Mechanisms of Scale Invariance in Embryonic Patterning Systems

Mechanisms of Scale Invariance in Embryonic Patterning Systems

A dissertation presented

by

Zachary M. Collins

to

The Division of Medical Sciences

in partial fulfillment of the requirement

for the degree of

Doctor of Philosophy

in the subject of

Biological and Biomedical Sciences

Harvard University

Cambridge, Massachusetts

April 2018

Mechanisms of Scale Invariance in Embryonic Patterning Systems

Abstract

Embryos pattern themselves with remarkable consistency and readily adjust their patterning programs to drastic changes in embryo size. This robustness of pattern formation, termed scale invariance, requires cells to determine their precise location within the organism. Recent technological advances in genetics, molecular biology, and imaging have enabled unprecedented insights into how cells send and receive patterning signals. In this dissertation, I examine how vertebrate embryos convey, interpret, and regulate positional information. In Chapter II, I use novel embryological techniques, genetic perturbations, and confocal fluorescence microscopy to explore how signaling by the morphogen Sonic Hedgehog enables scale-invariant patterning of the ventral spinal cord. We find that Sonic Hedgehog represses the positive signaling regulator Scube2 and explore its function. In addition, we demonstrate that this self-regulation of morphogen signaling is necessary for pattern scaling. In Chapter III, we uncover the gene expression of single cells during vertebrate development and map the cell state landscape of early patterning. We then focus on how cell state landscapes change when critical patterning cues are disrupted via targeted mutagenesis with CRISPR Cas9. In Appendix 3, my colleagues and I use the scale invariance of somite patterning to gain new insights into patterning mechanisms with live imaging, pharmacological interventions, and embryological manipulations. We find that Fibroblast Growth Factor signaling gradients in the presomitic mesoderm are scale-invariant and make important refinements to the existing clock and wavefront model. We then test the predictions of our mathematical model against proposed alternatives and observe “echos” in somite patterning, which are uniquely predicted by our model. Taken together, this work yields new insights into the mechanism of vertebrate patterning and provides a valuable genomic resource for the scientific community.

Acknowledgments

First and foremost, I'd like to thank my thesis advisor, Sean Megason, for the constant support, invaluable lessons, and guidance. Through learning from Sean, I have grown into a confident and engaged developmental biologist. The environment of the Megason lab that Sean built was always one of open scientific exploration where I could ask questions and do experiments only limited by our ability to come up with new ideas. I thank Kana Ishimatsu, Dan Wagner, Tony Tsai, and all of my other collaborators for their instrumental roles in building this body of work. I thank Ian Swinburne for sharing with me his immense technical expertise; I am convinced there is no question about biology he cannot answer. I thank Tom Hiscock, Akankshi Munjal, Dante D'India, Toru Kawanshi, Pengfei Xu, and all the other members of the Megason, DePace, and Klein labs with whom I overlapped for their encouragement, friendship, and helpful discussions. I thank the members of my Dissertation Advisory Committee for their insightful input and helpful suggestions: Lisa Goodrich, who has guided my growth as a scientist since my PQE, Rosalind Segal, and Wolfram Goessling. In addition, I thank Andrew Lassar, Angela DePace, and Marc Kirschner for their instrumental roles in teaching me how to think about developmental biology through my coursework. I thank Kate Hodgins and the entire BBS office for their diligent work in supporting all of us. I thank the DRB program for bringing together such a wonderful group of passionate developmental biologists and training the next generation of graduate students. I am thankful for the helpful discussion and support of all my classmates and the Department of Systems Biology. I thank the generations of committed of scientists that built the foundation of knowledge that underpins this work. Finally, none of this would be possible without the loving support of my family and friends.

Dedication

This work is dedicated to my loving wife, Paula, my parents Sarah and Rob, and my siblings, Kaitlin and Gareth. Paula, thank you for joining me on this adventure; without your love and steadfast support, I never would have made it to where I am today. To my parents, thank you for everything you have done to bring me into this world, to guide my journey, and for the sacrifices you have made to make my wonderful life what it is today. To Kaitlin and Gareth, thank you for being the best siblings a younger brother could ask and for all your lessons over the years.

Table of Contents

List of Figures

Chapter I.

Introduction

Authors: Zachary M. Collins [a]

[a] Department of Systems Biology, Harvard Medical School, Boston, MA 02115, USA.

Author contributions: Z.M.C. wrote the manuscript.

Part I: Scaling Pattern to Size with Morphogen Gradients

Scale invariance in pattern formation

Over 127 years ago, Hans Driesch first separated the blastomeres of the sea urchin embryo and found that each blastomere formed a separate larva that patterned normally at a reduced size. Several years later in 1895, Thomas Hunt Morgan separated the blastomeres of a frog embryo and found they, too, could form proportionally patterned embryos at a drastically reduced size (Morgan, 1895). Hans Spemann later showed in his famous constriction experiments that amphibian embryos could produce twins (Spemann, 1938). More recently, biologists have sought to quantify this patterning and found that surgically size-reduced embryos pattern their tissues with remarkable scale invariance (Cooke, 1981).

Similar findings have been reported in embryos with naturally varying embryo sizes and between related species of varying size (Gregor et al., 2008; Uygur et al., 2016). Systems that are capable of scale-invariant patterning have significant evolutionary advantages. Scale-invariant patterning programs would enable more rapid life history evolution when embryonic and larval sizes may need to adapt to new ecological niches. Moreover, these systems allow for correct patterning in stressful conditions when egg sizes are known to be more variable, or if blastomeres are separated in early development, as in monozygotic twins (Machin and Louis).

The ability of embryos to adjust patterning to tissue availability has long puzzled developmental biologists because scale-invariant patterning implies robust control of positional information during development. Cells must determine not only where they are compared to an absolute positional cue, but they must also be able to sense relative position. Revolutions in genetics and molecular biology, along with diligent work by developmental biologists, have revealed a great deal about the factors that embryos use to obtain positional information. However, how these patterning networks allow for cells to sense relative position during patterning remains unclear.

Scaling of morphogen gradients

Morphogen gradients are found throughout developing organisms and enable cells to sense their position in a tissue relative to a positional cue (Rogers and Schier, 2011; Wolpert, 2011). In morphogen-mediated patterning systems, a group of cells generates a positional cue termed a morphogen—often a secreted protein—that diffuses through the extracellular space and degrades as it crosses the tissue. These ligands are sensed by receiving cells in a dose-dependent manner, thus instructing cells on their position (Figure 1.1A). However, simple formulations of these models fail to pattern domains in a scale-invariant way (Figure 1.1B). Wolpert proposed that morphogen-mediated patterning systems may have sinks, an area of the tissue where degradation is extremely high to encourage pattern scaling. However, these models have a limited capacity for generating scale invariance (Ben-Zvi and Barkai, 2010). This led Cooke to question these formalizations of morphogen-mediated patterning after he measured the impressive scale invariance of Xenopus embryogenesis (Cooke, 1981).

Recently several theoretical models have emerged that aim to explain morphogen scaling. One such model—called the expander-repressor model—postulates that morphogen scaling can be achieved by a morphogen repressing a positive regulator of its own spread (Figure 1.1C). These models require that the “expander” acts cell-non-autonomously to either increase the diffusion or decrease the degradation of a morphogen (Figure 1.1G-H). The first proposed biological example of this mechanism was in patterning of the Xenopus dorsoventral axis (Ben-Zvi et al., 2008). In this system, gradients composed of Bone Morphogenetic Protein (BMP) are read out by cells of the early embryo as a ventral position cue. Establishment of the dorsal organizer is dependent on BMP antagonists, such as Chordin (Khokha et al., 2005). Previous work has argued that Chordin is not just responsible for antagonizing BMP signaling but is also necessary for concentrating it at the ventral pole through shuttling (Shimmi et al., 2005; Wang and Ferguson, 2005; Zee et al., 2006). Ben-Zvi and colleagues propose that Anti-Dorsalizing Morphogenic Protein (ADMP), which is repressed by BMP signaling, competes with

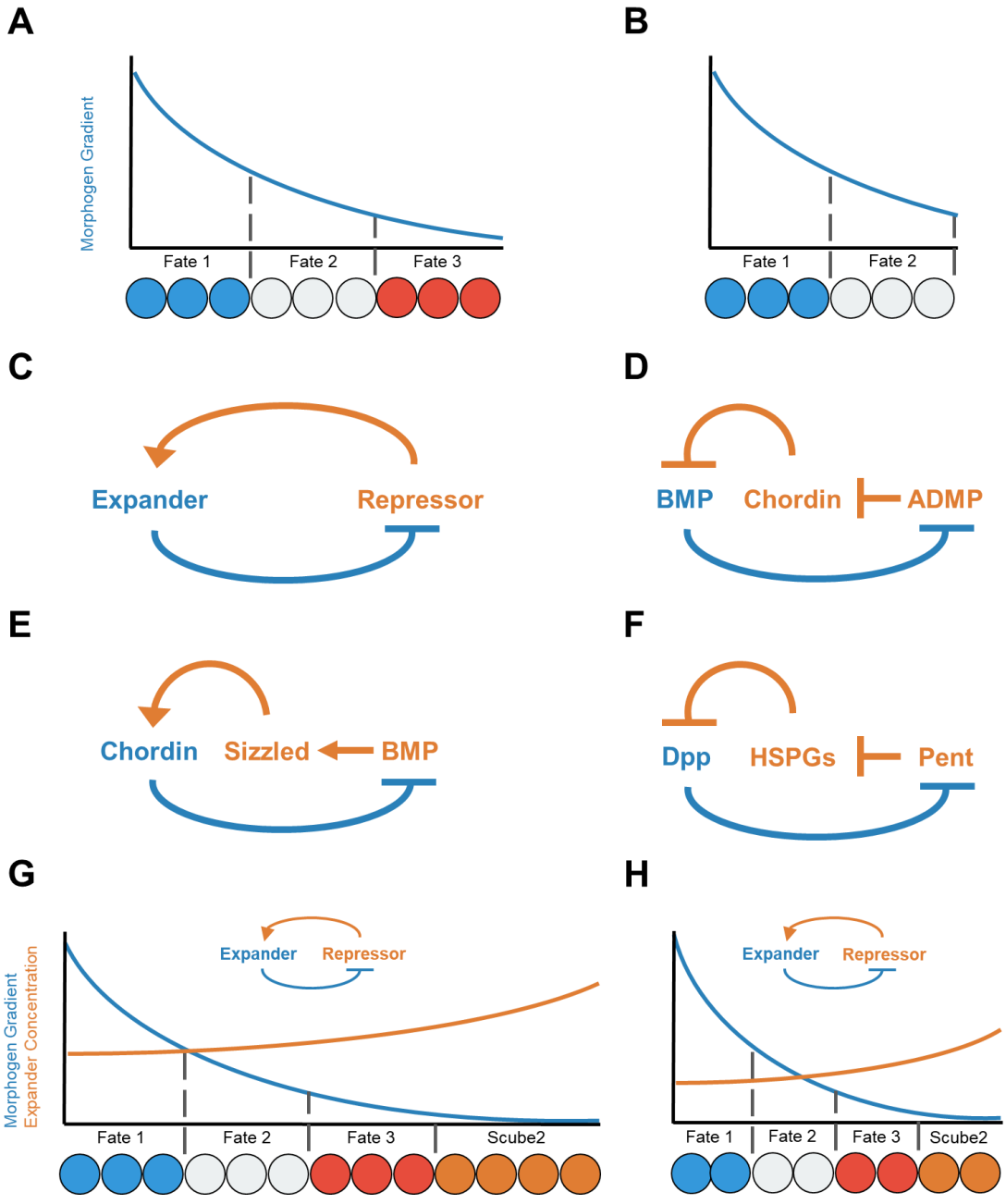


Figure 1.1 Morphogen gradient scaling

(A) A schematic of the classic French Flag Model of morphogen-mediated patterning. Morphogens are produced by cells on one side of the tissue and form a concentration gradient by transport (usually diffusion) and degradation. Cells then make threshold-like responses to

Figure 1.1 (Continued) perceived morphogen concentration and make cell fate decisions accordingly, as demonstrated by the blue, white, and red cell populations. (B) An example of aberrant patterning following size reduction of a simple morphogen-mediated patterning system. Disproportionate morphogen production and unchanged chemical properties causes a buildup of morphogen leading to incorrect spatial patterning. (C) Schematic of the expander-repressor model where a morphogen represses the expression of an expander that would broaden the signaling domain of the morphogen. (D) The initially proposed expander-repressor system thought to govern patterning of early D-V axis specification in Xenopus (Ben-Zvi, 2008). In this system, competition between ADMP and BMP for the binding of the inhibitor and proposed shuttle Chordin enables ADMP to expand distributions of BMP. ADMP is in turn repressed by BMP signaling. (E) Expander-repressor-like topology of the Sizzled-dependent scaling model (Inomata et al., 2013). In this model, inhibition of Chordin degradation by Sizzled serves to expand distributions of Chordin. BMP signaling inhibition by Chordin then lowers Sizzled expression as shown. (F) Scaling of Dpp gradients in the wing disc through the expander, Pent. In this system, Pent is repressed by Dpp signaling. Dpp ligand spread is promoted by Pent via Pent's promotion of HSPG internalization. HSPGs normally inhibit ligand diffusion. (G-H) Schematized models of expander-repressor systems whereby accumulation of the expander encodes a signaling equilibrium. This enables adjustment of morphogen spread to the size of the patterning domain. (H) Decreased expression of the expander following size reduction causes morphogen constriction and pattern scaling.

BMP for Chordin binding, thus allowing the BMP gradient to expand (Figure 1.1D) (Ben-Zvi et al., 2008).

More compelling data later showed that ADMP was dispensable for pattern scaling in early Xenopus dorsoventral patterning and proposed an alternative model (Inomata et al., 2013) (Figure 1.1E). This paper establishes that the metalloprotease inhibitor, Sizzled, is responsible for pattern scaling (Figure 1.1E) (Inomata et al., 2013). Chordin is required for graded BMP signaling and is degraded by metalloproteases during early patterning. Sizzled, which is upregulated in response to BMP signaling, inhibits these metalloproteases, thus expanding the distribution of Chordin (Lee et al., 2006). Interestingly, this topology is mathematically equivalent to the expander-repressor model, except in this regime Chordin can be thought of as the graded patterning cue—not BMP—and Sizzled the expander (Ben-Zvi et al., 2014).

Further evidence for the prevalence of expanders in scaling morphogen gradients was provided by two papers in the fly wing disc (Ben-Zvi et al., 2011a; Hamaratoglu et al., 2011). In this system, gradients of Dpp—the drosophila homologue of BMP—are known to scale during growth. These independent groups both found that Pentagone, a regulator of glypican internalization, fits all the qualifications for an expander of Dpp: it is repressed by morphogen signaling, cell-non-autonomously expands Dpp, and is required for pattern scaling (Figure 1.1F). Later work uncovered that Pent acts by promoting internalization of HSPGs, which normally inhibit Dpp transport (Norman et al., 2016).

Alternative models have also been proposed to underlie the scaling of pattern formation, the simplest being the flux optimization model. In this model, tuning morphogen production to overall embryo size may allow for “good enough” patterning where there is only mild error following size changes. Specification errors in these models are exaggerated in domains close to and far from the morphogen source and in patterning systems with many domains (Umulis and Othmer, 2012; Umulis and Othmer, 2013). However, these models require extremely

accurate control of morphogen production. If morphogen production does not scale with embryo size, no failsafes exist to correct patterning.

More commonly, it is proposed that opposing morphogen gradients are responsible for mediating pattern scaling. In this model, cells sense not just absolute levels of a morphogen, but also the ratio between two morphogens. These models—as with flux optimization models—are sensitive to minor level variations in individual morphogens (Ben-Zvi et al., 2011b). In these regimes, integration of morphogen signaling is thought to happen in the transcriptional networks that are downstream of morphogen signaling. While such opposing gradients are known to exist in some systems, it is unclear whether ratios between signals are used in normal embryos to determine fates over long ranges or set the boundaries between patterning regimes. Intermediate cell types between two domains are not formed in the vertebrate neural tube when significant levels of both the opposing morphogens Shh and BMP are present, casting some doubt on a simple ratio-sensing theory in this system (Pierani et al., 1999; Zagorski et al., 2017).

Scaling of morphogen gradients has been studied in many contexts: in species of varying sizes, during organ growth, and via experimental manipulation of embryo size within a species. But only the last of these yields insights into the robustness of initial patterning systems without interference from potential genetic or tissue level changes. To examine patterning robustness in genetically comparable embryos, we designed a system to reduce the size of zebrafish embryos surgically. This technique will be discussed in part II of this introduction.

Patterning of the neural tube

Patterning of neural subtype specification in the vertebrate neural tube is regulated by a series of morphogen gradients that are interpreted by a mutually repressive transcriptional network to select cell fates. Positional information for cells in the dorsal spinal cord cell is provided by BMP signaling, which specifies dorsal cell types in a dose/duration-dependent manner (Liem et al., 1995; Timmer et al., 2002). BMP ligands are secreted by the roof plate of

the neural tube, forming a signaling gradient from dorsal to ventral. BMP is only necessary for the induction of the dorsal-most cell types and suppresses ventral and intermediate neural cell type specification. High doses of BMP signaling induce Atoh1, which marks the D1 neuron domain, while intermediate doses induce NeuroG1 and D2 neurons (Figure 1.2). BMP signaling also promotes the expression of more broadly expressed transcription factors that mark dorsal neural cells including Pax2, Pax7, and Olig3.

A gradient of Shh signaling patterns the ventral neural tube and induces cell types in a dose-dependent manner. Shh ligands are secreted by the notochord and later by the floorplate of the developing neural tube. Prolonged high doses of Shh signaling induce Nkx2.2 and medium-high doses induce Olig2, and they in turn repress each other to specify the V3 and Mn domains respectively (Figure 1.2). Lower doses of Shh signaling then induce the V2 and V1 domains, which are defined by overlapping expression domains, and not by single transcription factors.

Specification of cells in the intermediate neural tube—V0 or D6 cells—are known to depend on retinoic acid which is produced in the somites, and not on Shh or BMP signaling (Pierani et al., 1999). Recent work, which used microfluidics to precisely control morphogen dose, indicated that cells of the intermediate neural tube need morphogen levels to be extremely low for the specification of intermediate cell fates (Zagorski et al., 2017). Interestingly, both dorsal and ventral patterning regimes will expand to fill the neural tube when the other is absent, indicating that they are both responsible for setting boundary conditions to inhibit expansion by the alternative program. How this expansion of progenitor specification is achieved remains unclear.

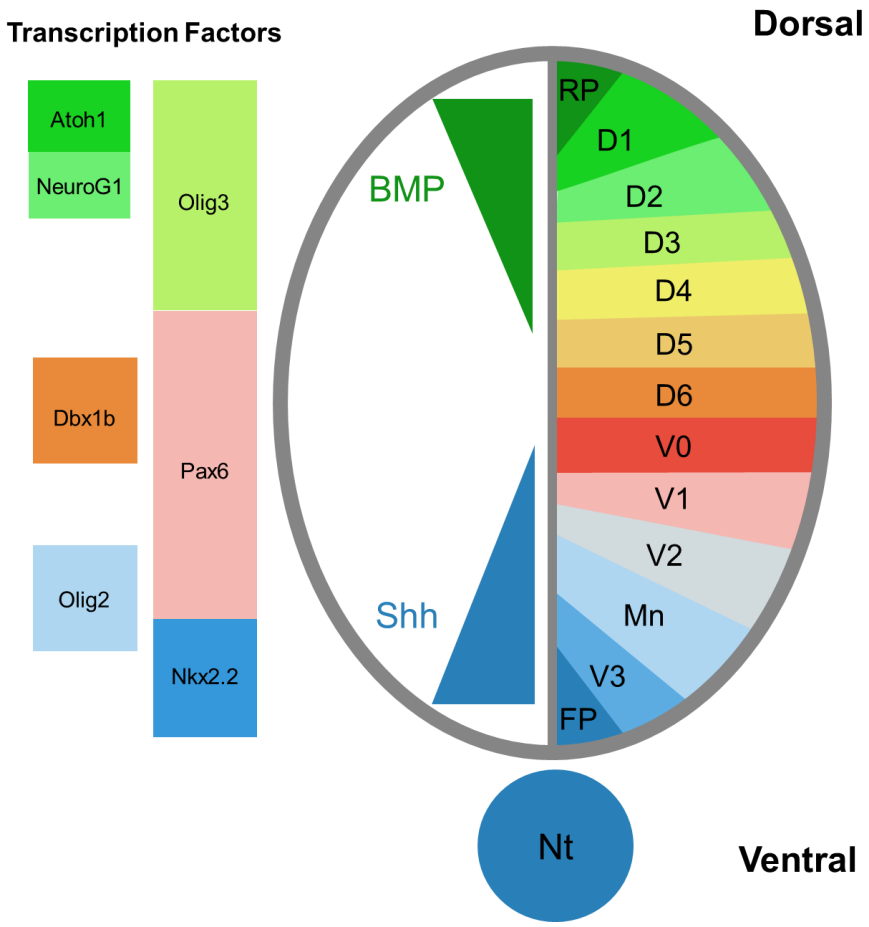


Figure 1.2 Morphogen-mediated patterning of the neural tube

Schematic representation of patterning of the neural tube by the morphogens Shh and BMP. Shh ligands are secreted from the notochord and floorplate and specify ventral neural progenitors in a dose-dependent manner. V3 interneuron progenitors receive the highest dose of Shh and initially express both *olig2* and *nkx2.2*. High levels of Shh signaling cause *nkx2.2* to be induced at high enough levels to repress *olig2*. Olig2+ progenitors lacking *nkx2.2* expression go on to form the motor neuron domain of the ventral spinal cord. V1 and V2 interneurons depend on low levels of Shh signaling for their specification and are defined by overlapping expression domains of multiple transcription factors, including Pax6. V0 and D6 neuron populations are thought to be independent of Shh and BMP signaling and rely on retinoic acid signaling for their specification. V0 and D6 interneurons are marked by the transcription

Figure 1.2 (Continued) factor *dbx1*. Dorsal interneurons are then specified by a network of transcription factors, simplified here to show only *atoh1* which defines the D1 domain, *neurog1* which defines the D2 domains, and *olig3,* which is expressed in the D1-D4 domains.

Shh signaling and ligand transport

Hedgehog proteins are a family of dually-lipid-modified secreted signaling proteins that are conserved throughout metazoans (Adamska et al., 2007). The Hedgehog pathway plays a central role in development and disease as both a differentiation and growth factor. Shh proteins begin as a full length amino acid chain with an N-terminal signaling domain, ShhN, and c-terminal intein-like domain. The C-terminal intein-like domain catalyzes its own removal and in the process covalently attaches cholesterol to the C-terminus of ShhN (Figure 1.3.1) (Porter et al., 1996a; Porter et al., 1996b). In addition to its C-terminal processing, the N-terminus is processed to remove the signal sequence and is then palmitoylated by Hedgehog Acetyltrasferase (Figure 1.3.2) (Buglino and Resh, 2008; Pepinsky et al., 1998). This unique dual lipid modification of Shh has complicated their study as long range signaling ligands. Without a secondary mechanism, Shh ligands ought to be tightly associated with cell membranes.

Release of Shh proteins from producing cells is known to require Dispatched (Figure 1.3.3) (Burke et al., 1999; Kawakami et al., 2002). The activity of Dispatched in Shh release depends on its cholesterol modification, which is counterintuitively required for its long range signaling (Burke et al., 1999; Lewis et al., 2001). Given the homology of Dispatched to RND efflux pumps—which are known to expel hydrophobic molecules from bacterial membranes—some have hypothesized that Dispatched may act by a similar mechanism (Petrov et al., 2017; Tseng et al., 1999). Recent work has shown that Dispatched alone is not sufficient for full Shh release, which requires the presence of Scube family proteins; this will be addressed in the next section (Creanga et al., 2012; Petrov et al., 2017; Tukachinsky et al., 2012). As expected by its function, Dispatched is commonly co-expressed with Shh, and its expression in Shh+ cells is required for release (Nakano et al., 2004; Tian et al., 2005).

Soluble multimeric Shh is known to require lipid modifications for its formation and is necessary for long range Shh activity *in vivo* (Chen et al., 2004; Zeng et al., 2001). However,

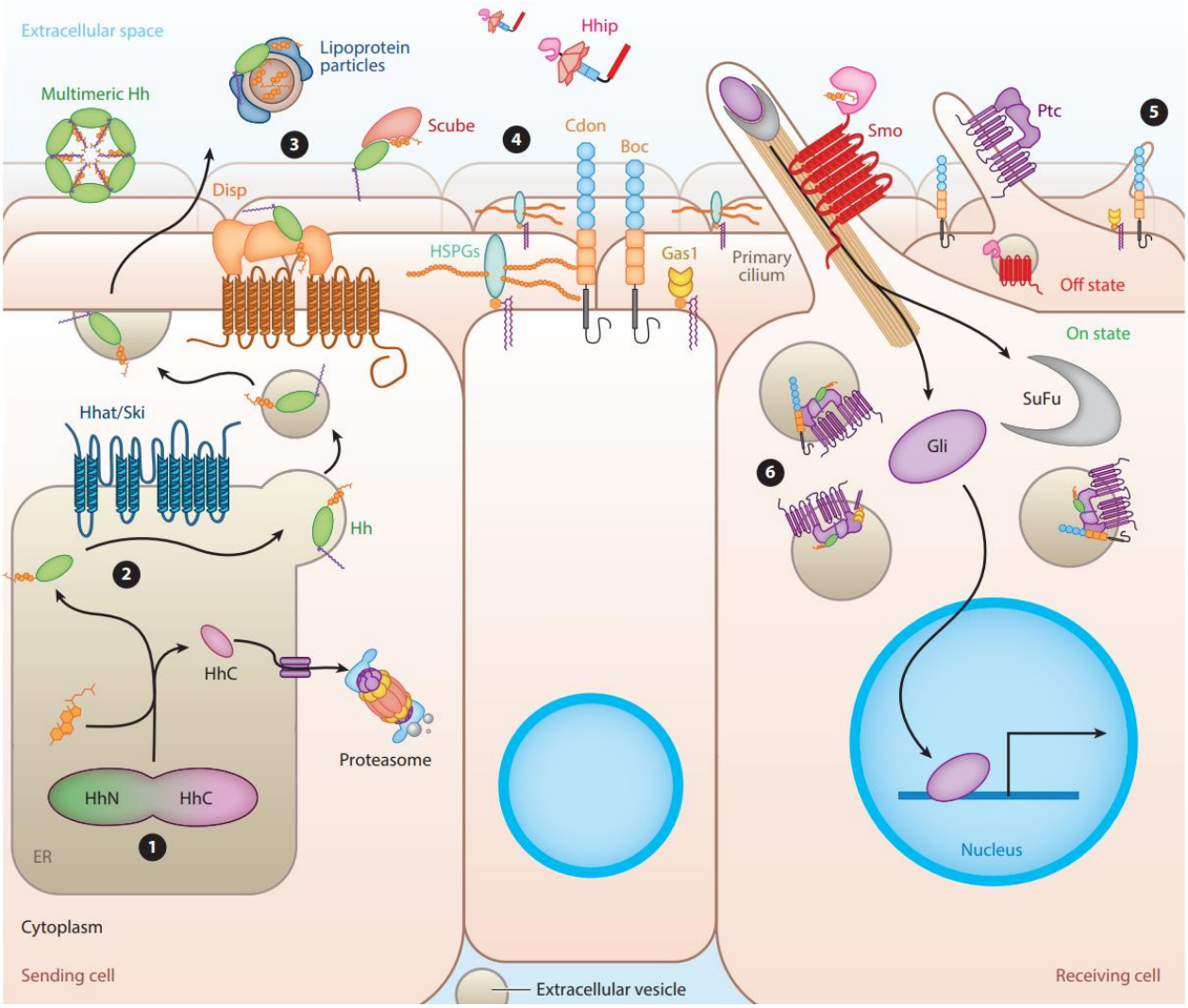


Figure 1.3 Molecular mechanisms of Shh signaling

Figure adapted from “Sending and Receiving Hedgehog Signals” by Petrov, K., Wierbowski, B.M. and Salic, A., as published in Annual Review of Cell and Developmental Biology (2017). (1) Autocatalytic removal of the c-terminus of Hh ligands catalyzes the addition of a cholesterol modification. (2) Hhat-mediated palmitate modification of Hh. (3) Release of Hh ligands by several proposed mechanisms, including Dispatched-and Scube2-dependent solubilization and transport on lipoprotein particles. (4) Released Hh ligands interact with a variety of Hh-signaling modifying agents including HSPGs, the Hh co-receptors Boc, Cdo, and Gas1, and Hhip, a secreted antagonist. (5) In the absence of Hh ligands, Patched inhibits Smoothened by

Figure 1.3 (Continued) regulating inhibitory sterols or sterol trafficking to the cilium. Hh co-receptors facilitate the binding of Hh to Patched. Hh ligand binding causes the internalization of Ptch, relieving its repression of Smoothened. In its active state, Smoothened then induces the activity of Gli family transcription factors.

alternative models of Shh transport via cytonemes and lipoprotein particles have been proposed (Gradilla et al., 2014; Matusek et al., 2014; Parchure et al., 2015; Sanders et al., 2013). Unlike soluble Shh, these methods of Shh transport have not been shown to be required for long range signaling, and lipoprotein particles containing Shh seem to have a low signaling activity (Palm et al., 2013).

Transport of Shh through the extracellular space is also facilitated by Heparin Sulfate Proteoglycans (HSPGs) (Figure 1.3.4) (Bellaiche et al., 1998; Rubin et al., 2002; The et al., 1999). Loss of certain HSPGs leads to the accumulation of Shh on producing cells and a failure of release (Bellaiche et al., 1998). Facilitation of Shh signaling by HSPGs is dependent on a Cardin-Weintraub motif within Shh that facilitates HSPG interactions (Rubin et al., 2002). Interestingly, HSPG-binding deficient mouse mutants have defects in Shh mitogenic responses but not patterning, indicating that HSPGs may have a more important role in regulating growth than fate specification (Chan et al., 2009). HSPGs modify a variety of morphogen distribution and are known as important regulators of BMP, FGF, and Wnt signaling (Yan and Lin, 2009).

The extracellular protein Hhip was thought to be a membrane-anchored inhibitor of Shh's binding to its receptor Patched (Chuang and McMahon, 1999; Chuang et al., 2003; Jeong and McMahon, 2005). However, recent work has revealed cell-non-autonomous activity for Hhip (Figure 1.3.4) (Holtz et al., 2015; Kwong et al., 2014). Interestingly, Hhip-Shh inhibitory complexes may assemble on HSPGs, potentially accounting for HSPG's observed role in Shh inhibition (Holtz et al., 2015).

Shh ligands are received by target cells through binding to its receptor, Patched (Figure 1.3.5) (Goodrich et al., 1996; Hooper and Scott, 1989; Nakano et al., 1989; Stone et al., 1996). Binding of Shh to Patched is facilitated by co-receptors including Gas1, Cdo, and Boc (Martinelli and Fan, 2007; Tenzen et al., 2006; Zhang et al., 2006). While Patched is a receptor for Shh, it is also a negative regulator of pathway activity. In the absence of Shh binding, Patched represses the activity of the transmembrane protein, Smoothened (Figure 1.3.5) (Taipale et al.,

2002). Recent work has indicated that cholesterol is responsible for Smoothened activation and that Patched may act by limiting the access of Smoothened to cholesterol or by producing an inhibitory sterol (Bidet et al., 2011; Huang et al., 2016; Roberts et al., 2016). Following Shh binding, Patched is internalized along with bound Shh, relieving its repression of Smoothened (Figure 1.3.6) (Briscoe et al., 2001; Chen and Struhl, 1996). Once this repression is relieved, Smoothened converts Gli family transcription factors from a transcriptionally repressive to activating form (Aza-Blanc et al., 2000; Wang et al., 2000). Gli family transcription factors have many targets in cells that are context dependent (Oosterveen et al., 2013; Vokes et al., 2007).

Feedback regulation of components of Shh signaling is known to modulate cellular responses to Shh (Ribes and Briscoe, 2009). Shh signaling directly upregulates expression of its receptor, Patched, which is often used as a proxy for Shh signaling. In addition to inhibiting Smoothened, Patched sequesters Shh, giving Patched two roles in inhibition of Shh signaling (Chen and Struhl, 1996). Hhip is also upregulated by Shh signaling in many contexts and inhibits further Shh signaling. Both Patched and Hhip have recently been suggested to have cell-non-autonomous negative effects on Shh signaling (Holtz et al., 2015; Kwong et al., 2014; Roberts et al., 2016). While Hhip is thought act over a distance by traveling through the extracellular space, Patched is a transmembrane protein, leading some to hypothesize that a secondary inhibitory cholesterol precursor produced by Patched is responsible for this effect (Roberts et al., 2016). As a rule, positive regulators of Shh signaling are usually down-regulated by pathway activation. The Shh co-receptors Cdo, Boc, and Gas1 are down-regulated by Shh signaling in the neural tube, and found expressed in the dorsal neural tube (Allen et al., 2007; Tenzen et al., 2006). These feedback networks are thought to buffer Shh signaling to variability in morphogen dose, but the molecular logic and design principles for these relationships remain unclear.

Scube2's expression and function during patterning

Scube (Signal peptide-CUB domain-EGF repeat containing) proteins were first identified from mouse cDNA libraries, and Scube2 was found to show localized expression to the dorsomedial neural tube during mouse development (Grimmond et al., 2000; Grimmond et al., 2001). The zebrafish "*you*" mutant, known for deficient Shh signaling, was mapped to Scube2 and was shown to be required for Shh signaling (Hollway et al., 2006; Woods and Talbot, 2005). Unlike other positive regulators of Shh signaling, Scube2 was identified to act cell non-autonomously in zebrafish; this was later confirmed by cell culture experiments (Creanga et al., 2012; Kawakami et al., 2005; Tukachinsky et al., 2012; Woods and Talbot, 2005). Initial reports concluded that Scube2 was a permissive factor in Shh signaling, as ubiquitous mRNA expression was capable of rescuing mutants, and overexpression seemed to have no phenotype (Woods and Talbot, 2005). Later, Scube family proteins were shown to be strictly required for Shh signaling in zebrafish when Scube1, 2, and 3 were simultaneously knocked down (Johnson et al., 2012). Of the Scube family, Scube2 has a dominant role in facilitating embryonic Shh signaling in zebrafish, as Scube1 and Scube3 single and double knockdowns have no detectable phenotype (Johnson et al., 2012).

Cell culture experiments have demonstrated that Scube2 is a potent stimulator of Shh ligand release (Creanga et al., 2012; Tukachinsky et al., 2012). Scube2 expression is not required in Shh-secreting cells, and Scube2 conditioned media is sufficient to stimulate release of Shh in producing cells. Later work has argued that Scube2 may promote cleavage of lipophilic domains off of Shh, but this disputes previous HPLC analysis and has yet to be reproduced by another group (Creanga et al., 2012; Jakobs et al., 2014). As with other positive regulators of Shh signaling, *scube2* expression is repressed in the ventral neural tube, though the localization of its expression is somewhat disputed in zebrafish (Grimmond et al., 2001; Hollway et al., 2006; Kawakami et al., 2002). Furthermore, a genome-wide screen for genes regulated by Shh identified that Scube2 was down-regulated by hyper-activation of the Shh

pathway (Xu et al., 2006). These features led me to hypothesize that Scube2 may act in a similar manner to previously described expanders in BMP signaling. In Chapter II, I explore the localization, regulation, and role of Scube2 during ventral neural patterning.

Part II: Dorsal-Ventral Axis Specification in Vertebrates

Symmetry breaking and organizer induction

In early development, multiple symmetry-breaking events are required for setting up the future body axes. The first symmetry-breaking event after fertilization is the induction of Spemann's organizer (Figure 1.4). In Xenopus, the position of this event is determined by the sperm entry site. After sperm entry, a cytoskeletal rearrangement termed "cortical rotation" occurs towards the position of sperm entry, and future dorsal is induced directly opposite (Gerhart et al., 1989; Spemann, 1938). This clear demarcation of the future embryonic axis in amphibians made possible the early studies of the organizer by Spemann and his colleagues (Spemann, 1938). In zebrafish, the sperm entry point is fixed to the position of the micropyle, a narrow pore at the animal pole (Amanze, 1990). Thus, sperm entry position is unable to influence the position of organizer induction. Some studies have argued that division orientation sets the future embryonic axis in zebrafish, but this was falsified by analysis of *janus* mutants (Abdelilah et al., 1994). Zebrafish *janus* mutants have defects in the first embryonic division that cause the separation of the blastomeres in early development. If the initial embryonic division was responsible for inducing organizer position, these embryos would develop in a stereotyped manner. However, organizer formation in these mutants—and in all zebrafish—is random with respect to the division plane; this causes a spectrum of axis duplication and single axis embryos in *janus* mutants.

Microtubule polymerization is known to be necessary for organizer induction in both zebrafish and Xenopus (Gerhart et al., 1989; Jesuthasan and Stähle, 1997). Transport of the dorsal determinants, Wnt mRNAs, on microtubules is required for induction of the organizer (Figure 1.4A-B) (Lu et al., 2011; Nojima et al., 2004; Nojima et al., 2010). Perturbation of the early embryo by wounding the yolk in zebrafish or inverting/centrifuging Xenopus can disrupt the transport of dorsal determinants and inhibit organizer formation (Marikawa et al., 1997; Mizuno

et al., 1999; Neff et al., 1983). Wnt signaling then induces the embryonic organizer, which begins its transcriptional program as described in the next section (Grill, 2011; Lu et al., 2011; Nojima et al., 2004; Schneider et al., 1996).

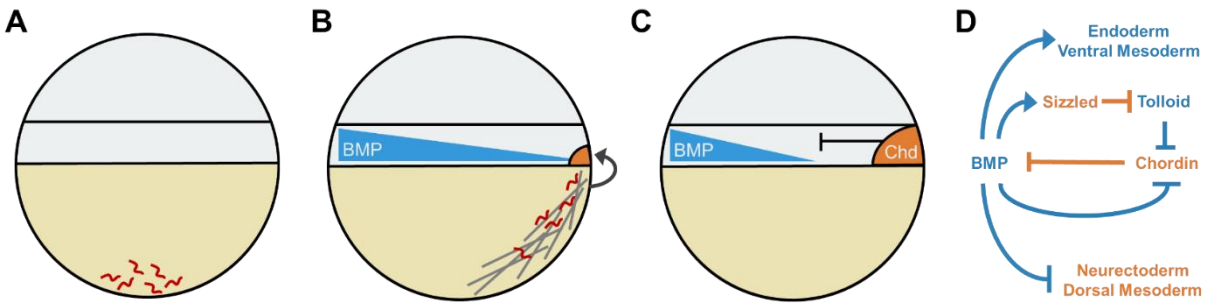


Figure 1.4 Induction and function of the organizer

(A) A schematic of the early zebrafish embryo at the one cell stage. Maternal mRNAs are deposited at the vegetal pole, including mRNA encoding the dorsal determinant (likely *wnt8a*). (B) Fertilization of the egg triggers seemingly random selection of the direction and orientation of microtubules that grow rapidly in the first 10 minutes of egg activation. Maternal mRNAs encoding the dorsal determinant are transported to the margin, where they are translated and induce the organizer (marked in orange). (C) As the organizer becomes signaling active, it expresses *chordin*, which inhibits BMP signaling that would otherwise grow to become ubiquitous in the embryo. This inhibition of BMP signaling enables formation of the dorsal tissue of the early embryo, including ectoderm and dorsal mesoderm. (D) A simplified topology of the regulation of BMP signaling following organizer formation.

The function and regulation of Spemann's organizer

The dorsal organizer has been known to be critical for developmental patterning since Hans Spemann's constriction experiments in the beginning of the 20th century (later reviewed by Spemann in 1938). Spemann's PhD student, Hilde Mangold, took these experiments further by

showing via transplantation that a small group of cells was capable of inducing—or "organizing"—an entirely new body axis. Since this time, we have elucidated many of the molecular mechanisms underlying this central patterning system.

Before organizer induction, the TGFβ family signaling ligands BMP2, BMP4, and BMP7 are expressed broadly in the early embryo. When Wnt ligands induce the organizer, downstream transcription factors turn on the expression of the BMP antagonists, Noggin, Chordin and Follistatin (Figure 1.4B) (Hemmati-Brivanlou et al., 1994; Lamb et al., 1993; Sasai et al., 1994). Their antagonism is required for the formation of the BMP signaling gradient that patterns the dorsal-ventral axis (Barth et al., 1999; Khokha et al., 2005). BMP antagonists are degraded by Tolloid metalloproteases that in turn control their level and range of inhibition (Piccolo et al., 1997). Interestingly, these proteases are capable of cleaving Chordin and leaving BMP intact in order to encourage further BMP signaling (Piccolo et al., 1997). Another layer of regulation is added by the protein, Sizzled, whose expression is induced by BMP signaling. Sizzled then promotes the activity of chordin by inhibiting cleavage by Tolloid metalloproteases (Lee et al., 2006). Understanding the function and specification of the organizer is critical to designing surgical techniques which manipulate the early embryo. These studies informed my development of a technique to reduce the size of zebrafish embryos.

Embryologically sound surgical size reduction in zebrafish

Inspired by work on the scaling of Xenopus embryos following bisection, a Megason lab post-doctoral fellow, Kana Ishimatsu, had already developed a technique for reducing the size of zebrafish embryos when I joined the lab (Figure 1.5A). In this method, yolk and cells of the blastula were removed at once in a single cut along the Animal-Vegetal axis. Using this technique, one could generate size-reduced embryos that patterned normally around one fifth of the time. Embryos that survived surgical size reduction with this technique showed highly variable phenotypes. Embryos were routinely somewhat or severely ventralized (Figure 1.5B-C).

This was, in part, due to the inability to visualize the organizer in early zebrafish development. Unlike amphibian embryos, zebrafish embryos are radially symmetric in appearance until the onset of gastrulation and lack a visual indicator of organizer location, such as the grey crescent.

In search of a better method, I experimented with a litany of ways to produce smaller embryos. One such method was to force the maturation of smaller oocytes using oocyte culture systems (Langdon and Mullins, 2011; Lokman et al., 2007; Nagahama and Yamashita, 2008). However, after lengthy experimentation, embryos produced by this method were not significantly smaller than naturally spawned embryos. Attempts at chopping later—when it is possible to visualize the organizer through fluorescent reporters at sphere stage in tg(*dharma:gfp)* or morphologically during shield formation—were also unsuccessful. Zebrafish embryos almost always failed to heal the yolk after wounding at these stages.

My research on organizer induction informed my development of an embryologically sound surgical method for reducing the size of zebrafish embryos. This method is performed by removing cells of the animal cap and wounding the yolk at the vegetal pole (Figure 1.5D). We found that wounding embryos at the mid-to-late blastula stage had the best rates of recovery. At this stage, embryos heal rapidly and wounding of the yolk does not affect transport of maternal organizer determining mRNAs. This method proved to be much more successful, producing healthy embryos without a spectrum of ventralization phenotypes. By avoiding damage to the organizer, which is being induced along the blastoderm margin at these time points, size reduced embryos showed better health and survival rates closer to 50-60%. Embryos reduced in size by this method showed remarkable scale invariance in their patterning, enabling my study in Chapter II and the work led by Kana Ishimatsu in Appendix 3.

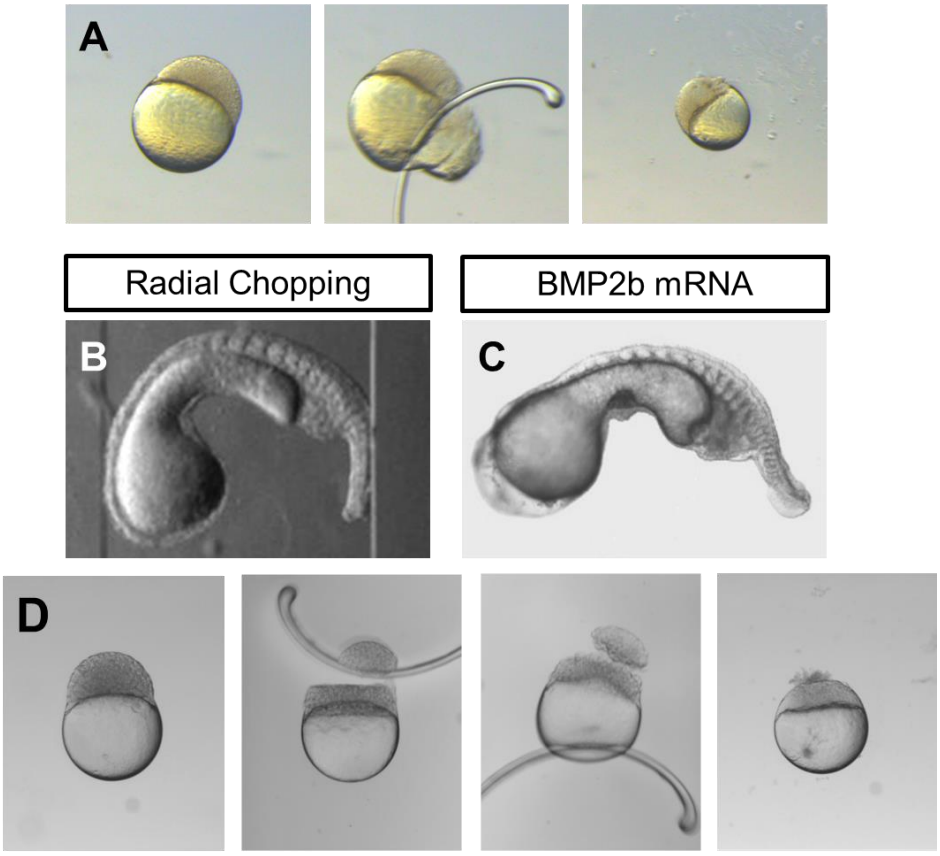


Figure 1.5 Initial radial chopping and revised lateral chopping

(A) Initial radial chopping technique for embryonic size reduction. In this method, a glass needle is used to remove yolk and cells of the blastula simultaneously in a longitudinal cut. (B) Ventralized embryo generated by radial chopping. (C) Ventralized embryo generated by injection of BMP2b mRNA (Nguyen et al., 1998). (D) Lateral chopping technique that I developed to increase consistency in embryonic size reduction. This technique avoids damaging signaling centers that are important for early D-V patterning which are located at the blastoderm margin.

Analysis of Chordin phenotypes using single cell transcriptomics and gene editing

Single cell sequencing technology is enabling new levels of insights into developing systems. With these technologies, we can quantitatively measure gene expression at the single cell level. When I started my PhD, these techniques were prohibitively expensive and enabled measurement of only a few hundred single cells. In 2015, a breakthrough approach that enables the rapid barcoding of individual cells for sequencing on a single sequencing lane was published by two labs at Harvard (Klein et al., 2015; Macosko et al., 2015). With these tools in hand, Dan Wagner, a post-doctoral fellow in the Megason and Klein labs, began developing zebrafish single cell sequencing protocols and assembled an atlas of single cell transcriptomes throughout early zebrafish development, which is detailed in Chapter III. In this work, we developed a map of the cell trajectory landscapes of early vertebrate development. We set out to explore how disruption of major patterning programs, such as D-V patterning, might alter this landscape.

Techniques for targeted mutagenesis have made rapid progress due to the CRISPR Cas9 system (Cong et al., 2013; Hwang et al., 2013; Jinek et al., 2012). When I began graduate school, CRISPR was just beginning to be applied to zebrafish with pioneering work that showed how inexpensive and rapid CRISPR mutant generation could be (Gagnon et al., 2014; Hwang et al., 2013). Remarkably, concentrated Cas9 ribonucleoprotein complexes are so efficient at mutagenesis in zebrafish that homozygous mutant phenotypes can be studied in injected embryos without waiting for germline genetic transmission (Burger et al., 2016; Gagnon et al., 2014). We used this method to reliably produce Chordin-null embryos and study the effects of Chordin loss on the landscape of single cell gene expression, explored in Chapter III. In addition, we performed a small-scale screen for novel regulators of organizer function based on co-expression of uncharacterized genes with known organizer factors. This work is not discussed in Chapter III and is still ongoing.

References

Abdelilah, S., Solnica-Krezel, L., Stainier, D. Y. and Driever, W. (1994). Implications for dorsoventral axis determination from the zebrafish mutation janus. Nature 370, 468–71.

Adamska, M., Matus, D. Q., Adamski, M., Green, K., Rokhsar, D. S., Martindale, M. Q. and Degnan, B. M. (2007). The evolutionary origin of hedgehog proteins. Curr. Biol. 17, R836-7.

Allen, B. L., Tenzen, T. and McMahon, A. P. (2007). The Hedgehog-binding proteins Gas1 and Cdo cooperate to positively regulate Shh signaling during mouse development. Genes Dev. 21, 1244–57.

Amanze, D. (1990). The micropyle: a sperm guidance system in teleost fertilization. Development 109, 495–500.

Aza-Blanc, P., Lin, H. Y., Ruiz i Altaba, A. and Kornberg, T. B. (2000). Expression of the vertebrate Gli proteins in Drosophila reveals a distribution of activator and repressor activities. Development 127, 4293–301.

Barth, K., Kishimoto, Y., Rohr, K., Seydler, C., Schulte-Merker, S. and Wilson, S. (1999). Bmp activity establishes a gradient of positional information throughout the entire neural plate. Development 126, 4977–4987.

Bellaiche, Y., The, I. and Perrimon, N. (1998). Tout-velu is a Drosophila homologue of the putative tumour suppressor EXT-1 and is needed for Hh diffusion. Nature 394, 85–88.

Ben-Zvi, D. and Barkai, N. (2010). Scaling of morphogen gradients by an expansion-repression integral feedback control. Proc. Natl. Acad. Sci. U. S. A. 107, 6924–9.

Ben-Zvi, D., Shilo, B.-Z., Fainsod, A. and Barkai, N. (2008). Scaling of the BMP activation gradient in Xenopus embryos. Nature 453, 1205–11.

Ben-Zvi, D., Pyrowolakis, G., Barkai, N. and Shilo, B.-Z. (2011a). Expansion-repression mechanism for scaling the Dpp activation gradient in Drosophila wing imaginal discs. Curr. Biol. 21, 1391–6.

Ben-Zvi, D., Shilo, B.-Z. and Barkai, N. (2011b). Scaling of morphogen gradients. Curr. Opin. Genet. Dev. 21, 704–710.

Ben-Zvi, D., Fainsod, A., Shilo, B.-Z. and Barkai, N. (2014). Scaling of dorsal-ventral patterning in the Xenopus laevis embryo. Bioessays 36, 151–6.

