population. Dashed white line in all three insets delineates the tumor border.

Upon mapping these barcodes on our H&E stained cryosection, we observed that the proliferative cell clusters were mostly concentrated in and around the tumor center, while the invasive cells existed in smaller clusters and were spread out mostly towards the tumor edges (Fig 1I). In all melanoma sections analyzed, we also found a significant fraction of cells (254 spots) that had high expression of both proliferative and invasive genes, thus corresponding to the "Dual" cell population identified by scATAC-seq. Interestingly, the location of these proliferative-invasive cells was consistently biased towards the tumor periphery, suggesting that the spatial position of these cells may be relevant to their unique phenotypic state. We further validated this observation by immunostaining for FOSL1 and MITF in our tumor cryosections. This revealed a small population of cells located near the tumor border, which displayed high nuclear signals for both of these marker TFs (Fig 1I). The staining also confirmed that overall MITF-high cells were found near the tumor center while FOSL1 positive invasive cells populated the periphery. Altogether, combining scATAC-seq and spatial transcriptomics uniquely enabled us to identify and spatially map the distinct melanoma phenotypic states within the tumor.

NCSC identity is associated with two distinct phenotypic states in melanoma tumor:

Epigenomic profiling of melanoma xenografts revealed a spectrum of cell states driven by the considerable regulatory heterogeneity of tumor cells. To identify which of these cell states most closely resemble NCSCs, we attempted to directly compare the chromatin landscape of human neural crest cells (hNCCs) with that of melanoma tumor cells. We employed an *in vitro* protocol for deriving hNCCs from hESCs by treatment with Wnt agonist CHIR for three days, followed by an additional two days of culture in basal media (Fig 2A)(Gomez et al., 2019a). This protocol resulted in the induction of a highly pure population of hNCCs, which robustly expressed genes associated with pre-migratory and migratory states of neural crest cells *in vivo*. To assess the dynamic changes in chromatin accessibility that drive NC formation, we performed bulk ATAC-seq on cells collected from each day of the induction protocol (D0-D5).

PCA-based dimensionality reduction showed that the samples segregated according to days of induction along the first principal component, representing 40% of the variance (Fig 2B). Pairwise differential peak analysis with Diffbind across all days of induction identified ~6000 highly variable peaks (p . adj value < 0.01) which clustered into four distinct peak sets (Fig 2C). Of these, Cluster2 comprised of ~1100 genomic regions that were closed in hESCs but gradually opened during induction with the highest accessibility in D5 hNCs (D5 peaks) (Fig 2C). Given the dynamics of these D5 peaks, we reasoned that these regulatory elements are likely consequential to NC identity, and thus, mapping their accessibility in the melanoma scATAC-seq data can help identify cell clusters with similar chromatin landscape to NCSCs. To this end, we generated “pseudo-bulk” ATAC-seq datasets from each of the 13 melanoma cell clusters (C0-C12) and quantified the accessibility of D5 peaks in each cluster (Fig 2D). Surprisingly, this analysis revealed that hNCC-specific peaks had high accessibility in four different melanoma cell clusters, including the “dual” Cluster7 and the invasive Clusters9,6 and 12 (Fig 2D).

Next, as an orthogonal approach, we performed chromVar analysis on the bulk ATAC-seq datasets to identify TFs motifs that were enriched in D5 hNCCs. This resulted in a list of ~80 TFs (NC motif set) which included motifs for several factors previously implicated in NC development. We then computed the combined activity score for the NC motif set in individual melanoma cells and projected it on the scATAC UMAP (Fig 2E). Consistent with our previous observations, we found that the “NC motif score” was strongest in cells of Cluster7, 9, 6, and 12, further confirming that the epigenomic and regulatory landscape of these clusters’ mirrors that of hNCCs (Fig 2E). Indeed, melanoma cells having high NC motif activity (NCSC-like cells, n = 2500) represented a distinct epigenomic state within the tumor with about ~2000 genomic loci (NCSC specific peaks) uniquely accessible in this cell population. Furthermore, differential analysis of gene scores in NCSC-like cells revealed *Ngfr* and *Aqp1* gene scores to be enriched in this population (Fig 2F). This is in agreement with several published studies which have implicated these genes as markers of NCSC population in melanoma tumors (Boiko et al., 2010; Fazio et al., 2021; Kasemeier-Kulesa et al., 2018; Sun et al., 2019). Gene scores for factors associated with cell migration and metastasis (*Cldn4*, *Podxl*, *Creb5*, *Tgfb1*) were also highly enriched in the NCSC-like cells, while those for melanocyte differentiation markers (*Mitf*, *Dct*, *Trpm1*) were depleted from this population (Fig 2F). This suggest that the NCSC-like cells identified in our scATAC dataset are phenotypically invasive and de-differentiated, which supports previous characterizations of this cell state in murine and human melanoma tumors (Fazio et al., 2021; Rambow et al., 2018; Sun et al., 2019).

In line with the above analysis, we next wanted to spatially locate these NCSC-like cells in our spatial transcriptomics map of the melanoma tumor. Using factors whose gene scores were enriched in the NCSC population (Fig 2F) as markers, we identified a total of 488 barcoded spots (17%), which had elevated expression ($\log_2FC \geq 2.5$) for the combined NC gene set. Expectedly, a significant proportion of these NCSC-like spots displayed Dual or Invasive phenotypic states (Fig 2G). Consistent with this, the NCSC-like melanoma cells were preferentially situated at the tumor periphery, similar to the dual and invasive populations (Fig 2G). Indeed, inside the melanoma tumor, the NCSC-like cells clustered together within apparent spatial “niches,” in close proximation to each other and other invasive/ dual cells. We

further confirmed this unique spatial location of NCSC-like melanoma cells by performing immunostaining with two distinct NC markers, the cell surface protein NGFR and the TF ETS1 (Fig 2H). This revealed high signal for both factors at the tumor border, with most NGFR+ cells showing ETS1 nuclear localization (Fig 2H). Together, these analyses provide preliminary evidence that paracrine signaling between NCSC-like cells and cells of the tumor microenvironment at the border region is likely essential for sustaining the progenitor cell state in melanoma.

Our scATAC-seq and spatial transcriptomic data indicate that the NCSC-like cell population is phenotypically heterogeneous and comprises both dual (Cluster7) and invasive (Cluster 9,6 &12) melanoma cells. This prompted us to further examine the regulatory programs underlying these two phenotypic states in melanoma. Differential TF activity analysis between the dual and invasive clusters revealed ~100 TF motifs that were variable in the two phenotypic states (Fig 2I). In particular, Cluster7 cells displayed high activity of pluripotency TFs such as SOX2, MYC, MAX, and GATA factors which have also been implicated in neural crest multipotency (Kerosuo and Bronner, 2016; Shu et al., 2015). In contrast, the activity of TFs associated with the regulation of developmental EMT, such as TWIST1, FOSL1, JUN, SMAD2::SMAD3, and RREB1, were higher in the invasive clusters (Fig 2I) (Hao et al., 2019; Kang and Massague, 2004; Su et al., 2020). This observation led us to postulate that the regulatory programs of the dual and invasive clusters recapitulate that of early multipotent and late migratory NC cells, respectively. To test this possibility, we quantified the deviations in the activity of “Dual specific” and “Invasive specific” TF motifs through the course of hNCCs induction. Intriguingly, while both sets of TF motifs have the highest activity in D5 hNCCs, the “Dual specific” motifs also displayed significant activity in hESCs (D0) (Fig 2J). Since several components of the hESC program are repurposed to promote multipotency, this data supports the hypothesis that NCSC-like cells in Cluster7 resemble multipotent NC cells, which could account for their phenotypic plasticity. In contrast, NCSC-like invasive melanoma cells re-activate the migratory NC program, which likely underlies the metastatic potential of this population.

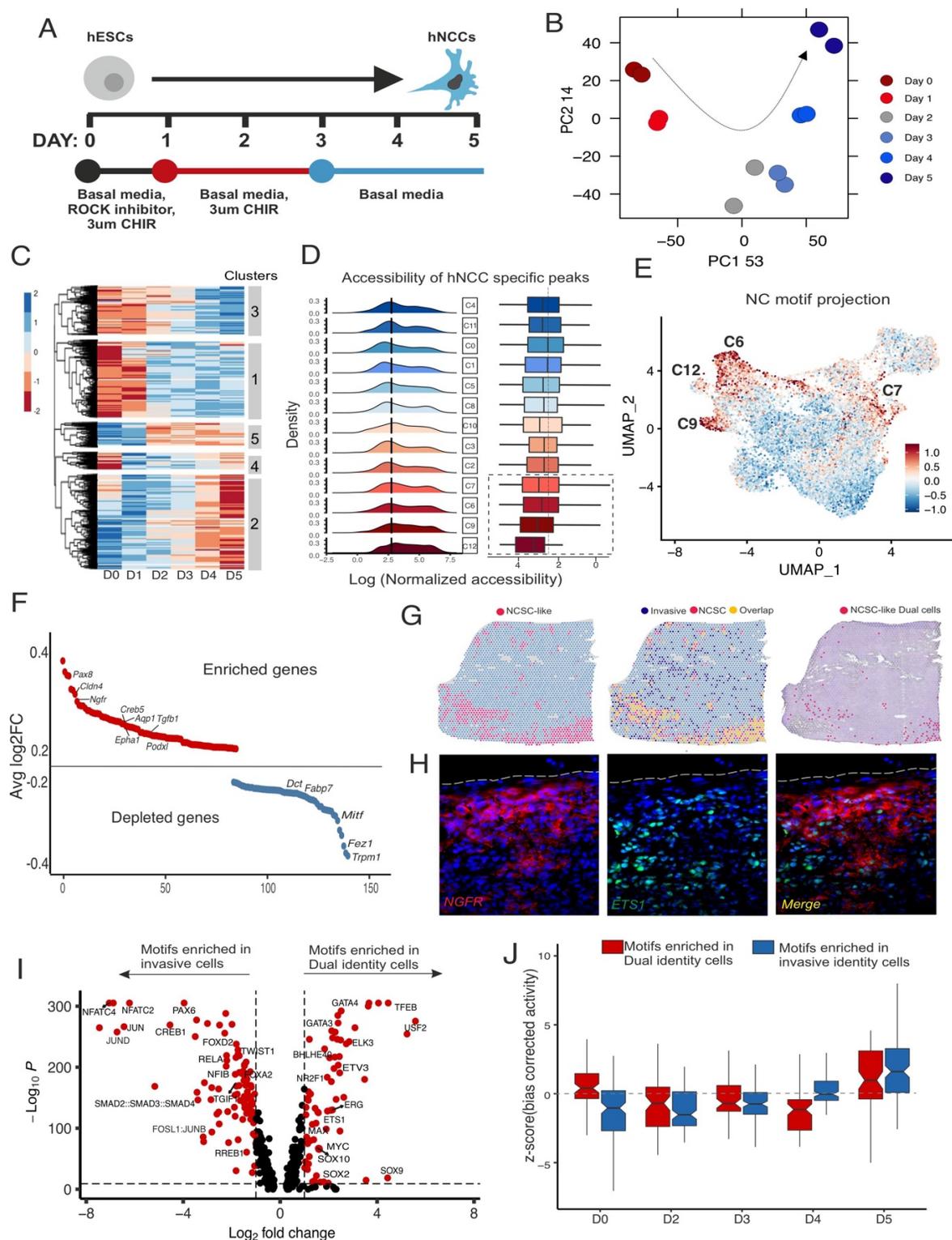


Figure 2: A sub-population of melanoma tumor cells reflect the epigenomic state of neural crest stem cells
(A) Schematic of the 5-day protocol used for obtaining induced human neural crest cells (hNCCs) from human embryonic stem cells (hESCs). WA01 cells were treated for 3-day with 3μM CHIR99021 followed by two-day

incubation in basal media to obtain hNCCs. **(B)** Principal Component analysis (PCA) based dimensionality reduction shows that PC1 represents the progression of hESCs to hNCCs during the course of the 5-day protocol. **(C)** Heatmap depicting the hierarchical clustering of differential ATAC peaks across the different days of hNCC induction. Peaks in Cluster2 are considered “hNCC-specific” as they display highest accessibility in D5 samples. **(D)** Normalized accessibility signal at “hNCC-specific peaks” in the pseudo-bulk ATAC-seq datasets generated from each of the scATAC-seq cell clusters. The dashed lines correspond to the normalized mean accessibility across all clusters. **(E)** Combined activity score of “D5 hNCC motifs” projected on the scATAC-seq UMAP. **(F)** Average Log2 Fold change of factors whose gene scores are enriched or depleted from the NCSC-like melanoma cell population (cells having hNCC motif score ≥ 1.0). **(G)** *Left*: Spatial transcriptomics map showing the barcoded spots (magenta) that correspond to NCSC-like melanoma cells within the tumor. Light blue spots signify background barcodes. *Center*: Spatial RNA-seq profile of melanoma tumors showing the overlap (yellow) between spots corresponding to Invasive (dark blue) and NCSC-like (magenta) cell states. *Right*: Projection of the NCSC-like dual identity spots on the H&E-stained melanoma tumor section. **(H)** Immunostaining for neural crest markers NGFR and ETS1 in melanoma tumor sections. Individual insets are as follows: *Left*: overlap of Dapi (blue) and NGFR (red) channels, *Center*: overlap of Dapi (blue) and ETS1 (green) channels, *Right*: merged image of all three channels. The dashed white line in each inset demarcates the tumor border. **(I)** Volcano plot depicting the TF motifs that display differential activity in Dual (Cluster7) vs invasive (Clusters 6,9, &12) cells. Motifs labelled in red have a $\log_2FC \geq |1|$ and $p\text{-value} < 10^{-6}$. **(J)** Scaled bias corrected activity of dual and invasive -specific TF motifs across different days of the NC induction protocol.

The Yap/Tead signaling system mediates re-activation of NC identity in melanoma tumors:

We next wanted to identify upstream *trans* factors that re-activate NC identity in discrete populations of melanoma tumor cells. Since NCSC-like melanoma cells are characterized by high accessibility of hNCCs specific regulatory elements (D5 peaks), we postulated that transcription factors that control these genomic regions during development might also be involved in the reopening of these loci in cancer cells. Accordingly, we analyzed TF motif enrichment in D5 hNCCs-specific peaks, which surprisingly uncovered TEAD factors to be the top-ranked motif in these regions (Fig 3 A, B). TEAD transcription factors are components of the Hippo signaling pathway, which upon binding to the co-activator protein Yap1, form an active complex that binds to DNA and induces gene expression (Zhao et al., 2008). The Yap/Tead signaling system is a crucial regulator of cancer metastasis and has also been implicated in promoting the invasive phenotypic state in melanoma tumors (Verfaillie et al., 2015; Zanconato et al., 2016; Zhang et al., 2020). Importantly, studies in multiple vertebrate models have identified essential roles of the Yap/Tead pathway in NC formation, migration, and differentiation *in vivo* (Bhattacharya et al., 2020; Hindley et al., 2016). Our observation that TEAD motifs are abundant in human NC-specific regulatory

regions underscores the evolutionarily conserved function of this signaling system in NC development. Further, given the relevance of this pathway in melanoma tumors, this finding led us to postulate that Yap/TeaD promotes phenotypic plasticity of melanoma tumors by re-activating the NC program in these cancer cells. To explore this possibility, we projected the Tead1 motif score of individual cells on the melanoma UMAP, which revealed that the activity pattern of the TF (Fig 3C) broadly recapitulated that of NC motif scores (Fig 2E). Indeed, mean normalized Tead1 and NC motif scores were strongly correlated ($r = 0.66$) for each cluster, with Clusters 7, 9, 6 and 12 displaying high values for both parameters (Fig 3D). Consistent with this, co-staining for the NCSC marker NGFR and Yap1 in melanoma tumor sections showed that most NGFR⁺ cells displayed a high nuclear signal for the co-activator protein (Fig 3E). Together, these analyses provide initial evidence that the Yap/TeaD signaling system may promote melanoma plasticity by re-activating the NC developmental program in these cells.

Next, to examine whether modules of the Yap/TeaD transcriptional program are shared between human NC and melanoma cells, we attempted to identify the direct targets of this pathway in these two biological contexts. We employed Cleavage Under Targets & Release Using Nuclease (CUT&RUN) (Skene and Henikoff, 2017) to profile the genomic occupancy of Yap1 in hNCCs and differentiated melanoblast cells (melanocyte precursors) as well as in four different melanoma cell lines (WM266.4, SKMEL-28, SKMEL-2, MEWO). This experimental design allowed us to directly assay the dynamics of the Yap/TeaD signaling during melanogenesis and tumorigenesis (Dupin and Le Douarin, 2003). CUT&RUN for the active form of Yap1 (actYap1) revealed strong binding of the factor at thousands of genomic loci (~8500 peaks) in hNCCs, which was completely lost upon differentiation of these cells to melanoblast (Fig 3F). Intriguingly, almost half of the actYap1 peaks in hNCCs regained Yap1 binding in the cancer cell lines (Fig 3F), suggesting that the pathway re-activates these NC-specific genomic elements in melanoma cells. We next assessed the relevance of Yap/TeaD activity in establishing the NCSC state in melanoma tumors by quantifying Yap1 binding at the “NCSC-specific peaks” identified in our scATAC-seq analysis. Consistent with our previous data, we observed significant Yap1 occupancy

at these regions in hNCCs and melanoma cell lines but almost complete absence of binding in melanoblasts (Fig 3G). This strongly evidence that the “NCSC-specific peaks” are essentially a subset of the Yap/TeaD targets in multipotent NC cells that are silenced upon differentiation but are re-activated downstream of the pathway in melanoma tumors. Consistent with this, the accessibility of the actYap1 bound regions that are shared between hNCCs and all melanoma cell lines (“shared Yap1 peaks”, n =2350) was highest in the NCSC-like clusters (Clusters 7, 6,9 &12) in our scATAC-seq data, reiterating the connection between Yap activity and NCSC fate in melanoma tumors (Fig 3H).