Bidet, M., Joubert, O., Lacombe, B., Ciantar, M., Nehmé, R., Mollat, P., Brétillon, L., Faure, H., Bittman, R., Ruat, M., et al. (2011). The Hedgehog Receptor Patched Is Involved in Cholesterol Transport. PLoS One 6, e23834.

Briscoe, J., Chen, Y., Jessell, T. M. and Struhl, G. (2001). A hedgehog-insensitive form of patched provides evidence for direct long-range morphogen activity of sonic hedgehog in the neural tube. *Mol. Cell* 7, 1279–91.

Buglino, J. A. and Resh, M. D. (2008). Hhat Is a Palmitoylacyltransferase with Specificity for N - Palmitoylation of Sonic Hedgehog. *J. Biol. Chem.* 283, 22076–22088.

Burger, A., Lindsay, H., Felker, A., Hess, C., Anders, C., Chiavacci, E., Zaugg, J., Weber, L. M., Catena, R., Jinek, M., et al. (2016). Maximizing mutagenesis with solubilized CRISPR-Cas9 ribonucleoprotein complexes. *Development* 143, 2025–37.

Burke, R., Nellen, D., Bellotto, M., Hafen, E., Senti, K.-A. A., Dickson, B. J. and Basler, K. (1999). Dispatched, a Novel Sterol-Sensing Domain Protein Dedicated to the Release of Cholesterol-Modified Hedgehog from Signaling Cells. *Cell* 99, 803–815.

Chan, J. A., Balasubramanian, S., Witt, R. M., Nazemi, K. J., Choi, Y., Pazyra-Murphy, M. F., Walsh, C. O., Thompson, M. and Segal, R. A. (2009). Proteoglycan interactions with Sonic Hedgehog specify mitogenic responses. *Nat. Neurosci.* 12, 409–417.

Chen, Y. and Struhl, G. (1996). Dual roles for patched in sequestering and transducing Hedgehog. *Cell* 87, 553–63.

Chen, M.-H., Li, Y.-J., Kawakami, T., Xu, S.-M. and Chuang, P.-T. (2004). Palmitoylation is required for the production of a soluble multimeric Hedgehog protein complex and long-range signaling in vertebrates. *Genes Dev.* 18, 641–59.

Chuang, P.-T. and McMahon, A. P. (1999). Vertebrate Hedgehog signalling modulated by induction of a Hedgehog-binding protein. *Nature* 397, 617–621.

Chuang, P.-T., Kawcak, T. and McMahon, A. P. (2003). Feedback control of mammalian Hedgehog signaling by the Hedgehog-binding protein, Hip1, modulates Fgf signaling during branching morphogenesis of the lung. *Genes Dev*. 17, 342–347.

Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P. D., Wu, X., Jiang, W., Marraffini, L. A., et al. (2013). Multiplex genome engineering using CRISPR/Cas systems. *Science* 339, 819–23.

Cooke, J. (1981). Scale of body pattern adjusts to available cell number in amphibian embryos. *Nature* 290, 775–778.

Creanga, A., Glenn, T. D., Mann, R. K., Saunders, A. M., Talbot, W. S. and Beachy, P. A. (2012). Scube/You activity mediates release of dually lipid-modified Hedgehog signal in soluble form. *Genes Dev. 26*, 1312–25.

Gagnon, J. A., Valen, E., Thyme, S. B., Huang, P., Ahkmetova, L., Pauli, A., Montague, T. G., Zimmerman, S., Richter, C., Schier, A. F., et al. (2014). Efficient Mutagenesis by Cas9

Protein-Mediated Oligonucleotide Insertion and Large-Scale Assessment of Single-Guide RNAs. *PLoS One* 9, e98186.

Gerhart, J., Danilchik, M., Doniach, T., Roberts, S., Rowning, B. and Stewart, R. (1989). Cortical rotation of the Xenopus egg: consequences for the anteroposterior pattern of embryonic dorsal development. *Development* 107 Suppl, 37–51.

Goodrich, L. V, Johnson, R. L., Milenkovic, L., McMahon, J. A. and Scott, M. P. (1996). Conservation of the hedgehog/patched signaling pathway from flies to mice: induction of a mouse patched gene by Hedgehog. *Genes Dev.* 10, 301–12.

Gradilla, A.-C., González, E., Seijo, I., Andrés, G., Bischoff, M., González-Mendez, L., Sánchez, V., Callejo, A., Ibáñez, C., Guerra, M., et al. (2014). Exosomes as Hedgehog carriers in cytoneme-mediated transport and secretion. *Nat. Commun.* 5, 5649.

Gregor, T., McGregor, A. P. and Wieschaus, E. F. (2008). Shape and function of the Bicoid morphogen gradient in dipteran species with different sized embryos. *Dev. Biol.* 316, 350–8.

Grill, S. W. (2011). Growing up is stressful: Biophysical laws of morphogenesis. *Curr. Opin. Genet. Dev.* 21, 647–652.

Grimmond, S., Larder, R., Van Hateren, N., Siggers, P., Hulsebos, T. J. M., Arkell, R. and Greenfield, A. (2000). Cloning, Mapping, and Expression Analysis of a Gene Encoding a Novel Mammalian EGF-Related Protein (SCUBE1). *Genomics* 70, 74–81.

Grimmond, S., Larder, R., Van Hateren, N., Siggers, P., Morse, S., Hacker, T., Arkell, R. and Greenfield, A. (2001). Expression of a novel mammalian epidermal growth factor-related gene during mouse neural development. *Elsevier.*

Hamaratoglu, F., de Lachapelle, A. M., Pyrowolakis, G., Bergmann, S. and Affolter, M. (2011). Dpp signaling activity requires Pentagone to scale with tissue size in the growing Drosophila wing imaginal disc. *PLoS Biol*. 9, e1001182.

Hemmati-Brivanlou, A., Kelly, O. G. and Melton, D. A. (1994). Follistatin, an antagonist of activin, is expressed in the Spemann organizer and displays direct neuralizing activity. *Cell* 77, 283–95.

Hollway, G. E., Maule, J., Gautier, P., Evans, T. M., Keenan, D. G., Lohs, C., Fischer, D., Wicking, C. and Currie, P. D. (2006). Scube2 mediates Hedgehog signalling in the zebrafish embryo. *Dev. Biol.* 294, 104–18.

Holtz, A. M., Griffiths, S. C., Davis, S. J., Bishop, B., Siebold, C. and Allen, B. L. (2015). Secreted HHIP1 interacts with heparan sulfate and regulates Hedgehog ligand localization and function. *J. Cell Biol.* 209, 739–758.

Hooper, J. E. and Scott, M. P. (1989). The Drosophila patched gene encodes a putative membrane protein required for segmental patterning. *Cell* 59, 751–65.

Huang, P., Nedelcu, D., Watanabe, M., Jao, C., Kim, Y., Liu, J. and Salic, A. (2016). Cellular Cholesterol Directly Activates Smoothened in Hedgehog Signaling. *Cell* 166, 1176–1187.e14.

Hwang, W. Y., Fu, Y., Reyon, D., Maeder, M. L., Tsai, S. Q., Sander, J. D., Peterson, R. T., Yeh, J.-R. J. and Joung, J. K. (2013). Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nat. Biotechnol.* 31, 227–9.

Inomata, H., Shibata, T., Haraguchi, T. and Sasai, Y. (2013). Scaling of dorsal-ventral patterning by embryo size-dependent degradation of Spemann's organizer signals. *Cell* 153, 1296–311.

Jakobs, P., Exner, S., Schürmann, S., Pickhinke, U., Bandari, S., Ortmann, C., Kupich, S., Schulz, P., Hansen, U., Seidler, D. G., et al. (2014). Scube2 enhances proteolytic Shh processing from the surface of Shh-producing cells. *J. Cell Sci.* 127, 1726–37.

Jeong, J. and McMahon, A. P. (2005). Growth and pattern of the mammalian neural tube are governed by partially overlapping feedback activities of the hedgehog antagonists patched 1 and Hhip1. *Development* 132, 143–54.

Jesuthasan, S. and Stähle, U. (1997). Dynamic microtubules and specification of the zebrafish embryonic axis. *Curr. Biol.* 7, 31–42.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A. and Charpentier, E. (2012). A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science* (80-.). 337, 816–821.

Johnson, J.-L. F. A., Hall, T. E., Dyson, J. M., Sonntag, C., Ayers, K., Berger, S., Gautier, P., Mitchell, C., Hollway, G. E. and Currie, P. D. (2012). Scube activity is necessary for Hedgehog signal transduction in vivo. *Dev. Biol.* 368, 193–202.

Kawakami, T., Kawcak, T., Li, Y.-J., Zhang, W., Hu, Y. and Chuang, P.-T. (2002). Mouse dispatched mutants fail to distribute hedgehog proteins and are defective in hedgehog signaling. *Development* 129, 5753–65.

Kawakami, A., Nojima, Y., Toyoda, A., Takahoko, M., Satoh, M., Tanaka, H., Wada, H., Masai, I., Terasaki, H., Sakaki, Y., et al. (2005). The zebrafish-secreted matrix protein you/scube2 is implicated in long-range regulation of hedgehog signaling. *Curr. Biol.* 15, 480–8.

Khokha, M. K., Yeh, J., Grammer, T. C. and Harland, R. M. (2005). Depletion of Three BMP Antagonists from Spemann's Organizer Leads to a Catastrophic Loss of Dorsal Structures. *Dev. Cell* 8, 401–411.

Klein, A. M., Mazutis, L., Akartuna, I., Tallapragada, N., Veres, A., Li, V., Peshkin, L., Weitz, D. A. and Kirschner, M. W. (2015). Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. *Cell* 161, 1187–1201.

Kwong, L., Bijlsma, M. F. and Roelink, H. (2014). Shh-mediated degradation of Hhip allows cell autonomous and non-cell autonomous Shh signalling. *Nat. Commun.* 5, 4849.

Lamb, T. M., Knecht, A. K., Smith, W. C., Stachel, S. E., Economides, A. N., Stahl, N., Yancopolous, G. D. and Harland, R. M. (1993). Neural induction by the secreted polypeptide noggin. *Science* 262, 713–8.

Langdon, Y. G. and Mullins, M. C. (2011). Maternal and zygotic control of zebrafish dorsoventral axial patterning. *Annu. Rev. Genet.* 45, 357–77.

Lee, H. X., Ambrosio, A. L., Reversade, B. and De Robertis, E. M. (2006). Embryonic Dorsal-Ventral Signaling: Secreted Frizzled-Related Proteins as Inhibitors of Tolloid Proteinases. *Cell* 124, 147–159.

Lewis, P. M., Dunn, M. P., McMahon, J. A., Logan, M., Martin, J. F., St-Jacques, B. and 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.

Liem, K. F., Tremml, G., Roelink, H. and Jessell, T. M. (1995). Dorsal differentiation of neural plate cells induced by BMP-mediated signals from epidermal ectoderm. *Cell* 82, 969–979.

Lokman, P. M., George, K. A. N., Divers, S. L., Algie, M. and Young, G. (2007). 11-Ketotestosterone and IGF-I increase the size of previtellogenic oocytes from shortfinned eel, Anguilla australis, in vitro. *Reproduction* 133, 955–67.

Lu, F.-I., Thisse, C. and Thisse, B. (2011). Identification and mechanism of regulation of the zebrafish dorsal determinant. *Proc. Natl. Acad. Sci. U. S. A.* 108, 15876–80.

Machin, G. and Louis, K. Biology of twins and other multiple pregnancies. In An atlas of multiple pregnancy, p. p13–24. New York: Parthenon Publishing.

Macosko, E. Z., Basu, A., Satija, R., Nemesh, J., Shekhar, K., Goldman, M., Tirosh, I., Bialas, A. R., Kamitaki, N., Martersteck, E. M., et al. (2015). Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. *Cell* 161, 1202–1214.

Marikawa, Y., Li, Y. and Elinson, R. P. (1997). Dorsal Determinants in the Xenopus Egg Are Firmly Associated with the Vegetal Cortex and Behave like Activators of the Wnt Pathway. *Dev. Biol.* 191, 69–79.

Martinelli, D. C. and Fan, C.-M. (2007). Gas1 extends the range of Hedgehog action by facilitating its signaling. *Genes Dev*. 21, 1231–1243.

Matusek, T., Wendler, F., Polès, S., Pizette, S., D'Angelo, G., Fürthauer, M. and Thérond, P. P. (2014). The ESCRT machinery regulates the secretion and long-range activity of Hedgehog. *Nature* 516, 99–103.

Mizuno, T., Yamaha, E., Kuroiwa, A. and Takeda, H. (1999). Removal of vegetal yolk causes dorsal deficencies and impairs dorsal-inducing ability of the yolk cell in zebrafish. *Mech. Dev.* 81, 51–63.

Morgan, T. H. (1895). Half embryos and whole embryos from one of the first two blastomeres. *Anat. Anz.* 10, 623–638.

Nagahama, Y. and Yamashita, M. (2008). Regulation of oocyte maturation in fish. *Dev. Growth Differ.* 50 Suppl 1, S195-219.

Nakano, Y., Guerrero, I., Hidalgo, A., Taylor, A., Whittle, J. R. S. and Ingham, P. W. (1989). A protein with several possible membrane-spanning domains encoded by the Drosophila segment polarity gene patched. *Nature* 341, 508–513.

Nakano, Y., Kim, H. R., Kawakami, A., Roy, S., Schier, A. F. and Ingham, P. W. (2004). Inactivation of dispatched 1 by the chameleon mutation disrupts Hedgehog signalling in the zebrafish embryo. *Dev. Biol*. 269, 381–392.

Neff, A. W., Malacinski, G. M., Wakahara, M. and Jurand, A. (1983). Pattern Formation in Amphibian Embryos Prevented from Undergoing the Classical " Rotation Response " to Egg Activation. *Dev. Biol*. 97, 103–112.

Nguyen, V. H., Schmid, B., Trout, J., Connors, S. A., Ekker, M. and Mullins, M. C. (1998). Ventral and Lateral Regions of the Zebrafish Gastrula, Including the Neural Crest Progenitors, Are Established by abmp2b/swirl Pathway of Genes. *Dev. Biol.* 199, 93–110.

Nojima, H., Shimizu, T., Kim, C.-H., Yabe, T., Bae, Y.-K., Muraoka, O., Hirata, T., Chitnis, A., Hirano, T. and Hibi, M. (2004). Genetic evidence for involvement of maternally derived Wnt canonical signaling in dorsal determination in zebrafish. *Mech. Dev.* 121, 371–386.

Nojima, H., Rothhämel, S., Shimizu, T., Kim, C.-H., Yonemura, S., Marlow, F. L. and Hibi, M. (2010). Syntabulin, a motor protein linker, controls dorsal determination. *Development* 137, 923–33.

Norman, M., Vuilleumier, R., Springhorn, A., Gawlik, J. and Pyrowolakis, G. (2016). Pentagone internalises glypicans to fine-tune multiple signalling pathways. *Elife* 5,.

Oosterveen, T., Kurdija, S., Enstero, M., Uhde, C. W., Bergsland, M., Sandberg, M., Sandberg, R., Muhr, J. and Ericson, J. (2013). SoxB1-driven transcriptional network underlies neural-specific interpretation of morphogen signals. *Proc. Natl. Acad. Sci.* 110, 7330–7335.

Palm, W., Swierczynska, M. M., Kumari, V., Ehrhart-Bornstein, M., Bornstein, S. R. and Eaton, S. (2013). Secretion and Signaling Activities of Lipoprotein-Associated Hedgehog and Non-Sterol-Modified Hedgehog in Flies and Mammals. *PLoS Biol.* 11, e1001505.

Parchure, A., Vyas, N., Ferguson, C., Parton, R. G. and Mayor, S. (2015). Oligomerization and endocytosis of Hedgehog is necessary for its efficient exovesicular secretion. *Mol. Biol. Cell* 26, 4700–17.

Pepinsky, R. B., Zeng, C., Wen, D., Rayhorn, P., Baker, D. P., Williams, K. P., Bixler, S. A., Ambrose, C. M., Garber, E. A., Miatkowski, K., et al. (1998). Identification of a Palmitic Acid-modified Form of Human Sonic hedgehog. *J. Biol. Chem.* 273, 14037–14045.

Petrov, K., Wierbowski, B. M. and Salic, A. (2017). Sending and Receiving Hedgehog Signals. *Annu. Rev. Cell Dev. Biol.* 33, 145–168.

Piccolo, S., Agius, E., Lu, B., Goodman, S., Dale, L. and De Robertis, E. M. (1997). Cleavage of Chordin by Xolloid metalloprotease suggests a role for proteolytic processing in the regulation of Spemann organizer activity. *Cell* 91, 407–16.

Pierani, A., Brenner-Morton, S., Chiang, C. and Jessell, T. M. (1999). A sonic hedgehog-independent, retinoid-activated pathway of neurogenesis in the ventral spinal cord. *Cell* 97, 903–15.

Porter, J. A., Ekker, S. C., Park, W.-J., von Kessler, D. P., Young, K. E., Chen, C.-H., Ma, Y., Woods, A. S., Cotter, R. J., Koonin, E. V, et al. (1996a). Hedgehog Patterning Activity: Role of a Lipophilic Modification Mediated by the Carboxy-Terminal Autoprocessing Domain. *Cell* 86, 21–34.

Porter, J. A., Young, K. E. and Beachy, P. A. (1996b). Cholesterol Modification of Hedgehog Signaling Proteins in Animal Development. *Science* (80-.). 274, 255–259.

Ribes, V. and Briscoe, J. (2009). Establishing and interpreting graded Sonic Hedgehog signaling during vertebrate neural tube patterning: the role of negative feedback. *Cold Spring Harb. Perspect. Biol.* 1, a002014.

Roberts, B., Casillas, C., Alfaro, A. C., Jägers, C. and Roelink, H. (2016). Patched1 and Patched2 inhibit Smoothened non-cell autonomously. *Elife* 5,.

Rogers, K. W. and Schier, A. F. (2011). Morphogen gradients: from generation to interpretation. *Annu. Rev. Cell Dev. Biol*. 27, 377–407.

Rubin, J. B., Choi, Y. and Segal, R. A. (2002). Cerebellar proteoglycans regulate sonic hedgehog responses during development. *Development* 129, 2223–32.

Sanders, T. A., Llagostera, E. and Barna, M. (2013). Specialized filopodia direct long-range transport of SHH during vertebrate tissue patterning. *Nature* 497, 628–32.

Sasai, Y., Lu, B., Steinbeisser, H., Geissert, D., Gont, L. K. and De Robertis, E. M. (1994). Xenopus chordin: a novel dorsalizing factor activated by organizer-specific homeobox genes. *Cell* 79, 779–90.

Schneider, S., Steinbeisser, H., Warga, R. M. and Hausen, P. (1996). Beta-catenin translocation into nuclei demarcates the dorsalizing centers in frog and fish embryos. *Mech. Dev.* 57, 191–8.

Shimmi, O., Umulis, D., Othmer, H. and O'Connor, M. B. (2005). Facilitated Transport of a Dpp/Scw Heterodimer by Sog/Tsg Leads to Robust Patterning of the Drosophila Blastoderm Embryo. *Cell* 120, 873–886.

Spemann, H. (1938). Embryonic Development and Induction. Yale Univ.; New Haven:

Stone, D. M., Hynes, M., Armanini, M., Swanson, T. A., Gu, Q., Johnson, R. L., Scott, M. P., Pennica, D., Goddard, A., Phillips, H., et al. (1996). The tumour-suppressor gene patched encodes a candidate receptor for Sonic hedgehog. *Nature* 384, 129–134.

Taipale, J., Cooper, M. K., Maiti, T. and Beachy, P. A. (2002). Patched acts catalytically to suppress the activity of Smoothened. *Nature* 418, 892–896.

Tenzen, T., Allen, B. L., Cole, F., Kang, J.-S., Krauss, R. S. and McMahon, A. P. (2006). The Cell Surface Membrane Proteins Cdo and Boc Are Components and Targets of the Hedgehog Signaling Pathway and Feedback Network in Mice. *Dev. Cell* 10, 647–656.

The, I., Bellaiche, Y. and Perrimon, N. (1999). Hedgehog movement is regulated through tout velu-dependent synthesis of a heparan sulfate proteoglycan. *Mol. Cell* 4, 633–9.

Tian, H., Jeong, J., Harfe, B. D., Tabin, C. J. and McMahon, A. P. (2005). Mouse Disp1 is required in sonic hedgehog-expressing cells for paracrine activity of the cholesterol-modified ligand. *Development* 132, 133–42.

Timmer, J. R., Wang, C. and Niswander, L. (2002). BMP signaling patterns the dorsal and intermediate neural tube via regulation of homeobox and helix-loop-helix transcription factors. *Development* 129, 2459–2472.

Tseng, T. T., Gratwick, K. S., Kollman, J., Park, D., Nies, D. H., Goffeau, A. and Saier, M. H. (1999). The RND permease superfamily: an ancient, ubiquitous and diverse family that includes human disease and development proteins. *J. Mol. Microbiol. Biotechnol.* 1, 107–25.

Tukachinsky, H., Kuzmickas, R. P. P., Jao, C. Y. Y., Liu, J. and Salic, A. (2012). Dispatched and Scube Mediate the Efficient Secretion of the Cholesterol-Modified Hedgehog Ligand. *Cell Rep.* 2, 308–320.

Umulis, D. M. and Othmer, H. G. (2012). Scale invariance of morphogen-mediated patterning by flux optimization. In *2012 5th International Conference on BioMedical Engineering and Informatics*, pp. 1030–1034. IEEE.

Umulis, D. M. and Othmer, H. G. (2013). Mechanisms of scaling in pattern formation. *Development* 140, 4830–43.

Uygur A, Young J, Huycke TR, Koska M, Briscoe J, Tabin CJ. (2016). Scaling Pattern to Variations in Size during Development of the Vertebrate Neural Tube. *Dev Cell* 37(2):127–135.

Vokes, S. A., Ji, H., McCuine, S., Tenzen, T., Giles, S., Zhong, S., Longabaugh, W. J. R., Davidson, E. H., Wong, W. H. and McMahon, A. P. (2007). Genomic characterization of Gli-activator targets in sonic hedgehog-mediated neural patterning. *Development* 134, 1977–1989.

Wang, Y.-C. and Ferguson, E. L. (2005). Spatial bistability of Dpp–receptor interactions during Drosophila dorsal–ventral patterning. *Nature* 434, 229–234.

Wang, B., Fallon, J. F. and Beachy, P. A. (2000). Hedgehog-regulated processing of Gli3 produces an anterior/posterior repressor gradient in the developing vertebrate limb. *Cell* 100, 423–34.

Wolpert, L. (2011). Positional information and patterning revisited. *J. Theor. Biol*. 269, 359–65.

Woods, I. G. and Talbot, W. S. (2005). The you gene encodes an EGF-CUB protein essential for Hedgehog signaling in zebrafish. *PLoS Biol.* 3, e66.

Xu, J., Srinivas, B. P., Tay, S. Y., Mak, A., Yu, X., Lee, S. G. P., Yang, H., Govindarajan, K. R., Leong, B., Bourque, G., et al. (2006). Genomewide expression profiling in the zebrafish embryo identifies target genes regulated by Hedgehog signaling during vertebrate development. *Genetics* 174, 735–52.

Yan, D. and Lin, X. (2009). Shaping morphogen gradients by proteoglycans. *Cold Spring Harb. Perspect. Biol.* 1, a002493.

Zagorski, M., Tabata, Y., Brandenberg, N., Lutolf, M. P., Tkačik, G., Bollenbach, T., Briscoe, J. and Kicheva, A. (2017). Decoding of position in the developing neural tube from antiparallel morphogen gradients. *Science* 356, 1379–1383.

Zee, M. v. d., Stockhammer, O., Levetzow, C. v., Fonseca, R. N. d. and Roth, S. (2006). Sog/Chordin is required for ventral-to-dorsal Dpp/BMP transport and head formation in a short germ insect. *Proc. Natl. Acad. Sci.* 103, 16307–16312.

Zeng, X., Goetz, J. A., Suber, L. M., Scott, W. J., Schreiner, C. M. and Robbins, D. J. (2001). A freely diffusible form of Sonic hedgehog mediates long-range signaling. *Nature* 411, 716–20.

Zhang, W., Kang, J.-S., Cole, F., Yi, M.-J. and Krauss, R. S. (2006). Cdo Functions at Multiple Points in the Sonic Hedgehog Pathway, and Cdo-Deficient Mice Accurately Model Human Holoprosencephaly. *Dev. Cell* 10, 657–665.

Chapter II.

A Scube2-Shh feedback loop links morphogen release to morphogen signaling to enable scale invariant patterning of the ventral neural tube

Authors: Zachary M. Collins[a], Kana Ishimatsu[a], Tony Y.C. Tsai[a], and Sean G. Megason[a]

[a] Department of Systems Biology, Harvard Medical School, Boston, MA 02115, USA.

Author contributions

Z.M.C. conducted experiments and data analysis. Z.M.C and S.G.M. conceived the study, designed the experiments, and wrote the paper. K.I aided in the development of the size reduction technique. T.Y.C.T helped develop the image analysis technique and generated the tg(*shha:memcherry*) reporter line. S.G.M. supervised the overall study.

Abstract

Sonic Hedgehog (Shh) signaling serves as a key patterning cue in early development. To enable robust patterning Shh signaling is tightly regulated. Here we explore how a Shh morphogen gradient in the ventral neural tube enables proportional patterning in embryos of varying sizes. Using a surgical technique to reduce the size of zebrafish embryos and quantitative confocal microscopy, we find that patterning of neural progenitors remains proportional after size reduction. We show that this robust patterning depends on Scube2, a protein expressed on the opposite side of the neural tube which is known to enable Shh release. *scube2* expression levels control the intensity of Shh signaling during ventral neural patterning. We find that Scube2 is highly diffusible and spreads broadly from producing cells during patterning. Membrane tethering experiments suggest, that Scube2's mobility is required for its full activity. In addition, we demonstrate that Shh signaling represses the expression of Scube2, thereby constricting its own signaling. By returning to surgical size reduction experiments, we show that Scube2 is disproportionately downregulated in size-reduced embryos. Moreover, this regulatory feedback is necessary for pattern scaling, as demonstrated by a loss of scaling in Scube2 overexpressing embryos. We conclude that feedback between Shh release and *scube2* expression enables proportional patterning in the ventral neural tube by encoding a tissue size dependent morphogen gradient.

Introduction

When Lewis Wolpert first posed the "French Flag Problem", he was seeking the answer to this fundamental question: What systems enable proportional patterning in embryos independent of embryo size? By the time Wolpert formalized his hypothesized answer to the French Flag Problem and the role of morphogens in positional information, developmental biologists had long known that embryos scale down patterning programs in response to alterations in embryo size (Wolpert, 1969). Sea urchin larva pattern normally from a single

blastomere up to the four-cell stage, and amphibian embryos can survive bisection and often pattern proportionally at a reduced size (Cooke, 1981; Driesch, 1892; Morgan, 1895; Spemann, 1938). Significant scaling of pattern formation to tissue availability is seen as a near universal property of developing organisms. Yet, as we pass the 50-year anniversary of Wolpert's work, how morphogen gradients scale to pattern domains of varied sizes remains unclear in many systems.

Recent theoretical studies have proposed mechanisms that could account for scaling of morphogen-mediated patterning (Ben-Zvi and Barkai, 2010; Umulis and Othmer, 2013). Amongst the most prominent of these is a model termed expander-repressor integral feedback control (Ben-Zvi and Barkai, 2010). Under this regime, a morphogen represses the expression of another gene that affects the range of the morphogen itself cell-non-autonomously, known as the expander. In such models, morphogen signaling will expand until it has reached an encoded equilibrium. This equilibrium is controlled by the morphogen's repression of the expander, thus enabling "measurement" of the size of the domain in need of patterning. The first proposed biological example of this mechanism was in Xenopus axial patterning. In this model, ADMP expands BMP signaling by binding Chordin and inhibiting shuttling of BMP towards the ventral side (Francois et al., 2009). However, more recent experimental work implicated another factor, Sizzled, which regulates the rate of Chordin degradation in a mathematically equivalent manner to expander-repressor systems (Ben-Zvi et al., 2014; Inomata et al., 2013). Expander-like relationships have also been proposed to regulate scaling of Dpp gradients during wing disc growth and even scaling of synthetic patterns in bacterial colonies (Ben-Zvi et al., 2011a; Cao et al., 2016; Hamaratoglu et al., 2011).

Though scaling of early axis patterning following size reduction has been extensively studied, the molecular mechanisms through which organs subsequently scale their patterning in these embryos has received relatively little attention (Ben-Zvi et al., 2008; Inomata et al., 2013). Patterning organs in size-reduced embryos face the same challenge as those of earlier axial

patterning: a need to compensate for a changed patterning domain size using signals with the same chemical properties. Scaling of patterning during organ growth has mostly been considered in the fly wing disc, which grows remarkably in size while maintaining proportion (Averbukh I et al., 2014; Ben-Zvi et al., 2011a; Hamaratoglu et al., 2011). In vertebrates, the developing neural tube has been a powerful model to study morphogen-mediated patterning. While neural tube patterning does not scale over time with growth, embryos of different species maintain consistent embryonic proportions with vastly different sizes at initial patterning (Kicheva et al., 2014; Uygur et al., 2016).

The vertebrate ventral neural tube is patterned by the morphogen Sonic Hedgehog (Shh). Shh is produced by the notochord and floorplate and induces ventral cell fates over a long range in a dose-dependent manner (Briscoe et al., 2001; Zeng et al., 2001). Shh ligands themselves are dually lipid-modified and are highly lipophilic (Pepinsky et al., 1998; Porter et al., 1996a; Porter et al., 1996b). While mechanisms of Shh transport have long been disputed, biochemical evidence supports soluble Shh as a primary component of long-range signaling (Chen et al., 2004; Zeng et al., 2001). Release of Shh ligands from cell membranes is thought to be critical for their long-range signaling, and thus gradient formation. Shh release was largely thought be achieved by the protein, Dispatched, but recent work has identified Scube2 as a more potent factor in promoting Shh release (Burke et al., 1999; Creanga et al., 2012; Kawakami et al., 2002; Tukachinsky et al., 2012).

Scube2 is a Signal sequence containing proteins with a CUB domain and EGF-like repeats. Scube2's role in Shh signaling was first identified from work using the zebrafish *you* mutant (Hollway et al., 2006; Kawakami et al., 2005; van Eeden et al., 1996; Woods and Talbot, 2005). Interestingly, while *scube2* mutants have defects in ventral patterning, *scube2* is predominantly expressed in the dorsal and intermediate neural tube in both mice and zebrafish (Grimmond et al., 2001; Kawakami et al., 2005; Woods and Talbot, 2005). Additionally, epistasis experiments indicated that Scube2 acts upstream of Patched to stimulate Shh signaling (Woods

and Talbot, 2005). This effect was also found to be cell-non-autonomous, as mosaic injection of *scube2* mRNA was capable of rescuing Shh-signaling defects over a long range (Hollway et al., 2006; Woods and Talbot, 2005). Studies in cell culture then demonstrated that Scube2 releases Shh from secreting cells in a cell-non-autonomous manner (Creanga et al., 2012; Tukachinsky et al., 2012). Recent work has argued that Scube2 may be responsible for catalyzing the shedding of lipids from the Shh ligands, but this model conflicts with previous findings that released Shh is dually lipid-modified (Creanga et al., 2012; Jakobs et al., 2014; Jakobs et al., 2016; Tukachinsky et al., 2012). Scube2's cell non-autonomous role in Shh release and unexpected expression pattern led us to wonder what role it plays in the regulation of Shh signaling. In this work, we use quantitative imaging of cell fate specification in zebrafish to investigate the scaling of ventral neural patterning and the regulatory role of Scube2.

Results

Ventral neural patterning scales with embryo size

Scaling of pattern formation to differences in tissue size occurs in many developmental and evolutionary contexts. However, studying mechanisms of scaling during growth of an organism or between species of different sizes is difficult because many aspects of the system change over developmental and evolutionary time. (Ben-Zvi et al., 2011a; Hamaratoglu et al., 2011; Kicheva et al., 2014; Uygur et al., 2016). To study scaling of pattern formation in embryos with comparable genetic backgrounds at matched time points, we developed a technique to reduce the size of zebrafish embryos inspired by classical work in amphibians (Ishimatsu et al., 2017; Morgan, 1895; Spemann, 1938). To avoid damaging signaling centers crucial to early D-V patterning, two lateral cuts are made across the blastula stage embryo. First, we removed cells from the animal cap, followed by removal of a limited amount of yolk at the vegetal pole (Figure 2.1A). With this technique, a significant fraction of embryos pattern normally and develop at a reduced size (Figure 2.1B).

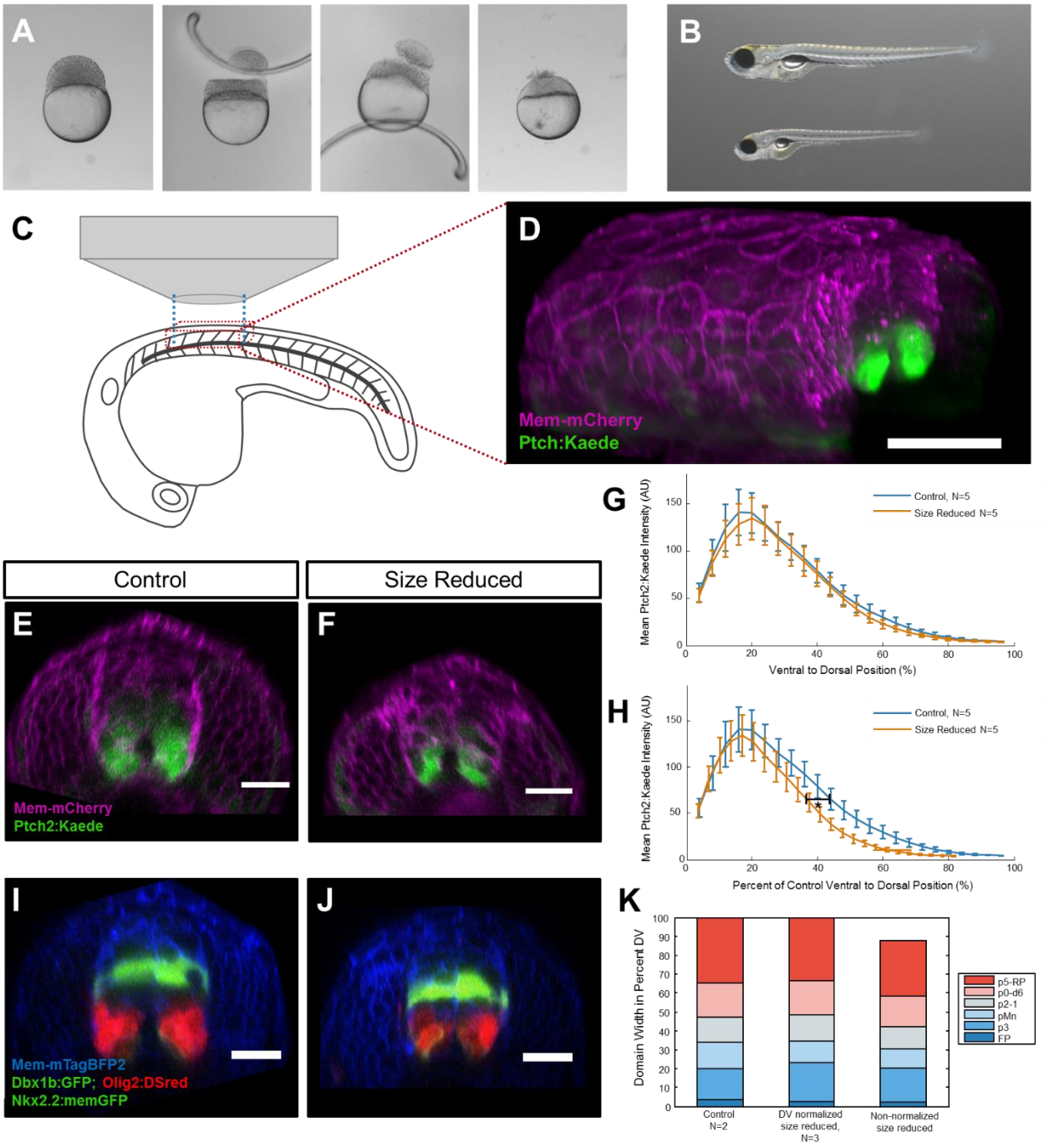


Figure 2.1 Neural tube patterning scales following embryonic size reduction

(A) Surgical size reduction technique during which cells and yolk are removed at the 128-256 cell stage. (B) Example of a resulting size-reduced larva at 6 DPF on the bottom, with a normal-sized sibling above. (C) Schematic of an embryo mounted for imaging; anterior-posterior length

Figure 2.1 (Continued) of the imaging window is shown with blue lines. Red lines indicate the 3-D extent of the imaging window. (D) 3-D rendering of a confocal z-stack on an example *ptch2:kaede mem-mcherry* image volume. (E-F) Transverse view of 20 hpf tg(*ptch2:kaede*) control (E) and size-reduced (F) embryos, which have been injected with *mem-mcherry* mRNA. (G) Dorsal-ventral height-normalized *ptch2:kaede* intensity profiles demonstrating scaling of the response gradient. (H) Non-D-V height-normalized *ptch2:kaede* intensity profiles from G. Statistically significant shift in position of 50% control max intensity is shown with an asterisk (unpaired t-test p=0.0100; Control N=5, Size Reduced N=5). (I-J) Transverse view of 22 hpf tg(*nkx2:mgfp; olig2:dsred; dbx1b:GFP*) control (I) and size-reduced (J) embryos. (K) Quantification of mean reporter intensity versus ventral-to-dorsal position of embryos from I-J. When compared relative to their individual D-V heights, statistically significant shifts are seen in the extent of dorsal-ventral patterning, as demonstrated by a ventral shift in the upper boundaries of the p1-p2 and p0-d6 domains (unpaired t-test p^{p2}=0.0495 and p^{d6} =0.0309, Control N=2, Size Reduced N=3).

To quantitatively measure scaling of neural patterning in the size-reduced embryos, we utilized recently developed imaging methods and designed an image analysis pipeline (Megason SG, 2009; Xiong et al., 2013). At 18-22 Hours Post Fertilization (hpf), embryos were anesthetized in tricaine and mounted for confocal imaging in the dorsal-mount, allowing high resolution imaging of the spinal cord (Figure 2.1C). During each imaging session, confocal z-stacks were collected with identical settings from a matched somite level in control and experimentally-perturbed embryos. Imaging volumes were analyzed by manually selecting the dorsoventral axis and width of the neural tube along the length of the dataset (see methods for more details). Imaging data from each left-right half of the neural tube were then recovered and quantified in a set number of bins along the D-V axis (Figure S2.1). This system allowed for the quantitative and unbiased comparison of 3-4 somite lengths of neural imaging data from multiple embryos.

Using our imaging platform, we compared the expression of *patched2*—a direct transcriptional target of Shh—using the tg(*ptch2:kaede*) reporter in stage-matched wild type and size-reduced embryos (Huang et al., 2012). When quantified relative to total neural tube dorsal-ventral height, Patched2 response gradients maintained nearly identical intensity distributions. Neural tube height in size-reduced embryos was diminished on average by 12% in this dataset (N=2), indicating that Shh responses scale following size reduction (Figure 2.1E-G). When viewed in their endogenous dimensions, *ptch2:kaede* intensity distributions no longer scaled with D-V height (Figure 2.1H). To quantify this effect at the level of cell fate specification, we utilized a triple transgenic imaging strategy based on reporter lines marking Nkx2.2a (p3 progenitors), Olig2 (pMn and some p3 progenitors), and Dbx1b (p0, d6 progenitors) (Gribble et al., 2007; Jessen et al., 1998; Kinkhabwala et al., 2011; Kucenas et al., 2008). Average intensity profiles were then collected and segmented to form cell fate profiles (see supplementary methods and Figure S2.2). Using this method, we generated average cell fate profiles which can be compared between embryos (Figure S2.2). After normalizing for their altered D-V height,

average cell fate profiles of size-reduced embryos were virtually indistinguishable from controls (Figure 2.1I-K). Furthermore, discrepancies between progenitor boundary positions were visible when size normalization was removed (Figure 2.1K). Statistically significant shifts were seen in the p0 upper boundary and d6 upper boundary positions relative to controls only when compared in their endogenous coordinates (Figure 2.1K). This demonstrates that ventral neural patterning adjusts to changes in total D-V height.

Scube2 levels control Shh signaling

Scube2's role in the cell-non-autonomous regulation of Shh release and dorsal expression pattern led us to hypothesize a potential role of Scube2 in enabling scaling of Shh gradients. This hypothesis depends on *scube2* expression levels modulating Shh signaling activity. However, previous work has argued that Scube2 is only required for Shh signaling as a permissive factor (Kawakami et al., 2005; Woods and Talbot, 2005). To examine the role of Scube2 in ventral neural patterning, we performed a morpholino knockdown of *scube2* in tg(*ptch2:kaede*) reporter embryos using a previously validated translation inhibiting morpholino (Figure 2.2A-C) (Woods and Talbot, 2005). We observed a decrease in Shh signaling following morpholino injection, as demonstrated by a statistically significant suppression of maximum *ptch2:kaede* intensity (Figure 2.2C) (Woods and Talbot, 2005). Additionally, quantification of Nkx2.2a, Olig2, and Dbx1b domain sizes in embryos injected with *scube2* morpholino showed a contraction of ventral progenitor domains (Figure 2.2D-F). Ventral shifts in the upper boundaries were statistically significant, due in part to near complete elimination of the *nkx2.2a*+ p3 domain (Figure 2.2F). Previous work concluded that Scube2 was a permissive factor based on a qualitative lack of amplification of Shh signaling following overexpression. Our data demonstrate that injection of *scube2* mRNA causes the expansion of Shh signaling, as shown by higher intensities and broader distributions of tg(*ptch2:kaede*) fluorescence (Figure 2.2G-I). Embryos injected with *scube2* mRNA showed significant increases in maximum tg(*ptch2:kaede*)

intensities (Figure 2.2H). In addition, *scube2* overexpression affected cell type specification in the ventral neural tube, as measured in triple transgenic Nkx2.2a, Olig2, and Dbx1b embryos (Figure 2.2J-L). Quantification of cell fate profiles in these embryos revealed large increases in p3 and pMn domain sizes, a decrease in the size of the p2-p1 domains, and unchanged patterning of the p0-d6 domains and more dorsal cell types. Ventralization was measured by comparing dorsal boundaries of the p3, pMn, and d6 domains, which were all statistically significantly shifted (Figure 2.2L). These data indicate that not only is Scube2 required for long range Shh signaling, but that *scube2* overexpression also amplifies endogenous Shh signaling. Additionally, this suggests that Scube2-stimulated Shh release may not be saturated in normal development.

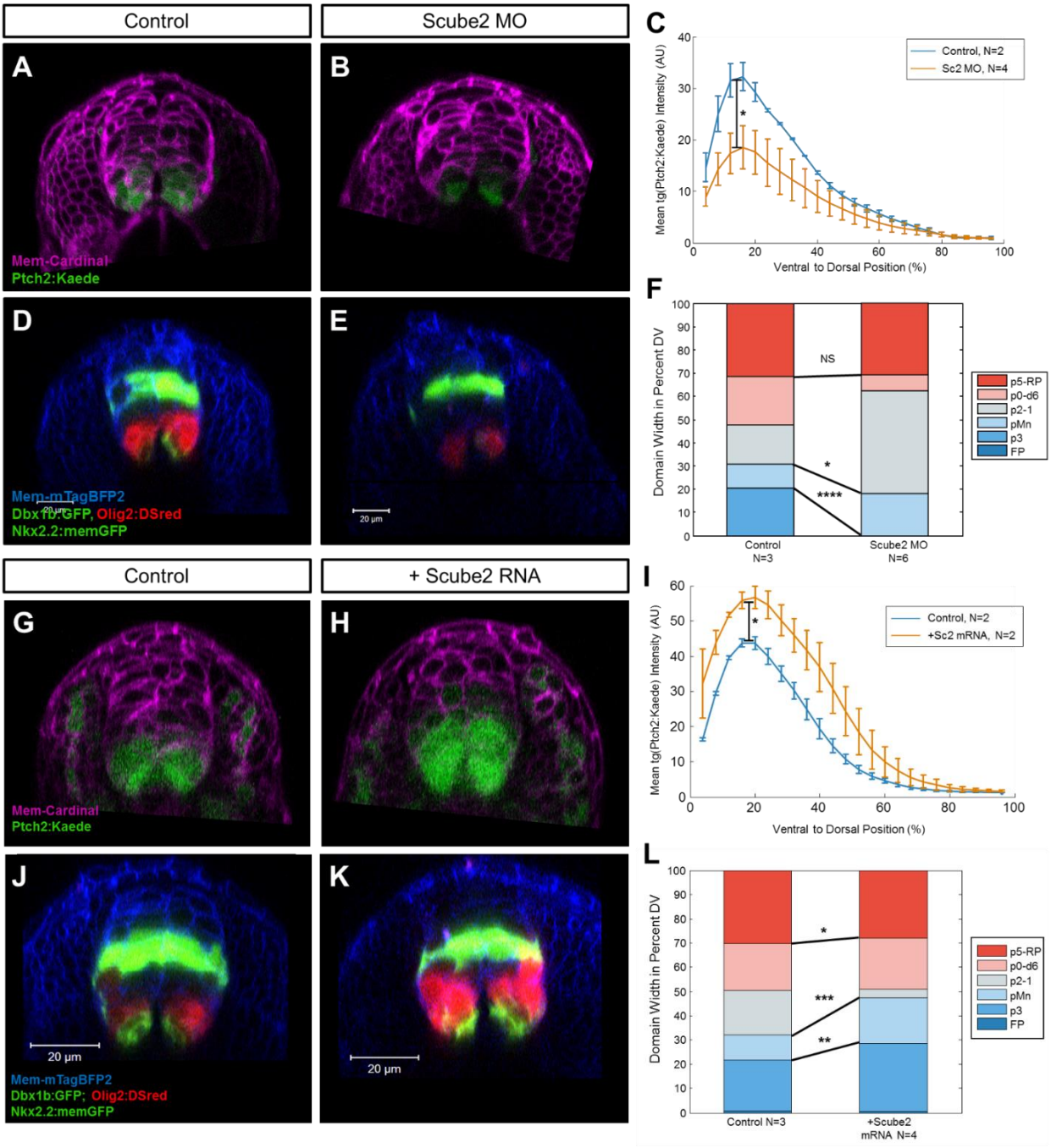


Figure 2.2 Scube2 expression levels regulate Shh signaling in the ventral neural tube

(A-B) Transverse view of a confocal z-stack of 22 hpf tg(*ptch2:kaede*) reporter line embryos injected with (A) *mem-mcardinal* mRNA alone or (B) co-injected with *scube2* morpholino. (C) Quantification of mean intensity distributions in segmented neural tissue from z-stacks of

Figure 2.2 (Continued) embryos as treated in A-B. Maximum intensities of morpholino treated embryos were statistically significantly reduced compared to controls (p= 0.0183). (D-E) Transverse view of 20 hpf tg(*dbx1b:gfp, olig2:dsred, nkx2.2a:memgfp*) reporter line embryos injected with (D) mem-*mtagbfp2* mRNA alone or (E) co-injected with *scube2* morpholino. (F) Results of automated segmentation of progenitor domain sizes using a custom algorithm with embryos treated as in D-E. Statistical comparisons of progenitor domain boundaries are shown with connected lines and significance marked by number of asterisks. Changes in the d6 upper boundary were not significantly different between control and morpholino populations (p= 0.2718); the upper boundary of pMn and p3 domain were both significantly contracted in morpholino injected embryos (p^{Mn}= 0.0158 and p^{p3}= 9.8729e-09). (G-H) Transverse view of 20 hpf tg(ptch2:kaede) reporter line embryos injected with (G) *mem-mcardinal* mRNA alone or (H) co-injected with *scube2* mRNA. (I) Quantification of mean intensity distributions of embryos as treated in G-H. Maximum intensity values of *scube2* mRNA-injected embryos is statistically significantly increased over controls (p=.0367). (J-K) Transverse view of 20 hpf tg(*dbx1b:GFP, olig2:dsred, nkx2.2:mgfp*) reporter line embryos injected with (J) *mem-mtagbfp2* mRNA alone or (K) co-injected with *scube2* mRNA. (F) Mean results of automated progenitor domain segmentation J-K. Statistical comparisons of progenitor domain boundaries are shown with connected lines and significance marked by number of asterisks. Changes in upper boundary of d6 was somewhat shifted while pMn and p3 domains were drastically shifted dorsally in *scube2* mRNA injected embryos (p^{d6}= 0.0196, p^{Mn}=8.6748e-04, and p^{p3}=0.0034 respectively).

Shh negatively regulates Scube2 expression over a long-range

To study Scube2's expression, we developed the tg(*scube2:moxng*) reporter line containing 7.6KB of the endogenous regulatory sequences driving the extremely bright moxNeonGreen fluorescent protein (Figure 2.3A). The expression of tg (*scube2:moxng*) we observed is consistent with previously reported *in situ* hybridizations (Grimmond et al., 2001; Kawakami et al., 2005; Woods and Talbot, 2005). Tg(*scube2:moxng*) embryos showed down-regulated expression close to the sources of Shh in the floor plate and notochord—as visualized with a transgenic *shh:memCherry* reporter line—and high levels of expression in the medial neural tube (Figure 2.3A-C). Time lapse imaging of tg(*scube2:moxng*) embryos revealed weak mesodermal expression in the early embryo, which faded during the onset of neurulation and was replaced by high levels of expression in the dorsal and intermediate neural tube (Figure S2.3).

A key prediction of expander-repressor based models of scaling is that the expander is repressed by signaling downstream of the morphogen. To test whether Scube2 is similarly downregulated by Shh signaling, we injected mRNA encoding a potent activator of the Shh pathway, dnPKA, at the single cell stage and observed the resulting embryos. Embryos injected with *dnpka* mRNA showed near complete ablation of neural tg(*scube2:moxng*) expression (Figure 2.3 D-F). To test whether Shha ligands themselves were capable of suppressing Scube2 expression at a distance, we mosaically overexpressed *shha* in tg(*scube2:moxng*) embryos by injecting them at the 16-cell stage with either *mem-mtagbfp2* alone or with *shha* mRNA (Figure 2.3 G-I). We expected local inhibition of Scube2 reporter activity near secreting cells. Surprisingly, tg(*scube2:moxng*) expression was nearly completely eliminated in these embryos as well, showing cell-non-autonomous repression of Scube2 by Shh, as quantified by a highly significant reduction of peak intensities (Figure 2.3 I). To test whether Shh's inhibition of Scube2 was required for its ventral downregulation, we treated embryos with cyclopamine, a potent Smoothened antagonist starting at the late gastrula stage. Resulting embryos showed

expanded *scube2* expression towards the floor plate and notochord (Figure 2.3J-L). Shifts in ventral boundaries were quantified by measuring the D-V position at which 50% of the maximum intensity of the control population was reached. These measurements were statistically significantly shifted in cyclopamine-treated embryos relative to controls, indicating that endogenous Shh signaling is responsible for a lack of ventral *scube2* expression (Figure 2.3L). To further probe the transcriptional regulation of Scube2's expression, we performed a small scale CRISPR screen. Pax6a and Pax6b mosaic CRISPR mutants showed significant downregulation of tg(*scube2:moxng*) relative to embryos injected with sgRNA targeting *tyrosinase* an unrelated pigment gene (Figure S2.4). Whether Shh directly regulates *scube2* expression or acts only by repressing *pax6* related transcription factors remains to be determined.

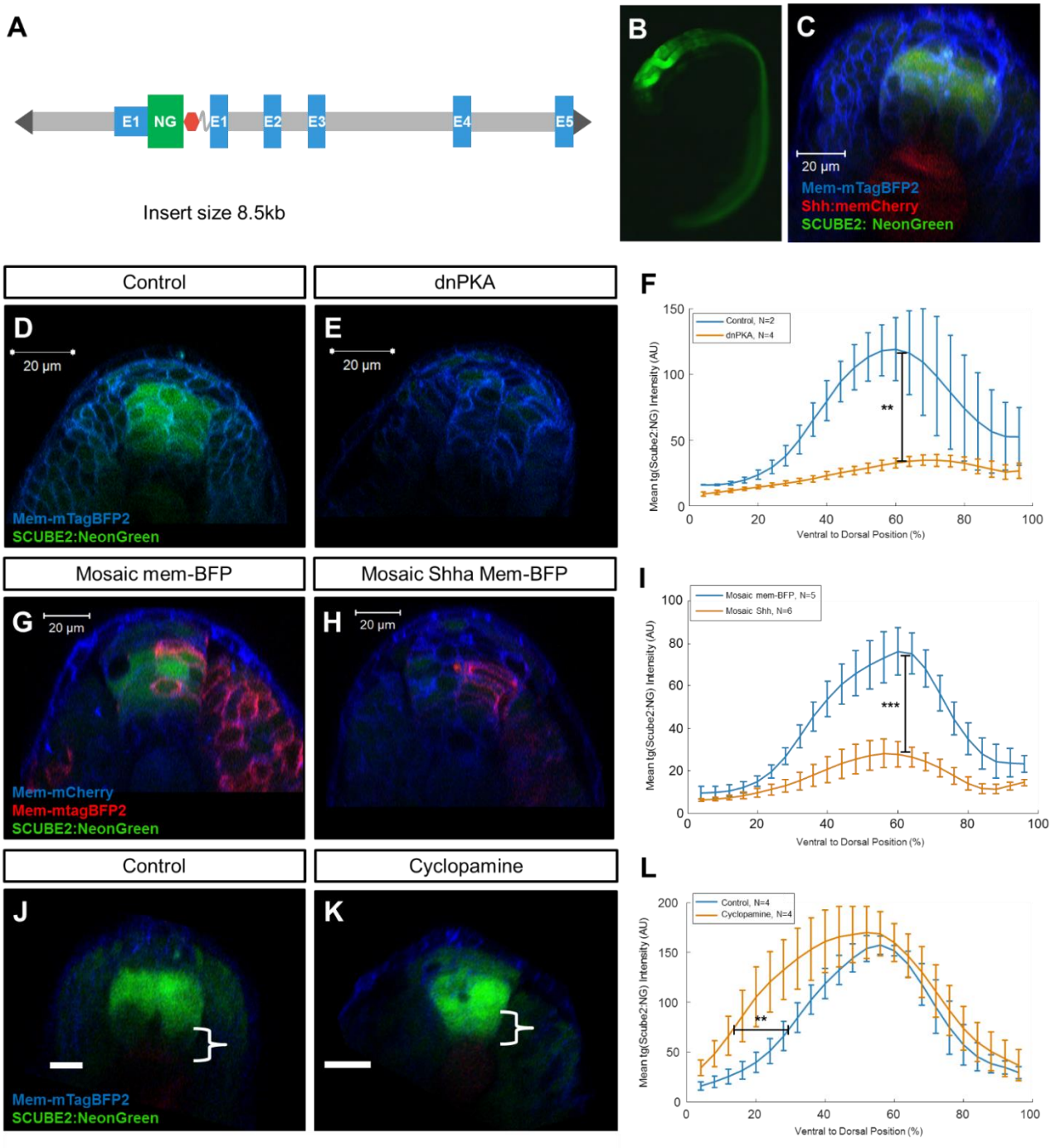


Figure 2.3 Scube2 is expressed distantly from Shh secreting cells and Shh signaling represses Scube2 expression

(A) Schematic of the *scube2:moxng* transgenic expression reporter construct used to generate the tg(*scube2:moxng*) line. (B) Wide-field fluorescence image of tg(*scube2:moxng*) embryos at

Figure 2.3 (Continued) 20 hpf. (C) Transverse view of *mem-mtagbfp2*-injected tg(*scube2:moxng; shh:mem-mcherry*) embryos at 20 hpf. (D-E) Transverse view of 18 hpf tg(*scube2:moxng*) reporter line embryos injected with (D) *mem-mtagbfp2* mRNA alone or (E) co-injected with *dnpka* mRNA. (F) Quantification of mean reporter intensity of embryos as treated in D-E. Maximum *scube2:moxng* intensity values were significantly reduced in *dnpka* mRNA-injected embryos (p= 0.0014). (G-H) Transverse view 20 hpf tg(*scube2:moxng*) reporter line embryos injected at the single cell stage with *mem-mcherry* mRNA and then injected at the 8-16 cell stage with either (G) *mem-mtagbfp2* mRNA alone or (H) co-injected with *shha* mRNA. (I) Quantification of mean reporter intensity of embryos as treated in G-H. Maximum *scube2:moxng* reporter intensity is significantly reduced in *shha* injected embryos (p=.0000091439). (J-K) Transverse view 20 hpf tg(*scube2:moxng*) reporter line embryos injected with *mem-mtagbfp2* mRNA and (J) treated with an ethanol sham or (K) treated with Cyclopamine. (L) Quantification of mean reporter intensity of embryos as treated in J-K. The black bar marks the position of 50% of control maximum intensity which was used for statistical testing. These values were statistically significantly shifted ventrally in *scube2* overexpressing embryos relative to control (p= 0.0045).

Scube2 diffuses during patterning which is necessary for its cell non-autonomous activity

While Scube2 is known to act cell non-autonomously from transplantation experiments and Scube2-conditioned media has a potent Shh release stimulating effect *in vitro*, Scube2's localization during development had yet to be explored (Woods and Talbot, 2005; Creanga et al., 2012). *In vitro*, Scube2 is thought to associate with Heparin Sulfate Proteoglycans, and Scube2's diffusion from expressing cells *in vivo* has been disputed (Kawakami et al., 2005; Hollway et al., 2006; Jakobs et al., 2016). To examine Scube2's localization, we developed Scube2 fluorescent fusion proteins by tagging the C-terminus based on previously validated Scube2 tagging approaches (Figure 2.4A) (Creanga et al., 2012). The resulting Scube2-Citrine fusion proteins were functional and rescued Scube2 CRISPR mutants at comparable rates to wildtype Scube2 (Figure S2.5). Mosaic injection of *scube2-citrine* mRNA at the 32-64 cell stage revealed that Scube2-Citrine diffuses distantly from injected cells (Figure 2.4B). Following single cell mRNA injection, Scube2-Citrine fusions were secreted and did not remain associated with cell membranes (Figure 2.4C). In addition, Scube2-Citrine fusions recovered rapidly after photobleaching, further supporting their solubility in the extracellular space (Figure 2.4D).

To observe distributions of Scube2 during development, we generated a transgenic line expressing the full length Scube2 protein under control of Scube2 regulatory sequences (Figure 2.4E-H). Tg*(scube2:scube2-moxng*) embryos showed broad distributions of Scube2 during patterning, but localized expression (Figure 2.4E-H). Distributions of Scube2-moxNeonGreen in early patterning show Scube2 present near cells marked by tg(*shha:mem-mCherry*) (Figure 2.4E-F). These data suggest that Scube2's long range of effect can be explained by diffusion from secreting cells to the floor plate and notochord to aid in the release of Shh.

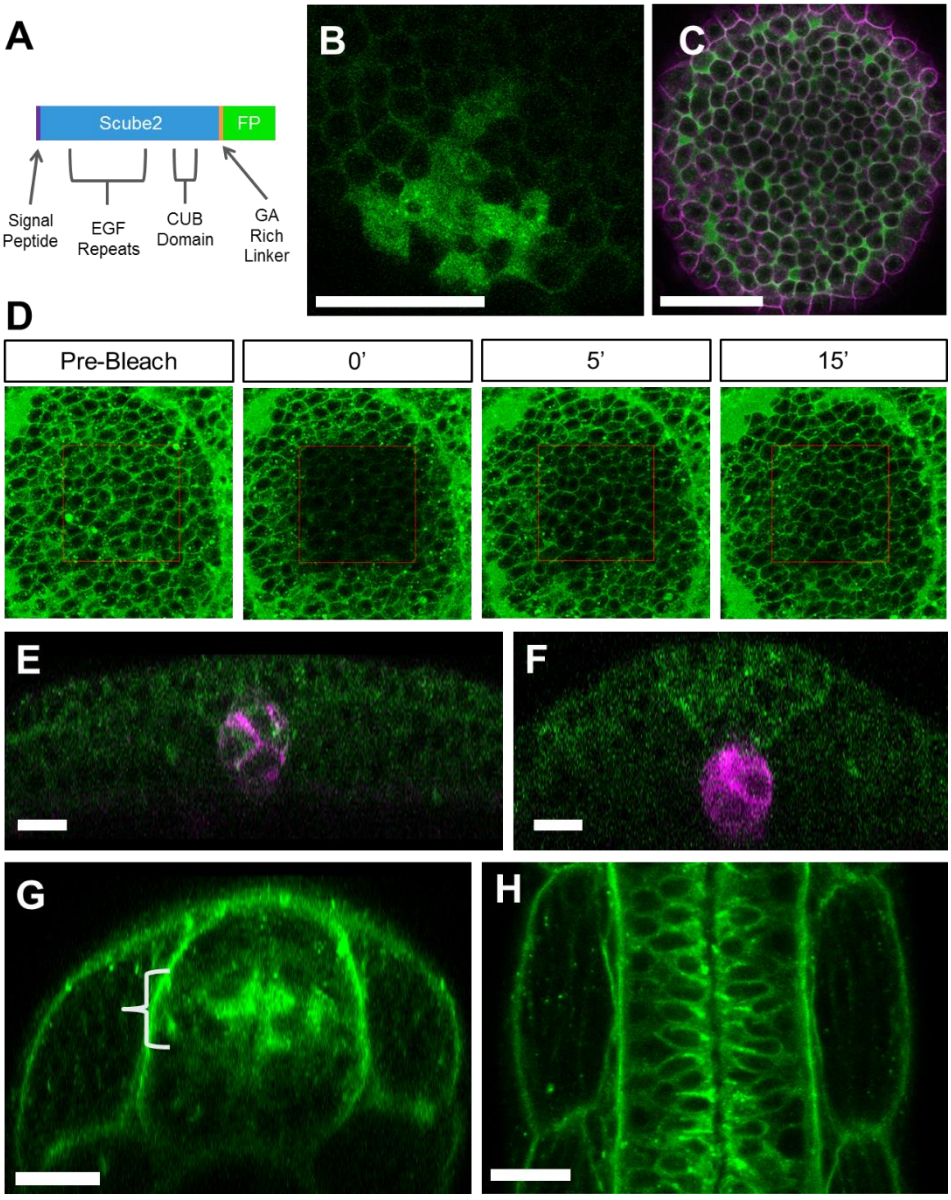


Figure 2.4 Scube2 diffuses from secreting cells and is broadly distributed during patterning

(A) Schematic of Scube2-Citrine fluorescent fusion protein design. (B-C) Scale bar represents 100 μm. (B) Scube2-Citrine fluorescence at the sphere stage from embryos injected at the 64-cell stage with *scube2-citrine* mRNA. (C) Scube2-Citrine fluorescence from embryos injected at the single cell stage *scube2-citrine* and *membrane mCherry* mRNA. (D) Fluorescence recovery after photobleaching at the neural plate stage of Scube2-Citrine. (E-F) Scale bar represents

Figure 2.4 (Continued) 20 µm. (E) Transverse view of an 11.5 hpf tg(*sc2:sc2-moxng;shh:mem-mCherry*) embryo. (F) Transverse view of a 14 hpf tg(*sc2:sc2-moxng;shh:mem-mCherry*) embryo. (G) Transverse view of a 22 hpf tg(*scube2:scube2-moxng*) embryo. (H) Horizontal view of the z-stack from D.

To assess whether Scube2's mobility is necessary for its function, we generated Morpholino-Resistant Scube2-Citrine (MR-Sc2-Cit) and Morpholino-Resistant Scube2-Citrine-TransMembrane (MR-Sc2-Cit-TM) constructs using the CD4 transmembrane domain (Figure 2.5A). As expected, MR-Sc2-Cit-TM failed to diffuse from secreting cells and localized to the plasma membrane (Figure 2.5B). Co-injection of *mr-sc2-cit* mRNA alongside *scube2* morpholino nearly fully rescued endogenous patterning in triple transgenic cell-fate reporters, with some residual defects in patterning of the *dbx1b:gfp*+ domain (Figure 2.5C-D,I). Single cell injection of *mr-sc2-cit-tm* mRNA showed partial patterning rescue (Figure 2.5C-E,I). Interestingly, MR-Sc2-Cit-TM-rescued embryos showed *olig2* expression levels and pMn/p2 boundary positions nearly identical to wildtype, which were significantly higher than embryos injected with morpholino only ($p=0.0126$) (Figure 2.5I). However, p3 progenitors in MR-Sc2-Cit-TM-rescued embryos were markedly reduced relative to control, as shown by a statistically significant ventral shift in the position of the p3-pMn boundary (Figure 2.5I). Rescue experiments performed in tg(*ptch2:kaede*) embryos yielded similar results, with greater rescue of tg(*ptch2:kaede*) intensity following injection of MR-Sc2-Cit relative to MR-Sc2-Cit-TM (Figure 2.5F-H,J). Both transmembrane and freely diffusible Scube2 constructs showed statistically significant increases in maximum *ptch2:kaede* intensity relative to control (Figure 2.5I). These data indicate that transmembrane Scube2 is able to only partially rescue the activity of the freely diffusible form, suggesting that Scube2's diffusivity or mobility in the extracellular space may be necessary for its full function in Shh signaling.

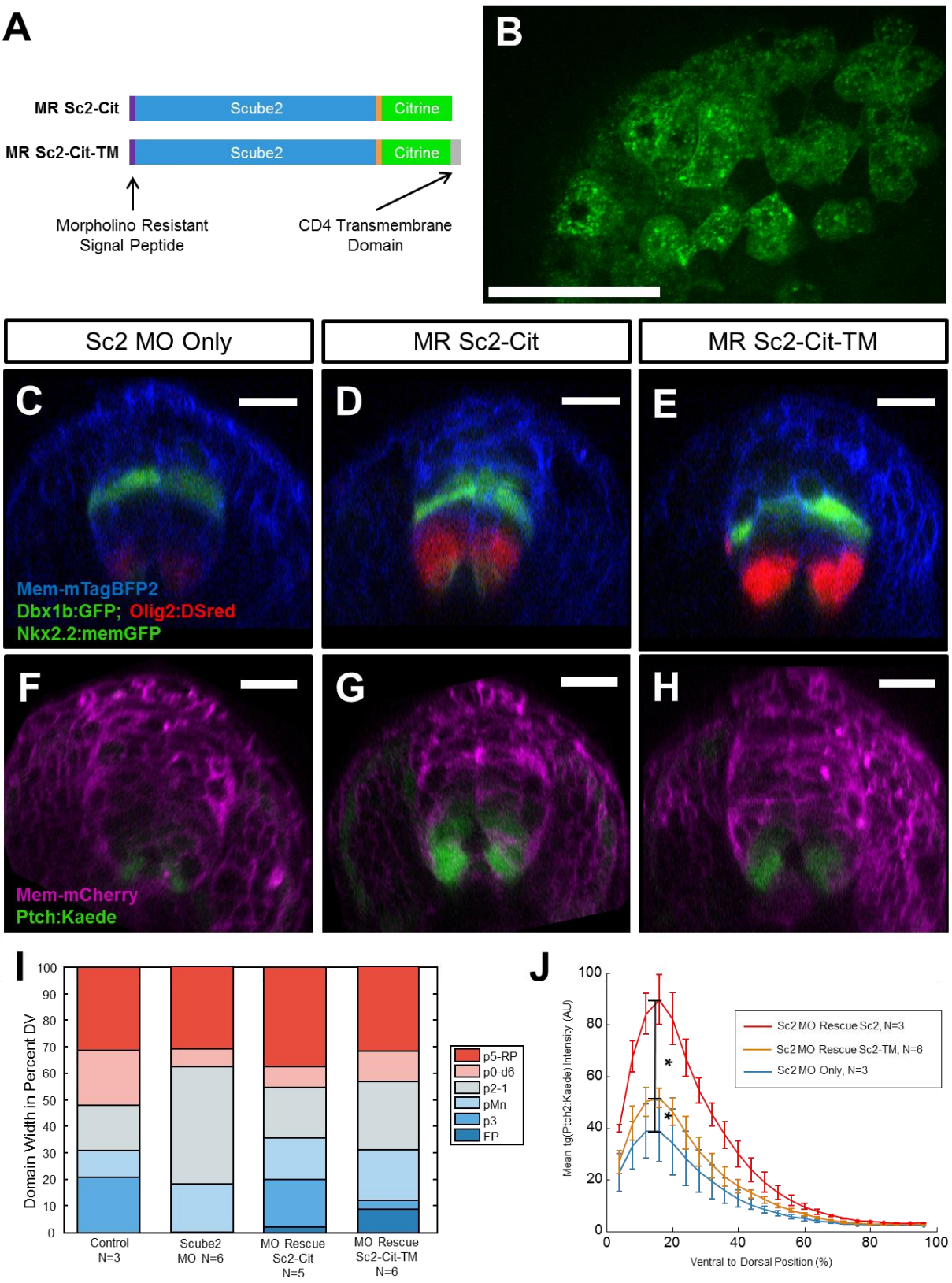


Figure 2.5 Membrane-tethered scube2 partially rescues Scube2 morphants

Figure 2.5 (Continued) (A) Schematic of Morpholino Resistant Scube2-Citrine (MR-Sc2-Cit) and Morpholino Resistant Scube2-Citrine-TransMembrane (MR-Sc2-Cit-TM) constructs. (B) Embryos mosaically injected with *scube2-citrine-tm* mRNA at the 8-16 cell stage, imaged at the sphere stage, and rendered as a maximum intensity projection. Scale bar represents 100 μm. (C-E) Transverse view of 20 hpf tg(*dbx1b:gfp, olig2:dsred, nkx2.2:memgfp*) reporter line embryos which were injected with *mem-mtagbfp2* and *scube2* morpholino (C) or co-injected with either (D) *mr-sc2-cit* or (E) *mr-sc2-cit-tm* mRNA at the single-cell stage. Scale bars represent 20 μm. (F-H) Transverse view of 20 hpf tg(*ptch2:kaede*) reporter line embryos which were injected with *mem-mtagbfp2* and *scube2* morpholino (F) or co-injected with either (G) *mr-sc2-cit* or (H) *mr-sc2-cit-tm* mRNA at the single-cell stage. Scale bars represent 20 μm. (I) Mean results of automated progenitor domain segmentation of embryos treated as in C-E compared to simultaneously imaged control (*mem-mtagbfp2* only) embryos. No statistically significant differences were found between control and morpholino rescued *scube2-cit* embryos in the position of p3, pMn, or d6 upper boundaries (p^{p3}=0.4891, p^{pMn}=0.1975, p^{d6}=0.0687). Upper boundaries of ventral progenitors, however, showed significant shifts when compared to embryos injected with only morpholino (p^{p3}=1.3517e-10, p^{pMn}=4.5318e-04). Significant lowering of the p3 upper boundary was observed in *sc2-cit-tm*-rescued embryos relative to controls, while pMn and d6 boundaries were not significantly shifted (p^{p3}=0.0280, p^{pMn}=0.3360, p^{d6}=0.2211). Transmembrane Scube2 rescued embryos also showed significant dorsal shift over morpholino alone in position of the pMn boundary, but no statistically significant change in either p3 or d6 upper boundary position (p^{p} =0.2415, p^{pMn}=0.0126, p^{d6}=0.1103). Both ventral progenitor domains show significant shifts in the calculated dorsal boundary in *sc2-cit* rescued embryos versus *sc2-cit-tm* rescued embryos, but no shift in d6 dorsal position (p^{p3}=0.0035, p^{pMn}=0.0154, p^{d6}=0.1444). (J) Mean distributions of *ptch2:kaede* fluorescence in *scube2* morpholino only, *scube2* morpholino rescued with *mr-sc2-cit*, or with *mr-sc2-cit-tm*. Maximum intensity for *mr-scube2-cit*-rescued embryos was found to be significantly higher than those injected with

Figure 2.5 (Continued) morpholino only (p=0.0056). Significant increases in maximum intensity over morpholino treatment alone were also seen in embryos rescued with *mr-sc2-cit-TM* (p=0.0435).

Feedback regulation of Scube2 levels is necessary for pattern scaling

To examine the regulation of Scube2 in size-reduced embryos, we performed our size reduction technique on tg(*scube2:moxng; shha:mem-mcherry*) embryos and imaged them at 20 hours post fertilization. Unlike other observed patterning genes, *scube2* expression levels did not scale in size-reduced embryos but were instead significantly reduced (Figure 2.6A-C). This finding is consistent with an expander-repressor-like model of Scube2-Shh. In this regime, inhibition of *scube2* expression would then contract Shh signaling, enabling adjustment of Shh signaling for a decreased tissue size.

We next examined whether feedback control of *scube2* expression levels by Shh signaling is required for pattern scaling by saturating *scube2* expression. Saturation of *scube2* expression was performed by injecting high doses of *scube2* mRNA into *ptch2:kaede* reporter lines and performing size reduction. If Scube2 is responsible for adjusting Shh signaling during scaling, we would expect Scube2-overexpressing size-reduced embryos to have expanded Shh response gradients compared to controls following normalization for differences in D-V height (Figure 2.6G-J). If scaling of ventral patterning is not dependent on Scube2, we would expect maintenance of pattern scaling with a proportional increase in *ptch2:kaede* distributions. We found that size-reduced Scube2-overexpressing embryos showed a disproportionate expansion of the Ptch signaling gradient compared to normal-sized Scube2-overexpressing embryos (Figure 2.6D-F). Dorsal expansion of Shh signaling is quantified using the position of 50% of maximum control peak height, which is statistically significantly shifted dorsally in size-reduced embryos (Figure 2.6 F). We thus conclude that control of *scube2* expression levels is required for scaling the Shh response gradient, as schematically represented in Figure 2.6G-J.

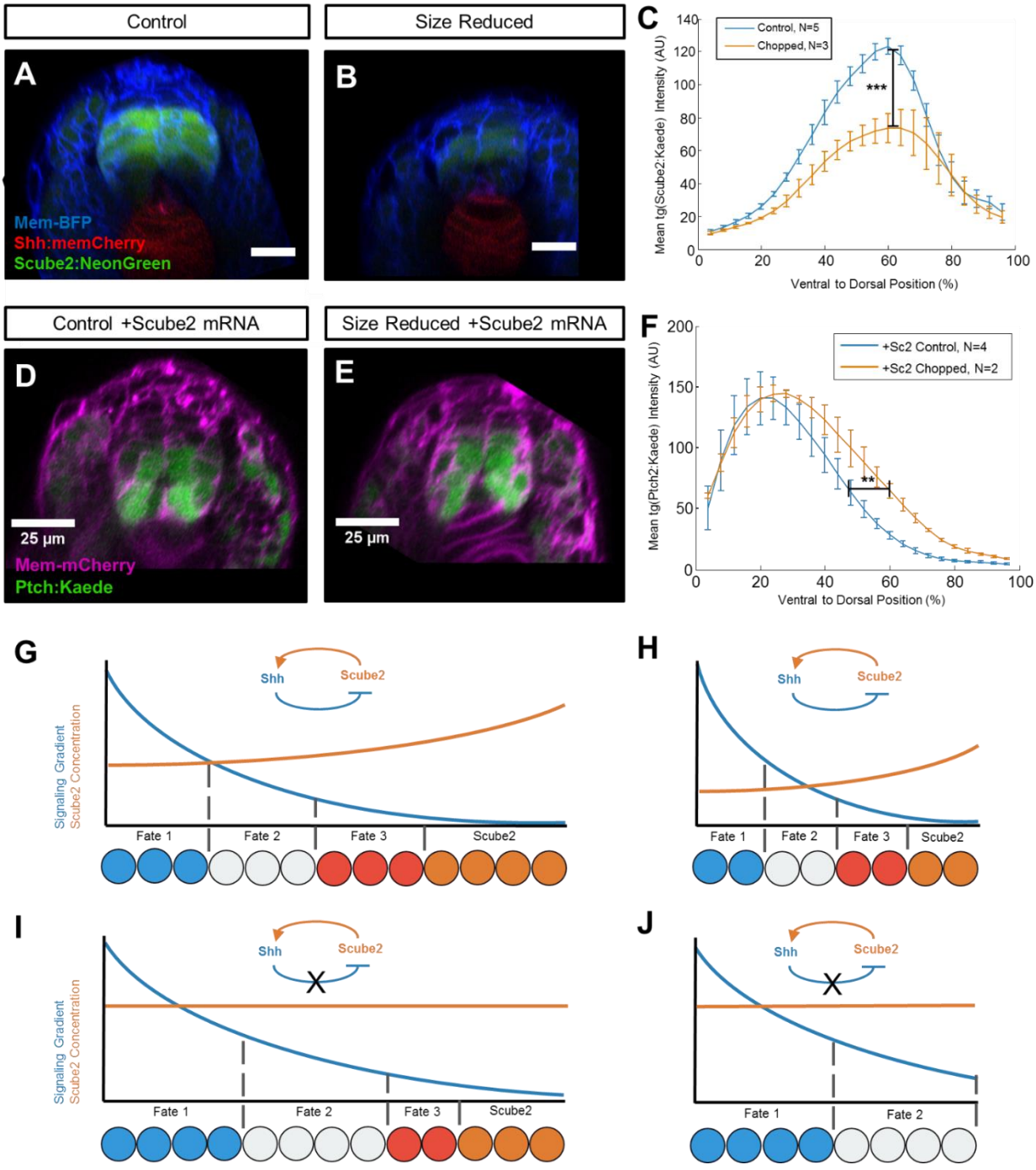


Figure 2.6 Scube2 expression is size-dependent and control of Scube2 levels is required for pattern scaling

(A-B) Transverse view of *mem-mtagbfp2* mRNA-injected tg(*scube2:moxng; shh:mem-mcherry*) control (A) or size-reduced (B) embryos at 20 hpf. (C) Quantification of mean tg(*scube2:moxng*)

Figure 2.6 (Continued) intensity versus ventral-to-dorsal position of embryos as treated in A-B. Maximum intensity values are statistically significantly reduced in treated embryos (p=.00030306). (D-E) Transverse view of 20 hpf tg(*ptch2:kaede*) control (D) and size-reduced (E) embryos injected with *mem-mcherry* and *scube2* mRNA. (F) Quantification of mean tg(*ptch2:kaede*) intensity versus ventral-to-dorsal position of embryos treated as in D-E. Statistically significant shifts are observed in the dorsal position of 50% of the maximum intensity value (p= 0.0092). (G-H) Schematic of expander-repressor-like feedback control of Shh signaling by Scube2 and its ability to enable pattern scaling. Repression of Scube2 by Shh encodes an equilibrium level of Shh signaling across the tissue by linking morphogen spread to tissue size. (I-J) Schematic representation of the experiment as described in D-E, where Scube2 levels are at saturation due to overexpression, and size-reduced embryos (J) are disproportionately affected.

Discussion

Our work uncovers a mechanism by which the morphogen Sonic Hedgehog can self-regulate to enable scale-invariant patterning by linking morphogen signaling to inhibition of Scube2, which is reported to promote Shh release (Creanga et al., 2012; Tukachinsky et al., 2012). We discovered that patterning of the neural tube adjusts to tissue availability following surgical size reduction in zebrafish embryos. Using overexpression experiments we demonstrated that Scube2's activity during patterning is not just permissive—overexpression of Scube2 instead enhances Shh signaling (Woods and Talbot, 2005). Utilizing a transgenic reporter line which we developed, we characterized the expression of Scube2 during neural patterning and found that Shh signaling is responsible for its repression in the ventral neural tube. Using Scube2 fluorescent fusion proteins we found that Scube2 is broadly distributed from secreting cells, explaining its previously reported cell non-autonomous activity (Ben-Zvi and Barkai, 2010; Creanga et al., 2012; Woods and Talbot, 2005). Unlike other patterning genes, Scube2 responds to changes in neural tube height by disproportionately decreasing its expression, and overexpression of Scube2 inhibits scaling of the Shh signaling gradient. Scube2's expression thus can be seen as comparable to the "size-dependent factor" Sizzled, which is thought to enable pattern scaling in early D-V patterning by tuning its expression levels in response to excessive Chordin production (Inomata et al., 2013).

The relationship between Scube2 and Shh has important similarities to proposed "expander-repressor" models of morphogen scaling (Barkai and Ben-Zvi, 2009; Ben-Zvi and Barkai, 2010; Inomata et al., 2013). As with expanders in these models, Scube2 is repressed by Shh signaling, acts cell non-autonomously, and enhances Shh pathway activity. However, Scube2's reported role in morphogen release is distinct from the proposed mechanism of expanders. Expanders-like molecules promote the diffusion or inhibit the degradation of morphogens (Ben-Zvi and Barkai, 2010). While release of Shh ligands from secreting cells would support their transport, the irreversible nature of this effect and local action at the

morphogen source would make distinct predictions for Scube2's effects on morphogen distributions. Nonetheless our study marks the first observation of an expander-repressor-like relationship outside of the BMP/Dpp signaling pathway. This finding raises the possibility that expander-repressor-like relationships may be common motifs in the regulation of morphogen gradients.

However, scaling of neural patterning is unlikely to be achieved by regulation of Shh signaling alone. Graded BMP signaling in the dorsal neural tube is known to pattern dorsal progenitors. Scaling of BMP signaling in neural patterning may be achieved via a similar mechanism to its scaling in early D-V axis patterning. In this system, both existing models propose expander-like relationships between elements of the BMP signaling pathway. The first model proposed ADMP as a scaling related factor, while more recent research has demonstrated that Sizzled has an indispensable role in scaling (Ben-Zvi et al., 2008; Ben-Zvi et al., 2014; Inomata et al., 2013). During neural patterning, the BMP antagonists Noggin, Follistatin, and Chordin are expressed in the notochord while BMP ligands are expressed in the roof plate. Intriguingly, while Sizzled does not seem to be expressed during neural patterning, ADMP is expressed in the notochord and thus may play a role in the scaling of BMP-mediated patterning of the dorsal neural tube (Willot et al., 2002).

BMP signaling is known to increase the thresholds for Shh-dependent cell fate specification, making signaling integration between these pathways a potential candidate regulator of scaling (Liem et al., 2000; McHale et al., 2006). Inhibition of either Shh or BMP signaling causes expansion of signaling by the alternative program. In normal patterning, cells do not measure ratios of BMP and Shh. In fact, Dbx1 positive progenitors in the medial neural tube require little to no Shh or BMP signaling present in order to be specified (Pierani et al., 1999). In addition, recent experiments with precise control of Shh and BMP concentrations in an explant system have shown that cells choose either ventral or dorsal fates in the presence of high BMP and Shh signaling (Zagorski et al., 2017). Regulation of *scube2* expression may be

another way to enable crosstalk between signaling pathways, as Scube2 is not expressed in the dorsal most cells of the spinal cord, suggesting repression by dorsal factors. Specification of the dorsal boundary of *scube2* expression may encode yet more information about the size of the tissue which would then affect Shh spread.

We began this work in part due to interest in the discrepancy between the area of Scube2's activity in the ventral neural tube and its expression exclusively in the dorsal neural tube. Our work with Scube2 fluorescent protein fusions revealed that Scube2 is highly diffusive and is distributed broadly from producing cells. Scube2's diffusion from producing cells could easily account for the distance between its expression domain and area of effect. Taken together with Scube2's known binding of Shh and release-promoting activity, this raises the possibility that Scube2 may serve as a chaperone for Shh during its transport (Tukachinsky et al., 2012). Our transmembrane Scube2 rescue experiments may help answer this question. In that experiment we ubiquitously over expressed transmembrane Scube2 and found that it was only partially capable of rescuing Shh signaling. If Scube2 was required only at the surface of secreting cells, we would have expected full rescue of Shh signaling in embryos rescued by transmembrane Scube2. These findings cast some doubt on the hypothesized role of Scube2 in promoting Shh lipid shedding to enable its release (Jakobs et al., 2014; Jakobs et al., 2016).

Scube2 is one of several recently identified elements of the Shh signaling pathway that exerts cell non-autonomous effects. Recent work has shown that Hhip—initially characterized as a membrane-tethered hedgehog antagonist—acts over a long range that cannot be explained by ligand sequestration (Kwong et al., 2014). Additionally, the Hedgehog receptor, Patched, may also have cell non-autonomous inhibitory effects on Smoothened through regulating inhibitory sterols or sterol availability (Bidet et al., 2011; Roberts et al., 2016). Together with known feedback relationships and the diffusivity of Scube2 that we demonstrated here, these mechanisms interlink Shh signaling between neighboring cells and may enable tissue level properties, such as the scaling of pattern formation we observed.

Methods

Generation of Transgenic Lines

The construct used to make tg(*scube2:moxng*) was generated by isothermal assembly of PCR-amplified *scube2* regulatory elements obtained from the CHORI-211 BAC library. Regulatory elements were in part chosen based on annotations of H3K4me1 and H3K4me3 binding (Aday et al., 2011). Selected regulatory sequences spanned 1677bp of upstream intergenic sequence and 5962bp of the area spanning exons 1-5 *scube2*. Regulatory sequences were cloned into a pMT backbone by placing a zebrafish codon-optimized moxNeonGreen fluorescent protein and sv40 poly-A tail just downstream of the endogenous *scube2* Kozak sequence (Costantini et al., 2015). The construct used to make tg(*scube2:scube2-moxng*) was generated using the same regulatory sequences as tg(*scube2:moxng*), with the addition of exons 6-23 of the Scube2 coding sequence downstream of exon 5 and moxNeonGreen attached at the c-terminus with a 10 amino acid long GA rich linker. The construct used to make tg(*shh:mem-mcherry*) was derived from the previously reported tg(*shh:gfp*), by replacement of GFP with mem-mCherry (Megason, 2009; Shkumatava et al., 2004).

Transgenic lines were generated by injecting plasmid DNA for each construct along with Tol2 mRNA into wild type (AB) embryos at the single cell stage, as described previously (Kawakami, 2004). moxNeonGreen positive embryos were then selected for raising. Upon reaching sexual maturity, F0s were outcrossed and screened for founders. Founders were isolated and raised as single alleles. Monoallelic versions of each line are shown throughout the paper.

Zebrafish Strains

For wild type lines, AB fish were used. All fish were kept at 28°C on a 14-hour-light/10-hour-dark cycle. Embryos were collected from natural crosses. All fish-related procedures were carried out with the approval of Institutional Animal Care and Use Committee (IACUC) at Harvard University. tgBAC(*ptch2:kaede*) (Huang et al., 2012; renamed from *ptch1* due to a change in zebrafish gene nomenclature), tg(*nkx2.2a:mgfp*) (Jessen et al., 1998), tg(*olig2:gfp*) (Shin et al., 2003), tg(*olig2:dsRed*) (Kucenas et al., 2008), and tgBAC(*dbx1b:GFP*) (Kinkhabwala et al., 2011) have been described previously.

Size Reduction Technique.

Size reduction was performed as described in our previous report (Ishimatsu et al., 2017). Embryo sizes were reduced by sequentially removing ~1/3 of the cells from the animal cap, then wounding the yolk. These surgeries are performed in 1/3 ringers solution, and embryos are immobilized in a thin layer of 2% methyl cellulose. Surgery is performed either with glass needles – as previously described – or using a loop of thin stainless-steel wire that is inserted through a glass capillary tube and mounted on a halved chopstick. Healthy uninjected embryos show a maximum success rate of ~60% while embryos which have undergone injection or from relatively inbred or older females have significantly lower success rates. In each size reduction experiment, embryos are selected for analysis by their health and reduced size; those with morphological defects are discarded.

Construct Generation and Injections of mRNAs and Morpholinos

Scube2-Citrine, MR-Scube2-Citrine, and MR-Scube2-Citrine were all generated using cDNA obtained from the Talbot lab (Woods and Talbot, 2005). Fluorescent protein fusions were made by attaching Citrine or moxNeonGreen with a 10 amino acid GA rich linker to the c-terminus of Scube2. Membrane-mTagBFP2 constructs were generated using membrane localization tags

reported previously (Megason, 2009; Subach et al., 2011). These constructs were each sub-cloned into a pMTB backbone. mRNA for all experiments was synthesized from pCS or pMTB backbones via *in vitro* transcription using the mMESSAGE mMACHINE system (Ambion). Embryos were injected at the single cell stage using a Nanoject system set to 2.3nl of injection volume containing 92pg of RNA for each mRNA injected, unless otherwise specified. Injected embryos were then screened for brightness, and damaged embryos were removed. Scube2 morpholino injections were performed with 7ng of Scube2 MO2 and 3.5ng of p53 MO to control for phenotypic variability (Gerety and Wilkinson, 2011; Woods and Talbot, 2005).

Cyclopamine Treatment

Cyclopamine was dissolved in 100% ethanol to make 50mM stock solution and was diluted for treatment in egg water to 100 uM. Equal amount ethanol as used to suspend cyclopamine was used for controls. Treatment began at 7 hpf and continued until imaging at 18-20 hpf.

Confocal Imaging

For quantitative imaging, embryos were staged and mounted in our previously described dorsal mount (Kimmel et al., 1995; Megason, 2009; Xiong et al., 2013) in egg water with 0.01% tricaine (Western Chemical, Inc.). Embryos were manipulated for proper positioning with hair loops, before gently lowering the coverslip. Embryos were not depressed by the coverslip or impinged by the mold, enabling imaging of their normal proportions. Imaging was performed on embryos staged at 18-22 hpf, unless otherwise noted in corresponding figure legends. Live imaging was performed using a Zeiss 710 confocal/2-photon microscope, Zen image acquisition software, and C-Apochromat 40X 1.2 NA objective. For fluorescent protein excitation, 405 nm (BFP), 488 nm (GFP/moxNeonGreen), 514 nm (Citrine), 561 nm (mCherry/dsRed) and 594 nm (mCardinal) lasers were used. The imaging field was centered in each embryo on the somite 6/7 boundary for consistent positioning between images. For quantitative analysis, imaging datasets are only

compared between sibling embryos imaged on the same day with the same settings. This approach aims to avoid clutch effects or variability in detector sensitivity and laser power that occur over time. Typical imaging settings with the 40x objective were as follows: image size of 1024x1024 pixels with .21um per pixel and an interval of 1um in the Z direction. For display purposes, images are rendered in cross sectional views (X-Z axis) which are then rotated for display, with image intensities for co-injection markers adjusted evenly within datasets for brightness. Imaging for FRAP, early stage embryo imaging, and the tg(*sc2:moxng;shh:mem-mcherry*) time-lapse were performed using a 1.0 NA 20x objective. Brightfield and widefield fluorescence images of whole embryos were obtained using an Olympus MVX10 and a Leica MZ12.5 dissecting microscope.

Image Analysis

Images were analyzed using a custom MATLAB-based image analysis software that enables rapid segmentation of neural tube imaging data. Neural imaging data is segmented by the user sequentially from anterior to posterior. Over a set step size (usually 50 pixels), the user selects points at the base of the floor plate cell and top of the roof plate cell that divide the neural tube into its two halves (Figure S2.1A). The user then selects the widest point of the neural tube in each image. Imaging data from mature neurons, found laterally, and within the lumen of the neural tube, found medially, are disregarded using a set percentage of neural width (Figure S2.1B). Once these positions are recorded, imaging data is then recovered as average pixel intensity in 25 bins from ventral to dorsal. This binning and averaging strategy enables comparison of data between embryos that accounts for variations in neural tube D-V height. During the segmentation process, the researcher is blinded to the title of the dataset which contains information about its treatment condition. For distribution plots, binned intensities are reported for each embryo as the average intensities for each bin along the entire AP axis. Each

embryo's average intensity profile is then treated as an independent sample and averaged for displayed distribution profiles and standard deviations.

Progenitor domain segmentation is performed on average intensities profiles from each embryo in a dataset in the following manner: first, all intensity profiles in the data set undergo background subtraction and if *nkx2.2* signal intensities are too low, .95 gamma is applied uniformly across the dataset to enable automated segmentation. Intensity profiles are then fed to a peak finding algorithm to identify local maxima. Both *dbx1b*+ and *nkx2.2a*+ progenitor domains are found in the green channel, so a maximum of two peaks is allowed. In the red channel, only one peak is specified to identify *olig2:dsred* signal. Average peak intensity values for each domain are then calculated for the entire control dataset, and 50% of this value in the case of the *nkx2.2a* and *dbx1b* domains is used as the threshold for calculating domain width. Given its greater spread along the D-V axis, a threshold of 25% of peak height is used in calculating width of the *olig2*+ domain. Domain widths are then extracted from spline-interpolated intensity profiles to avoid errors introduced by rounding to the next bin. Segmented widths and positions of *nkx2.2a*, *olig2*, and *dbx1b* expression are then averaged. Domain plots are generated by assigning all *nkx2.2a*+ progenitors to the p3 fate, *olig2*+ progenitors lacking *nkx2.2a* expression overlap to the pMn fate, and *dbx1b*+ progenitors to the p0-d6 fate. These domain sizes and positions are then used to reconstruct domains in-between or flanking them, which include the p2-p1 domain between pMn and p0-d6, the floorplate below p3, and the d5-roofplate above p0-d6. These heights and positions are then used to generate the stacked bar plots shown. Occasionally, expansion of the "floorplate" domain is observed in stacked bar plots, this is due to the lowered intensity of *nkx2.2a*:*memgfp* expression and not likely expansion of actual floor plate cells.

Statistical Analysis

Statistical comparisons of maximum average intensity and position of 50% maximum intensity are performed by an unpaired T-test. Although each dataset contains hundreds of measurements of each binned intensity value over the A-P axis of a z-stack, only the average of these measurements for each embryo is treated as a data-point for calculation of the standard deviation and statistical significance tests. This is done to avoid oversampling that would exaggerate statistical significance. In all measurements, statistical significance is markedly increased if analysis is performed by treating all underlying intensity measurements as samples. Thresholds for calculating position of half maximum are determined from the average maximum of the corresponding control dataset for each experiment. Position is then determined from the fitted trend-line to avoid inaccuracies due to rounding. To calculate the significance of shifts in boundary positions, upper domain boundaries for each embryo were compared in an un-paired t-test between embryos from each population. When the progenitor domain segmentation algorithm finds there is no domain present, the boundary is set to 0.

CRISPR Screen for Scube2 Regulators

Candidate gene lists were generated by a review of transcription factors known to be important in patterning of the neural tube near the boundaries of *scube2* expression. In addition, *znf362b* and *sox21b* were targeted based on their correlation with *scube2* expression from single cell sequencing data (Chapter III). Cas9 protein was generated and purified in lab. Three guide RNA sequences targeting the first one-to-three exons of each gene were selected based on their quality using the web-tool CHOP-CHOP and synthesized using standard methods (Gagnon et al., 2014). Equivalent guide RNA and Cas9 protein concentrations were used in all samples for mosaic knockout. Phenotypes were assessed at 18-20 hpf by confocal microscopy.

Acknowledgments

We thank Lisa Goodrich, Rosalind Segal, and Wolfram Goessling for their comments and helpful discussion. The Scube2 construct was a gift from the Talbot lab.

Competing interest

The authors declare no competing or financial interests.

Funding

Z.M.C was supported in part by the program in Biological and Biomedical Sciences at Harvard University. S.G.M. and Z.M.C. were supported by R01-GM107733 and R01-DC015478.

References

Aday, A. W., Zhu, L. J., Lakshmanan, A., Wang, J. and Lawson, N. D. (2011). Identification of cis regulatory features in the embryonic zebrafish genome through large-scale profiling of H3K4me1 and H3K4me3 binding sites. *Dev. Biol.* 357, 450–462.

Averbukh I, Ben-Zvi D, Mishra S, Barkai N, Averbukh, I., Ben-Zvi, D., Mishra, S. and Barkai, N. (2014). Scaling morphogen gradients during tissue growth by a cell division rule. *Development* 141, 2150–6.

Barkai, N. and Ben-Zvi, D. (2009). “Big frog, small frog”--maintaining proportions in embryonic development: delivered on 2 July 2008 at the 33rd FEBS Congress in Athens, Greece. *FEBS J.* 276, 1196–207.

Ben-Zvi, D. and Barkai, N. (2010). Scaling of morphogen gradients by an expansion-repression integral feedback control. *Proc. Natl. Acad. Sci. U. S. A.* 107, 6924–9.

Ben-Zvi, D., Shilo, B.-Z., Fainsod, A. and Barkai, N. (2008). Scaling of the BMP activation gradient in Xenopus embryos. *Nature* 453, 1205–11.

Ben-Zvi, D., Pyrowolakis, G., Barkai, N. and Shilo, B.-Z. (2011). Expansion-repression mechanism for scaling the Dpp activation gradient in Drosophila wing imaginal discs. *Curr. Biol.* 21, 1391–6.

Ben-Zvi, D., Fainsod, A., Shilo, B.-Z. and Barkai, N. (2014). Scaling of dorsal-ventral patterning in the Xenopus laevis embryo. *Bioessays* 36, 151–6.

Bidet, M., Joubert, O., Lacombe, B., Ciantar, M., Nehmé, R., Mollat, P., Brétillon, L., Faure, H., Bittman, R., Ruat, M., et al. (2011). The Hedgehog Receptor Patched Is Involved in Cholesterol Transport. *PLoS One* 6, e23834.

Briscoe, J., Chen, Y., Jessell, T. M. and Struhl, G. (2001). A hedgehog-insensitive form of patched provides evidence for direct long-range morphogen activity of sonic hedgehog in the neural tube. *Mol. Cell* 7, 1279–91.

Burke, R., Nellen, D., Bellotto, M., Hafen, E., Senti, K.-A. A., Dickson, B. J. and Basler, K. (1999). Dispatched, a Novel Sterol-Sensing Domain Protein Dedicated to the Release of Cholesterol-Modified Hedgehog from Signaling Cells. *Cell* 99, 803–815.

Cao, Y., Ryser, M. D., Payne, S., Li, B., Rao, C. V. and You, L. (2016). Collective Space-Sensing Coordinates Pattern Scaling in Engineered Bacteria. *Cell* 165, 620–630.

Chen, M.-H., Li, Y.-J., Kawakami, T., Xu, S.-M. and Chuang, P.-T. (2004). Palmitoylation is required for the production of a soluble multimeric Hedgehog protein complex and long-range signaling in vertebrates. *Genes Dev.* 18, 641–59.

Cooke, J. (1981). Scale of body pattern adjusts to available cell number in amphibian embryos. *Nature* 290, 775–778.

Costantini, L. M., Baloban, M., Markwardt, M. L., Rizzo, M., Guo, F., Verkhusha, V. V. and Snapp, E. L. (2015). A palette of fluorescent proteins optimized for diverse cellular environments. *Nat. Commun.* 6, 7670.

Creanga, A., Glenn, T. D., Mann, R. K., Saunders, A. M., Talbot, W. S. and Beachy, P. A. (2012). Scube/You activity mediates release of dually lipid-modified Hedgehog signal in soluble form. *Genes Dev.* 26, 1312–25.

Driesch, H. (1892). Entwicklungsmechanische Studien: I. Der Werthe der beiden ersten Furchungszellen in der Echinogdermenentwicklung. Experimentelle Erzeugung von Theil- und Doppelbildungen. *Zeitschrift für wissenschaftliche Zool.*

Francois, P., Vonica, A., Brivanlou, A. H. and Siggia, E. D. (2009). Scaling of BMP gradients in Xenopus embryos. *Nature* 461, E1–E1.

Gagnon, J. A., Valen, E., Thyme, S. B., Huang, P., Ahkmetova, L., Pauli, A., Montague, T. G., Zimmerman, S., Richter, C., Schier, A. F., et al. (2014). Efficient Mutagenesis by Cas9 Protein-Mediated Oligonucleotide Insertion and Large-Scale Assessment of Single-Guide RNAs. *PLoS One* 9, e98186.

Gerety, S. S. and Wilkinson, D. G. (2011). Morpholino artifacts provide pitfalls and reveal a novel role for pro-apoptotic genes in hindbrain boundary development. *Dev. Biol.* 350, 279–89.

Gribble, S. L., Nikolaus, O. B. and Dorsky, R. I. (2007). Regulation and function of Dbx genes in the zebrafish spinal cord. *Dev. Dyn.* 236, 3472–83.

Grimmond, S., Larder, R., Van Hateren, N., Siggers, P., Morse, S., Hacker, T., Arkell, R. and Greenfield, A. (2001). *Expression of a novel mammalian epidermal growth factor-related gene during mouse neural development*. Elsevier.

Hamaratoglu, F., de Lachapelle, A. M., Pyrowolakis, G., Bergmann, S. and Affolter, M. (2011). Dpp signaling activity requires Pentagone to scale with tissue size in the growing Drosophila wing imaginal disc. *PLoS Biol.* 9, e1001182.

Hollway, G. E., Maule, J., Gautier, P., Evans, T. M., Keenan, D. G., Lohs, C., Fischer, D., Wicking, C. and Currie, P. D. (2006). Scube2 mediates Hedgehog signalling in the zebrafish embryo. *Dev. Biol.* 294, 104–18.