To determine if Yap1-bound genomic regions function as regulatory elements in hNCCs and melanoma cell lines, we performed CUT&RUN for the active chromatin mark H3K27ac in the cell types described above. PCA analysis based on H3K27ac levels at shared Yap1 peaks revealed that the hNCCs and melanoblast samples segregate along the first principal component (PC1 46%), but all four melanoma cell lines cluster in close proximity of hNCCs in the PCA space (Fig 3I). This visualization indicates that tumorigenesis reverses the epigenomic alterations that accompany differentiation, with H3K27ac at Yap1 peaks “cycling” back to an hNCC state in melanoma cell lines. Consistent with this, quantitation of H3K27ac signal at the shared Yap1 peaks confirmed the high activity of these regions in hNCCs, which was significantly reduced in differentiated melanoblasts but was reinstated in melanoma cells lines (Fig 3J). This activity dynamics of Yap1 peaks were further evident upon inspection of individual regulatory elements in the *Axl* gene loci, which were associated with H3K27ac and Yap1, specifically in hNCCs and the WM266.4 cell line (Fig 3K). These data provide empirical evidence that human NC and melanoma cells share thousands of regulatory regions, the activity of which is controlled by the dynamics of the Yap/TeaD signaling pathway.

Figure 3

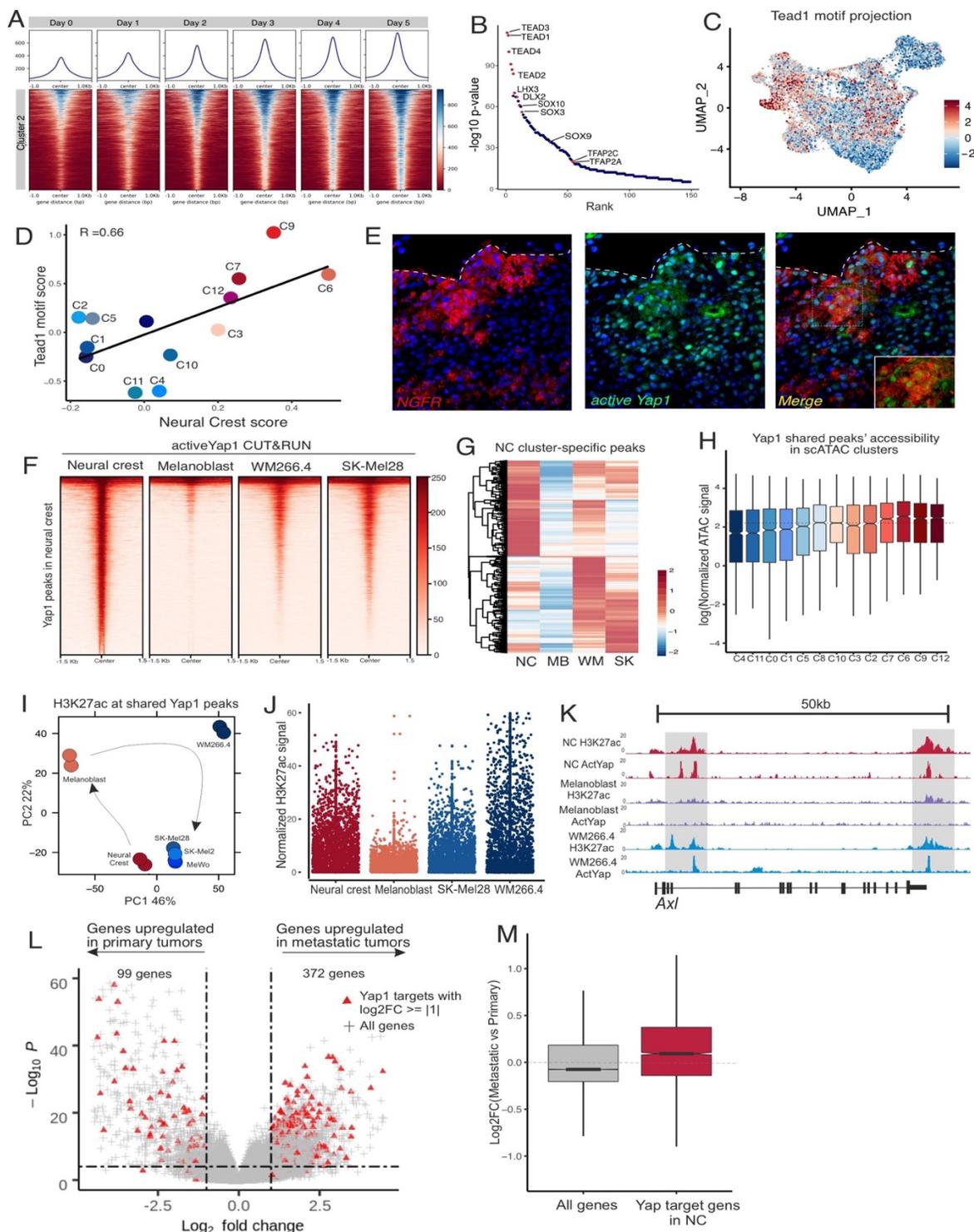


Figure 3: The Yap/Tead signaling pathway re-activates the neural crest regulatory program in human melanoma tumors

(A) Tornado plots showing the ATAC-seq signal at hNCC-specific peaks through the course of neural crest induction.

(B) A ranked-motif plot of TFs enriched in the hNCC-specific peaks (depicted in A) as determined through HOMER analysis. **(C)** Motif activity scores of TEAD1 projected on the scATAC-seq UMAP.

(D) Plot showing the correlation between NC motif and Tead1 activity score in individual scATAC-seq cell clusters. **(E)** Immunostaining for neural crest marker NGFR and the activated form of Yap1 (actYap1) in melanoma tumor sections. Individual insets: *Left*: Overlap of Dapi (blue) and NGFR (red) channels, *Center*: overlap of Dapi (blue) and actYap1 (green) and *Right*: Overlay of all three channels. The dashed rectangle in the third inset marks the region that is magnified to highlight that NGFR+ cells display high nuclear signal for Yap1. The dashed white line in each inset demarcates the tumor border. **(F)** Heatmaps displaying actYap1 (active-Yap1) CUT&RUN signal at hNCC Yap1 peaks in the different cell types. **(G)** Hierarchical clustering of NCSC-specific peaks based on actYap1 binding signal in hNCCs, melanoblast and the melanoma cell lines WM266.4 and SK-Mel28. **(H)** Boxplots showing the accessibility of actYap1 peaks shared between hNCCs and melanoma cells (shared peaks), in the pseudo-bulk chromatin accessibility data generated from each of the scATAC-seq cell clusters. **(I)** Principal component-based dimensionality reduction using H3K27ac signaling at shared actYap1 peaks shows that melanoma cell lines (WM-266.4, SK-Mel2, SK-Mel28 and MeWo) cluster close to hNCCs than melanoblast cells in the PCA space. **(J)** H3K27ac signal at shared actYap1 peaks in hNCCs, melanoblast and the two melanoma cells lines SK-Mel28 and WM266.4. **(K)** A genome browser view of actYap1 and H3K27ac CUT&RUN at the *Axl* locus. Highlighted areas in grey show a high signal for both the factors in hNCCs and melanoma cell line WM266.4 but loss of binding of the factors in melanoblast cells. **(L)** Volcano plot showing differentially expressed genes between metastatic and primary human melanoma tumors. The red triangles depict Yap1 target genes that are significantly upregulated or downregulated ($\text{Log}_2\text{FC} \geq |1|$) in metastatic tumor biopsies. **(M)** Boxplots depicting the Log_2 Fold Change in the expression of Yap1 target genes in metastatic vs. primary melanoma tumors. Compared to the entire transcriptome (all genes), Yap1 targets display higher expression overall in metastatic melanoma biopsies.