Inomata, H., Shibata, T., Haraguchi, T. and Sasai, Y. (2013). Scaling of dorsal-ventral patterning by embryo size-dependent degradation of Spemann's organizer signals. *Cell* 153, 1296–311.

Ishimatsu, K., Hiscock, T. W., Collins, Z. M., Sari, D. W. K., Lischer, K., Richmond, D. L., Bessho, Y., Matsui, T. and Megason, S. G. (2017). Size-reduced embryos reveal a gradient scaling based mechanism for zebrafish somite formation. *bioRxiv* 211615.

Jakobs, P., Exner, S., Schürmann, S., Pickhinke, U., Bandari, S., Ortmann, C., Kupich, S., Schulz, P., Hansen, U., Seidler, D. G., et al. (2014). Scube2 enhances proteolytic Shh processing from the surface of Shh-producing cells. *J. Cell Sci.* 127, 1726–37.

Jakobs, P., Schulz, P., Ortmann, C., Schürmann, S., Exner, S., Rebollido-Rios, R., Dreier, R., Seidler, D. G. and Grobe, K. (2016). Bridging the gap: heparan sulfate and Scube2 assemble Sonic hedgehog release complexes at the surface of producing cells. *Sci. Rep.* 6, 26435.

Jessen, J. R., Meng, A., McFarlane, R. J., Paw, B. H., Zon, L. I., Smith, G. R. and Lin, S. (1998). Modification of bacterial artificial chromosomes through Chi-stimulated homologous recombination and its application in zebrafish transgenesis. *Proc. Natl. Acad. Sci.* 95, 5121–5126.

Kawakami, K. (2004). Transgenesis and Gene Trap Methods in Zebrafish by Using the Tol2 Transposable Element.pp. 201–222.

Kawakami, T., Kawcak, T., Li, Y.-J., Zhang, W., Hu, Y. and Chuang, P.-T. (2002). Mouse dispatched mutants fail to distribute hedgehog proteins and are defective in hedgehog signaling. *Development* 129, 5753–65.

Kawakami, A., Nojima, Y., Toyoda, A., Takahoko, M., Satoh, M., Tanaka, H., Wada, H., Masai, I., Terasaki, H., Sakaki, Y., et al. (2005). The zebrafish-secreted matrix protein you/scube2 is implicated in long-range regulation of hedgehog signaling. *Curr. Biol.* 15, 480–8.

Kicheva, A., Bollenbach, T., Ribeiro, A., Valle, H. P., Lovell-Badge, R., Episkopou, V. and Briscoe, J. (2014). Coordination of progenitor specification and growth in mouse and chick spinal cord. *Science (80-.).* 345, 1254927–1254927.

Kimmel, C. B., Ballard, W. W., Kimmel, S. R., Ullmann, B. and Schilling, T. F. (1995). Stages of embryonic development of the zebrafish. *Dev. Dyn.* 203, 253–310.

Kinkhabwala, A., Riley, M., Koyama, M., Monen, J., Satou, C., Kimura, Y., Higashijima, S.-I. and Fetcho, J. (2011). A structural and functional ground plan for neurons in the hindbrain of zebrafish. *Proc. Natl. Acad. Sci. U. S. A.* 108, 1164–9.

Kucenas, S., Takada, N., Park, H.-C., Woodruff, E., Broadie, K. and Appel, B. (2008). CNS-derived glia ensheath peripheral nerves and mediate motor root development. *Nat. Neurosci.* 11, 143–151.

Kwong, L., Bijlsma, M. F. and Roelink, H. (2014). Shh-mediated degradation of Hhip allows cell autonomous and non-cell autonomous Shh signalling. *Nat. Commun.* 5, 4849.

Liem, K. F., Jessell, T. M. and Briscoe, J. (2000). Regulation of the neural patterning activity of sonic hedgehog by secreted BMP inhibitors expressed by notochord and somites. *Development* 127, 4855–4866.

McHale, P., Rappel, W.-J. and Levine, H. (2006). Embryonic pattern scaling achieved by oppositely directed morphogen gradients. *Phys. Biol.* 3, 107–120.

Megason, S. G. (2009). In toto imaging of embryogenesis with confocal time-lapse microscopy. *Methods Mol. Biol.* 546, 317–32.

Morgan, T. H. (1895). Half embryos and whole embryos from one of the first two blastomeres. *Anat. Anz.* 10, 623–638.

Pepinsky, R. B., Zeng, C., Wen, D., Rayhorn, P., Baker, D. P., Williams, K. P., Bixler, S. A., Ambrose, C. M., Garber, E. A., Miatkowski, K., et al. (1998). Identification of a Palmitic Acid-modified Form of Human Sonic hedgehog. *J. Biol. Chem.* 273, 14037–14045.

Pierani, A., Brenner-Morton, S., Chiang, C. and Jessell, T. M. (1999). A sonic hedgehog-independent, retinoid-activated pathway of neurogenesis in the ventral spinal cord. *Cell* 97, 903–15.

Porter, J. A., Ekker, S. C., Park, W.-J., von Kessler, D. P., Young, K. E., Chen, C.-H., Ma, Y., Woods, A. S., Cotter, R. J., Koonin, E. V, et al. (1996a). Hedgehog Patterning Activity: Role of a Lipophilic Modification Mediated by the Carboxy-Terminal Autoprocessing Domain. *Cell* 86, 21–34.

Porter, J. A., Young, K. E. and Beachy, P. A. (1996b). Cholesterol Modification of Hedgehog Signaling Proteins in Animal Development. *Science (80-.).* 274, 255–259.

Roberts, B., Casillas, C., Alfaro, A. C., Jägers, C. and Roelink, H. (2016). Patched1 and Patched2 inhibit Smoothened non-cell autonomously. *Elife* 5,.

Shin, J., Park, H.-C., Topczewska, J. M., Mawdsley, D. J. and Appel, B. (2003). Neural cell fate analysis in zebrafish using olig2 BAC transgenics. *Methods Cell Sci.* 25, 7–14.

Shkumatava, A., Fischer, S., Müller, F., Strahle, U. and Neumann, C. J. (2004). Sonic hedgehog, secreted by amacrine cells, acts as a short-range signal to direct differentiation and lamination in the zebrafish retina. *Development* 131, 3849–58.

Spemann, H. (1938). *Embryonic Development and Induction*. Yale Univ.; New Haven:

Subach, O. M., Cranfill, P. J., Davidson, M. W. and Verkhusha, V. V. (2011). An Enhanced Monomeric Blue Fluorescent Protein with the High Chemical Stability of the Chromophore. *PLoS One* 6, e28674.

Tukachinsky, H., Kuzmickas, R. P. P., Jao, C. Y. Y., Liu, J. and Salic, A. (2012). Dispatched and Scube Mediate the Efficient Secretion of the Cholesterol-Modified Hedgehog Ligand. *Cell Rep.* 2, 308–320.

Umulis, D. M. and Othmer, H. G. (2013). Mechanisms of scaling in pattern formation. *Development* 140, 4830–43.

Uygur A, Young J, Huycke TR, Koska M, Briscoe J, Tabin CJ. (2016). Scaling Pattern to Variations in Size during Development of the Vertebrate Neural Tube. *Dev Cell* 37(2):127–135.

van Eeden, F. J. J., Granato, M., Schach, U., Brand, M., Furutani-Seiki, M., Haffter, P., Hammerschmidt, M., Heisenberg, C. P. P., Jiang, Y. J. J., Kane, D. A. A., et al. (1996). Mutations affecting somite formation and patterning in the zebrafish, Danio rerio. *Development* 123, 153–164.

Willot, V., Mathieu, J., Lu, Y., Schmid, B., Sidi, S., Yan, Y.-L., Postlethwait, J. H., Mullins, M., Rosa, F. and Peyriéras, N. (2002). Cooperative Action of ADMP- and BMP-Mediated Pathways in Regulating Cell Fates in the Zebrafish Gastrula. *Dev. Biol.* 241, 59–78.

Woods, I. G. and Talbot, W. S. (2005). The you gene encodes an EGF-CUB protein essential for Hedgehog signaling in zebrafish. *PLoS Biol.* 3, e66.

Xiong, F., Tentner, A. R., Huang, P., Gelas, A., Mosaliganti, K. R., Souhait, L., Rannou, N., Swinburne, I. a, Obholzer, N. D., Cowgill, P. D., et al. (2013). Specified neural progenitors sort to form sharp domains after noisy Shh signaling. *Cell* 153, 550–61.

Zagorski, M., Tabata, Y., Brandenberg, N., Lutolf, M. P., Tkačik, G., Bollenbach, T., Briscoe, J. and Kicheva, A. (2017). Decoding of position in the developing neural tube from antiparallel morphogen gradients. *Science* 356, 1379–1383.

Zeng, X., Goetz, J. A., Suber, L. M., Scott, W. J., Schreiner, C. M. and Robbins, D. J. (2001). A freely diffusible form of Sonic hedgehog mediates long-range signalling. *Nature* 411, 716–20.

Chapter III.

Systematic mapping of cell state trajectories, cell lineage, and perturbations in the zebrafish embryo using single cell transcriptomics

Authors: Daniel E. Wagner[1], Caleb Weinreb[1], Zachary M. Collins[1], James A. Briggs[1], Sean G. Megason[1]*, Allon M. Klein[1]*

Affiliations:

[1]Department of Systems Biology, Harvard Medical School, Boston, MA 02115, USA

*Correspondence to: sean_megason@hms.harvard.edu or allon_klein@hms.harvard.edu

Author Contributions: D.E.W. designed and performed all single-cell experiments, processed and analyzed data, and generated figures. D.E.W., A.M.K., and S.G.M. conceived the study and wrote the manuscript with input from other authors. C.W. generated the web portal for viewing single-cell and coarse-grained graphs. Z.M.C designed sgRNA sequences and performed CRISPR/Cas9 injections. J.A.B. shared access to unpublished single-cell data.

Abstract

High-throughput mapping of cellular differentiation hierarchies from single-cell data promises to empower systematic interrogations of vertebrate development and disease. Here, we applied single-cell RNA sequencing to >92,000 cells from zebrafish embryos during the first day of development. Using a graph-based approach, we mapped a cell state landscape that describes axis patterning, germ layer formation, and organogenesis. We tested how clonally related cells traverse this landscape by developing a transposon-based barcoding approach ("TracerSeq") for reconstructing single-cell lineage histories. Clonally related cells were often restricted by the state landscape, including a case in which two independent lineages converge on similar fates. Cell fates remained restricted to this landscape in *chordin*-deficient embryos. We provide web-based resources for further analysis of the single-cell data.

Main Text

A major goal of developmental biology is to understand the progression of embryonic cell lineages from pluripotency to adulthood (Schier and Talbot, 2005). Fate mapping, and analysis of mutant phenotypes, have explained much of what we know of development, yet we still lack a systematic atlas of all cell states in a developing embryo. Owing to technical advances in single-cell RNA sequencing (scRNA-seq) (Hashimshony et al., 2016, Islam et al., 2014, Klein et al., 2015, Macosko et al., 2015, Mazutis et al., 2013) it is now possible to assemble comprehensive single-cell atlases describing complex and dynamic *in vivo* biological processes. Here, we utilized inDrops scRNA-seq (Klein et al., 2015, Zilionis et al., 2017) to collect over 92,000 single-cell transcriptomes from dissociated wild-type and mutant zebrafish embryos during the first 24 hours of embryonic development (Figure 3.1A-B and S3.1). For different developmental stages, we sampled 0.17x to 0.97x of the total cells per embryo, sufficient to detect cell states as rare as 0.1-0.5% of all cells (Figure S3.1C), including germ cells which were detected in all timepoints (Figure 3.1B and Table S3.2). From this dataset,

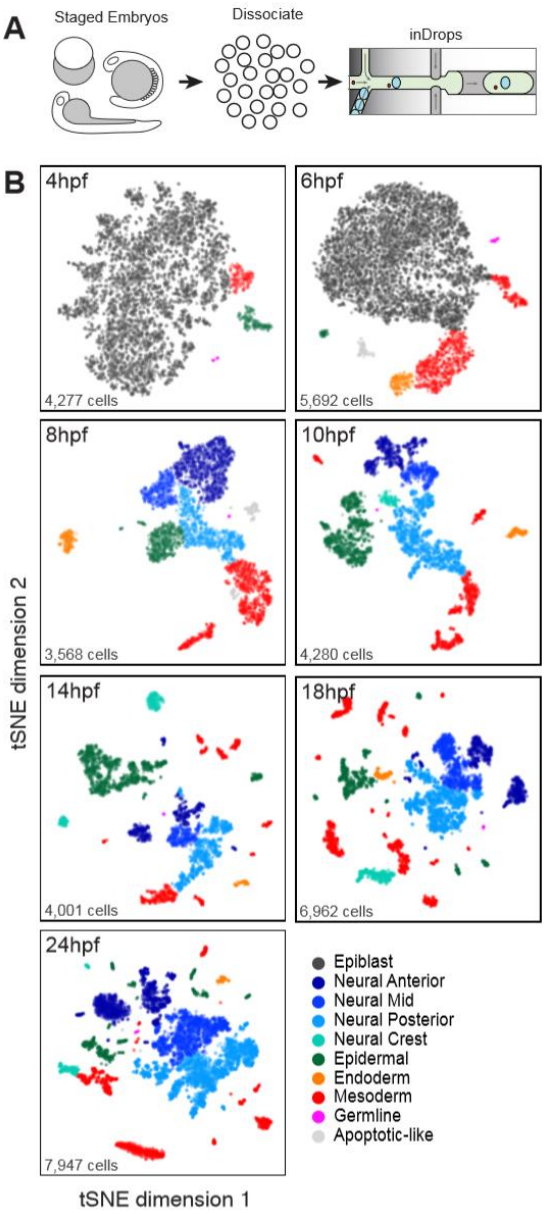


Figure 3.1. A single-cell transcriptional atlas of the zebrafish embryo.

(A) Experimental workflow: Single-cell suspensions were dissociated from staged zebrafish embryos and introduced into the inDrops microfluidic device. Single-cell transcriptome libraries were prepared and sequenced by RNA-seq. (B) tSNE maps for each timepoint, constructed in dimensionality-reduced PCA subspace defined by highly co-variable genes (see Methods). Cells are colored by germ layer identities inferred from expressed marker genes (see also Figure S3.2A and Table S3.2).

clustering of the wild-type transcriptomes revealed an expanding set of epidermal, neural, mesodermal, and endodermal cell states over developmental time, many of which could be specifically annotated based on expression of marker genes (Figure 3.1B, S3.2A and Table S3.2) (Thisse et al., 2001). We collected seven biological replicates for the final timepoint (24 hours post-fertilization, hpf) which demonstrated consistency of both transcriptional signatures and cell state proportions across independent specimens (Figure S3.2B-C).

A single-cell graph of cell state progression in the developing zebrafish embryo

We sought to map trajectories of cell state during development by linking cell states across time. Several computational approaches exist to infer orderings of asynchronous processes from scRNA-seq data (Trapnell et al., 2014, Bendall et al., 2014, Shin et al., 2015), typically by projecting all cells into a single low-dimensional latent space. Such strategies may be ill-suited to map gene expression in developing embryos, which exhibit dramatically increasing cell state dimensionality and continuous changes in the sets and numbers of cell state-defining genes (Figure S3.2D-E). To overcome these obstacles, we developed a graph-based strategy for locally embedding consecutive timepoints on the basis of biological variation that they share, rather than using a global coordinate system for all timepoints. This approach first constructs a single-cell k-nearest-neighbor graph for each timepoint t_i, with nodes representing cells and edges linking neighbors in a low-dimensional subspace; it then joins the graphs by identifying neighboring cells in pairs of adjacent time points, using a coordinate system learned from the future (t_{i+1}) timepoint (see Methods). The resulting graph spans all time points, and allows application of formal graph-based methods for data analysis. When applied to our zebrafish data, the full graph forms a branching network (Figure 3.2A). Inspection of numerous domain and cell-type specific transcriptional markers shows that major initial branches represent neural, epidermal, and mesendodermal states undergoing progressive and

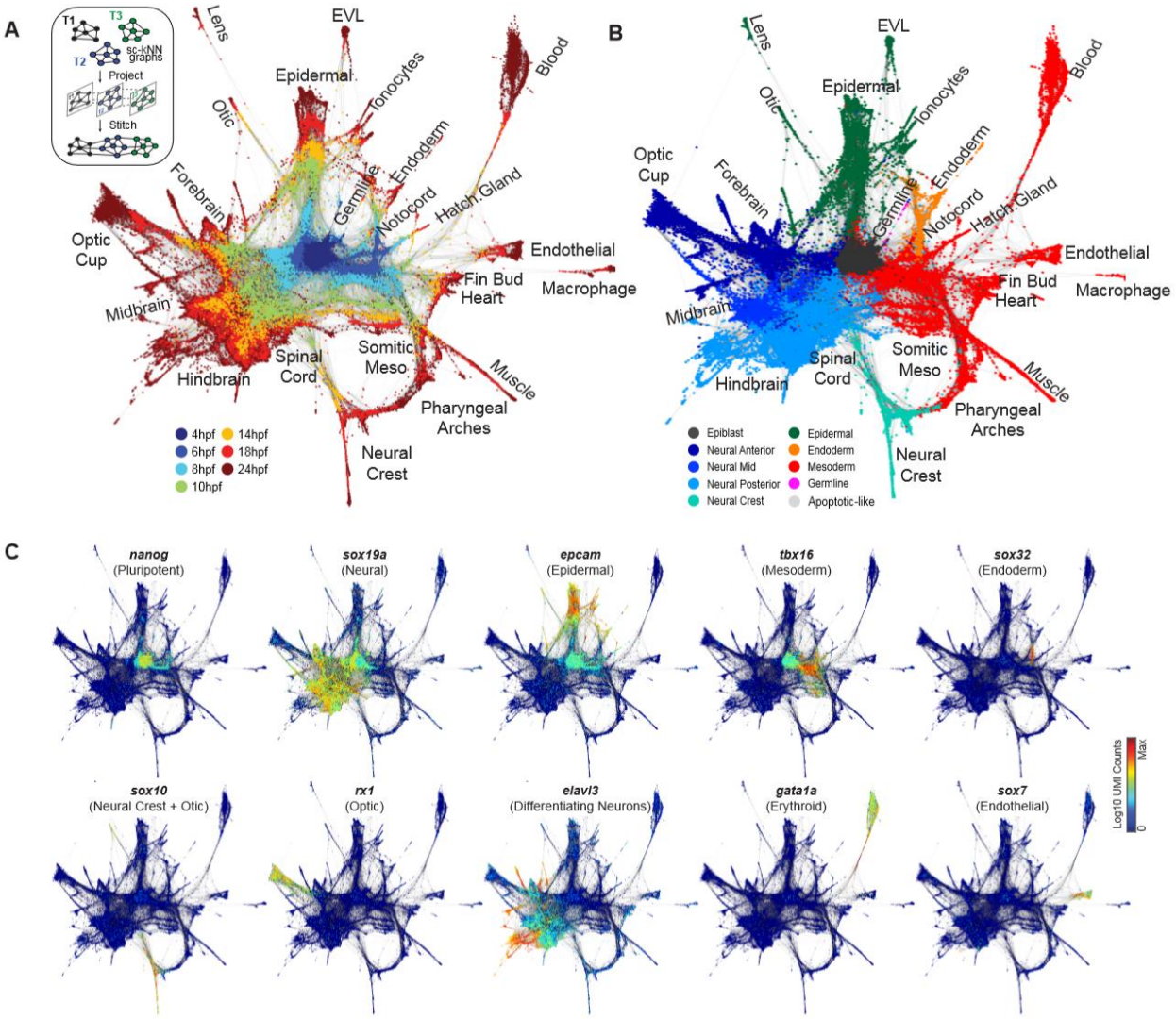


Figure 3.2. Single-cell graph reveals a continuous developmental landscape of cell states.

(A) Overview of graph construction strategy, and a force-directed layout of the resulting single-cell graph (nodes colored by collection timepoint). For each cell, up to 20 within- or between-timepoint mutual nearest neighbor edges are retained. (B) Single-cell graph, colored by germ layer identities inferred from differentially expressed marker genes (see Table S3.2). (C) Single-cell graphs, colored by log10 expression counts for indicated cell type-specific marker genes.

spatially restricted differentiation (Figure 3.2B-C, S3.3). We also noted distinct and early branching events for germline, notochord, enveloping layer (EVL) epidermis, and the prechordal plate.

To test whether this graph recapitulates known lineage relationships, we used a measure of graph distance (Diffusion Pseudotime or "DPT") (Haghverdi et al., 2016) to explore long-range temporal connections between cell states. Cell states of the early gastrula (shield stage, 6hpf) are defined largely by positional marker genes (Figure 3.3A), yet these cells are connected—through the single-cell graph—to tissue-specific states that emerge later (e.g. pharyngula stage, 24hpf). We found that the shield stage cells with the shortest mean graph distance to each particular 24hpf tissue were clustered, and expressed spatial marker genes predicted from previous *in vivo* fate mapping studies (Kimmel et al., 1990, Melby et al., 1996, Warga and Nusslein-Volhard, 1999, Woo and Fraser, 1995), e.g. 24hpf neural tissues mapped to the 6hpf dorsal anterior epiblast (Figures 3.3B, S3.4). Conversely, direct comparison of 6hpf and 24hpf gene expression states failed to capture lineage relationships (blue points, Figs. 3.3B, S3.4).

We next tested the extent to which the single-cell graph represents a simple tree-like hierarchy of discrete states. For this, we 'coarse-grained' the graph by collapsing groups of similar cells into state nodes; edges between state nodes were weighted by the number of original single-cell connecting edges. A spanning tree was then traced through the most densely weighted edges to a 4hpf root state (Figure 3.3C, S3.5A). This spanning tree (the 'state tree') reflects many specific aspects of early development. In the neural plate, we observe notable branch points for the optic cup, the diencephalon, telencephalon, mesencephalon, and rhombencephalon, with associated states for region-specific post-mitotic neurons (e.g. *eomes*+ and *dlx1*+ neurons in distinct forebrain branches). The neural plate also includes neural crest, which branches to include cell states for melanoblasts, iridoblasts and xanthoblasts. In the lateral plate / ventral mesoderm, the state tree encodes extensive branching into hematopoietic

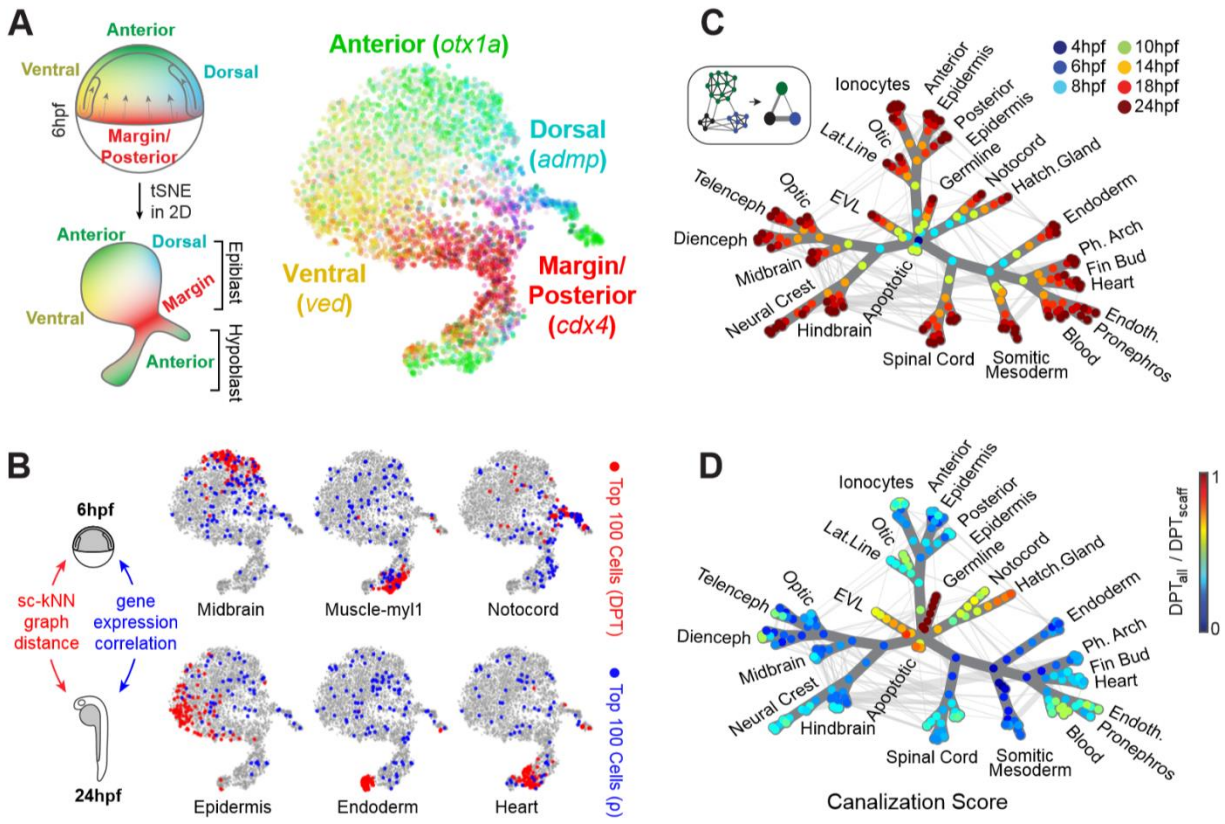


Figure 3.3. Single-cell and coarse-grained graphs encode progenitor-fate relationships.

(A) tSNE map of 6hpf epiblast and hypoblast states, colored by normalized transcript counts for select positional marker genes. Overlapping color gradients demonstrate continuous expression domains defined by position. Diagram relates positions of cells in the tSNE map to theoretical positions in the embryo. (B) In silico fate predictions for 6hpf embryo cells. The top 100 cells with predicted 24hpf fate outcomes are indicated for shortest graph diffusion distances (red) or direct single-cell gene expression correlation distances (blue) between 6hpf cells and 24hpf cluster centroids. (C) Construction and overview of the coarse-grained graph (See also Figure S3.5). Nodes indicate states (groups of transcriptionally similar cells), colored by timepoint. Weighted edges connect similar states within or between timepoints. Spanning tree edges connecting each node to the 4hpf root state through the top weighted edges are highlighted in dark grey. (D) Coarse-grained graph nodes are colored by a “canalization” score,

Figure 3.3 (Continued) defined as the ratio of diffusion distances between each node and the 4hpf root node through state tree edges only vs. through all graph edges. Highly canalized regions of the graph correspond to branches with the fewest off-tree edges.

cells, endothelial cells, heart, pharyngeal arches, the pronephritic duct, and fin buds. In the endoderm, two branch points give rise to cell states for pancreatic primordium (which includes *insulin*+ cells) and the pharyngeal pouch. In the epidermal lineage, branch points differentiate the otic placode, lateral line, ionocytes, and several states expressing markers for annotated "mucous-secreting" cells (Thisse et al., 2001). To facilitate data exploration, we developed web-based interfaces for the state tree and the full single-cell graph (www.tinyurl.com/scZfish2018). These tools permit interactive examination of: the inferred state hierarchy; expression for any gene of interest; and differential expression analysis between states, state combinations, or single cells.

Although many major cell state transitions are captured in the state tree, more complex features are evident in the coarse-grained and single-cell graphs. Off-tree interconnections between states, for example, were evident for (1) the neural crest and pharyngeal arches, (2) spinal cord and somitic mesoderm, (3) the neural plate, and others (Figure 3.3C, S3.5A). To formalize the degree to which the developmental landscape can be approximated as a hierarchy with discrete, non-looping branches, we defined a 'canalization score' (Figure 3.3D, see legend for definition), which reflects the off-tree connectivity of each coarse-grained state node. This analysis revealed widespread regions of 'low canalization', particularly in the neural plate and somitic mesoderm. These observations suggest that, in contrast to the classic notion of a cell lineage, the zebrafish cell state landscape cannot be fully represented as a tree.

Cell lineage history does not invariantly reflect cell state graph topology

While the single-cell and coarse-grained graphs represent an inferred landscape of developmental cell states, they do not reveal how individual cells traverse these states. A simple prediction would be that individual cell histories mirror graph topology. We tested this prediction by developing an inDrops-compatible strategy for recording *in vivo* lineage histories at the single-cell level: **Seq**uencing of **Tra**nscribed **C**lonally **E**ncoded **R**andom Barcodes

("TracerSeq"). TracerSeq utilizes the Tol2 transposase system (Kwan et al., 2007) to randomly integrate GFP reporter cassettes driven by the beta-actin promoter (*actb2*) into the zebrafish genome. To render each integration event unique and detectable by RNA-seq, we utilized Gibson assembly (Gibson et al., 2009) without subsequent amplification to introduce a random 20mer sequence barcode into the GFP 3' UTR (Figure 3.4A and S3.6). Because transgenic insertions can occur asynchronously over successive cell divisions, TracerSeq barcodes can facilitate the construction of lineage trees (Figure 3.4A). TracerSeq offers an advantage over related Cas9-based approaches (McKenna et al., 2016, JP et al., 2017), which can generate identical edits and/or large barcode deletions in independent lineages at non-trivial frequencies. By contrast, TracerSeq barcodes are uniformly distributed over a large sequence space (e.g., $4^{20} = 10^{12}$ unique sequences), facilitating straightforward calling of genetic clones (Figure S3.7). The small (20bp) locus size also greatly simplifies the construction, sequencing, and analysis of TracerSeq inDrops libraries.

The use of TracerSeq to analyze potentially small clones of cells (each restricted to a single embryo) requires high-efficiency tissue dissociation and transcriptomic barcoding methods. We therefore optimized a high-yield cell dissociation and recovery protocol for individual 24hpf zebrafish embryos (Figure S3.1D and Methods) and leveraged the high cell barcoding efficiency (>80%) of the inDrops platform (Zilionis et al., 2017). We then sequenced individual embryos (N=5) at 24hpf (Figure S3.7) that were injected at the 1-cell stage with the TracerSeq library, generating combined lineage+transcriptome datasets for 1,269 clonal barcodes distributed over 4,342 single cells (Figure S3.8). 2,361 of these cells (54%) were each marked by $\geq$2 distinct barcode integrations; 624 cells (14%) were marked by $\geq$5 integrations (Figure S3.8). Hierarchical clustering of TracerSeq barcodes organized these cells into over a hundred distinct founder clones with internal nested clone structures (Figure 3.4B, S3.9A-D). We then compared the lineage history and inferred transcriptional history of each founder clone by embedding its constituent cells onto the single-cell graph (Figure 3.4C). We

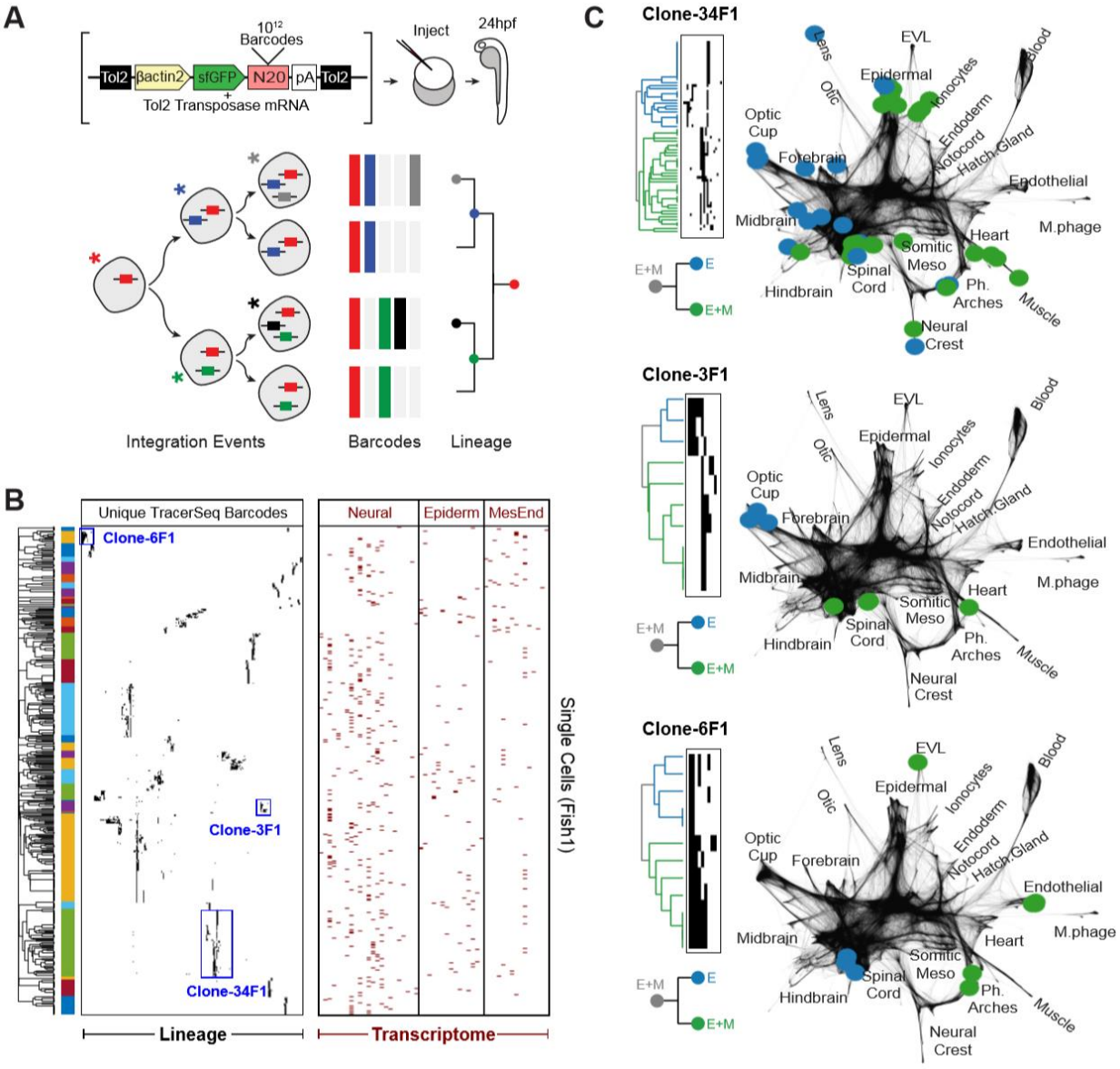


Figure 3.4. Single-cell transcriptomic barcoding of cell lineages using TracerSeq.

(A) Method overview. (B) Clustered heatmap for 1/5 TracerSeq embryos (See also Figure S9A-D) displaying lineage and transcriptome information for each cell. Heatmap rows are single cells for which both transcriptome and >1 TracerSeq barcodes were recovered. Columns denote unique TracerSeq barcodes (left, black squares: ≥1 UMI) and tissue identities (right, red squares) inferred from cluster annotations (Table S3.2). Heatmaps were clustered using Jaccard similarity and average linkage. (C) Examples of TracerSeq founder clones with positions of constituent cells (colored nodes) overlaid on the single-cell graph. Graph edges are

Figure 3.4 (Continued) shown in dark grey. Colors indicate the first lineage bifurcation within each founder clone. In the three cases shown, the founder clone included cells that differentiated into both ectodermal and mesodermal states, while one of the two first subclones was restricted to ectoderm.

found that the largest clones often marked a wide diversity of cell states. In multiple cases, however, additional barcode integrations in the same founder clone marked cells that were state-restricted. For example, one such clone (34F1), marked cells of the neural plate, epidermal tissues, and muscle, but contained a sub-clone restricted to anterior ectoderm. Similar lineage restriction events could be described for other founder clones (Figure 3.4C). These observations suggest that the current timing of TracerSeq integrations encompasses the transition from unrestricted pluripotency to the first fate restriction events appearing in the zebrafish embryo.

To investigate lineage relationships more systematically, we assessed the likelihood of recovering shared TracerSeq barcodes from all pairs of transcriptional states in the 24hpf zebrafish embryo. We first calculated a lineage coupling score (Figure S3.9E and Methods), defined as the number of shared barcodes relative to randomized data (z-score standardized), with values ranging from positive (coupled fates) to negative (anti-coupled fates). Hierarchical clustering of the pairwise correlation between coupling scores revealed structured groups of cell states (Figure 3.5A), which comprised related tissues and/or inferred germ layer derivatives. These included one distinct group that contained both mesodermal and endodermal derivatives, 4 groups containing ectodermal derivatives, and 2 groups containing mixtures of ectoderm and mesoderm. Several of these lineage groups are corroborated by prior fate mapping studies. We discuss here three examples. The first major lineage group, ('MesEndo'), includes derivatives of both lateral plate mesoderm and endoderm. These tissues originate from the marginal blastomeres of the early zebrafish gastrula, which involute first during gastrulation to form the hypoblast, and then rapidly migrate towards the animal pole (Kimmel et al., 1990, Warga and Nusslein-Volhard, 1999, Warga et al., 2009). The observed lineage isolation of these tissues is thus consistent with an early spatial partitioning of this region, further reflected in Figure 3.5A by negative lineage correlations to most other states. A second group, (Figure 3.5A 'Ecto III'), captures strong lineage couplings between anterior neural tissues including the optic cup,

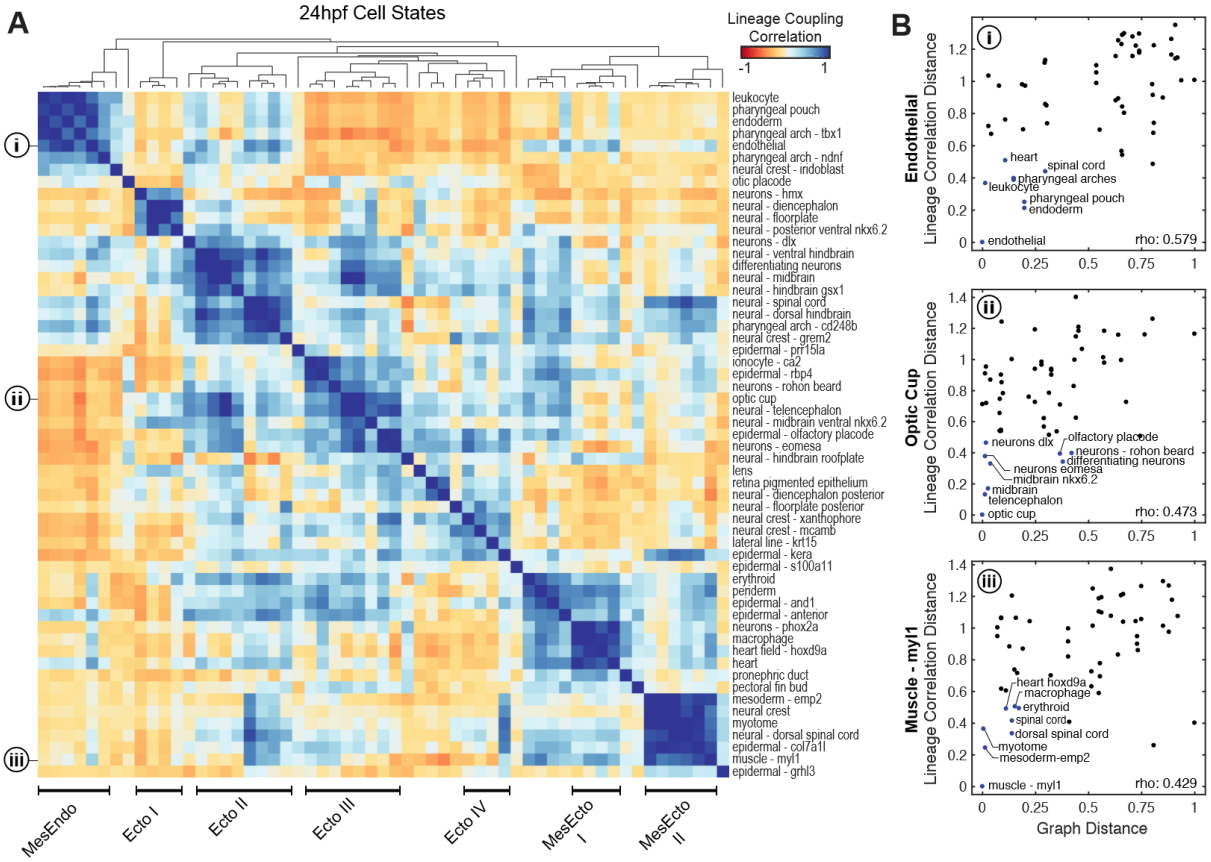


Figure 3.5. TracerSeq reveals systematic relationships between cell lineage and cell state.

(A) Heatmap of TracerSeq lineage coupling scores (see Methods) between pairs of 24hpf states, clustered by correlation distance and average linkage. Groups of states with similar lineage coupling signatures are annotated. (B) Quantitative relationships between lineage coupling correlation distances and scaled state tree diffusion distances for (i) endothelial, (ii) optic cup, and (iii) myl+ muscle states (see also S3.10A-F).

midbrain, and telencephalon (Woo and Fraser, 1995), and also to anterior epidermal derivatives such as the olfactory placode (Whitlock and Westerfield, 2000). These tissues are coupled to a lower degree with another group ('Ecto II'), which includes couplings between the hindbrain, spinal cord, and neural crest (grem2+). The third example we note is a group coupling ectoderm and mesoderm (Figure 3.5A 'MesEcto II), including muscle (*myl1*+), myotome, spinal cord, posterior neural crest, and epidermal states. These correlations mirror development of posterior body regions, which trace their origins to blastomeres proximal to the medial and ventral margin (Kimmel et al., 1990). These mesodermal-spinal cord couplings might also be explained by the presence of a later population of transient, multipotent neuromesodermal progenitor cells (NMPs) in the embryonic tailbud, which give rise to both of these populations (Tzouanacou et al., 2009, Davis and Kirschner, 2000, Kanki and Ho, 1997). Interestingly, these lineage groups tend to be organized by position (e.g. along the A-P axis) rather than strictly by germ layer/tissue origin (e.g. neural, epidermal, mesodermal).

We next questioned how clonal relationships compared with cell state relationships. A simplistic model of development is that cells progressively diverge in state as they diverge in lineage. Developing embryos, however, could violate this prediction in at least two ways: first, clonally distinct embryonic fields can give rise to similar cell types (i.e. 'convergent clones'); second, major transcriptional changes might drive related cells into qualitatively dissimilar states, possibly even late in development (i.e. ' divergent clones'). Overlaying TracerSeq lineage correlation scores on the cell state graph and comparing these scores to graph-derived state distances (Figure 3.5B, S3.10) revealed that some nearby states on the state graph were indeed clonally correlated, as expected by the simplistic model. However, nearby cell states also frequently displayed weak clonal correlations, suggesting convergent differentiation. These patterns were evident amongst state relationships for endothelial, optic cup, and muscle tissues (Figure 3.5B, S3.10A-F), and systematically when examining all states (Figure S3.10G).

We observed considerably fewer cases of divergent clonal behavior (Figure S3.10G). However, one notable example manifested as apparent looping of the neural crest into the pharyngeal arches, which originate in the graph from both neural plate and lateral plate mesoderm and merge at 18-24hpf (Figure 3.2A-B, S3.11A). While the contribution of neural crest to various mesenchymal tissues is well established (Le Douarin and Dupin, 2003, Le Douarin et al., 2004, Le Lievre and Le Douarin, 1975), the transcriptional information reflected by the graph loop alone does not reveal which annotated pharyngeal arch states arise from neural crest. TracerSeq data, however, provides a clear signature of distinct clonal patterns between pharyngeal arch states: one pharyngeal arch state (ph.arch-tbx1) is a member of the "MesEndo" lineage group with mesodermal clonal associations, while the second pharyngeal arch state (ph.arch-cd248b) is clonally related to neural crest and posterior neural states (Figs. 3.5A, S3.11B-F). These data indicate that cells in the ph.arch-cd248b state diverged from a neural plate lineage and subsequently converged with other lateral plate-derived states. The ability of embryonic clones to undergo dramatic converging/diverging behaviors thus underscores a continued need for independent measurements of both cell state and lineage in the mapping of cell fate hierarchies.

Robustness of cell type transcriptional programs following a signaling perturbation

Single-cell maps of vertebrate development can in principle facilitate unbiased, systematic analyses of mutant phenotypes and disease states. We used scRNA-seq to analyze the mutant phenotype for *chordin*, a well-studied developmental gene encoding a secreted BMP inhibitor expressed in the organizer and required for patterning the early dorsal-ventral axis (Sasai et al., 1994, Hammerschmidt et al., 1996, Schulte-Merker et al., 1997, Sasal et al., 1995, Piccolo et al., 1996). Chordin disruption leads to changes in gross embryo morphology, with an expansion of ventral tissues and a reduction of dorsal tissues (Hammerschmidt et al., 1996). scRNA-seq is uniquely suited to address how every tissue in the embryo changes in

abundance, and in gene expression, while also allowing detection of qualitatively new states, or combinations of states, if they occur.

We used CRISPR/Cas9 (Gagnon et al., 2014) to disrupt the *chordin* locus, resulting in highly penetrant clutches of mutant zebrafish embryos (Figure S3.12). inDrops profiling was performed on *chordin*-targeted and control embryos (*tyrosinase*-targeted, see Methods) in a narrow time series corresponding to ~14-16hpf (Figure 3.6A). After sequencing, we classified each of the *chordin*- and control-targeted cells to reference cell clusters of the 14hpf wild-type embryo (Figure S3.13 and Methods) and tested for altered gene expression. We reasoned that a qualitatively new cell state, if formed as a result of the aberrant patterning, would manifest as widespread changes in gene expression following mutation, with a magnitude comparable to the differences between wild-type embryonic states. Applying this criterion, we found no evidence of a qualitatively novel cell state following *chordin* depletion. Rather, the number of genes differentially expressed within states was modest compared to the differences defining the wild-type states of the 14hpf embryo (Figures 3.6B, S3.14A). Moreover, a tSNE mapping of CRISPR-targeted cells (Figure S3.13A-C) identified only a single cluster uniquely occupied by *chordin*-mutant cells (Figure S3.13D), distinguished primarily by a heat-shock/stress-like transcriptional signature. This same stress signature was elevated in multiple states in *chordin* targeted embryos (Figure S3.14A).

We next tested whether *chordin* disruption led to changes in abundance of particular classified cell types. As expected, expansion of states corresponding to ventral tissues (e.g., somitic mesoderm, epidermis, hatching gland, blood and endothelial tissues) at the expense of dorsal tissues (e.g., the neural plate and notocord) was observed (Figure S3.14A-B) (Hammerschmidt et al., 1996, Leung et al., 2005). Additional features could be appreciated by projecting the CRISPR datasets directly onto the wild-type single-cell graph (Figure 3.6C-D). For example, a sharp boundary bisected the lateral plate mesoderm into two compartments of opposing *chordin*-sensitivity, separating the heart and fin bud progenitor fields. Similar

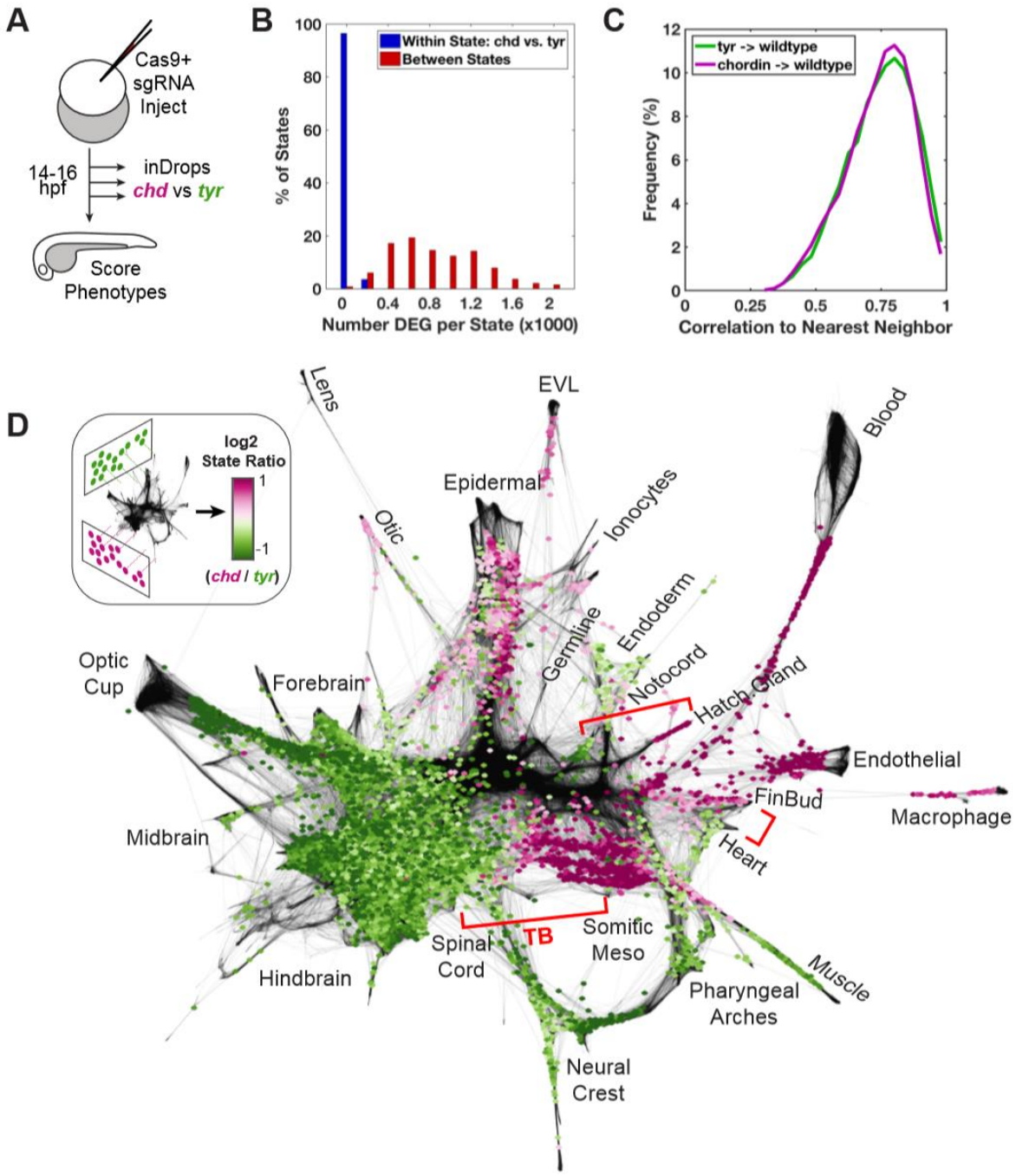


Figure 3.6. Regulatory features of the developmental landscape identified by genetic perturbation

(A) Left: Overview of the CRISPR experiment. Three pairs of chordin and tyrosinase (control) targeted samples were prepared and processed by inDrops ~14-16hpf. (B) Histogram

Figure 3.6 (Continued) depicting numbers of differentially expressed genes (DEG) identified in *chordin* vs. control (t*yrosinase*) cells for each state (blue bars), compared to DEG numbers when comparing between all state pairs (red bars). DEG were identified by Wilcoxon rank-sum test (adj. p-value < 0.01, absolute log2 fold change >1, average expression > 25 transcripts per million). (C) Histogram of Pearson correlation similarities (after PCA-projection) between each chordin/*tyrosinase* cell and its nearest neighbor from 10hpf, 14hpf, and 18hpf wild-type datasets (see Methods). (D) Log2 ratios of cell states with significant differential abundance (FDR < 0.25) in the *chordin* vs. *tyrosinase* samples. Purple and green regions correspond to wild-type cell states that are over- or under-represented in the *chordin* mutant, respectively. Adjacent graph domains with opposing *chordin* sensitivity are highlighted by brackets. TB: tailbud region (see *cdx4* expression in Figure S3.3).

juxtaposed domains of opposing *chordin* sensitivity were evident in the axial mesoderm, partitioning notocord from hatching gland, and in the tailbud separating spinal cord from somitic mesoderm (Figure 3.6D). Strikingly, each of these pairs of phenotypic domains appeared to be organized downstream of an inferred branchpoint in the cell state landscape. These domain pairs, therefore, likely reflect binary fate choices that are tuned by BMP signaling in wild-type embryos.