Lastly, we wanted to examine whether Yap/Tea driven re-activation of the NC transcriptional program is consequential to melanoma progression and disease outcome. To this end, we assessed the expression of the NC-specific Yap/Tea target genes in the TCGA transcriptome data of 469 subcutaneous melanoma patient biopsies. This analysis revealed that hundreds of Yap1 target genes are preferentially upregulated in metastatic melanoma tumors compared to primary tumors (355 vs. 99 genes) (Fig 3M). Indeed, Yap target genes in hNCCs overall had a significantly higher expression in metastatic melanoma tumors (Figure 3N). Taken together, our data evidence that downstream of Yap/Tea signaling, dedifferentiation of melanoma cells to a phenotypically plastic state promotes the metastatic program in this tumor type.

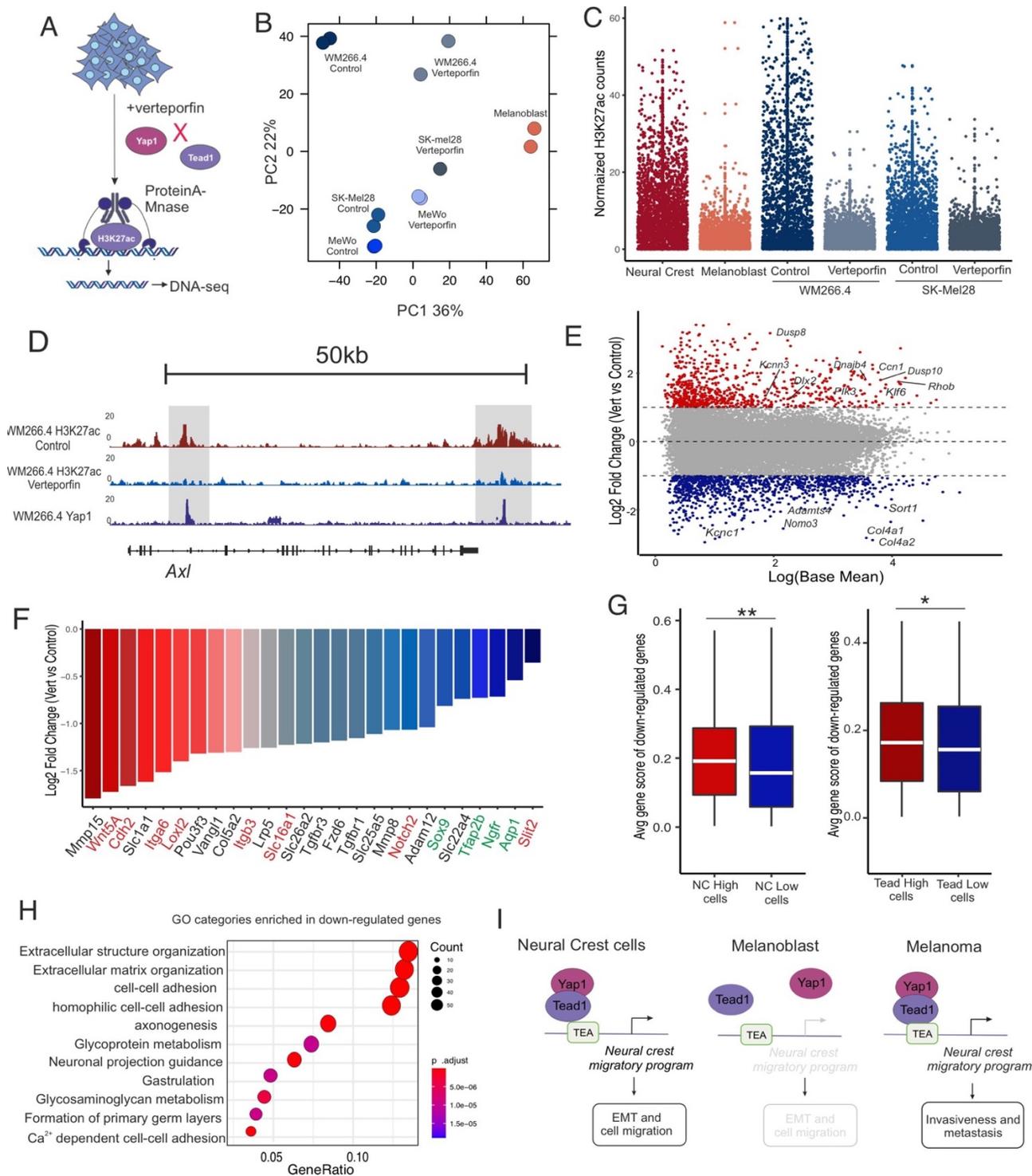


Figure 4: Inhibition of Yap/Tead signaling reverts melanoma dedifferentiation

(A) Experimental scheme for assaying the regulatory changes caused by the inhibition of Yap/Tead pathway in melanoma cells. The different cancer cell lines were treated with the inhibitor Verteporfin for 24 hours, and the H3K27ac landscape of control and drug-treated cells was profiled by CUT&RUN.

(B) Principal component analysis based on H3K27ac signal at shared Yap1 peaks in control and inhibitor-treated cells show that verteporfin samples cluster closer to differentiated melanoblast cells along PC1.

(C) Scatter plot showing H3K27ac signal at shared Yap1 peaks in hNCCs, melanoblast, and control and drug-treated melanoma cell lines WM266.4 and SK-mel28. **(D)** Genome browser view of H3K27ac CUT&RUN at the *Axl* locus in control and Verteporfin treated WM266.4 cells. The highlighted areas in grey, which correspond to ActYap peaks, show loss of H3K27ac signal in the inhibitor-treated sample. **(E)** MA plot highlighting the genes whose expression are significantly altered ($\log_2FC \geq |1|$, $p\text{-value} > 10^{-3}$ in Verteporfin treated samples). Red and blue dots correspond to genes that are respectively upregulated and downregulated in inhibitor-treated cells. **(F)** Barplots highlighting the decreased expression of invasive and NC marker genes in Verteporfin treated melanoma cell lines. Genes labelled in red are part of Hoek's invasive gene signature, while those labeled in green are NC marker genes. **(G)** Boxplots depicting the average gene scores of downregulated transcripts in the NC^{high} vs. NC^{low} cells and the $Tead^{\text{high}}$ vs. $Tead^{\text{low}}$ cells identified by scATAC analysis. **(H)** Dot plots showing the top enriched GO term categories associated with genes downregulated upon Verteporfin treatment. **(I)** A model for reactivation of neural crest identity in melanoma cells. During embryonic development, Yap/Tead signaling transcriptionally regulates the neural crest migratory program to promote EMT. Upon melanoblast differentiation, Yap/Tead signaling is inactivated, which results in silencing of the neural crest migratory program. However, in melanoma cells, activation of the Yap/Tead pathway downstream of oncogenic cues results in the re-activation of the neural crest migratory program, which promotes melanoma invasion and metastasis.

Inhibition of Yap/Tead signaling rewires the epigenome and transcriptome of human melanoma cells

Our findings thus far show that the Yap/Tead pathway re-activates thousands to NC-specific regulatory elements to sustain a progenitor state in melanoma tumors. This led us to postulate that disruption of this signaling should be sufficient to reverse melanoma dedifferentiation. To test this, we sought to inhibit the pathway using the small molecule inhibitor Verteporfin, which prevents Yap-Tead complex formation (Feng et al., 2016; Lui et al., 2019). To assess the effect of Yap inhibition on the activity of NC-specific enhancers in melanoma cells, we performed CUT&RUN for H3K27ac in control, and Verteporfin treated samples of three cancer cell lines (Fig 4A). PCA analysis using H3K27ac counts at the shared Yap1 peaks showed that compared to their corresponding controls, the inhibitor-treated samples were shifted closer to differentiated melanoblast cells along PC1 (Fig 4B). Consistent with this, we found that upon Verteporfin treatment, H3K27ac at shared Yap1 peaks was reduced to levels akin to that observed in melanoblast cells. This data indicates that disruption of Yap/Tead signaling results in the silencing of the developmental enhancers regulated by this pathway in melanoma cells. This was further evident upon examining the *Axl* gene locus, which showed complete loss of H3K27ac at actYap1 peaks in inhibitor-treated WM266.4 melanoma cells (Fig 4D).