In a final analysis, we searched for the putative identity of the cells responding to *chordin* in the tailbud, as this is the site showing the largest expansion (somitic mesoderm) and loss (spinal cord) after perturbation. In zebrafish, *chordin* is expressed in the embryonic shield, transiently in the neural plate, adaxial cells, and also in the posterior tailbud region (Miller-Bertoglio et al., 1997). All of these expression patterns were confirmed in our single-cell graphs (Figure S3.15A). Furthermore, in contrast to its earlier expression in the shield, continued expression of chordin in the tailbud was distinct among a large panel of known BMP inhibitor genes (Figure S3.15A) and was tightly apposed by expression domains for multiple *bmp* transcripts (Figure S3.15B). These expression characteristics might explain the elevated *chordin* sensitivity of posterior body regions. To examine this region in greater detail, we isolated a subgraph of tailbud and descendent cells, Consistent with previous studies, two cell state trajectories branching from a common neuromesodermal-like *brachyury*+;*sox2*+ progenitor state were identified, each expressing markers of neural fates (*sox3, sox19a, pax6a, neurog1*), or somitic fates (*tbx16, tbx6, tbx24, msgn1, myod1*) (Figure S3.16A-C) (Row and Kimelman, 2009, Kanki and Ho, 1997, Row et al., 2016, Gouti et al., 2017). Strikingly, the neural-mesodermal branchpoint coincided with the boundaries of both *chordin* expression and sensitivity (Figure 3.16D-E). The *chordin* expressing cells in this region of the single-cell graph exhibited a distinct expression profile (Figure S3.17), including a cadherin (*cdh11*), early neurogenic markers (*her3*, *her8a*, *sox19a*), and several relatively uncharacterized genes (*gig2g*, *foxb1b*, *foxb1a*). We

hypothesize that these cells represent a key transition state at which point tailbud cells initiate a posterior neurogenic program in a *chordin*-dependent manner.

Discussion

Our study demonstrates a graph-based approach for mapping whole-embryo developmental landscapes, over time, from scRNA-seq data. The graph was constructed with minimal assumptions about development, and describes individual cell states transitioning from pluripotent blastomeres to a large array of cell types and tissues during the first day of zebrafish embryogenesis. This dataset can now be mined to identify temporal and tissue associations for any gene, cell type, or biological process of interest. As with genome annotation efforts over the years, we expect that the annotation of identified cell states may undergo refinement with community input.

As single-cell atlases and landscapes of embryo development become routinely available, one is challenged to reconsider the relationship between a cell lineage (by definition, a tree), and the considerably more topologically complex gene expression landscape through which these cells traverse. Using TracerSeq, we confirmed that differentiating cells of the zebrafish embryo do not invariantly follow tree-like hierarchies. Instead, we observed both widespread convergence in cell states for clonally distant cells and instances in which clonally related cells diverged into distant states. Non-tree like convergence of cell states could be explained by the differentiation of well-separated spatial domains of the embryo into the same basic cell types (e.g. along the A-P axis), while divergence could involve mechanisms such as asymmetric cell division or exposure to spatially varying signals (Gonczy, 2008). We anticipate that the synthesis of single-cell lineage and transcriptome information will continue to be crucial for deciphering how cells traverse state trajectories with complex topologies (e.g. loops or continua).

Single-cell mapping of genetic perturbation data presents a powerful framework for identifying regulatory features of a developmental landscape. Following deletion of the BMP inhibitor, *chordin*, we showed that the defining transcriptional features of the landscape remained mostly unchanged, yet cell state abundances could be dramatically and reciprocally altered, as if the landscape were “tilted” but cell fates remain canalized. Future systematic mapping of signaling perturbations could be used to reveal the complete signaling logic of the embryo, as cells are specified toward their final fates. Together, these studies demonstrate the power, modularity, and quantitative benefits of unbiased scRNA-seq-based interrogations of embryonic development. We anticipate that similar large-scale datasets will facilitate explorations of additional developmental stages, tissues, and species.

Materials and Methods

Zebrafish

Both AB and TU wild-type strains were used. Embryos were generated by natural spawning and the time of fertilization was used to stage each clutch. Later stages were confirmed using morphological criteria (Kimmel et al., 1995). Embryos were incubated at 28.5C for all wild-type time course experiments and processed for inDrops at the indicated times. All zebrafish were housed in a facility overseen by the Harvard Medical Area Standing Committee on Animals (our IACUC) which performs regular inspections and under which we have an approved protocol for all animal procedures.

Cell Preparation

Zebrafish embryos were grown to the indicated times and chorions were removed by incubating in 1mg/mL Pronase (Sigma P5147-1G) for 3-4 min followed by washing in 0.3X Danieau Buffer. [10X Danieau Buffer = 174 mM NaCl, 2.1 mM KCl,1.2 mM $MgSO_4$, 1.8 mM

$Ca(NO_3)_2$, 15 mM HEPES, pH 7.6]. Dissociation of embryonic tissues was performed similarly as previously described (Manoli and Driever, 2012) with the following modifications. For wild-type time course experiments, 50-100 embryos were used to generate each sample. Embryo tissues were triturated to homogeneity in 1-5mL FACSmax cell dissociation solution (Genlantis T200100) and incubated for 4-5 minutes at room temperature. Cells were then filtered through a 40µm cell strainer mesh (Fisher 352340), and centrifuged in a swinging bucket rotor at 310g for 5 minutes. Cell pellets were resuspended in 1X DPBS (no Ca/Mg, Life Technologies 14190-144) containing 1% BSA (Sigma A3311-100G), and subjected to 2-3 additional rounds of centrifugation and resuspension. After washing, cells were resuspended in 0.05% BSA / DPBS containing 18% optiprep density medium (Sigma D1556-250ML). Cell density was quantified manually using INCYTO™ C-Chip™ Disposable Hemacytometers (Fisher 22-600-100), and adjusted to ~100,000 cells per mL. For single-embryo dissociations, all FACSmax and wash volumes were reduced to a volume 0.5 mL and were carried out in 0.5mL LoBind microcentrifuge tubes (Eppendorf 022431005) that had been pre-coated with 10% BSA/DPBS for 15 minutes at room temperature.

Single-cell Microfluidic Droplet Barcoding

Single-cell transcriptomes were barcoded using inDrops (Klein et al., 2015), as previously reported (Zilionis et al., 2017). Following the within-droplet reverse transcription step, emulsions were split into batches of approximately 1,000-2,000 cells, frozen at -80C, and subsequently processed as individual RNA-seq libraries (see Table S3.1).

Preparation of RNA-Seq Libraries

Standard transcriptome RNA-seq libraries were processed as previously reported (Zilionis et al., 2017). For TracerSeq experiments, TracerSeq-targeted RNA-Seq libraries were also prepared. These targeted libraries were reverse-transcribed from the product of the linear

amplification *in vitro* transcription (IVT) reaction (step 143 of the Zilionis et al 2017 protocol, ref 7) without prior RNA fragmentation. Non-fragmented IVT product (5 uL) was mixed with 4 uL water, 1 uL 10mM dNTPs, and 1 uL of 10uM primer TracerRT. This reaction was incubated at 70C for 3 min, then moved to ice. To this reaction was added 4 uL 5X PrimeScript Buffer, 3.5 uL water, 1 uL RNASE-OUT (Thermo-Fisher 10777-019), and 0.5 uL PrimeScript Enzyme (Clontec 2680A). Reverse transcription (RT) was performed by incubating the reaction first at 30C for 10 min, followed by 42C for 1 hour, and inactivated at 70C for 15 min. RT products were purified using 1.2X AMPureXP beads (Beckman Coulter A63881), and eluted in 15 uL of RE Buffer (10 mM Tris pH 7.5, 0.1mM EDTA). TracerSeq cDNA was then PCR-amplified as follows: 5 uL RT product was mixed with 3 uL of water, 1 uL each of inDropsTracerF and inDropsTracerR primers (10uM), and 10 uL of 2X Phusion Master Mix (Thermo-Fisher F548L). This reaction was thermal cycled: 98C for 30 sec; 10 cycles of [98C 10 sec; 63C 20 sec; 72C 30sec]; 72C for 3 min, and then cleaned up using 1.2X AMPureXP beads and eluting in 12 uL RE buffer. Illumina sequencing adapters and sample indices were then incorporated by resuming the standard inDrops library preparation protocol (step 157 of Zilionis et al 2017). Final libraries for each multiplexed sample index were quantified using the KAPA library quantification kit (Kapa KK4844) and pooled at equimolar ratios (2mM final concentration) prior to sequencing.

Sequencing and Read Mapping

All inDrops transcriptome and TracerSeq libraries were sequenced on an Illumina NextSeq 500 using the NextSeq 75 High Output Kits according to the following sequencing specifications. V2 libraries used custom sequencing primers and 35 cycles for Read1, 51 cycles for read2, and 6 cycles for IndexRead1 and included 15-25% PhiX spike-in. V3 libraries used standard Illumina sequencing primers and 61 cycles for Read1, 14 cycles for Read2, 8 cycles each for IndexRead1 and IndexRead2. Raw sequencing data (i.e. FASTQ files) were processed using the inDrops.py bioinformatics pipeline available at github.com/indrops/indrops.

Transcriptome libraries were mapped to a zebrafish reference transcriptome built from the zebrafish GRCz10 genome assembly (Assembly Accession: GCF_000002035.5). Bowtie version 1.1.1 was used with parameter –e 200; inDrops.py UMI quantification was performed with parameter –u 2 (counts were ignored from UMIs split between more than 2 genes). For TracerSeq libraries, sequencing reads were filtered and sorted by inDrops.py and then processed by custom barcode filtering pipeline (see below).

Cell Filtering and Data Normalization

inDrops data were filtered to only include UMIs originating from abundant cell barcodes. This determination was made by manually inspecting a weighted histogram of UMI counts for each cell barcode, and thresholding only the top ~95% of the largest (and often the only) mode of the distribution. Transcript UMI counts from multiplexed libraries originating from each biological sample then concatenated into a single genes x cells table and adjusted by a total-count normalization.

Identification and Filtering of Variable Genes

For each normalized UMI counts table corresponding to a single biological sample, highly variable genes were identified by first computing gene Fano factors, and ranking all genes by an above-Poisson noise statistic, as previously described (Klein et al., 2015). The top 2000 variable genes according to this statistic were then filtered to include only genes whose single-cell transcript counts were minimally correlated (correlation coefficient > 0.2) to at least one other variable gene. A set of cell cycle and housekeeping-associated genes were then excluded from downstream analyses. This set was generated by “growing” a list including any gene that was similarly expressed (single-cell correlation coefficient > 0.4) to any of the following genes. Cell Cycle: *cdk1, mcm2, mcm7, rrm2, cenpa, cdc6, ccnf, cdca4, ccnd1, kif4;* Housekeeping: *hmgb1b, hmgb3a, hspd1, hspa9, rplp0, hnrnpaba, rps2, rps12, rpl12, rps13,*

rps14, rps15a, rpl10, rps3a, rpl31, rpl37, rps6, rpl9, rpl11, rpl34, rpl13, rpl36a, rpl26, rps8a, rpl21, rps27.1, rpl27a, cirbpb. The resulting list was then subjected to a second "growing" round, and all associated genes were discarded.

Low Dimensional Embedding and Clustering

Normalized gene x cell counts data were projected into low dimensional subspace by first standardizing counts for each gene (by z-score) and performing principal component analysis (PCA). The number of significant PCA dimensions was then estimated by comparing the eigenvalue distribution of cell principal components to the eigenvalue distribution of randomized data, as previously described (Klein et al., 2015). Non-significant principal components were removed from subsequent steps of the analysis. Two-dimensional t-distributed stochastic neighbor embeddings (tSNE) (Van der Maaten and Hinton, 2008) were then generated from cell PC scores using a perplexity setting of 30, with 1000-5000 iterations. Datasets containing >10,000 cells implemented a Barnes-Hut approximation (Van der Maaten, 2014). Groups of related cells in the resulting tSNE maps were then identified by density clustering (Rodriguez and Laio, 2014). Cell cluster annotations were assigned according to known cell type and tissue expression patterns (https://zfin.org) and/or specific marker genes. For timepoints 6hpf, 14hpf, 18hpf, and 24hpf, a single small cluster was initially identified whose defining genes consisted solely of housekeeping factors (see "Identification of Variable Genes"). Cells contributing to these clusters were inferred to represent dead/unhealthy cells and were removed from subsequent analyses. In some cases, examination of cell or tissue-specific marker expression on the tSNE map revealed sub-groups of transcriptionally distinct cells that were not captured by the first round of density clustering. In these cases, clusters were isolated and sub-clustered. Cases in which two adjacent clusters failed to display any differentially expressed genes were merged into a single cluster.

Identification of Differentially Expressed Genes

Cluster-defining transcripts were identified by “Model-based Analysis of Single-cell Transcriptomics” / “MAST” (Finak et al., 2015) or a Wilcoxon rank-sum test, as indicated. Cluster-specific marker genes (Table S3.2) were identified by comparing cells of each cluster to cells from all other clusters in the same timepoint. Genes were considered differentially expressed based on fold-change, minimum expression, and adjusted p-value cutoffs, as indicated. Tests were implemented in Matlab (rank-sum) or R / Seurat 2.2.0 (rank-sum and MAST). Unless otherwise noted, p-values were adjusted for multiple hypotheses by either a Bonferroni correction (Seurat 2.2.0), or Storey (Storey, 2002) correction (Matlab/mafdr).

Automated Annotation of Cell States

Cells collected from *chordin* and *tyrosinase*-targeted embryos were classified using the fitcknn and predict functions in Matlab. A kNN classifier object was first trained using the PCA-projected wild-type dataset (14hpf) and corresponding cell-state assignments using the “exhaustive” search algorithm. The following parameters were chosen automatically via the “optimize hyperparameters” subroutine: Euclidean distance, and 5 nearest neighbors. Ties were broken by selecting the class with the nearest neighbor. Cell state assignments were then predicted for chordin and tyrosinase datasets by first standardizing counts for each gene (z-score) and projecting the data into PCA space defined by the 14hpf wild-type dataset. The predict function was implemented with default settings.

Construction of Single-Cell Graphs

A nearest-neighbor graph of cells represents the manifold of cell states observed in scRNA-Seq data by a set of nodes (cells) connected to their nearest neighbors by edges. Graph representations of single cell data have been used before, for example using k-nearest-neighbor (knn) graphs (Weinreb C, 2016). The knn graph construction has the property that it allows

regions of the graph to have variable neighborhood sizes. For scRNA-Seq embryo data, this means that each tissue can have a different natural scale for similarity between cells. However, such simple graph constructions are not suitable for the complexity of whole embryo time series data, because of two challenges: (1) different time points can reside in different sub-spaces of gene expression; (2) the absolute number of cells sampled from different tissues can vary significantly, such that the optimal connectivity of the graph required to identify features is not uniform across tissues. In addition to these two constraints, scRNA-Seq data still presents differing neighborhood sizes. To address constraint (1), we construct a single cell graph in a step-wise manner using a subspace defined by cells from consecutive timepoints. In this case, we find that projecting cells into a future time point is sensible because each progressive time point is of higher complexity, but sufficiently close to justify a unidirectional choice. Other situations may warrant projecting cells back in time point or defining a consensus subspace shared by two or even multiple time points. To address constraint (2), one can invoke the advantages of more than one graph construction approach in order to account both for differences in neighborhood size and in the abundance of cells sampled (Zhu Q, 2016, Ting D, 2011). We made sequential use of four neighbor-selection approaches: initially non-mutual k-nearest neighborhoods, then locally self-tuning neighborhoods, a globally-tuning neighborhood, and mutual k-nearest neighborhoods. This sequential approach is heuristic but succeeds in resolving the major aforementioned constraints. The specific steps used are as follows: (1) each cell in time point t_i forms an outgoing edge to its 200 nearest neighbors from all cells in time points (t_i, t_{i-1}), where all cells are projected into the non-trivial PC subspace defined by the cells in t_i alone (see **Low Dimensional Embedding** section above for definition of the subspace). This strategy forces any edges between timepoints to directly compete with edges within a timepoint. Correlation was used as the distance metric. (2) Edges are then subjected to local neighborhood restriction (a self-tuning graph construction (Ting D, 2011)): an outgoing edge from cell *i* to *j* was kept if the distance d_{ij} was less than a local threshold $e_{ij}=3*\min_{k\neq i}(d_{ik})$, i.e.

retained neighbors were at most 3-fold as far as the cell's closest neighbor. (3) To avoid very sparse neighborhoods becoming connected to distant regions of the graph, edges were then subject to a global neighborhood restriction: edges were kept if they were below the average edge distance across all cells between time points (t_i, t_{i-1}), or if they were within 1 standard deviation of the average edge distance within the same time point t_i. (4) The graph was further reduced by retaining at most 20 mutual nearest neighbor edges.

Construction of Coarse-Grained Graphs

A coarse-graining procedure to abstract the major features of the single-cell graph was performed as follows. First, single-cell nodes belonging to the same annotated tSNE cluster ID were collapsed into a single state node. Edges between each pair of state nodes were then weighted by calculating the Jaccard index of original shared single-cell edges (i.e. the ratio of shared single-cell edges to the total number of outgoing edges for that node pair). State edges were then discarded if they received a Jaccard index weight < 0.01. Finally, a spanning tree was traced through the weighted edges as follows. Beginning with the final timepoint, edges for all nodes were ranked according to weight. Edges then were then removed recursively, starting with the weakest edges, unless doing so would increase the total number of graph connected components. This process was then repeated for each timepoint. The resulting spanning tree connects all nodes to a single 4hpf "root" node defined by all cells of the first timepoint.

Graph Visualization

Single-cell graphs were visualized using a force-directed layout (Jacomy et al., 2014), implemented in Gephi (0.9.1) (Bastian M., 2009). Coarse-grained graph layouts were also rendered in Gephi, using the Yifan Hu Proportional algorithm (Y, 2006) and were based only on tree edges. The tailbud single-cell subgraph was visualized using the Matlab "subspace" layout subroutine implementing the visualization algorithm described in (Y, 2009).

Generation of TracerSeq Embryos

TracerSeq experiments were performed by co-injecting 2-4 nL of TracerSeq library (~20ng/uL) together with Tol2 transposase mRNA (~50 ng/uL) in water and containing a 1:20 dilution of Phenol Red solution (Sigma P0290). Embryos were screened for mosaic GFP fluorescence (see Figure S3.7A) the following day, and the brightest embryos were dissociated into single-cell suspensions and processed by inDrops.

Preparation of TracerSeq inDrops Libraries

TracerSeq barcode libraries were generated by isothermal assembly (Gibson et al., 2009) (ISO) of two PCR-amplified dsDNA fragments via a single-stranded "bridge" oligonucleotide containing a stretch of 20 randomized bases (Figure S3.6A). The initial TracerSeq libraries were based on a pMTB vector containing a superfolder-GFP gene driven by the *actb2* promoter and flanked by Tol2 sites. The two ISO fragments were PCR-amplified from the pMTP-sfGFP plasmid using Phusion Polymerase (Thermo-Fisher F548L) as follows: 2 uL pMTB-sfGFP plasmid (~ng/uL), 18uL water, 2.5 uL each of forward and reverse primers (see Table S3), and of 2X Phusion Master Mix (Thermo-Fisher F548L). Reactions were cycled: 98C for 30 sec; 25 cycles of [98C 10 sec; 66C 20 sec; 72C 30sec]; 72C for 3 min. PCR products were then cleaned up and concentrated with 1.0X AMPureXP beads and eluted in DS buffer (10 mM Tris pH 8.0, 0.1mM EDTA). Isothermal assembly was performed by combining 0.4 pmoles of each ISO fragment (1 and 2) with 1.2 pmoles of the TracerBridge oligo and water to a final volume of 5 uL. This mixture was then combined with ISO master mix and incubated at 50C for 30 min. The desired ~6.7kb final product was gel purified and then cleaned up using 1.0X AMPureXP beads and eluted in water. Aliquots of this final library were stored at -80C. To generate sufficient amounts of library for zebrafish embryo injections without the need for PCR amplification, the entire ISO procedure was scaled to 10-12X the recipe volumes listed above. All ISO reactions included a "no-Bridge" control in which the TracerBridge oligo was omitted

from the reaction (see Figure S3.6B). Some TracerSeq libraries were also generated by circularizing a single 6.7kb ISO fragment (amplified using pMTBF1_Reverse and pMTBF2_Forward primers) using the same TracerBridge oligo, and using the same reaction conditions as described above (Figure S3.6A). In these cases, gel purification was based on differential migration of the open circular vs. linear forms of similarly sized DNA molecules. Homemade ISO master mix was prepared as previously described (Gibson et al., 2009). Diversity of TracerSeq barcode libraries was initially assessed by Sanger sequencing. Briefly, the TracerSeq barcode junction was PCR-amplified from a fully assembled library, ligated into the pCR4-Blunt-TOPO cloning vector, and transformed into One-Shot Top10 competent cells. Sanger sequences from each of 19 individual bacterial clones are shown in Figure S3.6C. TracerSeq library diversity was subsequently and more thoroughly assessed using RNA-seq data (Figure S3.7B-D).

Processing of TracerSeq Sequencing Reads

TracerSeq sequencing data were processed by inDrops.py to perform read filtering and inDrops cell barcode correction. Sorted FASTQ files for TracerSeq cDNA reads (with inDrops cell barcodes and UMI sequences in the header) were then processed as follows. First, sequences flanking both sides of the N20 barcode were identified and trimmed; reads that did not contain at least 12bp of flanking sequence were discarded. Second, all reads corresponding to unique UMI-cell barcode pairs were combined and used to generate a multiple-sequence alignment consensus for each uniquely detected TracerSeq barcode in each cell. To determine which TracerSeq barcodes were derived from the same clonal insertion event, pairwise sequence comparisons were performed between all unique transcript barcodes detected across all single-cell libraries associated with a single TracerSeq embryo. Barcode diversity for a typical experiment is illustrated in Figure S3.7B, in which 1,000 unique transcript barcodes were selected at random from the TracerSeq Fish1 embryo and subject to pairwise edit distance

comparisons. The “edit” distance between two sequences was based on the Levenshtein distance, except that each single-base substitution was weighted as 2, while each indel was weighted as 1. The resulting distance matrix was clustered and plotted as a heatmap, revealing subsets of barcodes with identical or near-identical sequences. This signature was not detected when performing comparisons amongst random 20mers. A histogram plot of all pairwise edit distances (excluding self-pairs) confirmed a bimodal distribution in which the majority of pairwise distances were consistent with those of the random distribution (Figure S3.7C). The smaller peak in the distribution, centered at an edit distance of zero, was only observed when comparing barcodes within the same embryo (Figure S3.7D). Based on these observations, TracerSeq barcodes were grouped into clones first by identifying and collapsing any barcodes with exact sequence matches. Remaining unique barcodes were then combined into clones if they were within an edit distance of “6” (e.g. up to 3 base substitutions) to a previously identified clone barcode. A UMI counts table of clones vs. cells was then generated for each TracerSeq embryo and subjected to downstream analysis. In order to directly compare UMI counts for both clones and transcripts for each individual cell, error-corrected inDrops cell barcodes (identified by inDrops.py) were retained for all cells. Analyses appearing in Figures 3.4-5 and Figures S3.9-11 were restricted to cell barcodes for which both TracerSeq and transcriptome data were recovered.

Calculation of TracerSeq Lineage Coupling Scores

Cells from all 5 TracerSeq embryos that were associated with both the clone and transcript datasets were used. First, the total number of shared TracerSeq clones was tabulated for each pair of 24hpf states. A TracerSeq clone was defined as “shared” when it contained at least 2 individual cells assigned to each state. Only clones with at least 5 total cells were considered. The number of shared clones was then compared to randomized data in which cell state assignments were permuted. A total of 20,000 random permutations were performed and

used to calculate a z-score for each original “shared” clone count with respect to the random distribution. Lineage coupling z-scores were clustered and plotted as a hierarchically clustered heatmap in Figure S3.9E. Positive z-scores indicate pairs of 24hpf states that shared significantly more TracerSeq clone barcode hits than expected by chance; a negative score indicates state pairs that were significantly less coupled than expected by chance. We then computed a correlation coefficient between z-scores for each pair of states. Lineage coupling correlations were plotted as a clustered heatmap in Figure 3.5A.

CRISPR

CRISPR experiments were performed by co-injecting 1-cell stage zebrafish embryos with *in vitro*-transcribed purified sgRNAs together with purified recombinant Cas9 protein. 2-4 nL injections were performed with sgRNAs at a combined concentration of ~100ng/uL (~3uM) and with Cas9 protein at ~7uM final concentration diluted in water and containing a 1:20 dilution of Phenol Red solution (Sigma P0290). Embryos that were damaged during the injection process (typically < 5%) were discarded. Embryos were incubated at 28C for 4-6 hours before moving to 23.5C so that they would develop to approximately the 10-14 somite stage (normally corresponding to ~14hpf) by the following morning. Endpoint stages for each clutch of embryos were confirmed as described above. For each CRISPR experiment, 20-30 individual *chordin* or *tyrosinase*-targeted embryos were dissociated into single cell suspensions and barcoded sequentially by inDrops as approximate stage-matched sample pairs. A total of 3 such pairs were sequenced and analyzed. The *tyrosinase* gene, when disrupted, generates a mild pigmentation defect that can serve as both a positive control for Cas9 activity and a negative control for injection-induced toxicity that can accompany CRISPR experiments (Jao et al., 2013). Embryos with disrupted *tyrosinase* develop otherwise normally and can be effectively treated as wild-type for the purposes of this experiment. *chordin* sgRNAs were designed using ChopChop (http://chopchop.cbu.uib.no/) to target exons 1-2 of the *chordin* locus. *tyrosinase* was

targeted using a previously reported sgRNA sequence (Jao et al., 2013). sgRNAs were synthesized *in vitro* as previously described (Gagnon et al., 2014) using primer sequences listed in Table S3.3. Recombinant His-tagged Cas9 protein was expressed in *e.coli* and purified as previously reported (Gagnon et al., 2014).

Projecting CRISPR inDrops datasets onto the Single-Cell Graph

A total of 6 inDrops datasets consisting of 3 biological replicates for both *chordin* and *tyrosinase* CRISPR-targeted samples were analyzed. To compare cells of CRISPR-targeted samples to those of the wild-type timecourse, each CRISPR dataset was indexed to include z-scored variable genes and projected into PCA subspaces that were defined by each of the following wild-type timepoints: 10hpf, 14hpf, and 18hpf. These timepoints were chosen to limit the analysis to wild-type states that were collected within a similar time frame as the CRISPR samples. A set of k-nearest neighbor edges from each CRISPR cell to one of the wild-type cells were then identified. The nearest neighbor number k was scaled to correct for different numbers of cells in each sample such that the total number of edges sought for each timepoint projection was set to 100 * (the number of wild-type cells in the target dataset). This process resulted in a vector of 6 projecting edge counts for each wild-type cell, 3 from each of the two CRISPR target genes. To identify cell state neighborhoods that were significantly over- or under-represented in *chordin*-targeted embryos (relative to *tyrosinase*-targeted embryos), a t-test was performed between the 3 embryo edge counts from each CRISPR target, for each cell node. Cell nodes participating in significantly different numbers of projecting edges were identified (FDR < 0.25), and their associated log2 mean edge count ratios were color-coded on the single-cell graphs in Figure 3.6D and Figure S3.16D.

Pseudo-Spatiotemporal Ordering and Identification of Dynamically Varying Genes

To assess the continuum of cell states comprised within the tailbud region of the single-cell graph, a subgraph was first extracted. The subgraph consisted of cells from the 8hpf-18hpf timepoints and included all cells assigned to posterior neural (hindbrain or spinal cord), tailbud, or pre-somitic mesoderm tSNE clusters. A pseudo-spatiotemporal ordering of cells along the resulting continuum was determined in a variation of Wanderlust (Bendall et al., 2014), as follows. First, approximately 200 cells at each “end point” of the continuum were manually selected. Next, a series of shortest paths were calculated between these two sets of cells through a version of the subgraph in which 50% of all edges were randomly deleted. This process was repeated for a total of 100 iterations in which different sets of edges were randomly deleted. All cells/nodes discovered during this process were ordered based on their average position over all shortest paths in which they appeared, resulting in a continuous ordering of cells. A “start point” zone was then inferred based on expression of the *ta* / *brachyury* transcript. Genes that varied dynamically along this trajectory were then identified similarly to as previously described (Macosko et al., 2015). Sliding windows of 100 cells were first scanned to identify two windows with maximum and minimum average expression levels for all genes, respectively. A t-test was then performed between these two sets of 100 expression measurements (FDR < 0.05). Gaussian-smoothened expression z-scores for significantly variable genes along the trajectory were then calculated. A subset of the significant genes identified are shown in Figure S3.16B.

References

Bastian M., H. S., Jacomy M 2009. Gephi: an open source software for exploring and manipulating networks. *International AAAI Conference on Weblogs and Social Media*.

Bendall, S. C., Davis, K. L., Amir El, A. D., Tadmor, M. D., Simonds, E. F., Chen, T. J., Shenfeld, D. K., Nolan, G. P. & Pe'er, D. 2014. Single-cell trajectory detection uncovers progression and regulatory coordination in human B cell development. *Cell,* 157**,** 714-25.

Davis, R. L. & Kirschner, M. W. 2000. The fate of cells in the tailbud of Xenopus laevis. *Development,* 127**,** 255-67.

Finak, G., Mcdavid, A., Yajima, M., Deng, J., Gersuk, V., Shalek, A. K., Slichter, C. K., Miller, H. W., Mcelrath, M. J., Prlic, M., Linsley, P. S. & Gottardo, R. 2015. MAST: a flexible statistical framework for assessing transcriptional changes and characterizing heterogeneity in single-cell RNA sequencing data. *Genome Biol,* 16**,** 278.

Gagnon, J. A., Valen, E., Thyme, S. B., Huang, P., Akhmetova, L., Pauli, A., Montague, T. G., Zimmerman, S., Richter, C. & Schier, A. F. 2014. Efficient mutagenesis by Cas9 protein-mediated oligonucleotide insertion and large-scale assessment of single-guide RNAs. *PLoS One,* 9**,** e98186.

Gibson, D. G., Young, L., Chuang, R. Y., Venter, J. C., Hutchison, C. A., 3rd & Smith, H. O. 2009. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods,* 6**,** 343-5.

Gonczy, P. 2008. Mechanisms of asymmetric cell division: flies and worms pave the way. *Nat Rev Mol Cell Biol,* 9**,** 355-66.

Gouti, M., Delile, J., Stamataki, D., Wymeersch, F. J., Huang, Y., Kleinjung, J., Wilson, V. & Briscoe, J. 2017. A Gene Regulatory Network Balances Neural and Mesoderm Specification during Vertebrate Trunk Development. *Dev Cell,* 41**,** 243-261.e7.

Haghverdi, L., Buttner, M., Wolf, F. A., Buettner, F. & Theis, F. J. 2016. Diffusion pseudotime robustly reconstructs lineage branching. *Nat Methods,* 13**,** 845-8.

Hammerschmidt, M., Pelegri, F., Mullins, M. C., Kane, D. A., Van Eeden, F. J., Granato, M., Brand, M., Furutani-Seiki, M., Haffter, P., Heisenberg, C. P., Jiang, Y. J., Kelsh, R. N., Odenthal, J., Warga, R. M. & Nusslein-Volhard, C. 1996. dino and mercedes, two genes regulating dorsal development in the zebrafish embryo. *Development,* 123**,** 95-102.

Hashimshony, T., Senderovich, N., Avital, G., Klochendler, A., De Leeuw, Y., Anavy, L., Gennert, D., Li, S., Livak, K. J., Rozenblatt-Rosen, O., Dor, Y., Regev, A. & Yanai, I. 2016. CEL-Seq2: sensitive highly-multiplexed single-cell RNA-Seq. *Genome Biol,* 17**,** 77.

Islam, S., Zeisel, A., Joost, S., La Manno, G., Zajac, P., Kasper, M., Lonnerberg, P. & Linnarsson, S. 2014. Quantitative single-cell RNA-seq with unique molecular identifiers. *Nat Methods,* 11**,** 163-6.

Jacomy, M., Venturini, T., Heymann, S. & Bastian, M. 2014. ForceAtlas2, a continuous graph layout algorithm for handy network visualization designed for the Gephi software. *PLoS One,* 9**,** e98679.

Jao, L. E., Wente, S. R. & Chen, W. 2013. Efficient multiplex biallelic zebrafish genome editing using a CRISPR nuclease system. *Proc Natl Acad Sci U S A,* 110**,** 13904-9.

JP, J., B, S., J, P.-M., A, A., B, H., M, F. & A, V. O. 2017. Massively parallel clonal analysis using CRISPR/Cas9 induced genetic scars. *bioRxiv*.

Kanki, J. P. & Ho, R. K. 1997. The development of the posterior body in zebrafish. *Development,* 124**,** 881-93.

Kimmel, C. B., Ballard, W. W., Kimmel, S. R., Ullmann, B. & Schilling, T. F. 1995. Stages of embryonic development of the zebrafish. *Dev Dyn,* 203**,** 253-310.

Kimmel, C. B., Warga, R. M. & Schilling, T. F. 1990. Origin and organization of the zebrafish fate map. *Development,* 108**,** 581-94.

Klein, A. M., Mazutis, L., Akartuna, I., Tallapragada, N., Veres, A., Li, V., Peshkin, L., Weitz, D. A. & Kirschner, M. W. 2015. Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. *Cell,* 161**,** 1187-201.

Kobitski, A. Y., Otte, J. C., Takamiya, M., Schafer, B., Mertes, J., Stegmaier, J., Rastegar, S., Rindone, F., Hartmann, V., Stotzka, R., Garcia, A., Van Wezel, J., Mikut, R., Strahle, U. & Nienhaus, G. U. 2015. An ensemble-averaged, cell density-based digital model of zebrafish embryo development derived from light-sheet microscopy data with single-cell resolution. *Sci Rep,* 5**,** 8601.

Kwan, K. M., Fujimoto, E., Grabher, C., Mangum, B. D., Hardy, M. E., Campbell, D. S., Parant, J. M., Yost, H. J., Kanki, J. P. & Chien, C. B. 2007. The Tol2kit: a multisite gateway-based construction kit for Tol2 transposon transgenesis constructs. *Dev Dyn,* 236**,** 3088-99.

Le Douarin, N. M., Creuzet, S., Couly, G. & Dupin, E. 2004. Neural crest cell plasticity and its limits. *Development,* 131**,** 4637-50.

Le Douarin, N. M. & Dupin, E. 2003. Multipotentiality of the neural crest. *Curr Opin Genet Dev,* 13**,** 529-36.

Le Lievre, C. S. & Le Douarin, N. M. 1975. Mesenchymal derivatives of the neural crest: analysis of chimaeric quail and chick embryos. *J Embryol Exp Morphol,* 34**,** 125-54.

Leung, A. Y., Mendenhall, E. M., Kwan, T. T., Liang, R., Eckfeldt, C., Chen, E., Hammerschmidt, M., Grindley, S., Ekker, S. C. & Verfaillie, C. M. 2005. Characterization of expanded intermediate cell mass in zebrafish chordin morphant embryos. *Dev Biol,* 277**,** 235-54.

Macosko, E. Z., Basu, A., Satija, R., Nemesh, J., Shekhar, K., Goldman, M., Tirosh, I., Bialas, A. R., Kamitaki, N., Martersteck, E. M., Trombetta, J. J., Weitz, D. A., Sanes, J. R., Shalek, A. K., Regev, A. & Mccarroll, S. A. 2015. Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. *Cell,* 161**,** 1202-1214.

Manoli, M. & Driever, W. 2012. Fluorescence-activated cell sorting (FACS) of fluorescently tagged cells from zebrafish larvae for RNA isolation. *Cold Spring Harb Protoc,* 2012.

Mazutis, L., Gilbert, J., Ung, W. L., Weitz, D. A., Griffiths, A. D. & Heyman, J. A. 2013. Single-cell analysis and sorting using droplet-based microfluidics. *Nat Protoc,* 8**,** 870-91.

Mckenna, A., Findlay, G. M., Gagnon, J. A., Horwitz, M. S., Schier, A. F. & Shendure, J. 2016. Whole-organism lineage tracing by combinatorial and cumulative genome editing. *Science,* 353**,** aaf7907.

Melby, A. E., Warga, R. M. & Kimmel, C. B. 1996. Specification of cell fates at the dorsal margin of the zebrafish gastrula. *Development,* 122**,** 2225-37.

Miller-Bertoglio, V. E., Fisher, S., Sanchez, A., Mullins, M. C. & Halpern, M. E. 1997. Differential regulation of chordin expression domains in mutant zebrafish. *Dev Biol,* 192**,** 537-50.

Piccolo, S., Sasai, Y., Lu, B. & De Robertis, E. M. 1996. Dorsoventral patterning in Xenopus: inhibition of ventral signals by direct binding of chordin to BMP-4. *Cell,* 86**,** 589-98.

Rodriguez, A. & Laio, A. 2014. Machine learning. Clustering by fast search and find of density peaks. *Science,* 344**,** 1492-6.

Row, R. H. & Kimelman, D. 2009. Bmp inhibition is necessary for post-gastrulation patterning and morphogenesis of the zebrafish tailbud. *Dev Biol,* 329**,** 55-63.

Row, R. H., Tsotras, S. R., Goto, H. & Martin, B. L. 2016. The zebrafish tailbud contains two independent populations of midline progenitor cells that maintain long-term germ layer plasticity and differentiate in response to local signaling cues. *Development,* 143**,** 244-54.

Sasai, Y., Lu, B., Steinbeisser, H., Geissert, D., Gont, L. K. & De Robertis, E. M. 1994. Xenopus chordin: a novel dorsalizing factor activated by organizer-specific homeobox genes. *Cell,* 79**,** 779-90.

Sasal, Y., Lu, B., Steinbelsser, H. & De Robertis, E. M. 1995. Regulation of neural induction by the Chd and Bmp-4 antagonistic patterning signals in Xenopus. *Nature,* 378**,** 419.

Schier, A. F. & Talbot, W. S. 2005. Molecular genetics of axis formation in zebrafish. *Annu Rev Genet,* 39**,** 561-613.

Schulte-Merker, S., Lee, K. J., Mcmahon, A. P. & Hammerschmidt, M. 1997. The zebrafish organizer requires chordino. *Nature,* 387**,** 862-3.

Shin, J., Berg, D. A., Zhu, Y., Shin, J. Y., Song, J., Bonaguidi, M. A., Enikolopov, G., Nauen, D. W., Christian, K. M., Ming, G. L. & Song, H. 2015. Single-Cell RNA-Seq with Waterfall Reveals Molecular Cascades underlying Adult Neurogenesis. *Cell Stem Cell,* 17**,** 360-72.

Storey, J. D. 2002. A direct approach to false discovery rates. Journal of the Royal Statistical Society. *J. Royal Stat. Soc.,* 64**,** 479-498.

Thisse, B., Pflumio, S., Fürthauer, M., Loppin, B., Heyer, V., Degrave, A., Woehl, R., Lux, A., Steffan, T., Charbonnier, X. Q. & Thisse, C. 2001. Expression of the zebrafish genome during embryogenesis. *ZFIN Direct Data Submission*.

Ting D, H. L., Jordan M 2011. An Analysis of the Convergence of Graph Laplacians. *arXiv:1101.5435*

Trapnell, C., Cacchiarelli, D., Grimsby, J., Pokharel, P., Li, S., Morse, M., Lennon, N. J., Livak, K. J., Mikkelsen, T. S. & Rinn, J. L. 2014. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat Biotechnol,* 32**,** 381-386.

Tzouanacou, E., Wegener, A., Wymeersch, F. J., Wilson, V. & Nicolas, J. F. 2009. Redefining the progression of lineage segregations during mammalian embryogenesis by clonal analysis. *Dev Cell,* 17**,** 365-76.

Van Der Maaten, L. 2014. Accelerating t-SNE using Tree-Based Algorithms. *Journal of Machine Learning Research,* 15**,** 3221-3245.

Van Der Maaten, L. & Hinton, G. 2008. Visualizing data using t-SNE. *Journal of Machine Learning Research,* 9**,** 2579-2605.

Warga, R. M., Kane, D. A. & Ho, R. K. 2009. Fate mapping embryonic blood in zebrafish: multi- and unipotential lineages are segregated at gastrulation. *Dev Cell,* 16**,** 744-55.

Warga, R. M. & Nusslein-Volhard, C. 1999. Origin and development of the zebrafish endoderm. *Development,* 126**,** 827-38.

Weinreb C, W. S., Klein Am 2016. SPRING: a kinetic interface for visualizing high dimensional single-cell expression data. *bioRxiv*.

Whitlock, K. E. & Westerfield, M. 2000. The olfactory placodes of the zebrafish form by convergence of cellular fields at the edge of the neural plate. *Development,* 127**,** 3645-53.

Woo, K. & Fraser, S. E. 1995. Order and coherence in the fate map of the zebrafish nervous system. *Development,* 121**,** 2595-609.

Y, H. 2006. Efficient, high-quality force-directed graph drawing. *Mathematica Journal,* 10**,** 37-71.

Y, K. 2009. Drawing Graphs by Eigenvectors: Theory and Practice. *Computers and Mathematics with Applications,* 49**,** 1867-1888.

Zhu Q, F. J., Huang J 2016. Weighted natural neighborhood graph: an adaptive structure for clustering and outlier detection with no neighborhood parameter. *Cluster Comput,* 19**,** 1385-1397.

Zilionis, R., Nainys, J., Veres, A., Savova, V., Zemmour, D., Klein, A. M. & Mazutis, L. 2017. Single-cell barcoding and sequencing using droplet microfluidics. *Nat Protoc,* 12**,** 44-73.

Acknowledgements

We thank A.Ratner for technical support, T.W. Hiscock, V.Savova, S.L.Wolock, S.Mekhoubad and R.M.W. for helpful discussions.

Funding: D.E.W. acknowledges support from an HHMI-LSRF postdoctoral fellowship and 1K99GM121852. A.M.K was supported by an Edward J Mallinckrodt Foundation Grant and a Burroughs Wellcome Fund CASI Award. S.G.M was supported by R01GM107733 and R01DC015478.

Competing Interests: S.G.M. D.E.W, C.W., Z.M.C., J.A.B.: none declared. AMK is a founder of 1Cell-Bio, Inc.

Data Availability: A web portal providing access to the single-cell and coarse-grained graphs is available at www.tinyurl.com/scZfish2018. Single-cell counts matrices and FASTQ files have been deposited in the National Center for Biotechnology Information's Gene Expression Omnibus (GEO), accession GSE112294.

Publication Information: This paper was published in its entirety in the journal Science on April 26th 2018, and can be found online at:

http://science.sciencemag.org/content/early/2018/04/25/science.aar4362

Chapter IV.

Concluding Perspectives

Authors: Zachary M. Collins [a]

[a] Department of Systems Biology, Harvard Medical School, Boston, MA 02115, USA.

Author contributions: Z.M.C. wrote the manuscript.

This thesis provides insights into the mechanisms of scale invariance in pattern formation and reveals new features of the molecular networks that govern patterning. In this work, we used new techniques in embryology, cutting edge genetic tools, and quantitative imaging to interrogate the mechanistic basis of patterning. Here I discuss the broader implications of this work and outline open questions that merit further exploration.

Signaling gradients and patterning robustness

Extracellular signaling molecules and their distributions are critical cues in embryonic patterning. Since the discovery of the first morphogen, a growing list of factors have been identified that regulate their signaling. The molecular logic of these regulators dictates fundamental mechanisms of patterning. In Chapter II we focused on the role of Scube2 in regulating Shh signaling during patterning of the ventral neural tube. Scube2 was particularly interesting to us due to the ventral patterning defects of the mutant and its unexpected expression in the dorsal neural tube. In our work we demonstrated that feedback regulation of Scube2 by Shh links morphogen signaling to morphogen spread. In expander-repressor systems, signaling downstream of the morphogen represses the expression of an expander, which non-cell-autonomously enhances spread of the morphogen (Ben-Zvi and Barkai, 2010; Shilo and Barkai, 2017). Expander-repressor-like systems are powerful contributors to signaling robustness, particularly of pattern scaling, as discussed in Chapter II.

Scube2 may have the most direct effect on morphogen spread of existing expander-like interactions, as it acts by binding Shh and promoting its release from cell membranes (Creanga et al., 2012; Tukachinsky et al., 2012). Previously reported expanders act by a variety of indirect mechanisms. The proposed expander ADMP was thought to lengthen BMP distributions by competing with BMP ligands for the binding of Chordin, which would inhibit ventral shuttling of the morphogen (Ben-Zvi et al., 2008). However, Sizzled was later determined to play a more prominent role in scaling early D-V axis patterning (Inomata et al., 2013). The mechanism of

Sizzled as an expander is even less direct than that of ADMP, so much so that to view Sizzled as an expander you must also view Chordin as the relevant graded signaling molecule (Ben-Zvi et al., 2014). The proposed Dpp expander Pentagone in the wing disc also acts on the morphogen indirectly, as it promotes turnover of Heparin Sulfate Proteoglycans (HSPGs) that normally inhibit Dpp diffusion (Ben-Zvi et al., 2011a; Hamaratoglu et al., 2011; Norman et al., 2016).

Outside of its role in scaling, Scube2's contribution to Shh signaling is unique. Scube2's solubility in the extracellular space is not found in other members of the Shh signaling pathway. Shh ligands themselves are known for being associated with the cell surface in the absence of release machinery (Petrov et al., 2017). In addition, many regulators of the Shh signaling pathway are thought to be membrane tethered or are transmembrane proteins, including HSPGs, Gas1, Boc, Cdo, and Patched. However, the hedgehog antagonist, Hhip, was recently found to act cell non-autonomously (Kwong et al., 2014). This long-range of effect was inhibited by tethering Hhip to the cell membrane, implying travel of Hhip away from secreting cells is important for this activity. Extracellular interactions with Hedgehog ligands are critical for their signaling. Interestingly, our data are consistent with a model in which Scube2 persistently binds Hedgehog ligands to facilitate their transport as a chaperone.

Scube2's diffusivity, combined with its binding of Hedgehog, makes Scube2 an excellent candidate Shh chaperone (Petrov et al., 2017). Scube proteins as chaperones may help to resolve the mismatch between Shh ligand's lipophilic nature and their long range of effect. While it is possible that Scube2 plays a role in lipid-shedding, this model is made less likely by the impaired function of membrane-tethered Scube2 we observed (Jakobs et al., 2014). Combined with the conflict of shedding observations with previous HPLC analysis and the findings of two independent groups, this hypothesis seems unlikely (Creanga et al., 2012; Tukachinsky et al., 2012). A model of Scube2 in which it acts only transiently at the cell surface of producing cells—either by enabling the formation of multimeric Shh complexes or lipid shedding—would have

interesting implications for its role as an expander. Expanders are often formalized as having a dose dependent reversible effect on morphogen spread, while a transient role of Scube2 in Shh multimeric complex formation or shedding would be localized and irreversible. Mathematical modeling may reveal interesting implications each proposed mechanism in Scube2-Shh feedback interactions during pattern scaling.

In my initial proposals, I aimed to characterize the interactions of Scube2 with Shh ligands more directly using Shh fluorescent fusion proteins. However, these experiments were hampered by the lack of functionality shown by Shh-FP fusions. Shh-moxNG knock-in zebrafish lines, which I developed based on previously reported semi-functional Shh-GFPs, were homozygotic lethal and seem to remain anchored at the surface of producing cells (Chamberlain et al., 2008).This was confirmed in mosaic mRNA injection experiments where Shha-Citrine ligands had starkly shorter effective ranges than unmodified Shha. To answer questions about the interactions between Hedgehog ligands and Scube2 proteins, more biochemical studies are required. Of particular interest would be a study on the persistency of Scube2-Shh complexes and interactions of Scube2 with other regulators of the Hedgehog signaling pathway.

Scube2's expression and role in patterning of other tissues also warrants further exploration. During my work, I observed *scube2* expression in the developing dorsal ear, where Shh signaling also instructs the specification of ventral fates. This expression pattern raises the possibility that Scube2 plays a similar role in ear patterning as what I uncovered in the neural tube. Tg(*scube2:scube2-moxNG*) was also found diffusing between cells of the ear and secreted into the lumen of the ear. This expression suggests that Scube2-Shh feedback relationships may be a common feature in Shh-patterning systems. Scube2 was also strongly expressed in the pharyngeal arches and neural crest cells, where Shh is also known to be important for patterning. Relationships between Scube2 and Shh may be important for shaping Shh signaling in each of these contexts.

Further exploration of the roles of additional Scube family proteins in Shh signaling is also warranted. While no triple mutant mouse has been made, mouse mutants of Scube1, Scube2, and Scube3 have each been analyzed independently (Fuchs et al., 2016; Lin et al., 2015; Tu et al., 2008). Scube2 and Scube3 mutant mice both have modest phenotypes in craniofacial development and their neural development has not been rigorously investigated (Fuchs et al., 2016; Lin et al., 2015). Interestingly, however, Scube1 has strong neural and craniofacial development phenotypes in mouse (Tu et al., 2008). The authors compared this mutant to the Noggin mutant and argued that Scube1's primary role may be in antagonizing BMP signaling. This was also the conclusion reached by one of the three original cloning papers for the zebrafish Scube2 mutant, but later findings showed that this was a misinterpretation of Scube2's role as a positive regulator of Hedgehog signaling. Notably, Noggin mouse mutants, to which Scube1 was compared, have defects in ventral neural patterning (McMahon et al., 1998). Defective ventral neural patterning and excessive BMP signaling in the neural tube is also the phenotype of mutant lines with a partial loss of Shh signaling, such as the zebrafish Scube2 mutant. Scube1 in mouse may serve a similar role to Scube2 in zebrafish, as it is expressed in only the dorsal somites during patterning of the posterior neural tube. This expression pattern suggestions Scube1 is repressed by Shh mediated patterning of the ventral mesoderm. Further exploration of the role Scube1 and other Scube family proteins play in Shh mediated patterning in mice is warranted. Given the dependence of zebrafish Shh signaling on Scube family proteins, it is likely that double or triple Scube mutant mice would show extremely severe Shh signaling defects.

In Chapter III, we explored the role of the secreted BMP antagonist, Chordin, in shaping the cell fate landscape through single cell RNA sequencing. As expected, we found that Chordin CRISPR mutants show large increases in ventral cell fate specification relative to controls. Additionally, we were able examine cell-fate specification in the shared neuro-mesodermal progenitor population in the tailbud and uncover the role of *chordin* expressing cells in this

system. Through this analysis we identified the transcriptional signature of both the Chordin-producing and Chordin-responding cells and generated a list of novel, uncharacterized genes that help define this sub population. As sequencing costs decrease, this strategy for characterizing mutant phenotypes may become a central mode of analysis. From a relatively simple loss of function experiment, we were able to gain new insights into the functions of a critical regulator of early patterning. This fine-grained analysis of cells in development opens a new window into the study of patterning mechanisms. Studying mutant lines in this way will be a powerful tool for disease modeling. For example, genetic causes of developmental disorders could be investigated in a model system by inducing the known mutation and measuring systemic effects on development.

One of the most interesting things we learned from these data was how cell type specification is canalized in embryogenesis. In this thesis, we have devoted significant time to discussing patterning robustness in the context of morphogen-mediated patterning systems and scale invariance. Another crucial component of patterning robustness is how transcriptional regulation and signaling networks constrain cells to certain specification paths. Our mapping of cell fate specification in Chordin mutants is an intriguing example of this. Although Chordin null embryos are deficient in a crucial process in early development, specification of the organizer, these embryos had no fundamentally “new” cell types. Instead, while the cell state landscape was shifted towards more ventral fates, they were still constricted to the same cell identities. This canalization of cell identity is another way by which embryos enforce robust patterning regimes and gives us insights into how the body plan may evolve. Constriction of cell fates along certain paths would permit mutations in evolution to readily affect cell proportion while still conserving the fundamental roles of cell types.

In Appendix 3, we set out to study the scaling of somite patterning. Our initial interest in this question was driven by a curiosity for how pattern scaling is achieved in a system less dependent on morphogen gradients. Instead, we illuminated how central a scaling morphogen

activity gradient is to the patterning of the somites. When we began characterizing the scaling of somite patterning, we expected that a more physical change, such as a decrease in tail elongation speed, would account for somite scaling. However, only PSM size, when a delay in morphological emergence is accounted for, correlates with somite size between embryos of different sizes and over time. This led us to explore how PSM size may inform somite size via known signaling gradients.

FGF gradients in the somites are thought to set the “wavefront” at which somites are specified, and their travel posteriorly is thought to be set by the speed of tail elongation. After our studies on the scaling of somite size with PSM size, we were curious whether this signaling gradient was dynamically scaling with PSM size. Staining for dpERK revealed that the FGF response gradient adjusts its length scale proportionally to the size of the PSM, both over time and between embryos of varying sizes. This dynamic scaling of the FGF signaling gradient helps refine the existing clock and wavefront model. By formalizing our predictions in a mathematical model, we were able to make predictions that recapitulated past experiments and enabled a novel test in the form of our “echo” experiment. These predictions and experiments confirmed the importance of a scaling gradient to somite patterning models. While we attempted to explore gradient scaling somewhat in Appendix 3, the mechanism of remained elusive. Further complicating our understanding was our recent observation that Wnt signaling gradients also scale with PSM size. Our experiments with somite transplantation led us to hypothesize that newly formed somites help regulate the FGF signaling gradient. Retinoic acid signaling was an obvious candidate for this role, as it is known to inhibit FGF signaling in the PSM and to be produced in the newly formed somites. However, we found that FGF gradients still scale in retinoic acid production morphants.

HSPGs are known to be important regulators of FGF and Wnt signaling (Balasubramanian and Zhang, 2016; Yan and Lin, 2007; Yan and Lin, 2009; Yu et al., 2009). Interestingly, the HSPG Glypican 5c is specifically and highly expressed in the anterior PSM

and newly formed somites in zebrafish (Gupta and Brand, 2013). Of all the Glypicans, Glypican 5 is the closest homologue of the fly gene Dally (Flybase). Dally is known to act as a co-receptor for Wnt and FGF; while Notum is known to cleave Dally causing the release of a mobile form of Dally which then inhibits morphogen signaling (Ayers et al., 2010; Giráldez et al., 2002; Kreuger et al., 2004). Moreover, Notum is known to cleave Glypicans and enable their release in vertebrates (Traister et al., 2008). During somite formation, Notum's expression is restricted to the posterior tailbud and is known to inhibit Wnt signaling in the PSM (Flowers et al., 2012). I hypothesize that Notum, a secreted enzyme, diffuses over the length of the PSM and cleaves Glypican 5c in the anterior tailbud. In this hypothesized system, Notum-mediated cleavage of Glypican5c would increase over time as the PSM shrinks, thereby sharpening the Wnt signaling gradient, and potentially the FGF gradient as well. This would provide a dynamic source-sink like system that could regulate the distributions of both Wnt and FGF ligands. Given that source-sink gradients are also capable of some scaling, such a mechanism would plausibly explain scaling of both the Wnt and FGF activation gradients over time and between embryos of different sizes (Ben-Zvi and Barkai, 2010). HSPGs have been proposed to act as a sink for FGF gradients in previous studies, but the specific mechanism by which HSPGs inhibit morphogen spread locally was not clear (Yu et al., 2009). Furthermore, this model would explain the local inhibitory effect of transplanted somites on FGF signaling while accounting for the scaling of FGF signaling in retinoic acid signaling deficient embryos, both of which we observed in Appendix 3.

The intersection of gene expression databases, imaging, and genetic tools as a platform for studying development

The accessibility of digital gene expression databases has been a defining driver of this experimental work. During my dissertation research, I spent countless hours scouring the Zebrafish Information Network (ZFIN) to analyze gene expression patterns and inform my

thinking about developmental mechanisms. The broad availability of this information in the digital age vastly accelerated the pace of my work, helped instruct my thinking, and informed the hypothesis at the foundation of our work with Scube2, discussed in Chapter II. In Chapter III, we showed how single-cell RNA sequencing will streamline gene expression analysis. My hope is that the resources developed in this work will enable developmental biologists to rapidly probe the gene expression of their cell types of interest. The trove of gene expression information we generated in Chapter III can be readily combined with recent advances in forward genetics, such as CRISPR Cas9. With these tools, it is now possible for researchers to identify genes known to be co-expressed in their cell population, generate guide RNAs, and begin phenotyping zebrafish mutants within a single week. The mechanisms of these effects can then be interrogated *in vivo* using sophisticated imaging tools, as used throughout this work and further developed by collaborative work in Appendix 4. These tools will increase the pace at which we uncover the genetic, molecular, and cellular mechanisms of development, but it will require generations of diligent work and critical thinking by developmental biologists to integrate these findings into a systems-level understanding of how animals are built. I am proud through this work to have made a small contribution to this endeavor of human understanding.

References

Ayers, K. L., Gallet, A., Staccini-Lavenant, L. and Thérond, P. P. (2010). The Long-Range Activity of Hedgehog Is Regulated in the Apical Extracellular Space by the Glypican Dally and the Hydrolase *Notum. Dev. Cell* 18, 605–620.

Balasubramanian, R. and Zhang, X. (2016). Mechanisms of FGF gradient formation during embryogenesis. *Semin. Cell Dev. Biol.* 53, 94–100.

Ben-Zvi, D. and Barkai, N. (2010). Scaling of morphogen gradients by an expansion-repression integral feedback control*. Proc. Natl. Acad. Sci. U. S. A*. 107, 6924–9.

Ben-Zvi, D., Shilo, B.-Z., Fainsod, A. and Barkai, N. (2008). Scaling of the BMP activation gradient in Xenopus embryos. *Nature* 453, 1205–11.

Ben-Zvi, D., Pyrowolakis, G., Barkai, N. and Shilo, B.-Z. (2011). Expansion-repression mechanism for scaling the Dpp activation gradient in Drosophila wing imaginal discs. *Curr. Biol*. 21, 1391–6.

Ben-Zvi, D., Fainsod, A., Shilo, B.-Z. and Barkai, N. (2014). Scaling of dorsal-ventral patterning in the Xenopus laevis embryo. *Bioessays* 36, 151–6.

Chamberlain, C. E., Jeong, J., Guo, C., Allen, B. L. and McMahon, A. P. (2008). Notochord-derived Shh concentrates in close association with the apically positioned basal body in neural target cells and forms a dynamic gradient during neural patterning. *Development* 135, 1097–106.

Creanga, A., Glenn, T. D., Mann, R. K., Saunders, A. M., Talbot, W. S. and Beachy, P. A. (2012). Scube/You activity mediates release of dually lipid-modified Hedgehog signal in soluble form. *Genes Dev.* 26, 1312–25.

Flowers, G. P., Topczewska, J. M. and Topczewski, J. (2012). A zebrafish Notum homolog specifically blocks the Wnt/β-catenin signaling pathway. *Development* 139, 2416–25.

Fuchs, H., Sabrautzki, S., Przemeck, G. K. H., Leuchtenberger, S., Lorenz-Depiereux, B., Becker, L., Rathkolb, B., Horsch, M., Garrett, L., Östereicher, M. A., et al. (2016). The First Scube3 Mutant Mouse Line with Pleiotropic Phenotypic Alterations. *G3* (Bethesda). 6, 4035–4046.

Giráldez, A. J., Copley, R. R. and Cohen, S. M. (2002). HSPG modification by the secreted enzyme Notum shapes the Wingless morphogen gradient. *Dev. Cell* 2, 667–76.

Gupta, M. and Brand, M. (2013). Identification and Expression Analysis of Zebrafish Glypicans during Embryonic Development. *PLoS One* 8, e80824.

Hamaratoglu, F., de Lachapelle, A. M., Pyrowolakis, G., Bergmann, S. and Affolter, M. (2011). Dpp signaling activity requires Pentagone to scale with tissue size in the growing Drosophila wing imaginal disc. *PLoS Biol*. 9, e1001182.

Inomata, H., Shibata, T., Haraguchi, T. and Sasai, Y. (2013). Scaling of dorsal-ventral patterning by embryo size-dependent degradation of Spemann's organizer signals. *Cell* 153, 1296–311.

Jakobs, P., Exner, S., Schürmann, S., Pickhinke, U., Bandari, S., Ortmann, C., Kupich, S., Schulz, P., Hansen, U., Seidler, D. G., et al. (2014). Scube2 enhances proteolytic Shh processing from the surface of Shh-producing cells. *J. Cell Sci.* 127, 1726–37.

Kreuger, J., Perez, L., Giraldez, A. J. and Cohen, S. M. (2004). Opposing Activities of Dally-like Glypican at High and Low Levels of Wingless Morphogen Activity. *Dev. Cell* 7, 503–512.

Kwong, L., Bijlsma, M. F. and Roelink, H. (2014). Shh-mediated degradation of Hhip allows cell autonomous and non-cell autonomous Shh signalling. *Nat. Commun*. 5, 4849.

Lin, Y.-C., Roffler, S. R., Yan, Y.-T. and Yang, R.-B. (2015). Disruption of Scube2 Impairs Endochondral Bone Formation. *J. Bone Miner. Res.* 30, 1255–67.

McMahon, J. A., Takada, S., Zimmerman, L. B., Fan, C. M., Harland, R. M. and McMahon, A. P. (1998). Noggin-mediated antagonism of BMP signaling is required for growth and patterning of the neural tube and somite. *Genes Dev*. 12, 1438–52.

Norman, M., Vuilleumier, R., Springhorn, A., Gawlik, J. and Pyrowolakis, G. (2016). Pentagone internalises glypicans to fine-tune multiple signalling pathways. *Elife* 5,.

Petrov, K., Wierbowski, B. M. and Salic, A. (2017). Sending and Receiving Hedgehog Signals. *Annu. Rev. Cell Dev. Biol.* 33, 145–168.

Shilo, B.-Z. and Barkai, N. (2017). Buffering Global Variability of Morphogen Gradients. *Dev. Cell* 40, 429–438.

Traister, A., Shi, W. and Filmus, J. (2008). Mammalian Notum induces the release of glypicans and other GPI-anchored proteins from the cell surface. *Biochem. J*. 410, 503–511.

Tu, C.-F., Yan, Y.-T., Wu, S.-Y., Djoko, B., Tsai, M.-T., Cheng, C.-J. and Yang, R.-B. (2008). Domain and Functional Analysis of a Novel Platelet-Endothelial Cell Surface Protein, SCUBE1. *J. Biol. Chem*. 283, 12478–12488.

Tukachinsky, H., Kuzmickas, R. P. P., Jao, C. Y. Y., Liu, J. and Salic, A. (2012). Dispatched and Scube Mediate the Efficient Secretion of the Cholesterol-Modified Hedgehog Ligand. *Cell Rep*. 2, 308–320.

Yan, D. and Lin, X. (2007). Drosophila glypican Dally-like acts in FGF-receiving cells to modulate FGF signaling during tracheal morphogenesis. *Dev. Biol*. 312, 203–216.

Yan, D. and Lin, X. (2009). Shaping morphogen gradients by proteoglycans. *Cold Spring Harb. Perspect. Biol.* 1, a002493.

Yu, S. R., Burkhardt, M., Nowak, M., Ries, J., Petrásek, Z., Scholpp, S., Schwille, P. and Brand, M. (2009). Fgf8 morphogen gradient forms by a source-sink mechanism with freely diffusing molecules. *Nature* 461, 533–6.

Appendix 1.

Supplemental Information for Chapter II

Supplemental Figures:

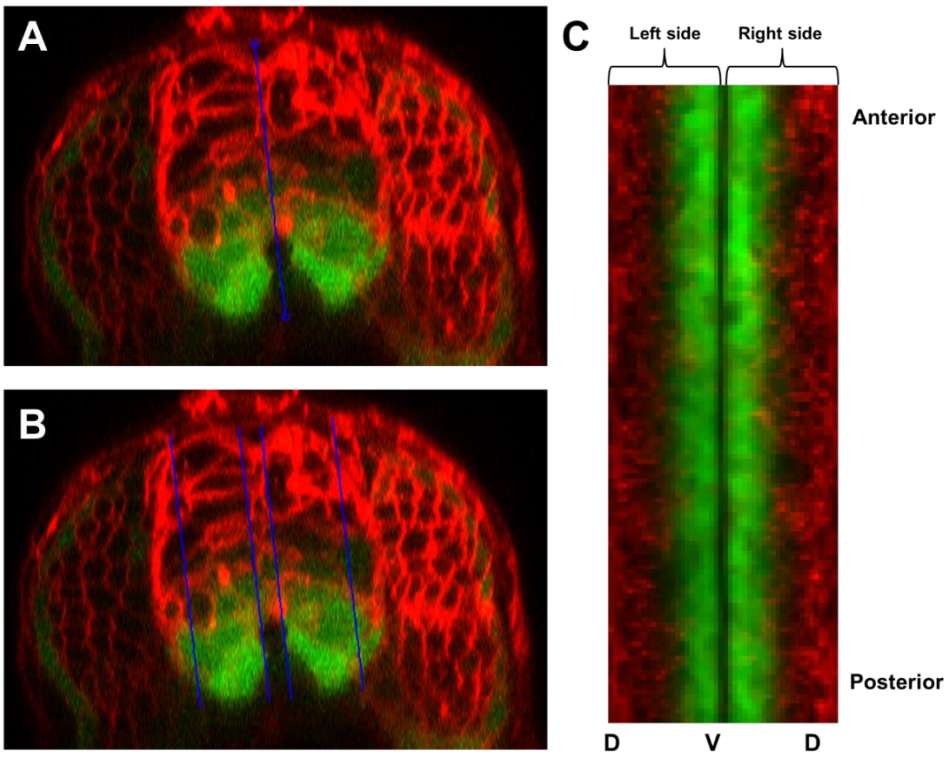


Figure S2.1 Segmentation and collection of neural imaging data

(A) Image of a 20 hpf tg(*ptch2:kaede*) reporter embryo undergoing selection of the "axis of reflection" which serves to mark a measurement of Dorsoventral height and separate the two halves of the neural tube. These positions are picked by the user, first by picking the bottom of the floor plate cell, then inputting the top coordinate of the roof plate cell. (B) Image of a 20 hpf tg(*ptch2:kaede*) reporter embryo after a user has selected the proper width of the spinal cord. The algorithm then calculates how much imaging data to collect based on a ratio which avoids mature neurons and the lumen of the spinal cord. (C) After collection of average intensities in each bin, data is stored as shown. Average profiles for generating distribution plots and

Supplemental Figure 2.1 (Continued) segmenting domains are gathered by averaging these data along the A-P axis for both halves of the neural tube.

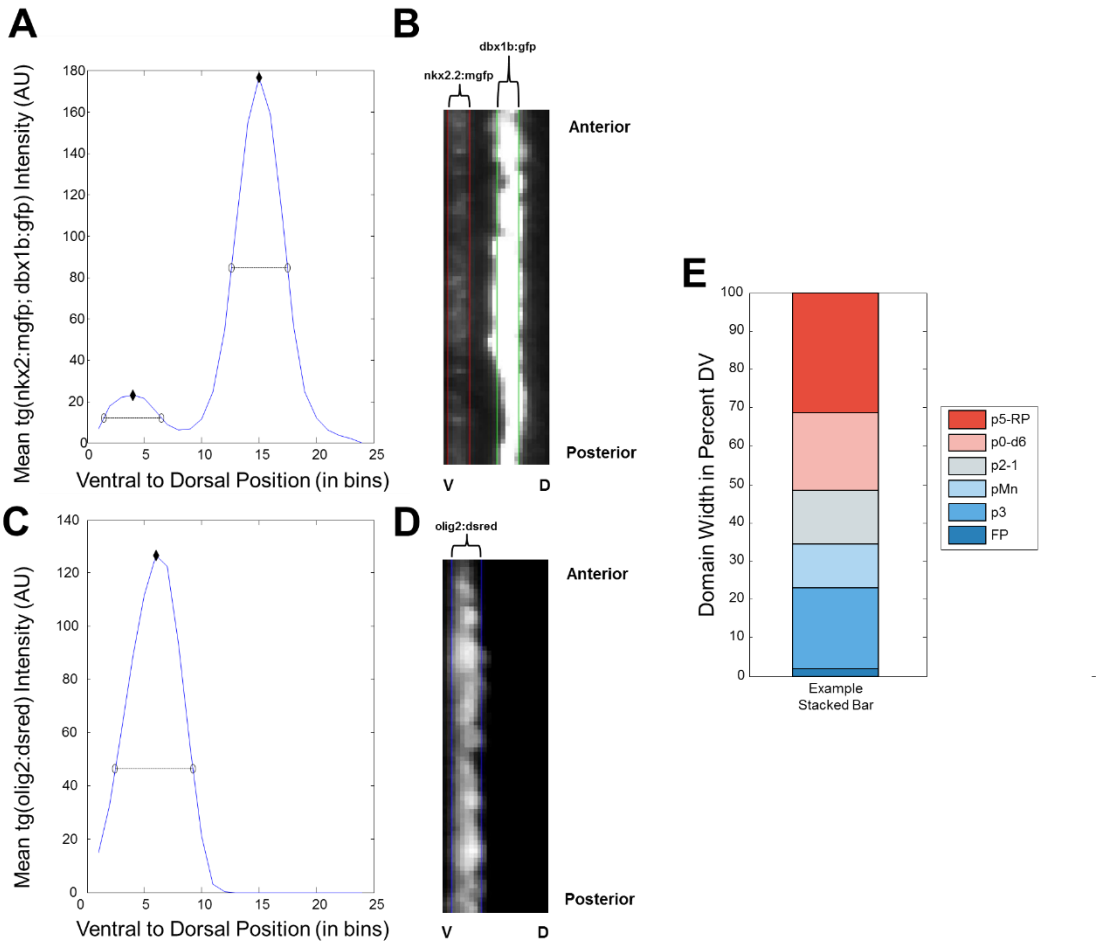


Figure S2.2 Progenitor domain width determination and bar plot generation

(A) Averaged image intensity profiles from the green channel of both sides of a segmented neural tube from a tg(*dbx1b:gfp;olig2:dsred;nkx2.2a:memgfp*) embryo. Black diamonds represent peaks found by a peak finding algorithm, while open circles and lines show the calculated domain boundaries and width for this embryo based on the universally applied threshold in this dataset. Thresholds are determined by 50% of average peak intensity of the control population for each dataset. (B) Example domain determination of *nkx2.2a* and dbx1b+

Supplemental Figure 2.2 (Continued) cells. Red lines mark the predicted *nkx2.2a* domain, which correlates with the boundary of their fluorescence. Green lines mark the predicted width of the *dbx1b* domain which correlates well with visible fluorescence of this domain. Some anterior-posterior variability in domain size is observed. (C) Formatted as in part A, this plots the averages *olig2:dsred*+ intensity, peak, and determined width. (D) Example domain determination of *olig2*+ cells. Blue lines mark the predicted *olig2:dsred* domain, which correlates with the boundary of their fluorescence. Thresholds are determined by 25% of average peak intensity of the control embryos in each dataset. (E) Example stacked bar plot generated only from this embryo using calculated domain positions to determine domain sizes.

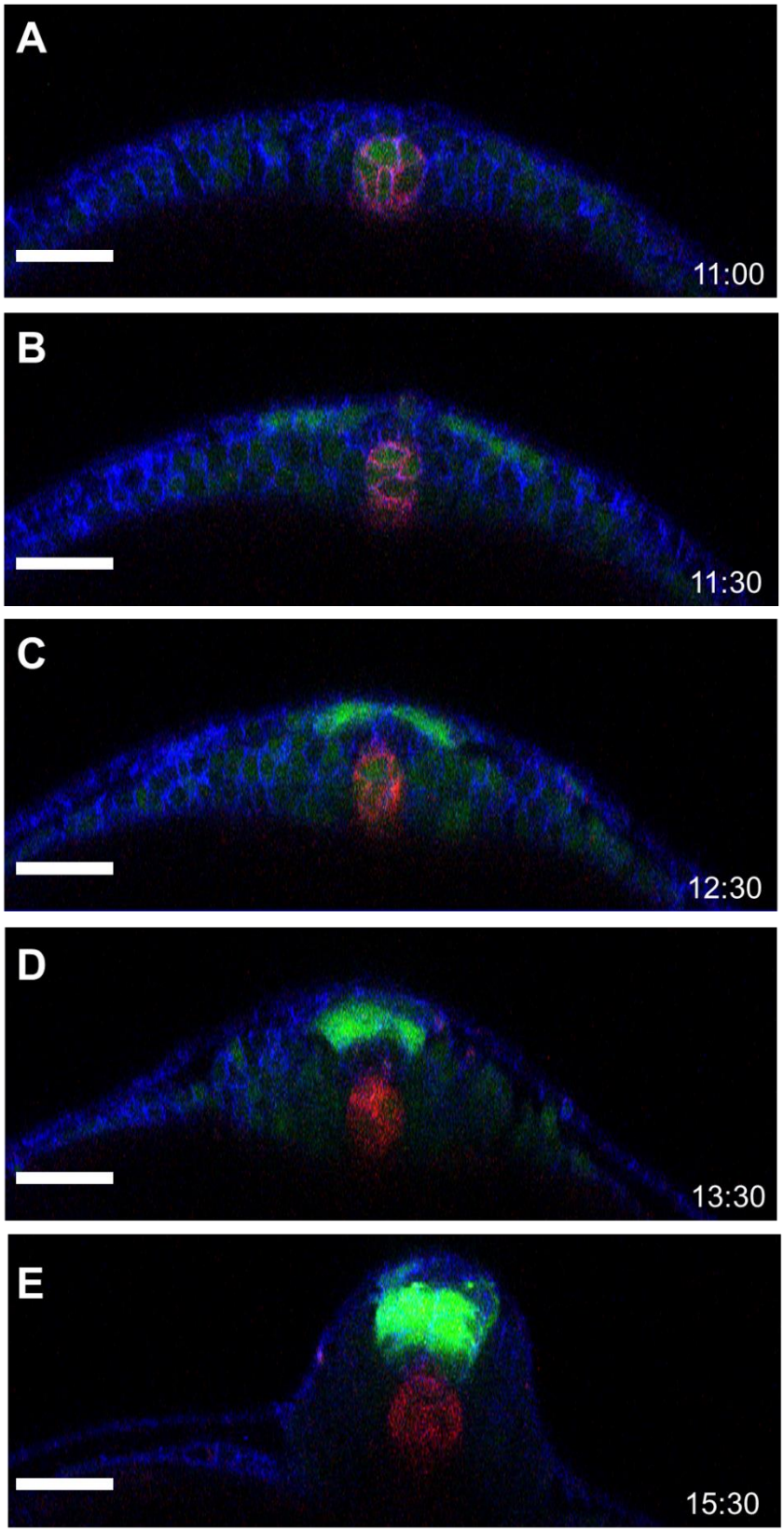


Figure S2.3 Time lapse imaging of tg(*scube2:moxng*)

Supplemental Figure 2.3 (Continued) Time-lapse images of tg(*scube2:moxng;shh:memCherry*) embryos which have been injected at the single cell stage with *membrane-mtagbfp2* mRNA. Time in hours post fertilization is displayed in the bottom right corner for each panel through the course of the movie. (A) At extremely early stages there is weak mesodermal expression of *scube2* in the notochord. (B) Expression of *scube2:moxng* is beginning to be seen in neural progenitors as the neural plate converges. (C) By 12.5 hpf a pronounced gap in inexpression of neural progenitors between *shh:mem-mcherry* and *scube2:moxng* cells is visible. (D) Expansion of the *scube2*+ domain dorsally is visible as cells continue to converge. (E) *scube2* expression is visibly constricted to the medial neural tube.

A

Gene Name	Observed Effect on Scube2 Expression
Pax6a	Suppresses SCUBE2
Pax6b	Suppresses SCUBE2
znf362b	Likely no effect
Irx3a	Unclear or likely no effect
Sox21b	Likely no effect
Nkx6.1	Unclear or no effect
Nkx6.2	Unclear or no effect
Nkx2.2a	Unclear or no effect
Nkx2.2b	Unclear or no effect
Pax7a	Unclear or no effect
Pax7b	Unclear or no effect

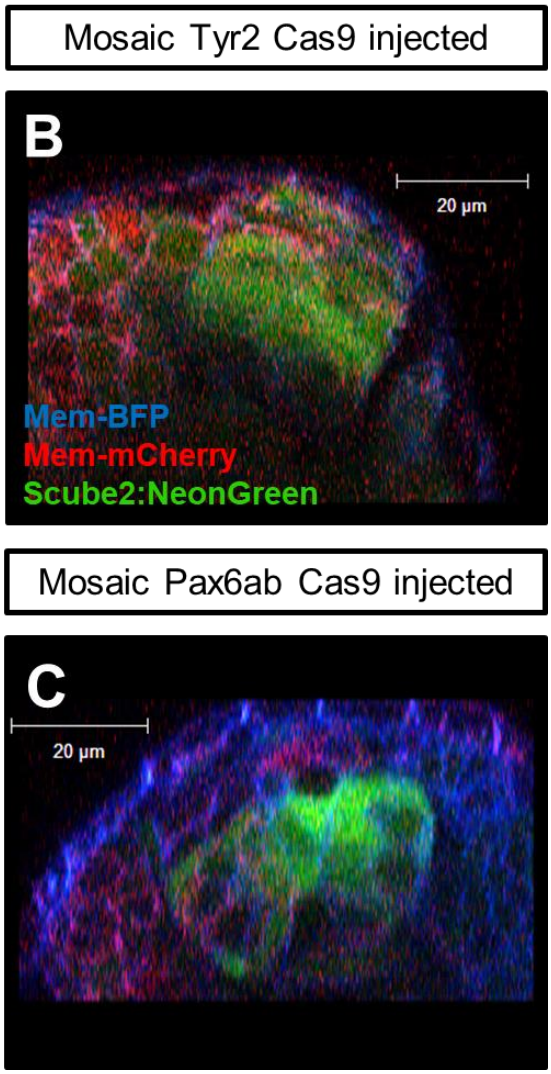


Figure S2.4 Results of a small scale CRISPR screen for regulators of scube2.

(A) List of genes selected for a screen for potential regulators of tg(*scube2:moxng*) and their phenotypes as gathered from imaging experiments. List was generated from known expression patterns, and in the case of *znf362b* and *sox21b*, correlation of expression with *scube2* in unpublished single cell sequencing data (discussed in Chapter III). (B-C) tg(*scube2:moxng*) embryos imaged at 18 hpf that were injected at the single cell stage with *mem-mtagbfp2* mRNA and injected at the 8-16 cell stage with *mem-mCherry* mRNA, Cas9 protein, and sgRNAs targeting either the tyrosinase pigment gene as a control (A) or *pax6a&b* (B).

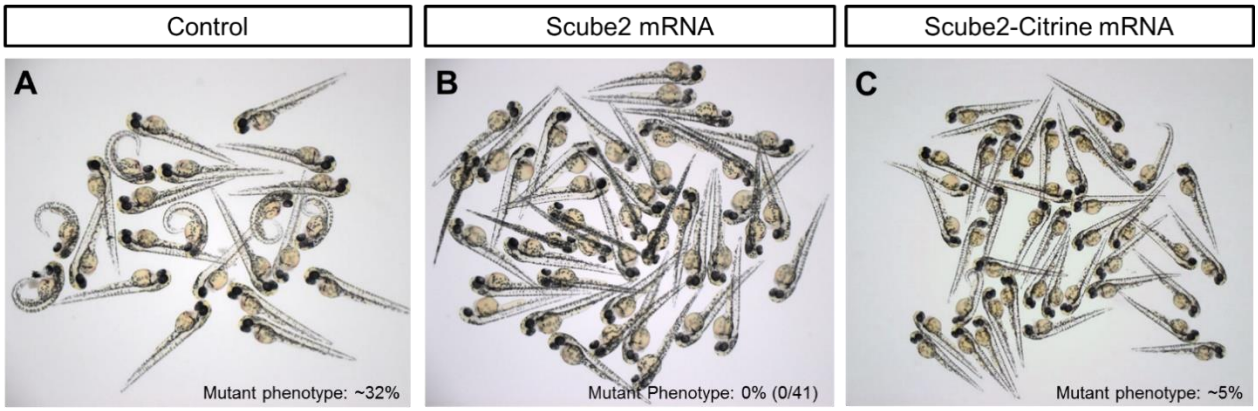


Figure S2.5 Rescue of scube2 CRISPR mutants with scube2 or scube2-citrine mRNA

(A) Results of a *scube2* mutant in-cross. The allele was generated by mutagenesis with CRISPR using three guides targeting *scube2* coding sequence (B) Embryos rescued by the injection of *scube2* mRNA co-injected with *mem-mcardinal* which were screened for being mem-mcardinal positive. (C) Embryos rescued by the injection of *scube2-citrine* mRNA co-injected with *mem-mcardinal* which were also screened for *mem-mcardinal* fluorescence.

Appendix 2.

Supplemental Information for Chapter III

Supplemental Tables Legends

Note: supplementary tables are included in supporting materials

Supplemental Table 3.1

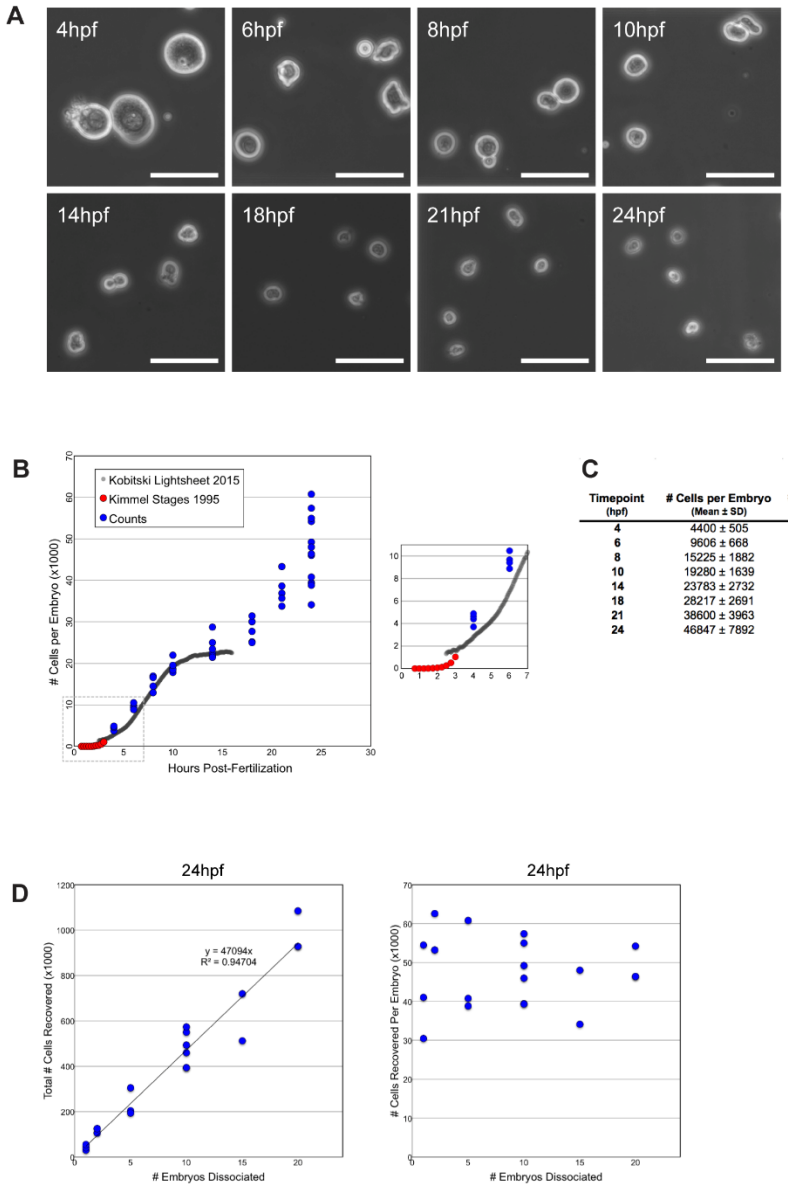
Summary of sequencing statistics for all inDrops RNA-seq libraries.

Supplemental Table 3.2

Table of significantly enriched marker genes and corresponding annotations for all 195 cell state clusters identified in the study. The top 20 positive differentially expressed genes (ranked by fold enrichment) determined by MAST (Finak et al., 2015) and Wilcoxon Rank Sum Test were determined by comparing cells of each cluster to all other cells from the same collection timepoint. Differentially expressed genes were identified using the FindAllMarkers routine in Seurat 2.2.0 according to the following criteria: (1) a log2-fold change >0.5, (2) Adjusted p-value <0.05. (3) >10% of cells in either test group must express at least one UMI.

Supplemental Table 3.3

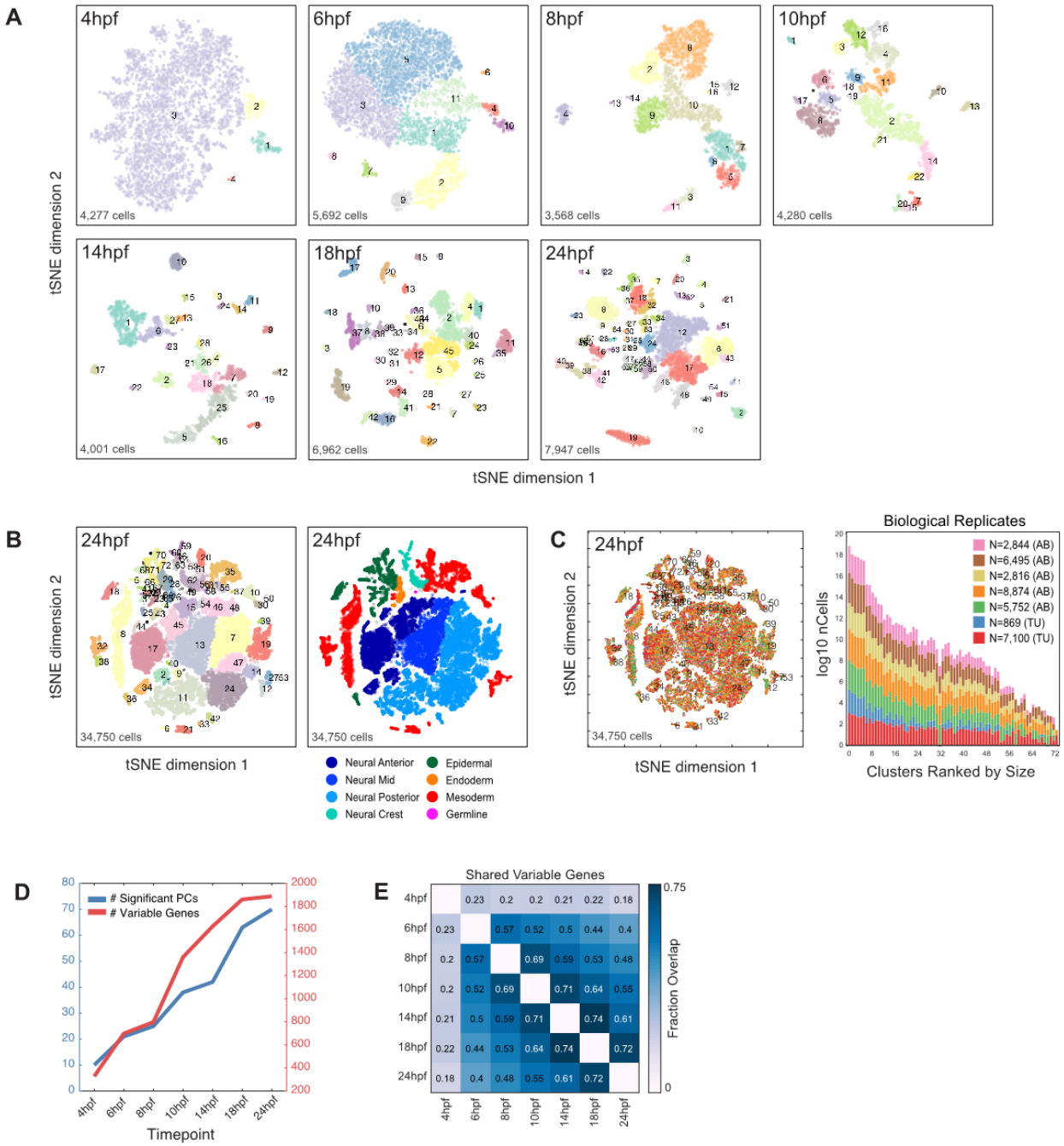
Sequences of primers used in this study.



Supplemental Figure 3.1

(A) Representative phase contrast images of cells dissociated from zebrafish embryos at the indicated timepoints. Scale bars: 50 μm. (B) Scatterplot of estimated total numbers of cells per embryo vs. time. Each blue dot represents a separate biological sample (cells dissociated from a set of 5-10 embryos, see Methods) manually quantified with a hemocytometer. Red and grey dots denote previously reported estimates from (Kimmel et al., 1995) and (Kobitski et al., 2015), respectively. (C) Estimation of scRNA-seq cell sampling depth. Table lists total cells per embryo

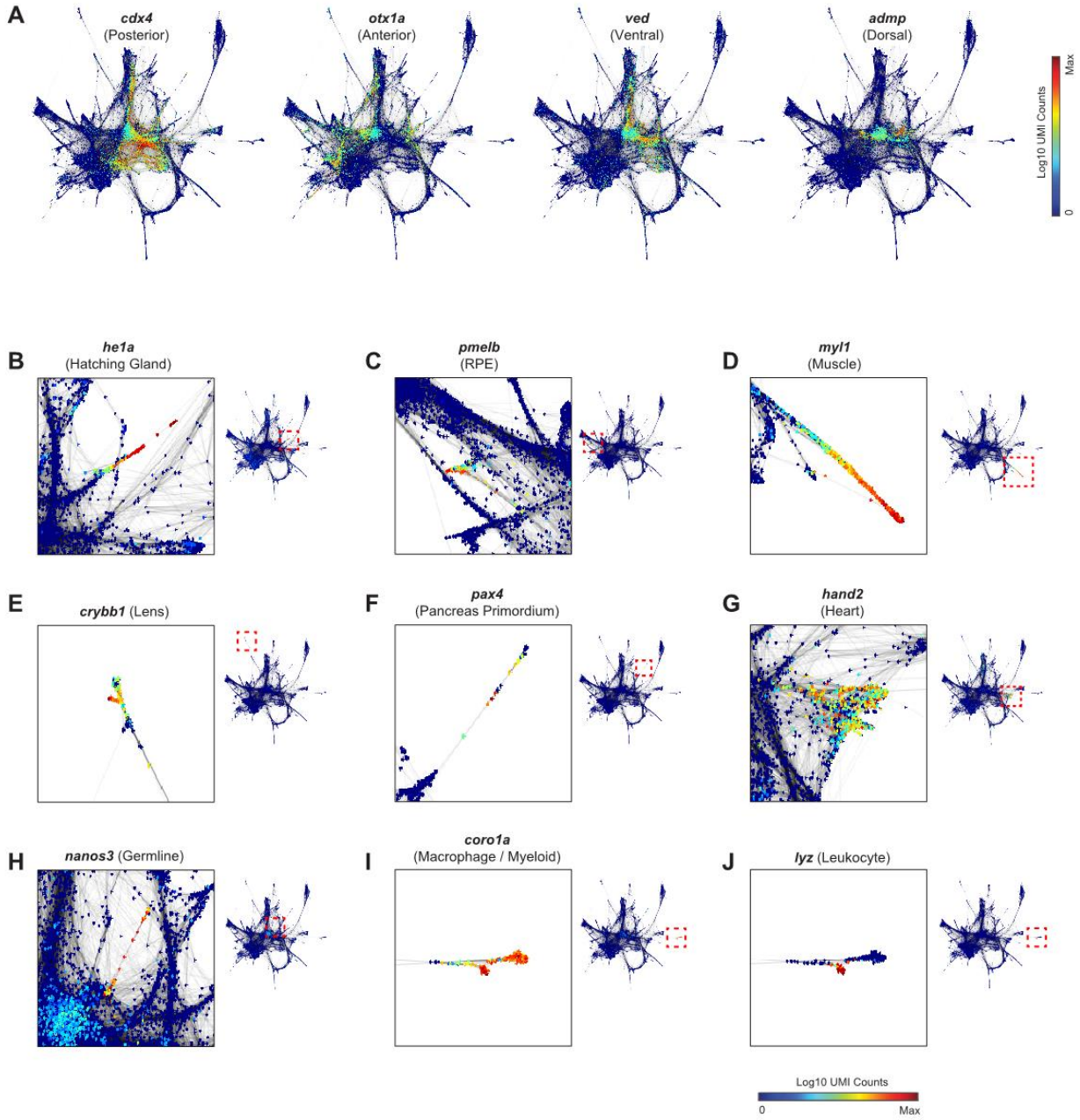
Supplemental Figure 3.1 (Continued) inferred from cell counts data in (B) alongside the total number of wild-type cells profiled by scRNA-seq, and associated sampling depths. (D) Demonstration of quantitative cell recovery. Estimated total cell yields resulting from the dissociation of sets of 1-20 embryos harvested at 24hpf.



Supplemental Figure 3.2

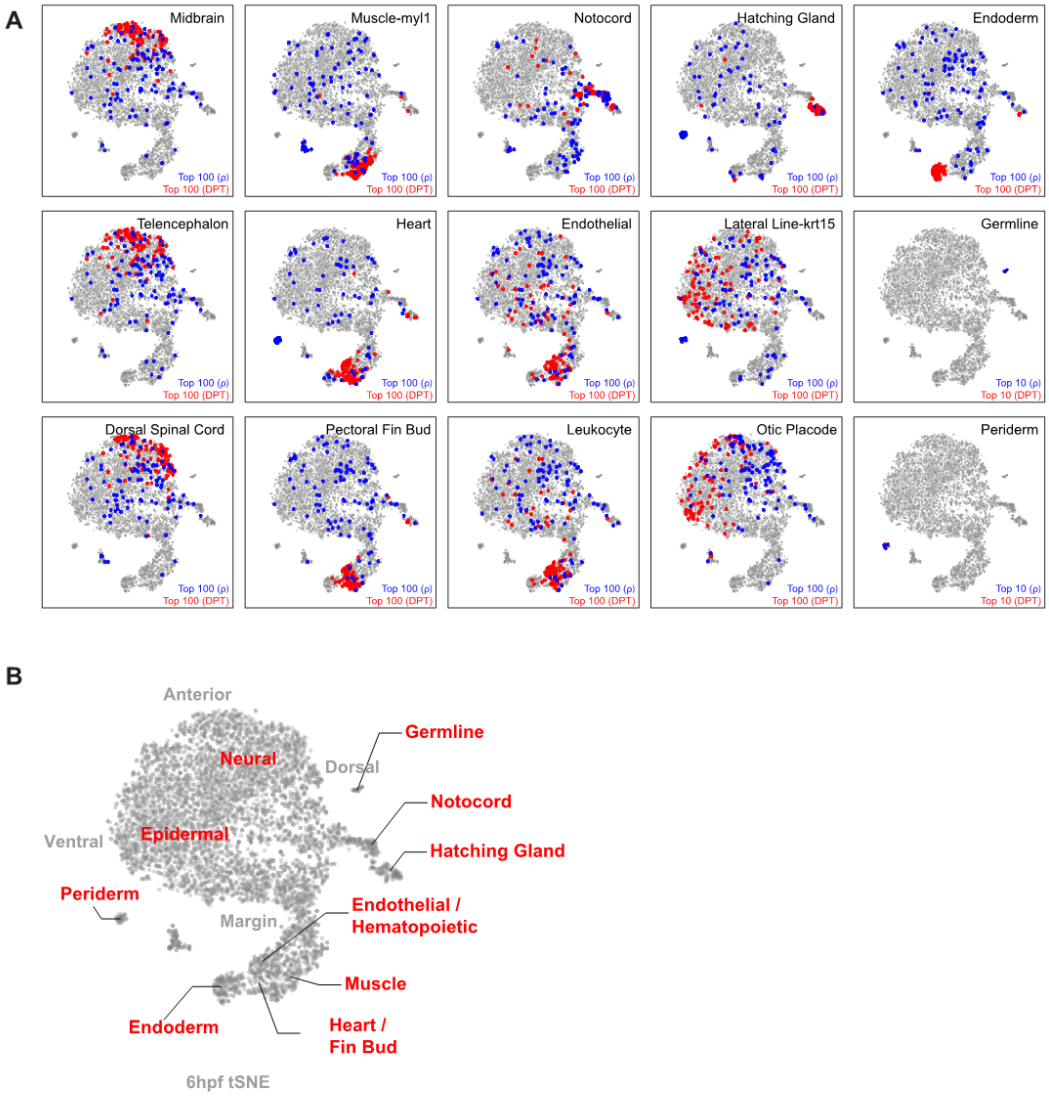
(A) tSNE embeddings of cells from 7 developmental timepoints, colored by local density cluster IDs. (B) tSNE embedding for the 24hpf timepoint, including additional biological replicates. Left: cells colored by local density cluster ID. Right: cells colored by inferred germ layer identity.(C) Left: tSNE embedding of cells for the 24hpf timepoint, colored by biological sample of origin.

Supplemental Figure 3.2 (Continued) Right: stacked heatmap depicting the contribution of each biological replicate to each cluster. Clusters are ranked in order of decreasing cell number.(D) Numbers of significant principal component dimensions and highly variable genes identified for each timepoint (see Methods). (E) Fraction of overlap in sets of identified variable genes between all timepoint pairs.



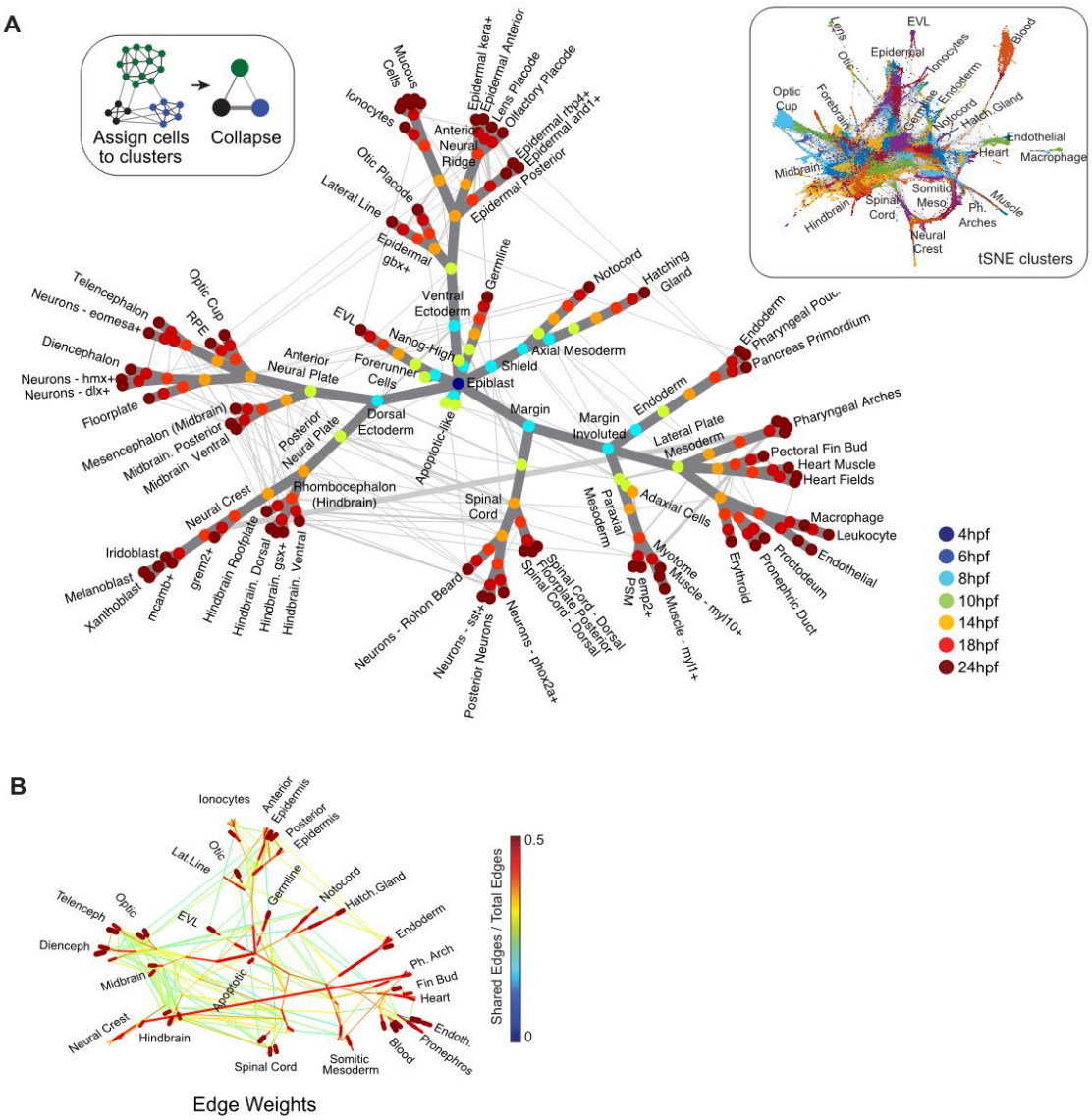
Supplemental Figure 3.3

Single-cell graphs (see also Figure 3.2). Nodes are colored by log10 expression counts for positional marker genes (A), and marker genes for particular cell/tissue types (B-J). Insets show zoomed regions marked by red boxes.