To interrogate how these regulatory changes impact the transcriptome of melanoma cells, we performed bulk RNA-seq in control and Verteporfin-treated cell lines. Differential gene expression analysis between the two conditions revealed hundreds of genes that were significantly downregulated ($n = 562$, $\log_2FC \geq |1|$, $p\text{-value} < 10^{-3}$) and upregulated ($n = 373$) upon inhibitor-treatment in all three melanoma cell lines (Fig 4E). In agreement with our previous data, we found that the expression of several invasive genes such as *Wnt5a*, *Itgb3*, *Cdh2*, etc., was considerably decreased upon Verteporfin treatment (Fig 4F). Importantly, we also observed lowered transcription of NC marker genes, including *Ngfr*, *Aqp1*, *Sox9*, and *Tfap2b* in inhibitor-treated melanoma cells (Fig 4F). To determine whether these downregulated genes characterized the NCSC-state identified in our scATAC-seq data, we computed the average gene scores for these transcripts in NC^{high} (NC motif score > 0.75) and NC^{low} (NC motif score < -0.75) cells of our dataset. This analysis showed that the gene scores for downregulated transcripts were significantly higher in the NCSC-like cells than their more differentiated counterparts, consistent with our hypothesis that Verteporfin treatment specifically affects this cell state in melanoma tumors (Fig 4G). We also observed a similar trend in Tead^{high} (Tead motif score ≥ 2.5) vs. Tead^{low} (Tead motif score ≥ 2.5), but the difference was less pronounced than that observed in the previous comparison (Fig 4G). Lastly, in agreement with a role of Yap/Tea signaling in promoting melanoma invasion, GO category analysis with the significantly downregulated genes revealed an enrichment of GO terms associated with cell adhesion, extracellular matrix reorganization, and developmental migration events like gastrulation and germ layer formation (Fig 4H). Taken together, these results support a model where the dynamic flux of Yap/Tea signaling controls the activity of thousands of developmental enhancers that are relevant for NC migration during embryogenesis and for the re-emergence of a stem cell state during melanoma formation (Fig 4I).

Discussion

This study presents the first single-cell ATAC-seq dataset characterizing the epigenomic heterogeneity of human melanoma tumors. To comprehensively delineate the *cis-regulatory* elements and TF modules that drive melanoma cell states, we profiled the chromatin landscape of over 20K individual cancer cells obtained from xenografts. Next, using differential ATAC peaks as markers of cell state, we identified tens of cell clusters that revealed a continuum of regulatory transitions. Defining the epigenomic signature of these diverse melanoma cell states allowed us to directly map the accessibility profile and regulatory program of hNCCs on this dataset. This served to unbiasedly identify the NCSC-like tumor population and further examine the molecular basis of their progenitor identity. First, we found that this cell population is phenotypically heterogeneous and comprises invasive cells as well as cells that display a “Dual” identity. This phenotypic diversity, driven by distinct modules of the embryonic NC GRN, demonstrated that NCSC-like melanoma cells recapitulate discrete stages of NC development within the tumor. Second, motif enrichment analysis and mapping DNA-protein interactions revealed that the Yap/Tead signaling pathway, a key regulator of NC formation during development, is also consequential for the re-activation of NCSC-identity in melanoma during tumorigenesis. Finally, we showed that inhibition of the Yap/Tead pathway is sufficient to promote a differentiated melanoblast-like state in melanoma cells, characterized by a low activity of this signaling system. These results indicate that the dynamics of the Yap/Tead pathway during embryogenesis and oncogenesis control NC formation and differentiation, as well as the re-emergence of NCSC identity in melanoma tumors.

In addition to characterizing the NCSC-like cell state, in this study, we also defined the *cis* and *trans-regulatory* factors driving intra-tumoral phenotypic heterogeneity in melanoma. Using TF modules driving chromatin accessibility differences as markers of proliferative or invasive melanoma, we classified the phenotypic states of all the cell clusters identified in our dataset. This revealed that, rather than just two discrete states, melanoma tumor cells display a spectrum of phenotypic identities that ranged from a

differentiated melanocytic fate to a highly mesenchymal invasive state. However, a large fraction of cells in our scATAC-seq dataset did not correspond to either of these extreme phenotypes but rather existed in intermediate states, characterized by high activity of specific subsets of proliferative and invasive TFs. These cell populations likely represent stable transitory states that can rapidly switch to either of the phenotypic extremes downstream of different microenvironmental cues. Additionally, we also identified a highly plastic “Dual” state that showed elevated activity of both proliferative and invasive markers and displayed strong enrichment for pluripotency TFs such as POU factors, SOX2 and MYC. Since a significant fraction of these dual identity cells were also “NCSC-like,” we postulate that this cluster comprises the most de-differentiated cells within the tumor, which have the potential to both sustain tumor growth and promote massive metastasis without undergoing a phenotypic switch. Interestingly a recent study by Tuncer and colleagues also identified a TGF- β regulated proliferative invasive cell population in murine melanoma tumors and human cell lines (Tuncer et al., 2019). However, in our scATAC-seq dataset, the TGF- β SMAD2/3 motif was not enriched in our dual identity cells but was highly active in the invasive clusters. While this suggests that the regulatory program controlling our Dual identity cluster is different from that described by Tuncer et al., we cannot rule out the possibility that disruption of TGF- β signaling affects the functionality of this unique phenotypic state *in vivo*.

This study also revealed a crucial role of Yap/Tead signaling in the re-emergence of NCSC identity in melanoma cells. The Yap/Tead pathway has been extensively studied in multiple cancer types, and several reports have implicated this signaling system as a general regulator of tumor metastasis. In the context of melanoma, overexpression of Yap promotes invasion at the expense of tumor growth, and metastatic melanoma-specific enhancers have an over-representation of Tead1 binding motifs (Verfaillie et al., 2015; Zhang et al., 2020). Consistent with this, in our scATAC-seq data, we observe robust activity of TEAD factors in the highly invasive melanoma clusters. However, unlike other metastasis promoting TFs such as FOSL1, JUN, SMAD2/3, etc., TEAD was the highest enriched motif in the Dual identity cells and also displayed basal activity in some of the transitory clusters. Thus, we propose that Yap signaling likely functions

upstream of other invasion-specific factors to mediate the early epigenomic changes that push proliferative cells towards more phenotypically heterogeneous states. Since transitioning away from the proliferative state involves progressive dedifferentiation (Arozarena and Wellbrock, 2019; Tsoi et al., 2018), this postulate is consistent with our finding that Yap/Tea activity is essential for the emergence of the NCSC identity in melanoma cells. Together, our data indicate that in addition to its general pro-metastatic function, Yap/Tea signaling promotes phenotypic plasticity in melanoma cells by re-activating embryonic stem cell programs specific to this cancer type.

Finally, our findings provide a general framework for understanding how developmental programs are co-opted during tumorigenesis. We show that the NCSC-like melanoma cells are not limited to expressing a few marker genes of the stem cell population, but in fact, re-use thousands of developmental enhancers to robustly deploy the NC GRN downstream of specific signaling cues. Thus, re-activation of progenitor identity during tumorigenesis is similar to cellular reprogramming in that it necessitates differentiated cells to considerably alter their chromatin landscape to retrace the developmental trajectory and attain a stable epigenomic state akin to embryonic cell types. While our data indicate that the Yap/Tea pathway mediates this regulatory transition in melanoma, other developmental signals such as the Wnt, Shh, and TGF- β pathway have also been implicated in the control of stem cell identity in different cancers (Kahn, 2018; Karamboulas and Ailles, 2013; Milla et al., 2012). This suggests that reactivation of embryonic signaling systems during oncogenesis is a general mechanism to induce the extensive epigenomic reprogramming required to transition differentiated cancer cells into a stem cell state in distinct tumor types.

Methods:

Cell culture

The melanoma cell line SK-MEL28 and MeWo were obtained from ATCC, while the WM266.4 and SK-MEL2 were a kind gift from Dr. Jeremy Baskin's lab (Shami Shah et al., 2021). The melanoma cell lines were grown in MEM media supplemented with 10%FBS, 1% Glutamax, and 1% anti-anti. All genomic experiments were performed with early passage cells not exceeding P10. The cell lines were grown for at least two passages post-thawing before being used in experiments.

To obtain GFP expressing WM266.4 cell lines, we integrated the pEGIP plasmid, a gift from Linzhao Cheng (Addgene plasmid #26777) (Zou et al., 2009) via lentiviral transduction in early passage WM266.4 cells at an MOI of 1. 48 hours post-transduction, puromycin selection (1ug/ul) was started, and the cells were selected for a total of four passages to ensure all cells were uniformly GFP positive. The fluorescent protein expression in the WM266.4-GFP was also verified by flow cytometry analysis.

Generation of induced neural crest cells(hNCC)

The induction of human neural crest cells from WA01 (H1) embryonic stem cells was performed as previously described (Gomez et al., 2019b). Briefly, H1 cells cultured in mTESR were dissociated with Accumex and plated on matrigel coated dishes at a density of 20K cells/cm² in DMEM/F12 media supplemented with 1% B27, 0.5%BSA and 3uM CHIR (induction media). For the first 24h, 10uM ROCK inhibitor (Y-27632) was added to the induction media to improve cell survival. After 72 hours (Day 3), induction media was substituted for basal media constituting DMEM/F12, 1% B27, and 0.5%BSA. The cells were grown in basal media for an additional 48 hours (Day 5), after which they were processed for downstream analysis. To obtain melanoblast cells from hiNCCs, we treated the induced neural crest cells we supplemented the basal media with 10ng/ul BMP4, 3uM CHIR, and 25uM Endothelin-3 peptides for an additional five days, which resulted in robust expression of differentiation markers such as MITF and TYR (Chambers et al., 2016).