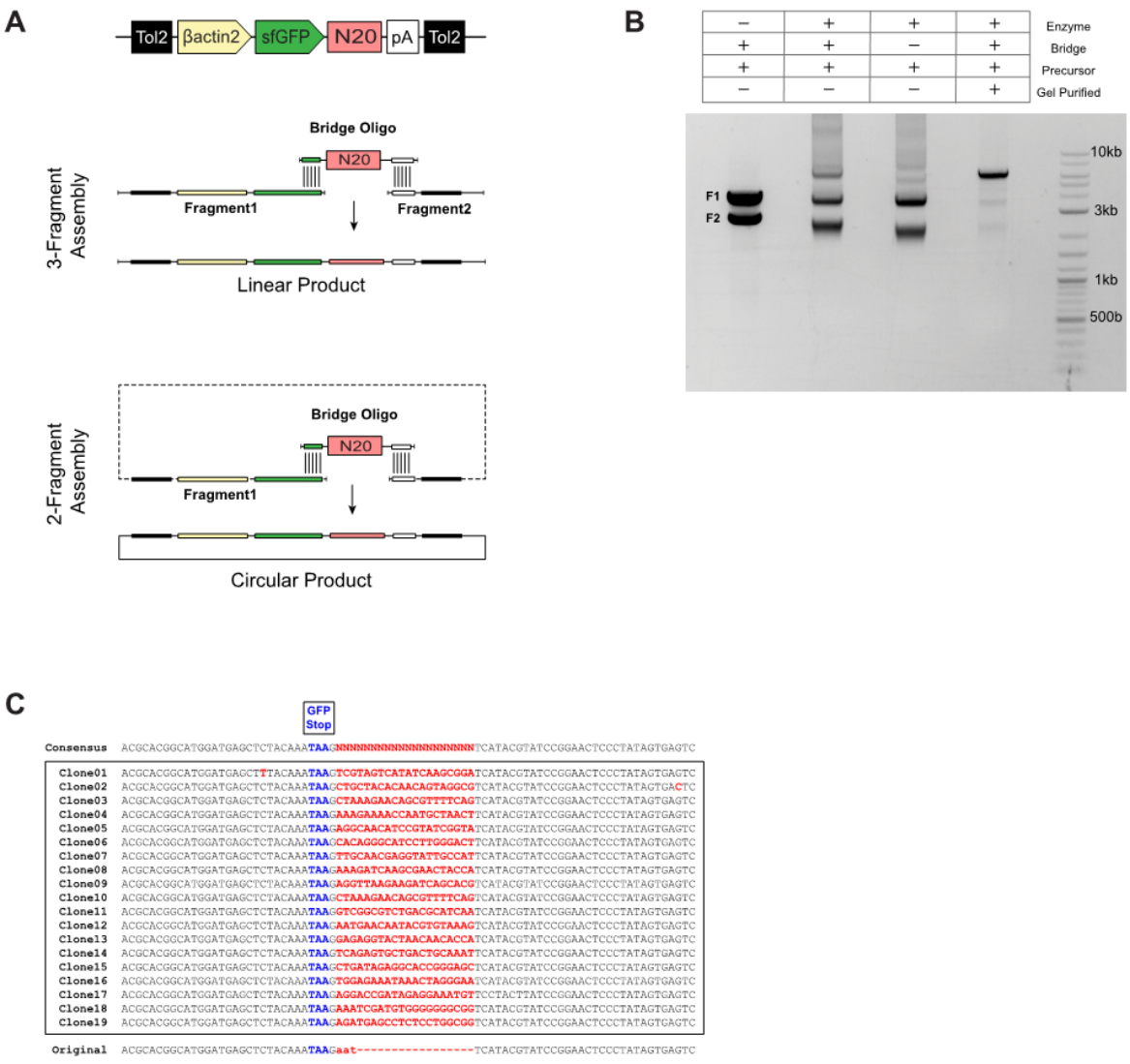
Supplemental Figure 3.4

(A) Additional tSNE maps of the 6hpf timepoint (see also Figure 3.3B), with the top 100 most proximal cells to specified 24hpf states highlighted in red (proximity calculated by single-cell graph diffusion distance) or blue (proximity calculated by correlation distance between gene expression profiles for each 6hpf cell and each 24hpf state). All regions of the original tSNE map are shown, including non-epiblast/non-hypoblast clusters. (B) Overview of the graph-predicted fate outcomes for each region of the 6hpf tSNE map.



Supplemental Figure 3.5

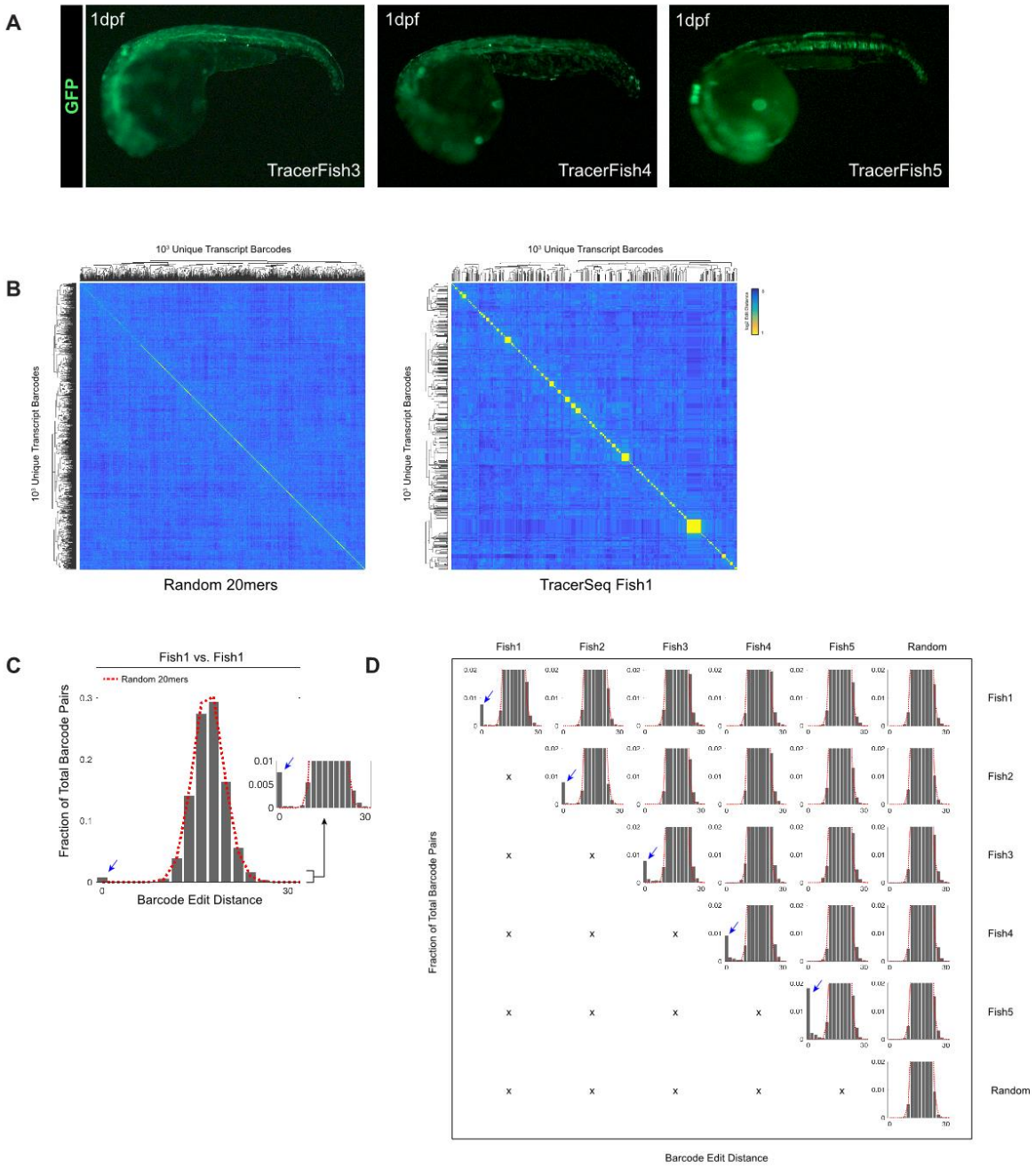
(A) Detailed overview of the coarse-grained graph depicted in Figure 3.3C. Inset: single-cell graph (same layout as Figure 3.2A) with nodes colored by local density tSNE clusters that were used for collapsing nodes. (B) Edge weights in the coarse-grained graph highlighted by color and edge thickness. Nodes are hidden from view.



Supplemental Figure 3.6

(A) Overview of isothermal / Gibson assembly strategies for constructing TracerSeq libraries. In a 3-fragment assembly, two dsDNA fragments are joined via a single-stranded “bridge oligo” that contains a random 20mer, flanked by two 20bp homology arms to fragments 1 and 2. In a 2-fragment assembly, a single fragment is circularized via assembly at an identical junction.(B) 1% TAE gel illustrating 2-fragment isothermal assembly of the ~6.7kb final product from the 4kb and ~2.7kb fragments.(C) Sanger sequencing of the fully assembled TracerSeq junction.

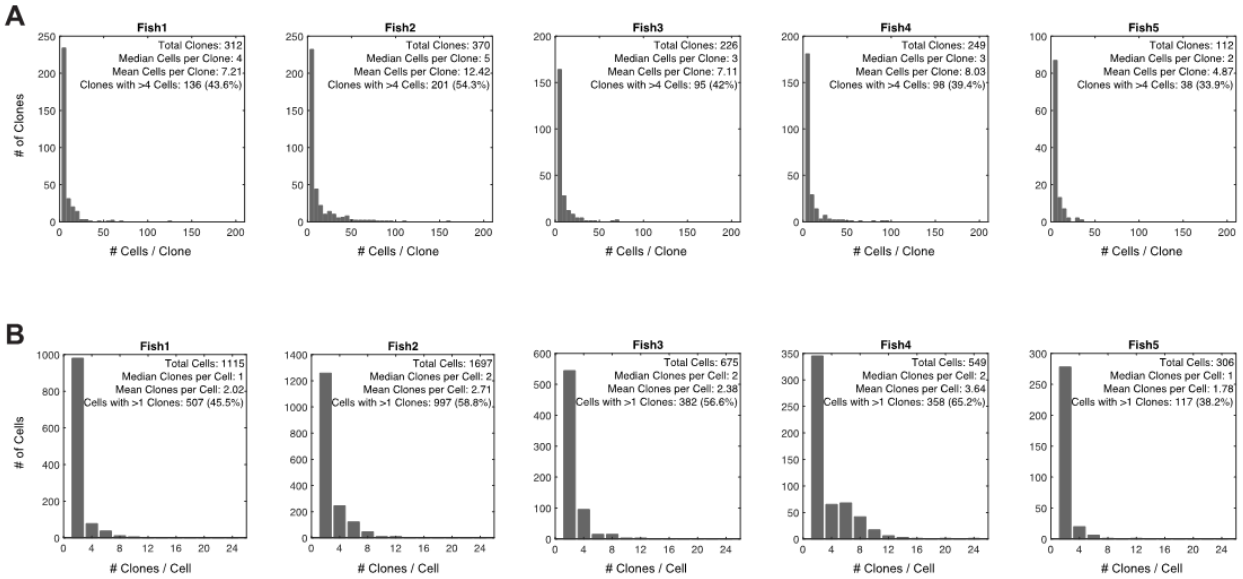
Supplemental Figure 3.6 (Continued) Sequencing reads from individual bacterial clones and their consensus sequence are indicated (note, the library is not passed through bacteria prior to use in fish). The N20 barcode region and GFP stop codon are marked in red and blue, respectively. Bottom, comparison to the original pMTB vector sequence.



Supplemental Figure 3.7

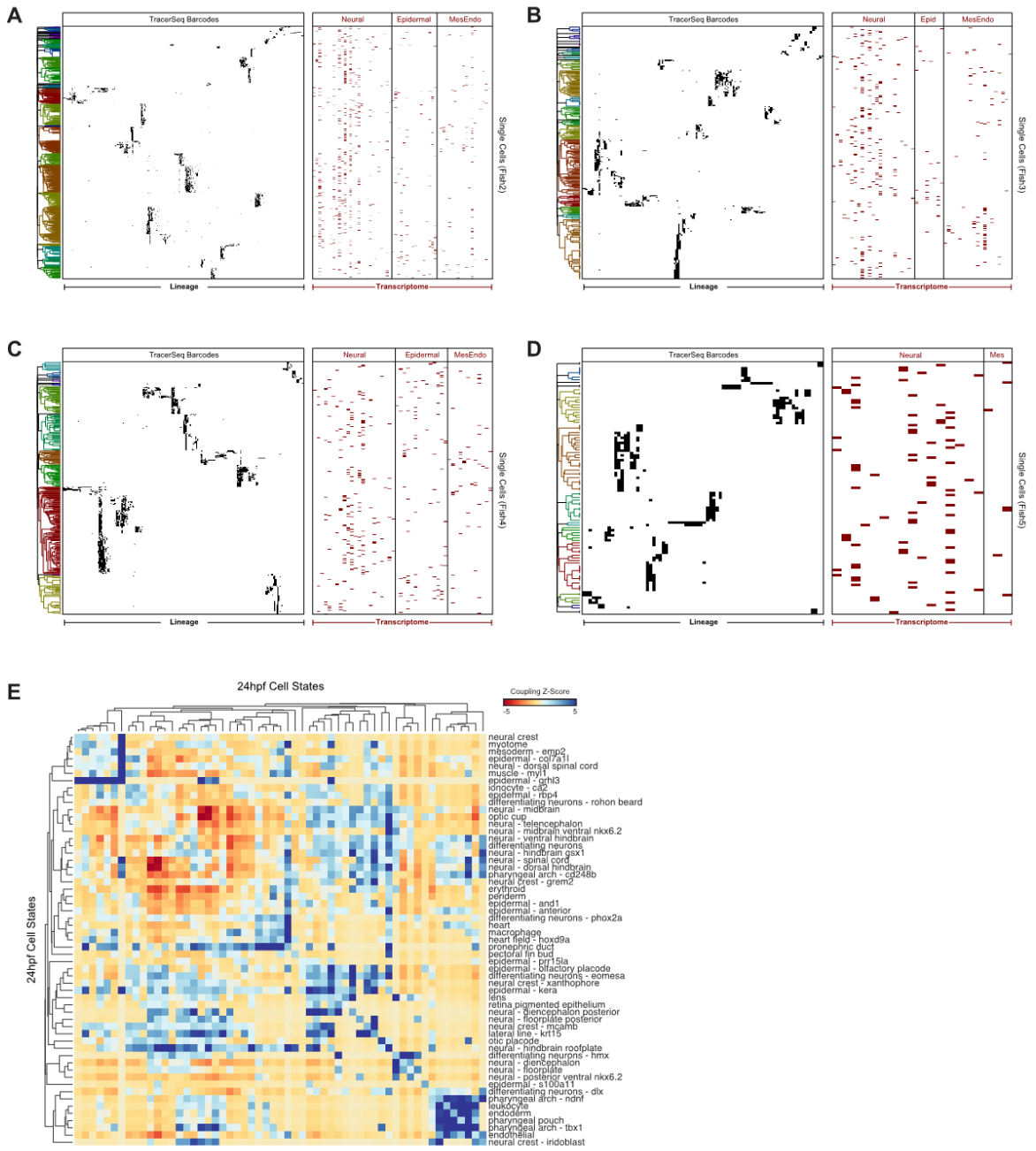
(A) Live fluorescent microscopy of representative TracerSeq embryos illustrating mosaic GFP expression. Anterior, left. (B) Heatmap of pairwise edit distances between 1,000 random 20mer sequences (left) and 1,000 randomly selected unique TracerSeq transcript barcodes from Fish1 (right). Heatmap rows and columns were hierarchically clustered using correlation distance and

Supplemental Figure 3.7 (Continued) average linkage. Clusters of sequences with small or zero edit distances (interpreted to derive from the same clonal insertion event) are evident for Fish1, but not for the random 20mers.(C) Histograms of all barcode edit distances plotted in (B), excluding self-barcode pairs. Grey bars denote pairwise distances between barcodes from Fish1, compared to distances between random 20mers (dotted red line). For Fish1, two histogram peaks are evident, with the first resembling the random distribution, and the second centered at an edit distance of zero (blue arrow). (D) Histograms comparing pairwise barcode edit distances within and between all 5 TracerSeq fish embryos vs. the random distribution. Peaks corresponding to identical or near identical barcodes (blue arrows) were only evident when comparing barcodes within a fish.



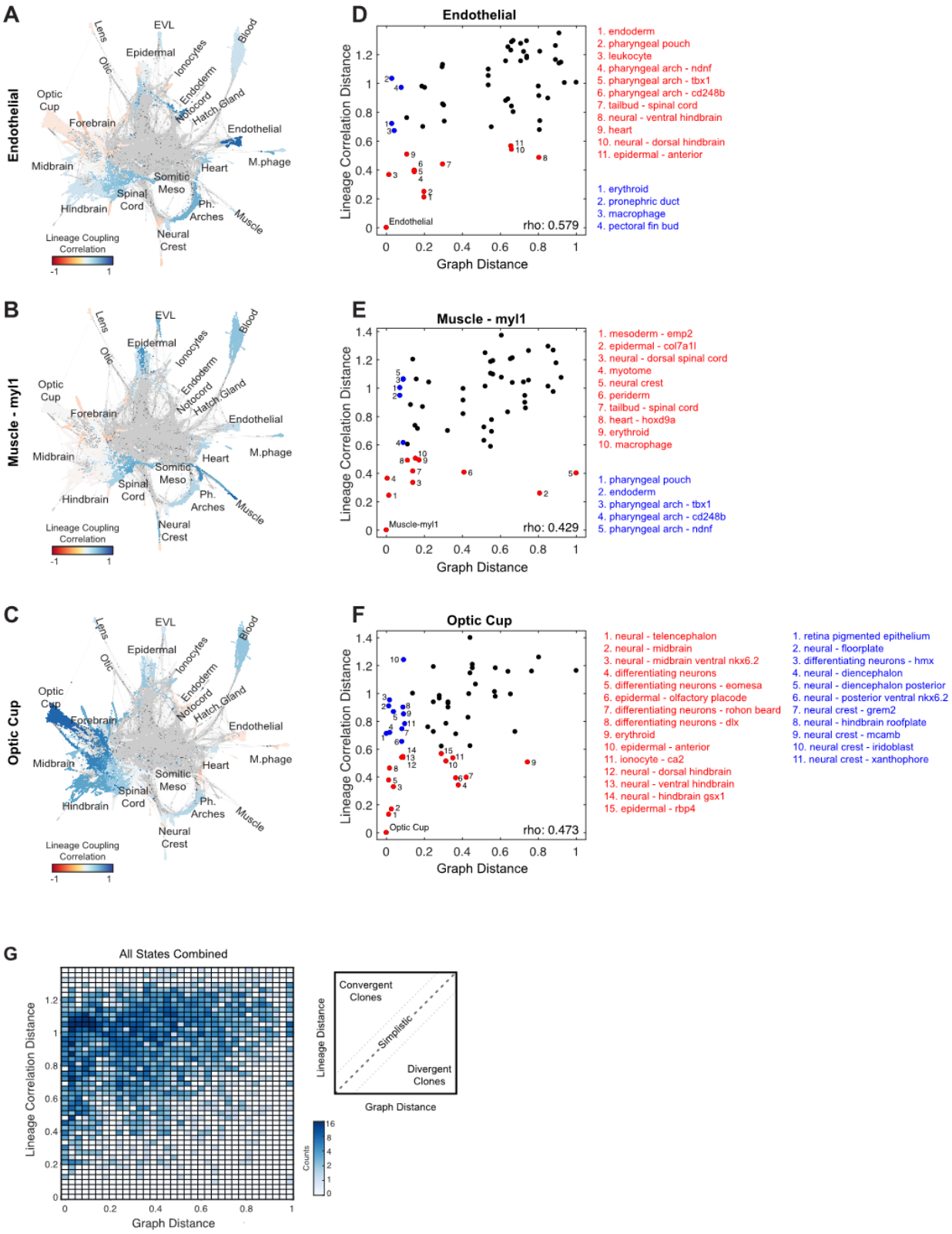
Supplemental Figure 3.8

(A) Histograms depicting the distribution in clone sizes (numbers of cells) over all 5 TracerSeq embryos. (B) Histograms depicting the distribution in numbers of clones detected per cell over all 5 TracerSeq embryos.



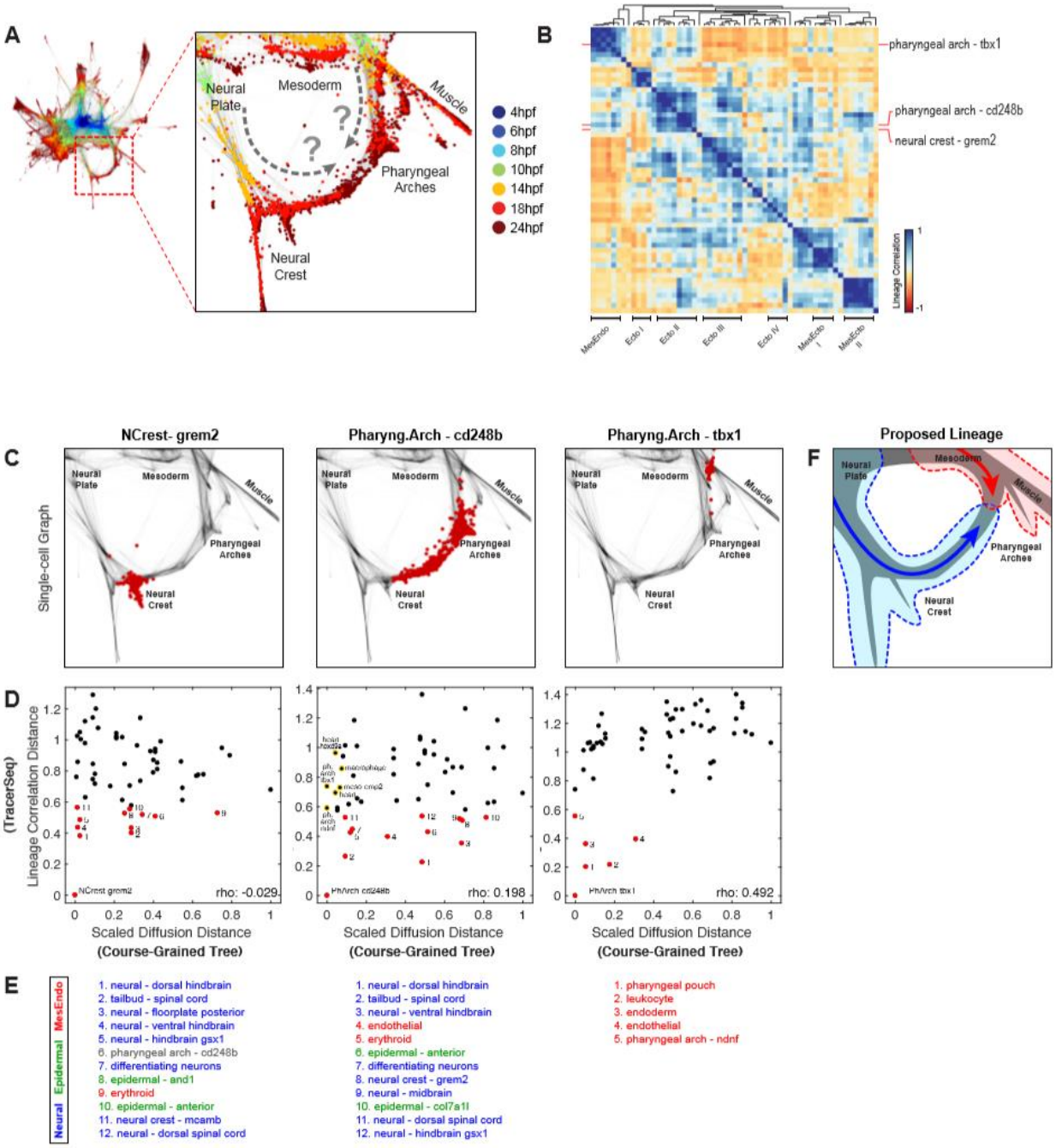
Supplemental Figure 3.9

(A-D) See also Figure 3.4B. Shown are additional heatmaps depicting dual lineage/transcriptome information for each cell in TracerSeq embryos 2-5. Heatmaps are clustered using Jaccard similarity and average linkage. (E) Heatmap of TracerSeq lineage coupling z-scores between each pair of 24hpf states. Z-scores are hierarchically clustered by correlation distance and average linkage.



Supplemental Figure 3.10

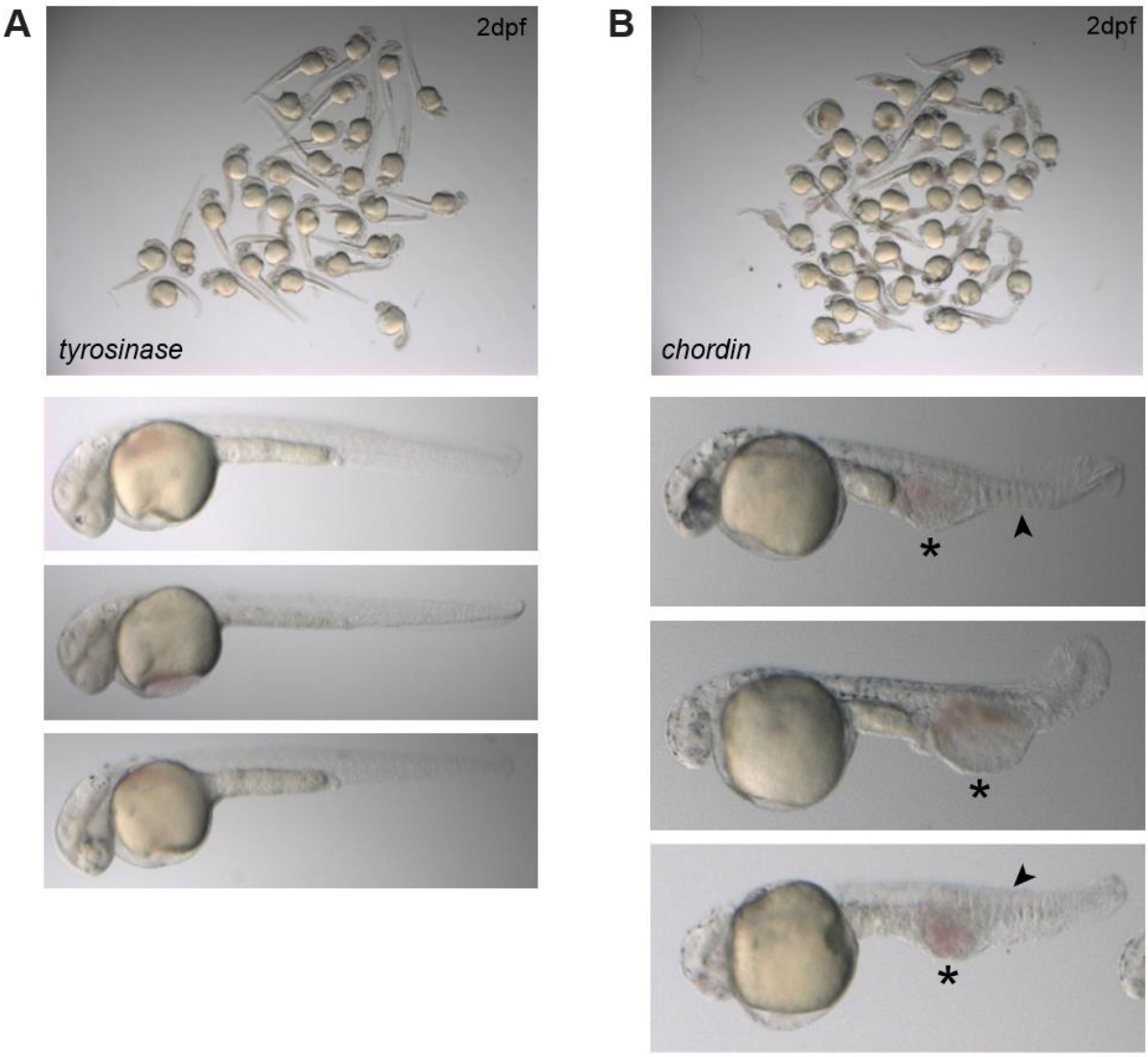
Supplemental Figure 3.10 (Continued) (A-C) Lineage coupling correlation scores mapped onto the single-cell graph. Color values correspond to a single row of the matrix in Figure 3.5A (D-F) Scatterplots comparing lineage coupling correlation distances vs. scaled diffusion distances predicted from the state tree. Plots match those appearing in Figure 3.5B, with additional points labeled. Red labels indicate states with significant non-zero TracerSeq lineage correlations (adj. p-value < 0.005), ranked by p-value. Blue labels indicate states in the upper left quadrant of the plot with short normalized graph diffusion distances (<0.1) and weak lineage correlations (correlation distance > 0.6), ranked by increasing diffusion distance. (G) Left, two-dimensional histogram of lineage correlation distance vs. diffusion distance scores over all state pairs. Number of scatter points appearing in each 2D bin is indicated by a heatmap. Right, proposed interpretation of various state-lineage relationships. A simplistic model, assuming tree-like relationships, predicts that lineage and state distances will be highly correlated. Cell state pairs that are related by lineage but separated by large distances in state space suggest the presence of “Divergent Clones”. Cell state pairs that are highly similar in state but unrelated by lineage suggest the presence of “Convergent Clones”.



Supplemental Figure 3.11

(A) Single-cell graph with close-up view of the pharyngeal arch / neural crest loop. Nodes are colored by collection timepoint. Arrows indicate two distinct hypothetical state trajectories into the loop with either neural plate or mesodermal origins. (B) Clustered heatmap of TracerSeq

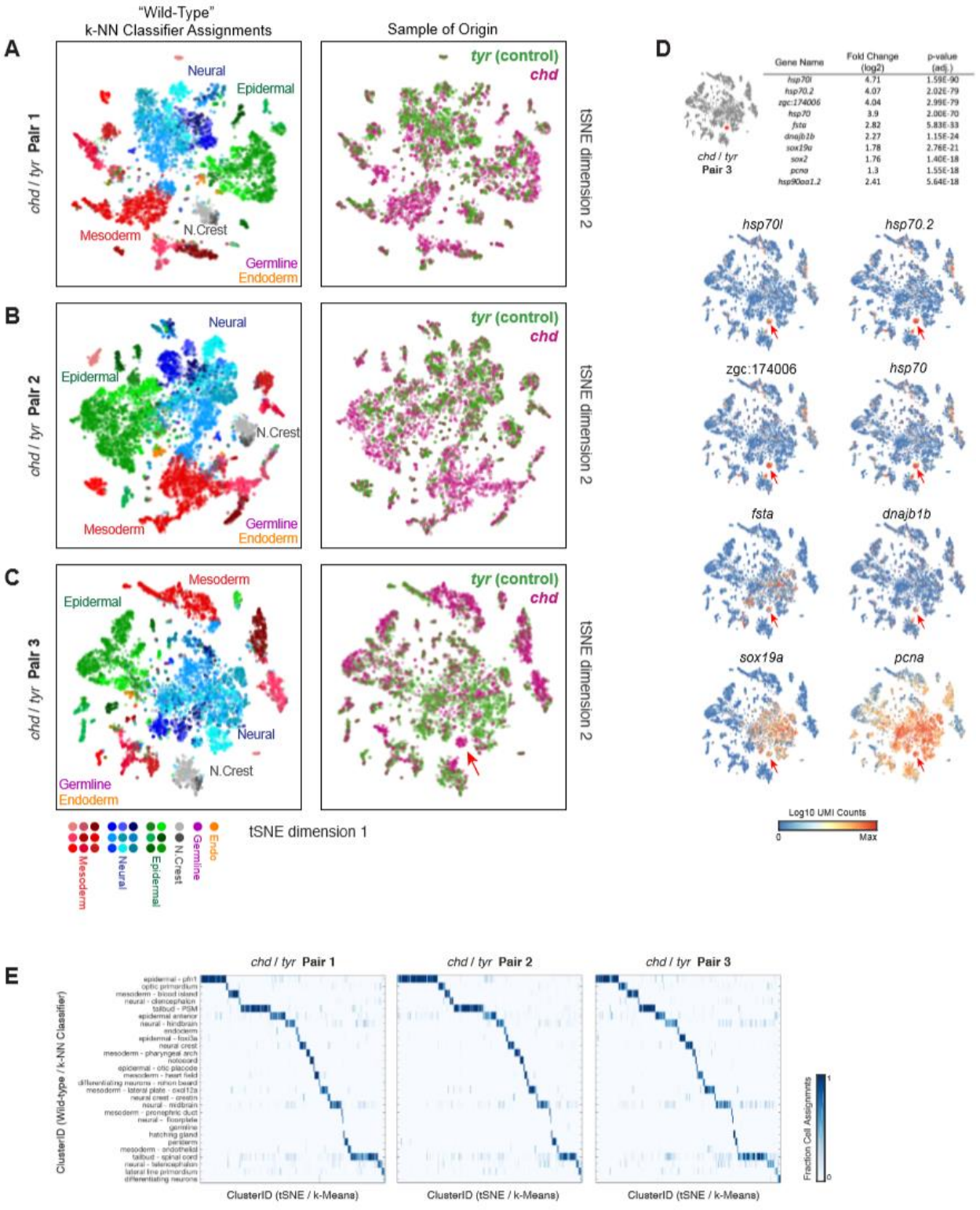
Supplemental Figure 3.11 (Continued) lineage correlation scores (from Figure 3.5A). State names are hidden except those for three select states participating in the pharyngeal arch / neural crest loop. Together these states span two independent TracerSeq groups (MesEndo and EctoII). (C) Close-up views of the single-cell graph. Graph edges are shown in dark grey. Red dots denote locations of cell nodes assigned to each state. (D) Scatterplots comparing lineage coupling correlation distances to scaled diffusion distances predicted from the state tree. Each plot is anchored on the states depicted in (B-C). Lineage correlation scores (y-axis) correspond to rows of the heatmap in (B). Middle panel: Additional relationships are highlighted in yellow between pharyngeal arch-cd248b and selected lateral plate mesodermal states. These states (plotted in the upper left quadrant) display weak lineage correlations to pharyngeal arch-cd248b, despite being in close transcriptional proximity on the state tree. (E) States with significant TracerSeq lineage correlations (adj. p-value < 0.005), ranked by p-value. Numbers correspond to labels in (D). State names are colored by germ layer of origin. (F) Proposed lineage model: A majority of cells in the pharyngeal arch / neural crest loop are neural-derived.



Supplemental Figure 3.12

Brightfield microscopy of live 2dpf embryos co-injected with CRISPR/Cas9 at the 1-cell stage with sgRNAs targeting *tyrosinase* (left) or *chordin* (right) genes. Following injection, damaged embryos were immediately identified and removed; all other embryos were retained. Embryos were chosen at random for inDrops sample preparation at stages corresponding to ~14-16hpf; remaining embryos were used to assess effectiveness of the CRISPR targeting. Top images: overview of the entire clutch including all embryos. Bottom images, representative examples of 3 different embryos from each condition. *tyrosinase*-targeted embryos displayed substantially

Supplemental Figure 3.12 (Continued) lower levels of visible pigment, but are otherwise normal. *chordin*-targeted embryos displayed the classic “ventralization” phenotype with small heads, enlarged somites (arrowheads), and overabundance of ventral mesodermal tissue (asterisks).



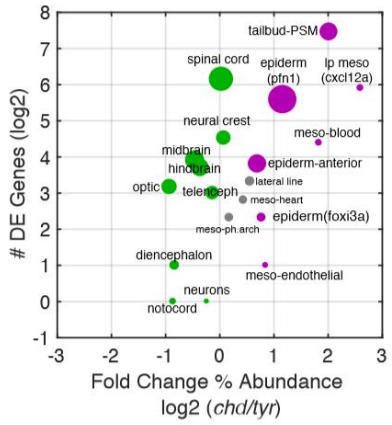
Supplemental Figure 3.13

Supplemental Figure 3.13 (Continued) (A-C) Single-cell tSNE maps for three stage-matched pairs of *chordin* and *tyrosinase* CRISPR-targeted samples (~14-16hpf). Left: cells colored by 14hpf (wildtype) state identities assigned by a kNN-classifier (see Methods). Close clustering of cells with similar color codes demonstrates the comparability of the *chordin* cell state landscape to the original wildtype landscape. Right: cells colored by genotype. Data points were down-sampled to an equal number of *chordin* and *tyrosinase* cells in each plot. Local changes in the proportions of the two genotypes demonstrate shifts in state abundances in the *chordin* and *tyrosinase* samples. Red arrow in (C) indicates a cluster in which no *tyrosinase* cells appeared. (D) Differential gene expression analysis of the "*chordin*-only" cell cluster identified in (C). Listed are the top 10 positively enriched genes for this cell cluster, determined by Wilcoxon rank-sum test (log2 fold change >1) ranked by adjusted p-value. Analysis was limited to transcripts with average expression level >25 transcripts per million (TPM). Counts overlays for selected differentially expressed genes confirm enrichment of transcripts associated with heat-shock/ stress response and a neural progenitor / dividing cell state (*sox19a*, *pcna*). (E) Structural similarities between *chordin*/*tyrosinase* and wild-type datasets, assessed by a confusion matrix. Rows: wild-type kNN-classifier assignments (n=28). Columns: k-Means cell cluster assignments derived from the tSNE map (k = 200, cityblock distance). Data are column-normalized and columns are sorted by row maxima. Cells within each k-Means cluster predominantly map to a single wild-type state.

A

Cell State Name	Relative Abundance (*chd/tyr*) Log2 FC	Adj. P-Value	# DE Genes State-Defining	*chd/tyr*	*chd/tyr* (Novel)	Top10 Novel Upregulated Genes (*chd/tyr*)
mesoderm - lateral plate - cxcl12a	2.6	0.0052	569	60	22 (36.7%)	*apoeb, ppp1r15a, id1, ddx21, mgst1.2, ta, gata3, histh1l, si:dkey-103d23.5*
tailbud - PSM	2.02	0.0001	712	176	80 (45.5%)	*apoeb, zgc:136826, zbtb16a, txnipa, ppp1r15a, id3, ier2, wu:fb55g09, id1, sesn3*
mesoderm - blood island	1.83	0.0187	1134	21	15 (71.4%)	*apoc1l, txnipa, apoeb, hsp70, ppp1r15a, h1f0, her6, fam46ba, tp53inp1, si:ch211-208m1.2*
epidermal - pfn1	1.16	0.0002	985	48	32 (66.7%)	*txnipa, h1f0, wu:fb55g09, fam46ba, hsp70, hsp70.2, ved, zgc:174710, pfkfb4l, LOC100330598*
hatching gland	1.13	0.0266	1407	0	0	
periderm	0.85	0.0184	854	0	0	
mesoderm - endothelial	0.85	0.0088	612	2	1 (50%)	*id1*
epidermal - foxi3a	0.77	0.0242	1032	5	2 (40%)	*h1f0, hsp70.2*
epidermal anterior	0.7	0.0077	340	14	13 (92.9%)	*txnipa, ppp1r15a, hsp70, hsp70.2, LOC103908986, h1f0, fam46ba, setd8b, tbx3a, szl*
lateral line primordium	0.56	0.0102	134	10	9 (90%)	*hsp70, hsp70.2, id3, ppp1r15a, h1f0, wu:fb55g09, hoxa9b, apoc1l*
differentiating neurons - rohon beard	0.5	0.0176	766	0	0	
epidermal - otic placode	0.48	0.0183	193	0	0	
mesoderm - heart field	0.44	0.033	476	7	6 (85.7%)	*hsp70.2, hsp70, ppp1r15a, fam46ba, LOC103911698, LOC103908986*
neural - floorplate	0.35	0.0314	276	0	0	
mesoderm - pharyngeal arch	0.17	0.0211	431	5	4 (80%)	*id1, hsp70.2, hsp70, LOC103908986*
germline	0.12	0.0557	430	0	0	
neural crest	0.07	0.0523	460	23	20 (87%)	*hsp70.2, hsp70, zgc:136826, txnipa, ppp1r15a, LOC103908986, drl, si:dkey-261j4.4, cebpb, mpp1*
neural crest - crestin	0.03	0.0549	204	0	0	
tailbud - spinal cord	0.03	0.05	491	71	43 (60.6%)	*zgc:136826, hsp70, hsp70.2, ppp1r15a, snai1a, ved, cdx1a, znf703, prdx1, vox*
endoderm	-0.09	0.049	264	0	0	
neural - telencephalon	-0.13	0.0223	501	8	4 (50%)	*hsp70.2, hsp70, ppp1r15a, dusp6*
mesoderm - pronephric duct	-0.24	0.0541	646	0	0	
differentiating neurons	-0.24	0.0504	1273	1	1 (100%)	*hsp70*
neural - hindbrain	-0.36	0.0091	695	13	9 (69.2%)	*hsp70, hsp70.2, ppp1r15a, fam46ba, pgam1a, trub2*
neural - midbrain	-0.46	0.0067	678	15	9 (60%)	*hsp70.2, hsp70, LOC101885394, pgam1a*
neural - diencephalon	-0.83	0.0188	429	2	2 (100%)	*hsp70, hsp70.2*
notocord	-0.86	0.0328	1771	1	1 (100%)	*ppp1r15a*
optic primordium	-0.93	0.015	653	9	7 (77.8%)	*yap1, ftr82, slit2, hsp70, ntn1a*

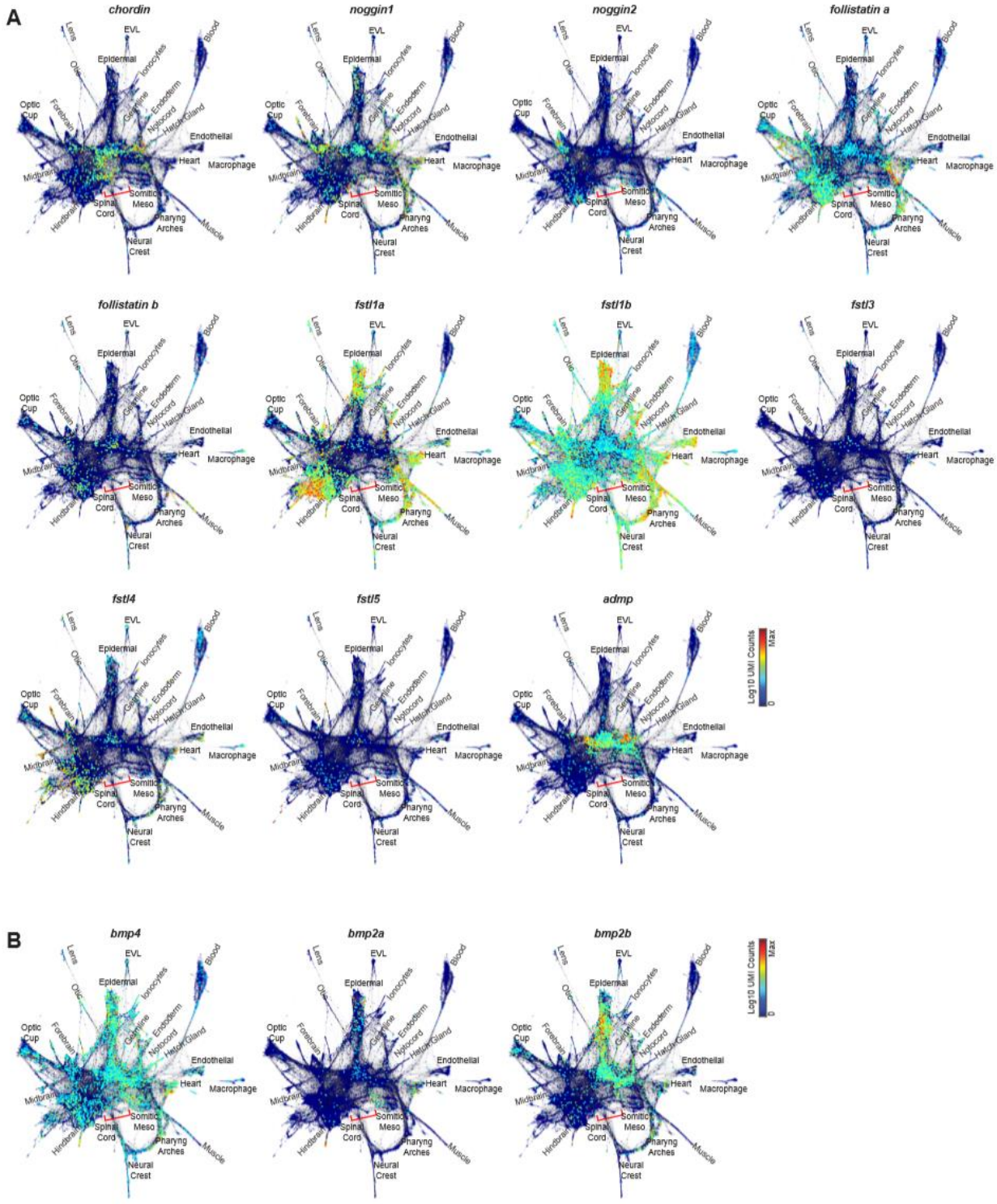
B



Supplemental Figure 3.14

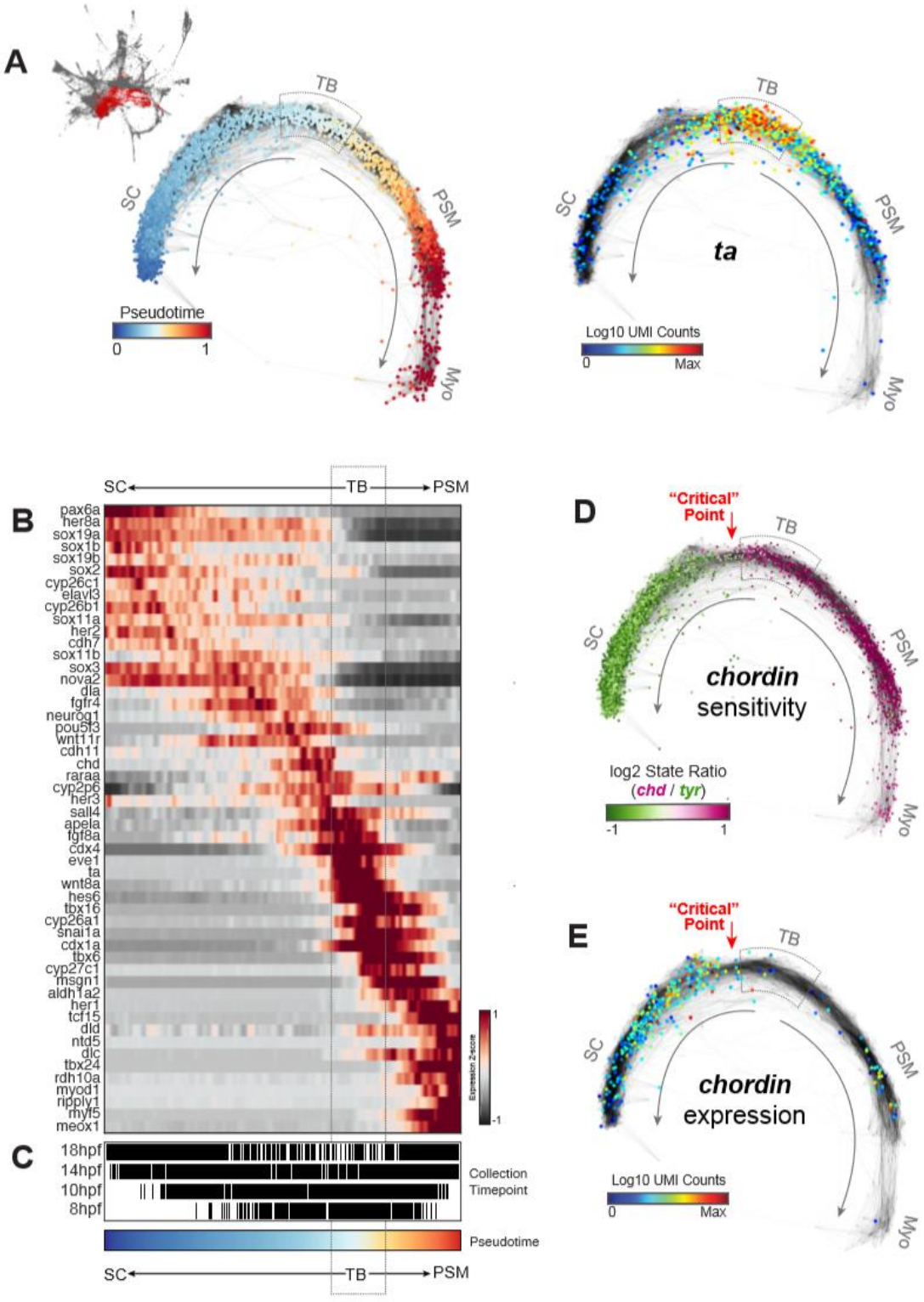
(A) State-by-state analysis of the *chordin* phenotype. Cells from 3 *chordin*-targeted and 3 control (*tyrosinase*-targeted) samples were assigned to one of the 28 wild-type annotated states by a kNN-classifier (see Figure S3.13 and Methods) and assessed for cell abundance changes and differential gene expression. After normalizing to total cell counts, the relative contribution of cells from the two CRISPR samples to each state was assessed by log2-fold change. Adjusted p-values report significant changes (2-tailed t-test). The extent to which *chordin* and control (*tyrosinase*) cells that were assigned to the same state differed in their transcriptional signatures was assessed by identifying differentially expressed genes (Wilcoxon rank-sum test, adj. pvalue

Supplemental Figure 3.14 (Continued) < 0.01; absolute log2 fold change >1). This analysis was limited to transcripts with average expression level >25 transcripts per million (TPM). For each state, the number of differentially expressed genes (DEG) between *chordin* and *tyrosinase* cells was compared to the number of "state-defining" DEG. State-defining DEG were identified by comparing cells of each state to cells of all other states, using the same rank-sum test criteria specified above. Differences between states were systematically associated with much larger DEG numbers than were *chd*/*tyr* differences. On average, >70% of all *chd*/*tyr* DEG were "novel" (i.e. not present in the state-defining DEG list). For each state, up to 10 "novel" up-regulated DE genes, ranked by adj. p-value, are listed. These genes include most predominantly a set of multiple heat-shock / stress associated transcripts (e.g. *hsp70*, *hsp70.2, tp53inp1*). (B) Quantitative trends of the *chordin* phenotype, depicted in a scatterplot. Each cell state is colored by known dorsal / ventral locations in the embryo (magenta: ventral tissues; green: dorsal tissues; grey: "intermediate" tissues). The size of each point reflects the % abundance of each tissue in the wild-type embryo. X-values report the relative changes in state abundance (log2) in *chordin* vs control (*tyrosinase*) embryos. Y-values report the number of differentially expressed genes identified between *chordin* vs control cells in each cluster, same rank-sum test criteria as in (A).