Tumor xenografts with WM266.4-GFP cell lines

On the day of injection, the dorsal side of 5-7 weeks old NSG mice (2 males + 2 females/experiment) were shaved to enable two injections per animal. The WM266.4-GFP cells were dissociated in Accumex, resuspended at a concentration of 1×10^6 in 100ul of PBS, and subcutaneously injected using a 28-gauge needle within 40 mins of their resuspension. The injected mice were monitored every two days. For all xenograft experiments with the WM266.4-GFP cell line, tumor formation was apparent within two weeks, and tumor endpoint (tumor volume = 1000mm^3) was attained within five weeks post-injection. Following endpoint attainment, the NSG mice were euthanized, and tumors were dissected from each flank and processed to downstream analysis. All mouse studies performed in this study were approved by the Cornell Institutional Animal Care and Use Committee.

scATAC-seq with melanoma xenograft cells

To obtain single-cell suspensions of tumor cells, dissected melanoma xenografts were minced with a blunt-ended scissor and dissociated in 2mg/ml collagenase solution for 1 hour at 37°C. The two tumors obtained from the same mice were pooled together, while those obtained from different mice were processed as individual samples. Following tumor dissociation, the cell suspensions were strained with a 100um filter and subsequently a 40um cell filter to remove debris and undigested tissues. The filtered cell suspension was centrifuged at 200g for 10 mins, resuspended in HANKS buffer, and transferred to FACS tubes. From individual mouse tumor samples 100,000 GFP+ tumor cells were sorted into HANKS buffer. Subsequently, the cells were spun down in a cold centrifuge at 200g for 10 mins, the supernatant was discarded, and the pellet resuspended in 75ul Cryostor media and was frozen down at -80°C. On the day of scATAC-seq library preparation, the frozen cells were quickly thawed at 37°C, and 1ml of ice-cold ATAC-RSB buffer was added to each sample to dilute out the freezing media. Samples were centrifuged at 500rcf for 5 minutes at 4°C, and the supernatant was removed. Next, cells were resuspended in 100µl of ATAC-RSB-LYSIS and kept on ice for 6 minutes. To stop lysis, 1mL of ATAC-RSB-WASH was added to each sample, which was again centrifuged for 5 minutes at 500rcf. The

quality and quantity of nuclei thus obtained was evaluated by trypan blue staining followed by cell counting using a hemocytometer, subsequent to which subset of the nuclei was then spun down and resuspended in 10X dilution buffer. These nuclei were then used for single-cell ATAC-seq library preparation using the Chromium Single Cell ATAC solution v1.0 kit (10x Genomics). Completed libraries were further quality checked using an ABI 3730xl DNA Analyzer. The Individual samples obtained from four different mice were were pooled at an equimolar ratio calculated based on Bioanalyzer concentration and sequenced on an Illumina NextSeq500 the using the 75.bp kit in a paired end configuration

RNA extraction, library preparation, and sequencing

For RNA-seq analysis of control and Verteporfin treated melanoma cells, each cell line (WM266.4, SK-Mel28, Mewo) was plated in duplicates and treated with DMSO or 10uM Verteporfin for 24 hours. Post-treatment ~100,000 cells were collected/replicate/sample and were lysed using the lysis buffer from the RNA Aqueous micro kit (ThermoFisher #AM1931). RNA was extracted according to the manufacturer's protocol and quantified using a Qubit RNA HS Assay (ThermoFisher #Q32852). A total of 100ng of RNA/sample was used for preparing libraries using the NEBNext® Ultra™ II Directional RNA Library Prep Kit (NEB #E7765) according to the manufacturer's protocol. The libraries were amplified for 12 cycles and were quantified using a Qubit DNA HS Assay (ThermoFisher #Q33230) and were also checked for fragment size distribution and quality on an ABI 3730xl DNA Analyzer. Individual samples were pooled at an equimolar ratio calculated using the KAPA Library Quantification Kit (Roche #07960336001) and sequenced in a paired-end configuration on an Illumina NextSeq500 using the High Output 37bp kit at the Biotechnology Resource Center at Cornell.

RNA-Seq Analysis

Raw sequencing reads were demultiplexed by the Biotechnology Resource Center and trimmed using CutAdapt (v2.10) (Martin, 2011) with the TruSeq sequencing adapter and a minimum read length of at

least 25. Reads were then aligned to the ENSEMBL hg38 genome using HiSat2 (Kim et al., 2019) using reverse strandedness and discarding unaligned reads. Counts were assigned to genes using featureCounts (Liao et al., 2014) at ENSEMBL gene annotation (v104). Downstream differential expression analysis and GO term enrichment analysis were performed using the R packages DEseq and clusterProfiler.

ATAC-Seq library preparation and sequencing

For time-course human ESC and hNCC bulk ATAC-Seq, cells were collected in duplicates on each day of induction (D0-D5), dissociated using Accumex, and counted. One hundred thousand cells were resuspended in CryoStor media and frozen at -80°C until library preparation. Cells were processed using the OMNI-ATAC-Seq protocol (Corces et al., 2017). Briefly, frozen cells were quickly thawed at 37°C, and ATAC-RSB buffer was added to a total volume of 1mL. Samples were centrifuged at 500rcf for 5 minutes at 4°C, and the supernatant was removed. Next, cells were resuspended in 100µl of ATAC-RSB-LYSIS and kept on ice for 4 minutes. To stop lysis, 1mL of ATAC-RSB-WASH was added to each sample, which was again centrifuged for 5 minutes at 500rcf. The supernatant was removed, and cells were resuspended in 50µl of OMNI-ATAC Mix (~100nM concentration of Illumina TDE1 enzyme). Cells were then tagmented on a mixing (500rpm) thermoblock at 37°C for one hour. Tagmented DNA was recovered using a Qiagen MinElute Kit (#28204), with 21µl of elution buffer warmed to 55°C. Library amplification PCR was performed with the NEBNext Ultra II Q5 2X Master Mix (NEB #M0544S) using Nextera primers for 12 cycles. DNA concentration was measured using a Qubit DNA HS Assay (ThermoFisher #Q33230), and stereotypical nucleosomal banding was observed using an ABI 3730xl DNA Analyzer. Individual samples were pooled at an equimolar ratio calculated based on Bioanalyzer concentration and sequenced on an Illumina NextSeq500 the using the 37bp kit in a paired-end configuration.

ATAC-Seq sequence processing

Raw sequencing reads were demultiplexed by the Biotechnology Resource Center and trimmed using CutAdapt (v2.10) in paired-end mode with the forward and reverse Nextera sequencing adapters and a minimum read length of at least 25. Next, we used Bowtie2 (Langmead and Salzberg, 2012) to align paired-end reads to the UCSC hg38 genomes. We used the options “--local --very-sensitive-local --no-unal --no-mixed --no-discordant” and excluded alignments > 850bp. PicardTools (<https://github.com/broadinstitute/picard>) was used to mark duplicates, which were then filtered from the BAM files using samtools (Li et al., 2009). MACS2 was used to call peaks genome-wide, with a q value of 0.05 and the arguments “-f BAMPE -g \$GENOME_SIZE --nomodel --shift 37 --extsize 73” (Zhang et al., 2008). Downstream analyses, including generating counts at and performing differential peak analysis, were done using the DiffBind program. Tornado plots were generated using DeepTools and motif enrichment analysis was performed with chromVar and HOMER, as described in the text.

CUT&RUN library preparation and sequencing

CUT&RUN was performed as previously described (Rothstein and Simoes-Costa, 2020). Briefly, human neural crest cells, melanoblast cells, or melanoma cells (WM266.4, SK-MEL28, SK-MEL2, and Mewo) were dissociated using Accumex and ~500,000 cells /sample were used for performing CUT&RUN. The cells were washed twice with 1ml Wash buffer, immobilized on BioMag Plus Concanavalin A magnetic beads (Bangs Laboratories, BP531), and incubated with rabbit anti-Active-Yap1 (Abcam, ab205270) (1:50) or anti-Histone H3 (acetyl K27) (Abcam, ab177178) (1:50) antibody overnight at 4°C. After washing away the unbound antibody, 2.5ul of CUTANA™ pAG-MNase was added to the beads and incubated for 1 h at 4°C. Cells were cooled to 0°C, and CaCl₂ was added to a final concentration of 2 mM to activate the MNase enzyme. MNase digestion was performed for 45 minutes and terminated by the addition of 2XSTOP buffer. The protein-DNA complexes were released by incubation at 37°C for 10 mins followed by digestion with proteinase K for 10 minutes at 70°C. DNA fragments were isolated via phenol-chloroform extraction and ethanol precipitation and were quantified using a Qubit DNA HS Assay. A total

of 20-40 ng of purified DNA was used by library preparation with the NEBNext Ultra II DNA Library Prep Kit (New England Biolabs, E7645) according to the manufacturer's protocol. Each library was amplified for 12-13 cycles, and Fragment analysis was performed with ABI 3730xl DNA Analyzer. Equimolar concentrations of the libraries were pooled using the KAPA Library Quantification Kit - ROX Low (Roche, 07960336001) and sequenced with paired-end 37bp reads on an Illumina NextSeq500 instrument.

CUT&RUN sequence processing

Raw sequencing reads were demultiplexed by the Biotechnology Resource Center and trimmed using CutAdapt (v2.10) in paired-end mode with the forward and reverse TruSeq sequencing adapters and a minimum read length of at least 25. Next, samples were aligned to ENSEMBL UCSC hg38 using Bowtie2. We used the following options "--local --very-sensitive-local --no-unal --no-mixed --no-discordant" and excluded alignments > 1000bp. PicardTools (<https://github.com/broadinstitute/picard>) was used to mark duplicates, which were then filtered from the BAM files using samtools. MACS2 was used to call peaks genome-wide, with a q value of 0.05 and the arguments "-f BAMPE -g \$GENOME_SIZE -q 0.05 --call-summits." Downstream analysis based on generating peak counts and PCA analysis was performed using R package DiffBind. Hierarchical clustering of peaks was performed with the pheatmap package in R, and tornado plots were generated with the python package DeepTools.

Immunocytochemistry

Immunocytochemistry for fresh-frozen melanoma tumor tissue sections was performed as previously described. Briefly, slides with OCT embedded sections were fixed in 4% PFA for 10 mins at RT. The slides were then washed thrice in PBST (0.1% Tween20) and blocked with 10% Donkey serum (diluted in PBST) for 1 hour at RT. Appropriate dilutions of the primary antibody was added to the tumor sections, which were then incubated overnight at 4°C. Primary antibodies were prepared in the blocking solution and used at the following dilutions: mouse anti Fra1 antibody (R&D Systems, MAB4935) (1:200), rabbit anti-MITF (Invitrogen, PA5-82074) (1:250), rabbit anti-Yap1 antibody (Abcam, ab205270)(1:500), mouse

anti-NGFR (SCBT, sc-13577) (1:200) and rabbit anti-ETS1 (CST, D8O8A). Following primary antibody incubation, the slides were washed thrice in PBST and incubated with appropriate Alexa-conjugated secondary antibodies diluted 1:1500 in PBST for 1 hour at RT. The slides were then washed in PBST, stained with DAPI, and mounted with Fluormount medium before imaging.

References:

- Arozarena, I., and Wellbrock, C. (2019). Phenotype plasticity as enabler of melanoma progression and therapy resistance. *Nat Rev Cancer* 19, 377-391.
- Bhattacharya, D., Azambuja, A.P., and Simoes-Costa, M. (2020). Metabolic Reprogramming Promotes Neural Crest Migration via Yap/Tead Signaling. *Dev Cell* 53, 199-211 e196.
- Boiko, A.D., Razorenova, O.V., van de Rijn, M., Swetter, S.M., Johnson, D.L., Ly, D.P., Butler, P.D., Yang, G.P., Joshua, B., Kaplan, M.J., *et al.* (2010). Human melanoma-initiating cells express neural crest nerve growth factor receptor CD271. *Nature* 466, 133-137.
- Borah, A., Raveendran, S., Rochani, A., Maekawa, T., and Kumar, D.S. (2015). Targeting self-renewal pathways in cancer stem cells: clinical implications for cancer therapy. *Oncogenesis* 4, e177.
- Capp, J.P. (2019). Cancer Stem Cells: From Historical Roots to a New Perspective. *J Oncol* 2019, 5189232.
- Caramel, J., Papadogeorgakis, E., Hill, L., Browne, G.J., Richard, G., Wierinckx, A., Saldanha, G., Osborne, J., Hutchinson, P., Tse, G., *et al.* (2013). A switch in the expression of embryonic EMT-inducers drives the development of malignant melanoma. *Cancer Cell* 24, 466-480.
- Chambers, S.M., Mica, Y., Lee, G., Studer, L., and Tomishima, M.J. (2016). Dual-SMAD Inhibition/WNT Activation-Based Methods to Induce Neural Crest and Derivatives from Human Pluripotent Stem Cells. *Methods Mol Biol* 1307, 329-343.
- Cheng, P.F., Shakhova, O., Widmer, D.S., Eichhoff, O.M., Zingg, D., Frommel, S.C., Belloni, B., Raaijmakers, M.I., Goldinger, S.M., Santoro, R., *et al.* (2015). Methylation-dependent SOX9 expression mediates invasion in human melanoma cells and is a negative prognostic factor in advanced melanoma. *Genome Biol* 16, 42.
- Corces, M.R., Trevino, A.E., Hamilton, E.G., Greenside, P.G., Sinnott-Armstrong, N.A., Vesuna, S., Satpathy, A.T., Rubin, A.J., Montine, K.S., Wu, B., *et al.* (2017). An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. *Nat Methods* 14, 959-962.
- Dupin, E., and Le Douarin, N.M. (2003). Development of melanocyte precursors from the vertebrate neural crest. *Oncogene* 22, 3016-3023.
- Fazio, M., van Rooijen, E., Dang, M., van de Hoek, G., Ablain, J., Mito, J.K., Yang, S., Thomas, A., Michael, J., Fabo, T., *et al.* (2021). SATB2 induction of a neural crest mesenchyme-like program drives melanoma invasion and drug resistance. *Elife* 10.
- Feng, J., Gou, J., Jia, J., Yi, T., Cui, T., and Li, Z. (2016). Verteporfin, a suppressor of YAP-TEAD complex, presents promising antitumor properties on ovarian cancer. *Onco Targets Ther* 9, 5371-5381.
- Friedmann-Morvinski, D., and Verma, I.M. (2014). Dedifferentiation and reprogramming: origins of cancer stem cells. *EMBO Rep* 15, 244-253.
- Gomez, G.A., Prasad, M.S., Wong, M., Charney, R.M., Shelar, P.B., Sandhu, N., Hackland, J.O.S., Hernandez, J.C., Leung, A.W., and Garcia-Castro, M.I. (2019a). WNT/beta-catenin modulates the axial identity of embryonic stem cell-derived human neural crest. *Development* 146.
- Gomez, G.A., Prasad, M.S., Wong, M., Charney, R.M., Shelar, P.B., Sandhu, N., Hackland, J.O.S., Hernandez, J.C., Leung, A.W., and Garcia-Castro, M.I. (2019b). WNT/ β -catenin modulates the axial identity of embryonic stem cell-derived human neural crest. *Development* 146, dev175604.
- Hao, Y., Baker, D., and Ten Dijke, P. (2019). TGF-beta-Mediated Epithelial-Mesenchymal Transition and Cancer Metastasis. *Int J Mol Sci* 20.
- Hindley, C.J., Condurat, A.L., Menon, V., Thomas, R., Azmitia, L.M., Davis, J.A., and Pruszk, J. (2016). The Hippo pathway member YAP enhances human neural crest cell fate and migration. *Sci Rep* 6, 23208.
- Hoek, K.S., Schlegel, N.C., Brafford, P., Sucker, A., Ugurel, S., Kumar, R., Weber, B.L., Nathanson, K.L., Phillips, D.J., Herlyn, M., *et al.* (2006). Metastatic potential of melanomas defined by specific gene expression profiles with no BRAF signature. *Pigment Cell Res* 19, 290-302.

- Johnsen, J.I., Dyberg, C., and Wickstrom, M. (2019). Neuroblastoma-A Neural Crest Derived Embryonal Malignancy. *Front Mol Neurosci* 12, 9.
- Kahn, M. (2018). Wnt Signaling in Stem Cells and Cancer Stem Cells: A Tale of Two Coactivators. *Prog Mol Biol Transl Sci* 153, 209-244.
- Kang, Y., and Massague, J. (2004). Epithelial-mesenchymal transitions: twist in development and metastasis. *Cell* 118, 277-279.
- Karamboulas, C., and Ailles, L. (2013). Developmental signaling pathways in cancer stem cells of solid tumors. *Biochim Biophys Acta* 1830, 2481-2495.
- Kasemeier-Kulesa, J.C., Romine, M.H., Morrison, J.A., Bailey, C.M., Welch, D.R., and Kulesa, P.M. (2018). NGF reprograms metastatic melanoma to a bipotent glial-melanocyte neural crest-like precursor. *Biol Open* 7.
- Kaufman, C.K., Mosimann, C., Fan, Z.P., Yang, S., Thomas, A.J., Ablain, J., Tan, J.L., Fogley, R.D., van Rooijen, E., Hagedorn, E.J., *et al.* (2016). A zebrafish melanoma model reveals emergence of neural crest identity during melanoma initiation. *Science* 351, aad2197.
- Kerosuo, L., and Bronner, M.E. (2016). cMyc Regulates the Size of the Premigratory Neural Crest Stem Cell Pool. *Cell Rep* 17, 2648-2659.
- Kim, D., Paggi, J.M., Park, C., Bennett, C., and Salzberg, S.L. (2019). Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat Biotechnol* 37, 907-915.
- Kulesa, P.M., Morrison, J.A., and Bailey, C.M. (2013). The neural crest and cancer: a developmental spin on melanoma. *Cells Tissues Organs* 198, 12-21.
- Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9, 357-359.
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and Genome Project Data Processing, S. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25, 2078-2079.
- Liao, Y., Smyth, G.K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* 30, 923-930.
- Lui, J.W., Xiao, S., Ogomori, K., Hammarstedt, J.E., Little, E.C., and Lang, D. (2019). The Efficiency of Verteporfin as a Therapeutic Option in Pre-Clinical Models of Melanoma. *J Cancer* 10, 1-10.
- Manoranjan, B., Venugopal, C., McFarlane, N., Doble, B.W., Dunn, S.E., Scheinemann, K., and Singh, S.K. (2012). Medulloblastoma stem cells: where development and cancer cross pathways. *Pediatr Res* 71, 516-522.
- Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnetjournal* 17, 10.
- Milla, L.A., Gonzalez-Ramirez, C.N., and Palma, V. (2012). Sonic Hedgehog in cancer stem cells: a novel link with autophagy. *Biol Res* 45, 223-230.
- Nwabo Kamdje, A.H., Takam Kamga, P., Tagne Simo, R., Vecchio, L., Seke Etet, P.F., Muller, J.M., Bassi, G., Lukong, E., Kumar Goel, R., Mbo Amvene, J., *et al.* (2017). Developmental pathways associated with cancer metastasis: Notch, Wnt, and Hedgehog. *Cancer Biol Med* 14, 109-120.
- Rambow, F., Marine, J.C., and Goding, C.R. (2019). Melanoma plasticity and phenotypic diversity: therapeutic barriers and opportunities. *Genes Dev* 33, 1295-1318.
- Rambow, F., Rogiers, A., Marin-Bejar, O., Aibar, S., Femel, J., Dewaele, M., Karras, P., Brown, D., Chang, Y.H., Debiec-Rychter, M., *et al.* (2018). Toward Minimal Residual Disease-Directed Therapy in Melanoma. *Cell* 174, 843-855 e819.
- Rothstein, M., and Simoes-Costa, M. (2020). Heterodimerization of TFAP2 pioneer factors drives epigenomic remodeling during neural crest specification. *Genome Res* 30, 35-48.
- Satija, R., Farrell, J.A., Gennert, D., Schier, A.F., and Regev, A. (2015). Spatial reconstruction of single-cell gene expression data. *Nat Biotechnol* 33, 495-502.
- Sensi, M., Catani, M., Castellano, G., Nicolini, G., Alciato, F., Tragni, G., De Santis, G., Bersani, I., Avanzi, G., Tomassetti, A., *et al.* (2011). Human cutaneous melanomas lacking MITF and melanocyte differentiation antigens express a functional Axl receptor kinase. *J Invest Dermatol* 131, 2448-2457.

- Shakhova, O. (2014). Neural crest stem cells in melanoma development. *Curr Opin Oncol* 26, 215-221.
- Shami Shah, A., Cao, X., White, A.C., and Baskin, J.M. (2021). PLEKHA4 Promotes Wnt/beta-catenin Signaling-Mediated G1/S Transition and Proliferation in Melanoma. *Cancer Res*.
- Shu, J., Zhang, K., Zhang, M., Yao, A., Shao, S., Du, F., Yang, C., Chen, W., Wu, C., Yang, W., *et al.* (2015). GATA family members as inducers for cellular reprogramming to pluripotency. *Cell Res* 25, 169-180.
- Simoës-Costa, M., and Bronner, M.E. (2015). Establishing neural crest identity: a gene regulatory recipe. *Development* 142, 242-257.
- Skene, P.J., and Henikoff, S. (2017). An efficient targeted nuclease strategy for high-resolution mapping of DNA binding sites. *Elife* 6.
- Stuart T, S.A., Lareau C, Satija R. (2020). Multimodal single-cell chromatin analysis with Signac. *bioRxiv*.
- Su, J., Morgani, S.M., David, C.J., Wang, Q., Er, E.E., Huang, Y.H., Basnet, H., Zou, Y., Shu, W., Soni, R.K., *et al.* (2020). TGF-beta orchestrates fibrogenic and developmental EMTs via the RAS effector RREB1. *Nature* 577, 566-571.
- Sun, Q., Lee, W., Mohri, Y., Takeo, M., Lim, C.H., Xu, X., Myung, P., Atit, R.P., Taketo, M.M., Moubarak, R.S., *et al.* (2019). A novel mouse model demonstrates that oncogenic melanocyte stem cells engender melanoma resembling human disease. *Nat Commun* 10, 5023.
- Theveneau, E., and Mayor, R. (2012). Neural crest delamination and migration: from epithelium-to-mesenchyme transition to collective cell migration. *Dev Biol* 366, 34-54.
- Tirosh, I., Izar, B., Prakadan, S.M., Wadsworth, M.H., 2nd, Treacy, D., Trombetta, J.J., Rotem, A., Rodman, C., Lian, C., Murphy, G., *et al.* (2016). Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science* 352, 189-196.
- Tsoi, J., Robert, L., Paraiso, K., Galvan, C., Sheu, K.M., Lay, J., Wong, D.J.L., Atefi, M., Shirazi, R., Wang, X., *et al.* (2018). Multi-stage Differentiation Defines Melanoma Subtypes with Differential Vulnerability to Drug-Induced Iron-Dependent Oxidative Stress. *Cancer Cell* 33, 890-904 e895.
- Tuncer, E., Calcada, R.R., Zingg, D., Varum, S., Cheng, P., Freiburger, S.N., Deng, C.X., Kleiter, I., Levesque, M.P., Dummer, R., *et al.* (2019). SMAD signaling promotes melanoma metastasis independently of phenotype switching. *J Clin Invest* 129, 2702-2716.
- Verfaillie, A., Imrichova, H., Atak, Z.K., Dewaele, M., Rambow, F., Hulselmans, G., Christiaens, V., Svetlichnyy, D., Luciani, F., Van den Mooter, L., *et al.* (2015). Decoding the regulatory landscape of melanoma reveals TEADS as regulators of the invasive cell state. *Nat Commun* 6, 6683.
- Wouters, J., Kalender-Atak, Z., Minnoye, L., Spanier, K.I., De Waegeneer, M., Bravo Gonzalez-Blas, C., Mauduit, D., Davie, K., Hulselmans, G., Najem, A., *et al.* (2020). Robust gene expression programs underlie recurrent cell states and phenotype switching in melanoma. *Nat Cell Biol* 22, 986-998.
- Zanconato, F., Cordenonsi, M., and Piccolo, S. (2016). YAP/TAZ at the Roots of Cancer. *Cancer Cell* 29, 783-803.
- Zhang, X., Yang, L., Szeto, P., Abali, G.K., Zhang, Y., Kulkarni, A., Amarasinghe, K., Li, J., Vergara, I.A., Molania, R., *et al.* (2020). The Hippo pathway oncoprotein YAP promotes melanoma cell invasion and spontaneous metastasis. *Oncogene* 39, 5267-5281.
- Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nussbaum, C., Myers, R.M., Brown, M., Li, W., *et al.* (2008). Model-based Analysis of ChIP-Seq (MACS). *Genome Biology* 9, R137.
- Zhao, B., Ye, X., Yu, J., Li, L., Li, W., Li, S., Yu, J., Lin, J.D., Wang, C.Y., Chinnaiyan, A.M., *et al.* (2008). TEAD mediates YAP-dependent gene induction and growth control. *Genes Dev* 22, 1962-1971.
- Zou, J., Maeder, M.L., Mali, P., Pruetz-Miller, S.M., Thibodeau-Beganny, S., Chou, B.K., Chen, G., Ye, Z., Park, I.H., Daley, G.Q., *et al.* (2009). Gene targeting of a disease-related gene in human induced pluripotent stem and embryonic stem cells. *Cell Stem Cell* 5, 97-110.