Supplemental Figure 3.15

Single-cell graphs, colored by normalized transcript counts for genes encoding select BMP inhibitors (A), and BMP proteins (B).



Supplemental Figure 3.16

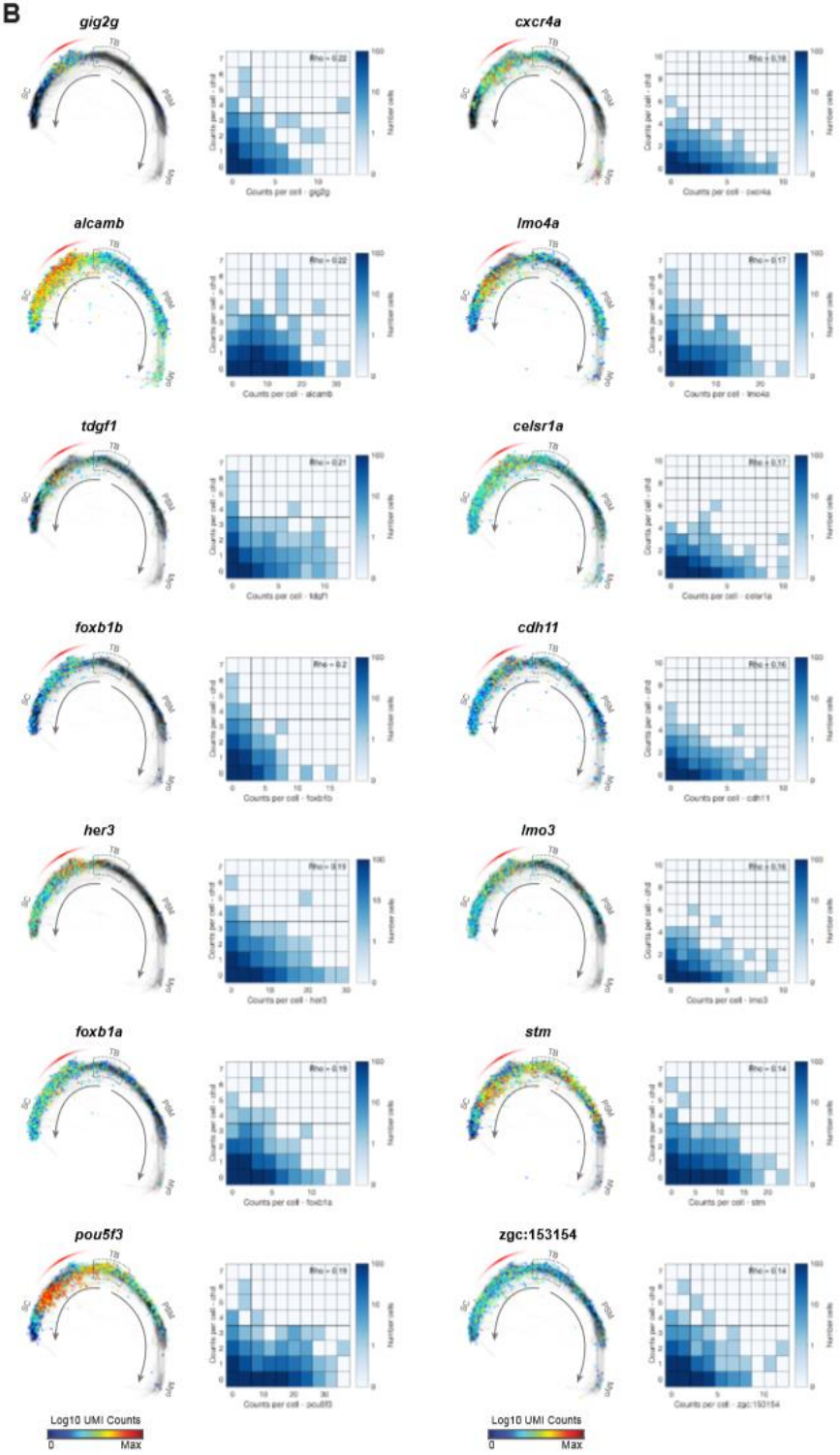
(A) Left: Subspace-projection layout of a subgraph (n=5,634 cells) corresponding to the tailbud

Supplemental Figure 3.16 (Continued) region of the zebrafish embryo, colored by pseudo-spatiotemporal ordering (see Methods). Colormap was centered on peak *brachyury/ta* expression. TB: tailbud; SC: spinal cord; PSM: presomitic mesoderm. Right: Expression counts for *ta* / *brachyury*. Heatmap of gaussian-smoothened expression z-scores for select dynamically expressed genes (see Methods). Distinct cascades of pro-neural or pro-mesodermal genes define two bifurcating trajectories emerging from the central tailbud region. Trajectory cells originated from multiple collection timepoints. Black bars indicate the relative contributions of each timepoint to different regions of the tailbud trajectory. Early timepoint cells dominate the center of the rajectory; later timepoints dominate the tips. Log2 ratios of differential cell state abundances in *chordin* vs. *tyrosinase* samples (same data as in Figure 3.6D). Red arrow hypothesizes a “critical point” in the branching trajectory. Overlay of normalized expression counts for *chordin*.

A

Gene Symbol	Rho
chd	1
gig2g	0.2225
alcamb	0.2198
tdgf1	0.2147
foxb1b	0.2043
her3	0.194
foxb1a	0.1891
pou5f3	0.1854
cxcr4a	0.1808
lmo4a	0.1732
celsr1a	0.1695
cdh11	0.1593
lmo3	0.1554
stm	0.1436
zgc:153154	0.1433
her8a	0.1433
si:ch211-152c2.3	0.1422
sp5l	0.142
sox19a	0.1383
hspb1	0.1382
ssbp4	0.1379
hmga1a	0.1374
cyp2p6	0.1334
foxj1b	0.1311
hoxb6b	0.1309
map1lc3b	0.1285
sp5a	0.1276
bida	0.1275
atp1b1a	0.1252
hmgb2a	0.1248
apela	0.1241
usp54a	0.1229
LOC101885649	0.1226
LOC101885973	0.1226
LOC101884747	0.1204
anp32e	0.1202
hoxb1b	0.1197
ptmaa	0.1185
nkx6.2	0.1157
stk40	0.1145
hoxc3a	0.1144

B



Supplemental Figure 3.17

Panel of genes co-expressed with *chordin* in the tailbud. (A) Pearson correlation coefficients were calculated comparing the normalized transcript counts for *chordin* to all other genes,

Supplemental Figure 3.17 (Continued) across all cells in the tailbud subgraph (Figure S3.16A). Listed are the top 40 most correlated genes. (B) Left: Gene expression overlays for the top 15 most correlated genes plotted for the tailbud subgraph. Red stripe indicates the graph region containing *chordin*+ cells. Right: Single-cell coexpression scatterplots of normalized transcript counts (vs. *chordin* counts) for the same 15 genes as in (A).

Appendix 3.

Size-reduced embryos reveal a gradient scaling-based mechanism for zebrafish somite formation

Authors: Kana Ishimatsu[a*], Tom W. Hiscock[a], Zachary M. Collins[a], Dini Wahyu Kartika Sari[b,c], Kenny Lischer[b], David L. Richmond[d], Yasumasa Bessho[b], Takaaki Matsui[b] and Sean G. Megason[a*]

[a] Department of Systems Biology, Harvard Medical School, Boston, MA 02115, USA.

[b] Gene Regulation Research, Nara Institute of Science and Technology, Nara 630-0101, Japan.

[c] Department of Fisheries, Universitas Gadjah Mada, Yogyakarta 55281, Indonesia.

[d] Image and Data Analysis Core, Harvard Medical School, Boston, MA 02115, USA.

Author contributions:

K.I. conceived the study and conducted experiments and data analysis. T.H. did modeling, simulations and wrote the program for wavelet analysis. Z.C. established size reduction technique with K.I. D.S., K.L. Y.B. and T.M. established ERK reporter fish and performed live imaging for this reporter. D.R. helped with image analysis. S.M. supervised the overall study. K.I., T.H. and S.M. wrote the paper with input from other authors.

Abstract

Little is known about how the sizes of animal tissues are controlled. A prominent example is somite size which varies widely both within an individual and across species. Despite intense study of the segmentation clock governing the timing of somite generation, how it relates to somite size is poorly understood. Here we examine somite scaling and find that somite size at specification scales with the length of the PSM despite considerable variation in PSM length across developmental stages and in surgically size-reduced embryos. Measurement of clock period, axis elongation speed, and clock gene expression patterns demonstrate that existing models fail to explain scaling. We posit a “clock and scaled gradient” model, in which somite boundaries are set by a dynamically scaling signaling gradient across the presomitic mesoderm. Our model not only explains existing data, but also makes a unique prediction that we experimentally confirm—the formation of periodic “echoes” in somite size following perturbation of the size of one somite. Our findings demonstrate that gradient scaling plays a central role both in progression and size control of somitogenesis.

Introduction

Scaling—matching organ size to body size—is a fundamental property of developing organisms. Even within the same species, developing embryos often vary in size, due to environmental and maternal variability. In addition, embryo size can change drastically across developmental stages. Nevertheless, embryos robustly develop with invariant proportions, suggesting that some mechanism of pattern scaling is encoded in the developmental program (Cooke, 1981). While the scaling of morphogen gradients has received significant attention, both theoretically and experimentally (Ben-Zvi and Barkai, 2010; Gregor et al., 2005; Gregor et al., 2008; Inomata et al., 2013; Lander et al., 2011; McHale et al., 2006; O'Connor et al., 2006), understanding has been limited for scaling of other patterning processes, such as somite segmentation.

During embryogenesis, somites provide the first body segments in vertebrates, eventually giving rise to tissues such as the vertebrae and axial skeletal muscles. Somite segmentation occurs sequentially in an anterior to posterior progression along the presomitic mesoderm (PSM), with temporal and spatial periodicity. Temporal periodicity (e.g. somites are formed in symmetric pairs every 25 min in zebrafish (Schroter et al., 2008)) is known to be generated by a system of coupled cellular oscillators (Delaune et al., 2012; Lauschke et al., 2013; Masamizu et al., 2006; Palmeirim et al., 1997), called the segmentation clock, which is driven and synchronized by complex signaling networks (Dequeant et al., 2006; Hubaud and Pourquie, 2014; Krol et al., 2011). Yet, how these oscillations relate to the spatially periodic pattern of the mature somites and how somite sizes are determined remains controversial (Akiyama et al., 2014; Cooke and Zeeman, 1976; Cotterell et al., 2015; Lauschke et al., 2013; Shih et al., 2015; Soroldoni et al., 2014; Takahashi et al., 2010; Tsiairis and Aulehla, 2016).

Somites were first documented to scale in *Xenopus* following bisection; the resulting embryos have smaller but the same number of somites when compared to intact control embryos (Cooke, 1975). Although this experiment was performed more than 40 years ago, the underlying mechanism for somite scaling has not been identified. In particular, the relationship between PSM length and somite size has been disputed: previous groups have reported that in intact developing embryos, somite size does not scale with PSM size (Gomez et al., 2008), while in ex vivo culture of PSM, somite length has been shown to linearly scale with PSM length (Lauschke et al., 2013).

In this study, we demonstrate that somite length indeed scales with PSM length and that gradient scaling underlies somite scaling, using both surgically size-reduced and normally developing zebrafish embryos, in combination with live imaging, quantitative measurement, and mathematical modeling. We demonstrate that previously reported discrepancies between somite size and PSM size can be explained by a time delay between somite size and

morphological boundary formation. The relationship between somite and PSM length is remarkably different when this delay is considered, revealing that somite length always scales with PSM length. This result led us to evaluate several variables that could potentially modulate somite length. We found that clock period, axis elongation speed, and clock gene expression patterns did not scale, whereas the Fgf activity gradient did scale with PSM length. Based on this observation, we developed a "clock and scaled gradient model" based on the original clock and wavefront model (Cooke and Zeeman, 1976) with a simple yet important refinement in which a gradient scaled to the PSM size sets the wavefront position. Using transplants, we show that somite derived signals can inhibit Fgf signaling providing a potential mechanism for gradient scaling. The clock and scaled gradient model not only explains existing experimental data but also inspired a novel experimental test with an unintuitive outcome—the creation of "echoes" in somite size following perturbation of the system. Together, we present the quantitative study of somite scaling as an experimental platform to test the feasibility of multiple theoretical models.

Results

Somite length at specification scales with PSM length throughout developmental time. Although somite length has been shown to scale with overall body length in *Xenopus* (Cooke, 1975), whether somite length scales with PSM size has been controversial (Gomez et al., 2008; Lauschke et al., 2013). To test this relationship we measured somite length and PSM length using live imaging. Initially we did not observe a clear relationship between PSM length and somite size (see Figure 1F). However, somite specification within the PSM occurs long before the appearance of the morphological boundaries (Akiyama et al., 2014; Bajard et al., ; Dubrulle et al., 2001; Elsdale et al., 1976; Giudicelli et al., 2007; Ozbudak and Lewis, 2008; Primmett et al., 1989; Roy et al., 1999) (Figure 1A), and thus we speculated that the inconsistency with

respect to somite scaling could be attributed to this delay. Although previous studies have shown the delay is around 4-5 cycles, the delay duration could vary along developmental stages. To examine if somite length scales with PSM length when this specification to formation delay is considered, we experimentally measured this delay using embryos from different developmental stages. Dual-specificity phosphatase inhibitor BCI is known to act immediately on Fgf signaling leading to an eventual reduction of somite size (Figure S1) (Akiyama et al., 2014). We transiently treated embryos at 5 somite stage (ss), 10ss, and 15ss with BCI and measured the length of the newly formed somites using live imaging for six subsequent cycles (Figure 1B and C). Regardless of the developmental stage for the pulse BCI treatment, we observed 4-cycle delay on average before a visibly smaller somite formed (Figure 1D). Our experimentally determined delay is similar, albeit slightly shorter, to what has been proposed in previous work (4-5 cycles) (Akiyama et al., 2014; Bajard et al., ; Dubrulle et al., 2001; Elsdale et al., 1976; Giudicelli et al., 2007; Ozbudak and Lewis, 2008; Primmett et al., 1989; Roy et al., 1999). Taking this 4-cycle delay into consideration, we reexamined the relationship between PSM length and somite size (comparing the size of the Nth somite with the PSM size at the N-4 ss, Figure 1E). Strikingly, we found that somite size indeed scales with PSM size when this 4-cycle delay is considered (Figure 1F). No clear relationship between somite and PSM length is apparent without the delay (Figure 1F). This relationship between PSM length and somite size was still observed with a 3 or 5 cycle delay, suggesting that minor fluctuations in the delay or measurement error would not affect the conclusion (Figure 1F). The delay between somite specification and formation is reflected in different peak positions in time course measurements of PSM and somite size (Figure 1G). Consideration of this delay may be necessary to assess scaling in previous data (Gomez et al., 2008; Schroter et al., 2008).

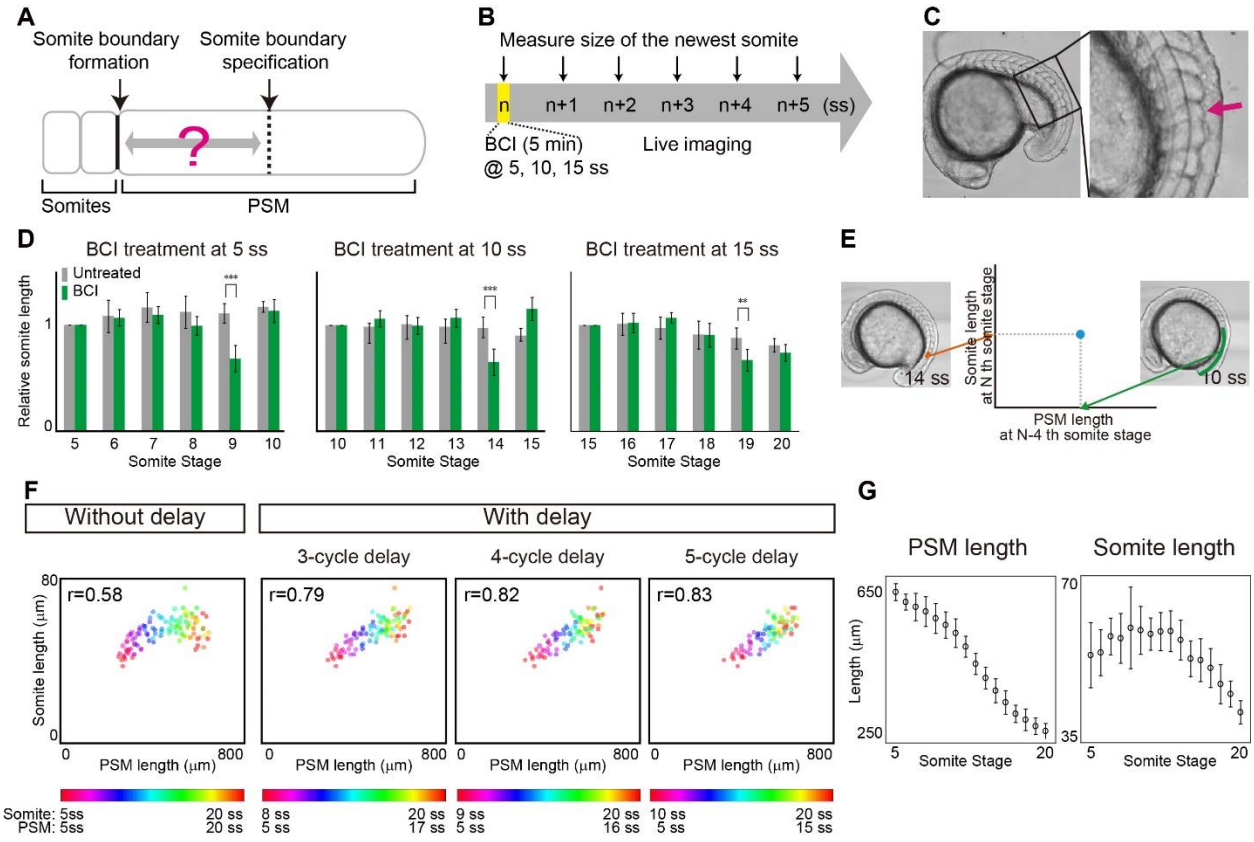


Figure 1. Somite scaling over time with time delay.

(**A**) Schematic illustration of time delay between somite boundary specification and somite boundary formation. (**B**) Schematic illustration of BCI experiment. The embryos were treated with BCI for 5 min and then subjected to live imaging in egg water without BCI. The BCI treatment was done at 3 different somite stages (5, 10, 15 ss), in case the delay time varies over time. (**C**) BCI treated embryos form smaller somites (magenta arrow). (**D**) Relative AP length of somites, normalized by the somite length of control embryos at somite stage of BCI treatment. At each somite stage, the smaller somite was formed 4 cycles after BCI treatment. Error bars denote SD. **P < 0.01 and ***P < 0.001. (n=5 for each condition) (**E**) Comparison of PSM length and somite length was made using PSM length at N-4 ss (e.g. 10 ss) and somite length at N ss (e.g. 14 ss), using live imaging data. (**F**) Somite size vs PSM size between control and chopped embryos with and without time delay (3, 4, 5 cycles). (**G**) Size dynamics of PSM

Figure 1 (Continued) and somites. Note the peaks appear at different somite stages. Error bars denote SD.

Somite length at specification scales with PSM length among individuals with different body sizes.

Given that somite size at specification scales with PSM length throughout developmental time, we then wondered whether somite length scales with PSM length between zebrafish embryos of varying sizes. Inspired by classic work in *Xenopus* (Cooke, 1975) on somite scaling in surgically size reduced embryos, we sought to apply this technique to zebrafish. We first attempted to cut zebrafish embryos at the blastula stage longitudinally (along the animal-vegetal axis) as was done in *Xenopus*. However, the resulting embryos had varying degrees of dorsalization or ventralization presumably due to dorsal determinants being portioned in unpredictable ways and were difficult to study quantitatively. We thus sought a method to reduce embryo size without perturbing D-V patterning. By using separate latitudinal cuts to remove cells near the animal pole and yolk near the vegetal pole at the blastula stage (Figure 2A left panel), we found that the resulting size-reduced embryos quickly recovered and a large percentage of them developed normally (Figure 2A). Total body size and organ size, including somites, of these chopped embryos were found to be smaller (Figure 2B and C). Consistent with previous work in *Xenopus* (Cooke, 1975), the chopped embryos had the same number of somites, each of which was smaller in size (33 in both control and chopped embryos at 1 day post-fertilization, n=5 for each. Somite number was counted using still images of the live embryos). Combining this size reduction technique and live imaging, we measured somite and PSM length, and found somite length scales with PSM length between embryos of varying sizes when the same 4-cycle delay is considered (Figure 2D, see also Figure S2). The scaling was observed throughout our timecourses (from 5 ss to 20 ss, Figure S3). Taken together, we conclude that somite length always scales with PSM length as long as the time delay between specification and morphological boundary formation is considered.

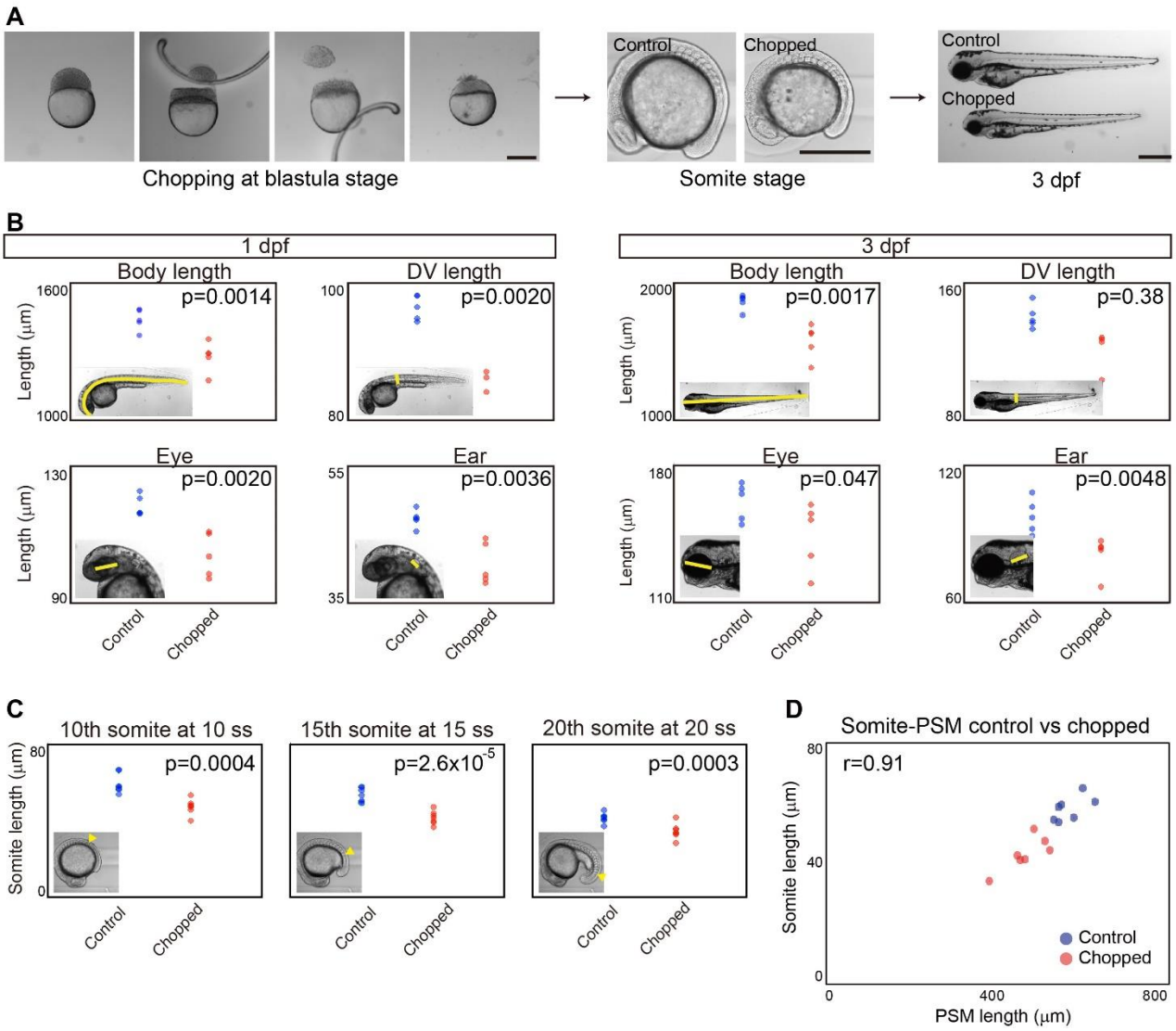


Figure 2. Somite scaling between individuals of different sizes.

(**A**) Size reduction technique. Scale bar, 500 µm. (**B**) Body and organ sizes comparison between control and chopped embryos. (**C**) Somite size comparison between control and chopped embryos. (**D**) Somite size vs PSM size between control and chopped embryos.

Clock period does not scale with PSM length.

Given our finding that somite length scales with PSM length both over time and among individuals with different sizes, we next asked what mechanism might link PSM size to somite size. For this purpose, we searched for a component of the known somite patterning system that scales with PSM length, both across developmental stages and among individuals. In the classic clock and wavefront model, somite length is the product of clock period and wavefront regression speed. We first measured the period of the segmentation clock both in control and chopped embryos over time, since it is known that a change in the period of clock gene expression causes a change in somite length (Herrgen et al., 2010; Kim et al., 2011; Schroter and Oates, 2010). We measured the clock period as the time between the formation of successive somite boundaries, and found no difference in this period between control and chopped embryos (Figure 3A, Figure S4) or between those at different developmental stages (Figure 3B) (Schroter et al., 2008), suggesting that scaling is not achieved by regulation of clock period.

Axis elongation speed does not scale with PSM length.

We next quantified the axis elongation speed, since slower axis elongation is known to lead to shorter somite length (Goudevenou et al., 2011; Rauch et al., 1997). One explanation for this comes from the clock and wavefront model, in which the wavefront speed (and hence somite size) has often been directly linked to axis elongation speed (Cooke and Zeeman, 1976; Dubrulle and Pourquie, 2004; Hubaud and Pourquie, 2014; Saga, 2012). This possibility is also consistent with the idea that a gradient of Fgf is established by mRNA decay coupled with axis elongation, and that this drives wavefront progression (Dubrulle and Pourquie, 2004). Therefore, we expected somites to be smaller in chopped embryos due to a decrease in the axis elongation speed (e.g. cells are incorporated into the PSM at the tailbud at a slower rate). We measured the change in axis length, defined by a distance between the posterior boundary of

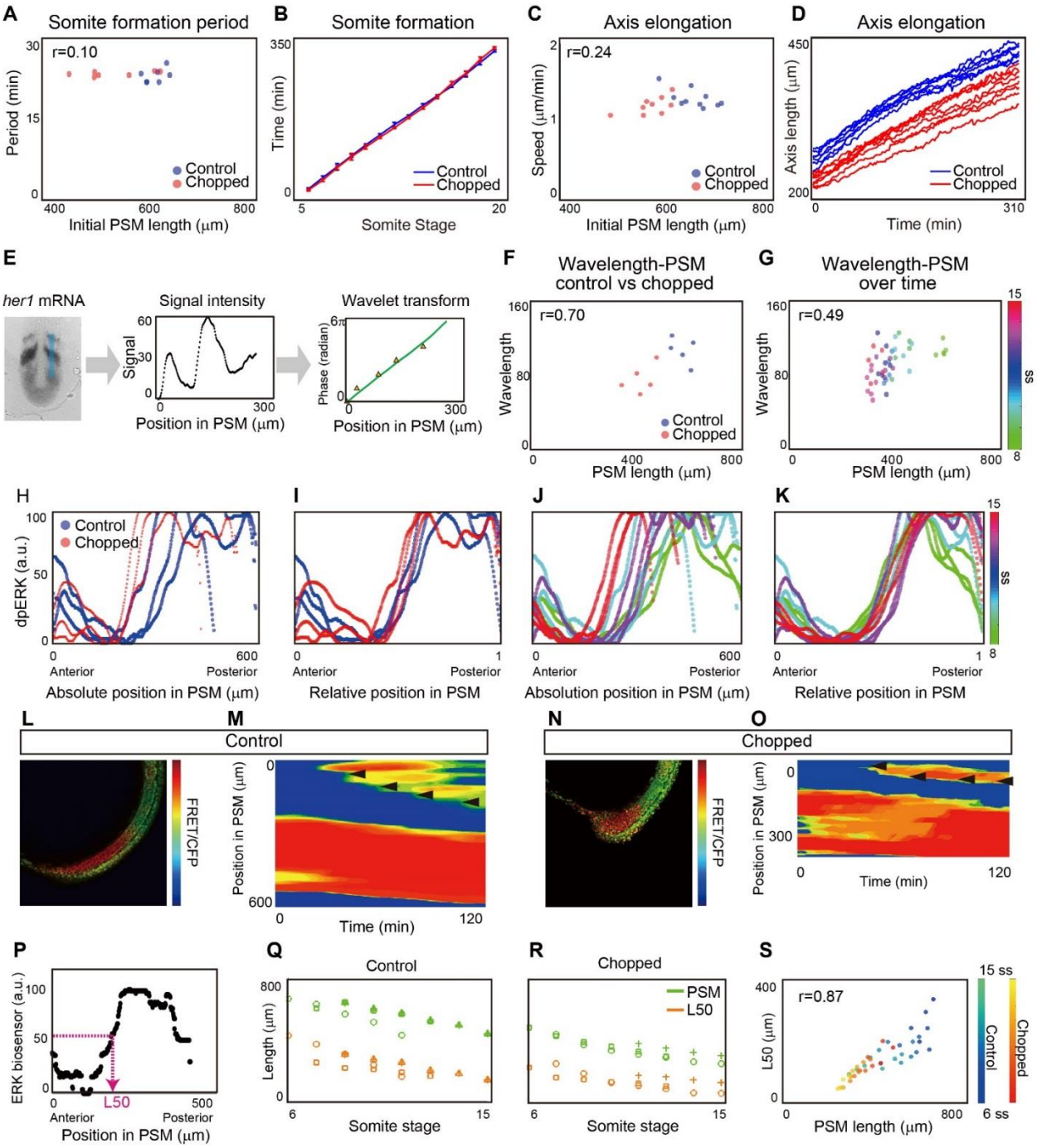


Figure 3 Determining which components of the somite formation system scale.

(**A**) Somite formation frequency vs. initial PSM size. No significant difference was found between control and chopped embryos at the 5% significance level, and the confidence interval on the difference of means (-1.78 - 0.66) includes the hypothesized value of 0. (**B**) Somite

Figure 3 (Continued) formation time of control and chopped embryos. The slope corresponds to somite formation period. Note the slopes do not change over time. Error bars denote SD. (**C**) Axis elongation speed vs. PSM size. No significant difference was found between control and chopped embryos at the 5% significance level, and the confidence interval on the difference of means (-0.10 - 0.15) includes the hypothesized value of 0. (**D**). Axis length vs. time. The slope represents the speed of axis elongation. (**E**) Quantification of her1 wavelength along the blue line in the first panel. Green line in the third panel shows the phase gradient obtained by wavelet transform. Orange triangles show manually measured wavelength. (**F**) Wavelength vs. PSM among individuals. (**G**) Wavelength vs. PSM size over time. (**H** to **K**) Quantification of Fgf activity based on dpERK immunostaining. (**H** and **I**) dpERK scaling between control and chopped embryos. (**J** and **K**) dpERK scaling across developmental stages. Both are shown by absolute position (**H** and **J**), and relative position (**I** and **K**). (**L** to **S**) Quantification of Fgf activity based on ERK biosensor mRNA-injected embryos. The manipulated embryos (**L** and **N**) were used to generate kymographs of ERK activity (**M** and **O**). Black arrowheads represent newly formed somites. LUT, high (red) to low (blue) reporter intensity. (**P**) Definition of L50. (**Q** and **R**) Change in PSM size and L50 position overtime, in control embryos (**Q**) (n=4) and chopped embryos (**R**) (n=3). Different marks correspond to different embryos. (**S**) L50 vs PSM length over time both in control and chopped embryos.

4th somite and the tail tip, over time(Bajard et al., 2014). Contrary to our expectation, we found that axis elongation speed did not differ between control and chopped embryos, at least for 5ss—15ss (Figure 3C, Figure S4); this seemingly confusing result can be explained if the major mechanism of axis elongation at these stages is, for example, convergence and extension, whose rate should not be size dependent (Steventon et al., 2016). Notably, the axis elongation speed was nearly constant over our experimental time window (Figure 3D), although PSM size decreased drastically. Since axis elongation speed neither changes over time as somites decrease in size nor between embryos of varying sizes, altered axis elongation speed cannot explain scaling of somite patterning.

Wavelength of her1 traveling waves does not scale with PSM length.

We then asked if the wavelength of the traveling wave pattern of a segmentation clock gene (e.g. *her1*) could explain scaling of somite formation. Canonical segmentation clock genes exhibit traveling waves; a stripe pattern that sweeps through the PSM from posterior to anterior due to a phase delay toward the anterior direction. While these traveling waves have not been experimentally shown to cause somite size alterations, a correlation between wavelength (spatial interval of the stripes) and somite length has been observed (Jorg et al., 2016; Lauschke et al., 2013). To determine whether *her1* traveling waves are involved in scaling, we generated and quantified phase maps from *her1 in situ* hybridization samples (Figure 3E). We extracted the phase information from signal intensities using a wavelet transform, then converted the approximately linear phase gradient into an effective wavelength, defined as the distance between peaks of *her1* intensity (Figure 3E). We measured the phase gradient from an area of PSM including B-4 (the presumptive position corresponding to a morphological boundary four cycles later, blue line in Figure 3E, left panel). We also measured the phase gradient manually, by identifying peaks and troughs in the intensity profile (orange triangles in Figure 3E, right panel). This manual measurement was found to correspond well with phases

obtained from the wavelet transform (green line in Figure 3E, left panel). We found that unlike somite size, wavelength does not always scale with PSM size: although the wavelength scales with PSM size following embryonic size reduction, it does not scale during embryonic development (Figure 3 F and G) (Holley et al., 2000). This is consistent with recent work demonstrating that the number of *her1* waves changes over time, confirming that the phase gradient does not scale with PSM size (Soroldoni et al., 2014). Since somite size scales with PSM size over developmental stages as well as among individuals of different size, this result indicates that it is unlikely that the somite scaling is achieved through regulation of the wavelength of *her1*. The conclusion is supported by a previous study which showed that repeated induction of *deltaC* expression in a *deltaC* mutant background can successfully rescue somite boundary formation, although the induced *deltaC* expression did not show the traveling wave pattern (Soza-Ried et al., 2014).

The Fgf activity gradient scales with PSM length.

Our final candidate feature that could relate somite size to PSM size was the FGF gradient (Akiyama et al., 2014; Dubrulle et al., 2001; Sawada et al., 2001). To measure FGF signaling we used whole mount immunohistochemistry against doubly phosphorylated ERK (dpERK), a downstream readout of Fgf activity, and extracted the signal intensity. We found that the gradient range varies considerably between embryos on an absolute length scale, but is quite consistent when plotted as a function of relative PSM length, both for control and chopped embryos (Figure 3H and I, Figure S5 and S6) and for embryos from different developmental stages (Figure 3J and K, Figure S5 and S6). We further tested if Fgf activity scales with PSM size in embryos carrying a FRET-based ERK biosensor (Figure 3 L-S). We calculated the PSM location where the relative intensity of FRET signal crosses 50% of the maximal intensity (L50) (Figure 3P). Time course analysis of L50 in both control and chopped embryos confirmed that the Fgf activity gradient scales with PSM size (Figure 3Q-S). L50 analysis was further

performed when Fgf activity was measured by dpERK and by *sprouty4* mRNA, a downstream gene of Fgf signaling, confirming Fgf activity scaling (Figure S7).

Since Wnt signaling is also known to form a gradient in PSM, we examined whether Wnt signaling scales with PSM length. We performed L50 analysis on expression patterns of *sp5l* mRNA, a downstream gene of Wnt signaling (Thorpe et al., 2005), and found Wnt activity also scales with PSM length (Figure S8). Although we cannot determine which signaling is upstream (we expect them to interacting with each other (Bajard et al., 2014; Stulberg et al., 2012; Wahl et al., 2007)), for simplicity, we will mainly focus on Fgf signaling in this paper.

A clock and scaled gradient model can explain somite scaling.

Given our observation of a dynamically scaling FGF activity gradient, we turned to modeling to see whether this feature is capable of explaining scaling of somite patterning. In the original clock and wavefront model, the timing of somite boundary specification is controlled by a clock and the positioning by the level of a signal that encodes a posteriorly moving wavefront. How the position of the wavefront is determined at each time point is unspecified in the original model. Importantly, our observations reveal that the activity of the signaling molecule linked with wavefront activity, Fgf or potentially Wnt, forms a dynamic gradient that scales with PSM size. We term this updated model the “clock and scaled gradient” model. In this model, scaling of the gradient to PSM size generates a posteriorly moving wavefront, when it is combined with axis elongation (which increases PSM size) and somite formation (which decreases PSM size) (Figure 4 A and B). We constructed a simple mathematical model to formalize these interactions (Supplementary Materials and Methods) and found that this model can successfully reproduce our biological results on somite size scaling (Figure 4C-F). Similar somite formation dynamics can be observed regardless of the precise shape of the gradient (Figure 4C and D; steep sigmoidal gradient, Figure 4E and F; linear gradient).

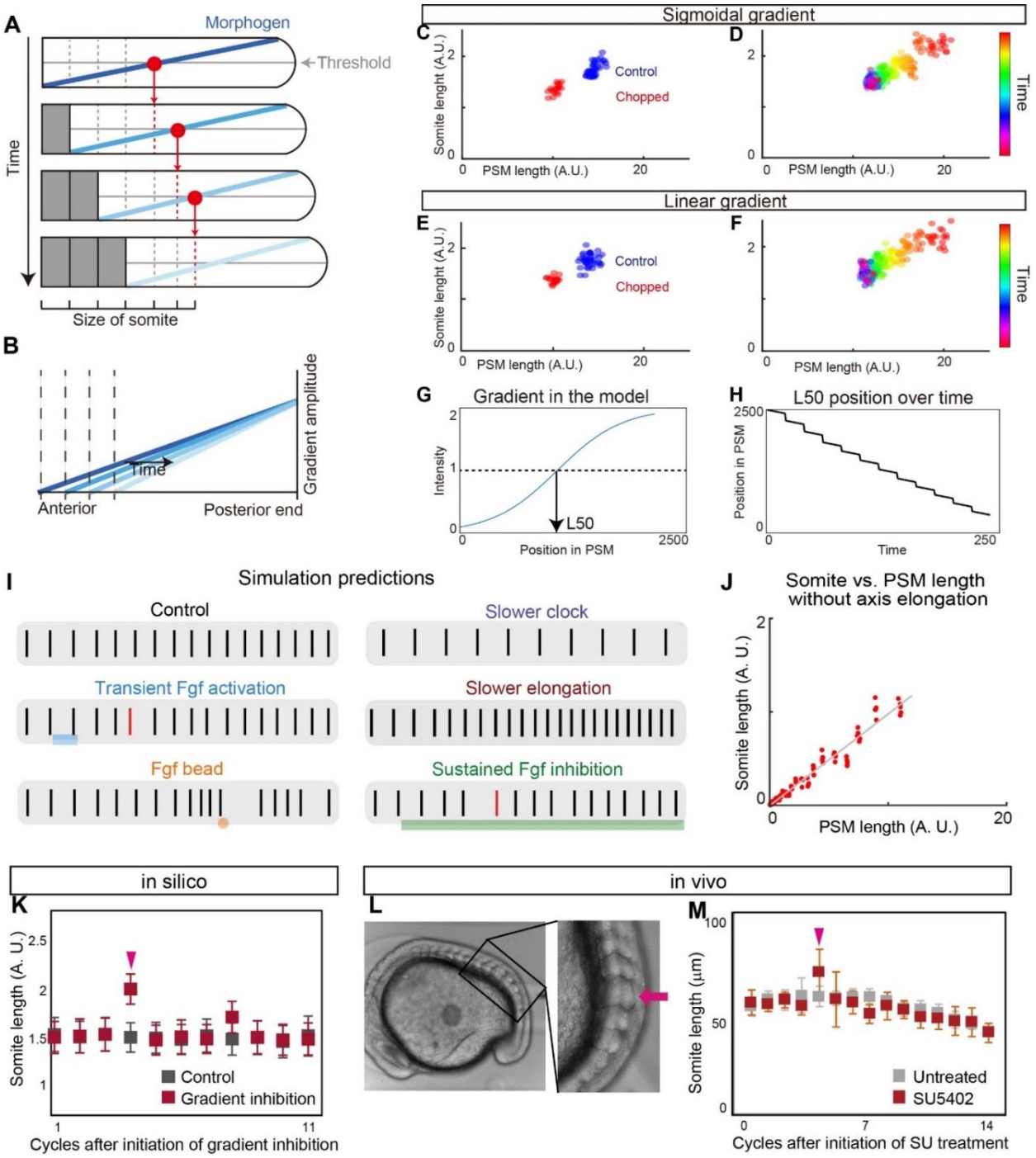


Figure 4. Clock and scaled gradient model.

(**A**) Schematic illustration of the clock and scaled gradient model. (**B**) Superimposition of the gradients from each time point in (**A**). (**C** and **D**) Simulation results using a sigmoidal gradient. (**E** and **F**) Simulation results using a linear gradient. (**C** and **E**) Simulation results of control and

Figure 4 (Continued) chopped embryos. (**D** and **F**) Simulation results of a single embryo over time. (**G** and **H**) Stepwise regression of the gradient in clock and scaled gradient model. (**G**) L50 in the model was determined similarly to Figure 3I. (**H**) Clock and scaled gradient model predicts stepwise regression of L50 position. (**I**) Simulation results for perturbation experiments for local or global inhibition/ activation of Fgf, slower clock and slower axis elongation. (**J**) Somite size versus PSM length shows perfect scaling in silico when axial elongation speed is zero, mimicking the results from the *in vitro* mPSM system(Lauschke et al., 2013). (**K**) Simulation results of long-term suppression of a gradient in the clock and scaled gradient model. Error bars denote SD. (**L** and **M**) Low concentration of SU5402 (16 μM) results in one or two larger somite(s) (n=7 for both SU5402 and untreated). Error bars denote SD.

We also observed step-wise regression of the L50 in our model, consistent with the recent report (Figure 4G and H) (Akiyama et al., 2014). Moreover using this model, we can also accurately predict the resulting changes in somite size following a wide range of additional perturbations (Figure 4 I and J): one smaller somite following transient Fgf activation (Akiyama et al., 2014) (Figure 4I); multiple smaller somites followed by one larger somite after Fgf bead transplantation (Dubrulle et al., 2001; Sawada et al., 2001) (Figure 4I); larger somites with a slower clock (Herrgen et al., 2010; Kim et al., 2011; Schroter and Oates, 2010) (Figure 4I); smaller somites with slower axis elongation (Goudevenou et al., 2011; Rauch et al., 1997)(Figure 4I); and scaling of somite and PSM size *in vitro* under culture conditions that do not permit axis elongation (Lauschke et al., 2013)(Figure 4J). We found that in all cases, the model's predictions were in agreement with experimental results.

The clock and scaled gradient model predicts one larger somite in long-term Fgf inhibition.

A simple perturbation to test our model is long-term Fgf inhibition. This experiment was recently carried out using chick embryos and multiple larger somites were shown to form during long-term Fgf inhibition (Cotterell et al., 2015). This result was contradictory to what the clock and wavefront model would predict, but consistent with a novel Turing framework for somitogenesis (Cotterell et al., 2015). We simulated the same perturbation using our clock and scaled gradient model and found that it predicts the same result as the clock and wavefront model: only one larger somite (Figure 4K). To test if the long-term Fgf inhibition has the same effect in zebrafish embryos, we treated zebrafish embryo with the Fgf inhibitor, SU5402 (Sawada et al., 2001), at a low concentration (16 μM) in which embryos grew until late stages. Unlike in chick (Cotterell et al., 2015), we observed one larger somite but not multiple larger somites following long-term SU5402 treatment (Figure 4L and M, for individual data, see Figure S14), consistent with our model. Moreover, we observed the same tendency under constant darkness, confirming the

result we obtained was not due to the light instability of SU5402 (10 out of 11) (Figure S10). These differences in results could potentially be explained by how acutely the drug can be administered: in zebrafish, embryos can be soaked in a vast excess of drug causing a rapid step up in drug levels followed by a plateau *in vivo*, whereas in chick the drug levels may rise more slowly. Simulations showed that increasing FGF inhibition over a few hours can cause multiple large somites in our model (Figure S11).

Newly formed somites play a critical role in Fgf gradient scaling.

One potential mechanism of gradient scaling is that newly formed somites modulate the Fgf gradient, for example, by secreting a negative regulator of the pathway. To examine whether the newly formed somite can modulate Fgf gradient, we transplanted a newly formed somite into the posterior PSM, and compared it to a control experiment in which PSM cells were transplanted to the same axial level (Figure 5A). From our model, we predicted that the ectopically transplanted somite would locally inhibit Fgf signaling. One to two cycles (0.5-1 hour) after transplantation, the embryos were fixed and stained for dpERK. We found that in the PSM surrounding the transplanted somite, the dpERK level was significantly decreased (Figure 5B), whereas the dpERK level in the PSM surrounding transplanted PSM cells was largely unaffected (Figure 5C). To quantify ERK activity, we normalized the dpERK signal near the transplant with that of the non-transplanted side of the same embryo at the same axial level (Figure 5A). We found the dpERK levels around the transplanted somite to be significantly lower than the control (Figure 5D). These data support our hypothesis that mature somites rapidly and potently modulate the Fgf activity gradient to effect gradient scaling.

One immediate candidate molecule that could contribute to Fgf scaling is retinoic acid, since it forms an opposing signaling gradient to Fgf and antagonizes Fgf signaling in posterior tailbud. To test the role of retinoic acid in Fgf scaling, we examined whether Fgf scaling holds in the absence of retinoic acid by knocking down retinoic acid synthetace (*raldh1*) using

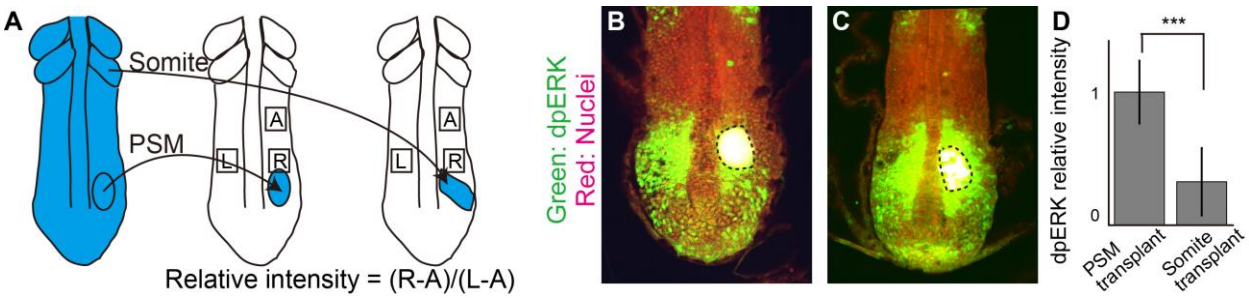


Figure 5 New somites inhibit FGF activity.

(**A**) Schematic illustration of somite transplantation. (**B** and **C**) dpERK immunostaining. Dashed line shows transplanted tissue. (**D**) Comparison of relative intensities between PSM transplanted samples (n=9) and somite transplanted samples (n=9). Error bars denote SD. ***P < 0.001.

morpholino. The Fgf activity was found to scale with PSM length in *raldh1* knocked down embryos, suggesting that retinoic acid does not play the central role in Fgf gradient scaling (Figure S12).

A unique prediction from the clock and scaled gradient model: an "echo effect" on somite size

We further sought a novel experimental test for which our model makes a unique prediction. Key aspects of the clock and scaled gradient model are the 4-cycle delay between somite specification and formation, and the feedback between newly formed somites and gradient length. We thus reasoned that if we experimentally created one larger somite, it would shorten the PSM and rescale the gradient in a jump, which would then result in another larger somite four cycles later, and this process would repeat creating "echoes" of larger somites with a ~4-cycle periodicity (Figure 6A). Simulations of our model supported this idea (Figure 6B

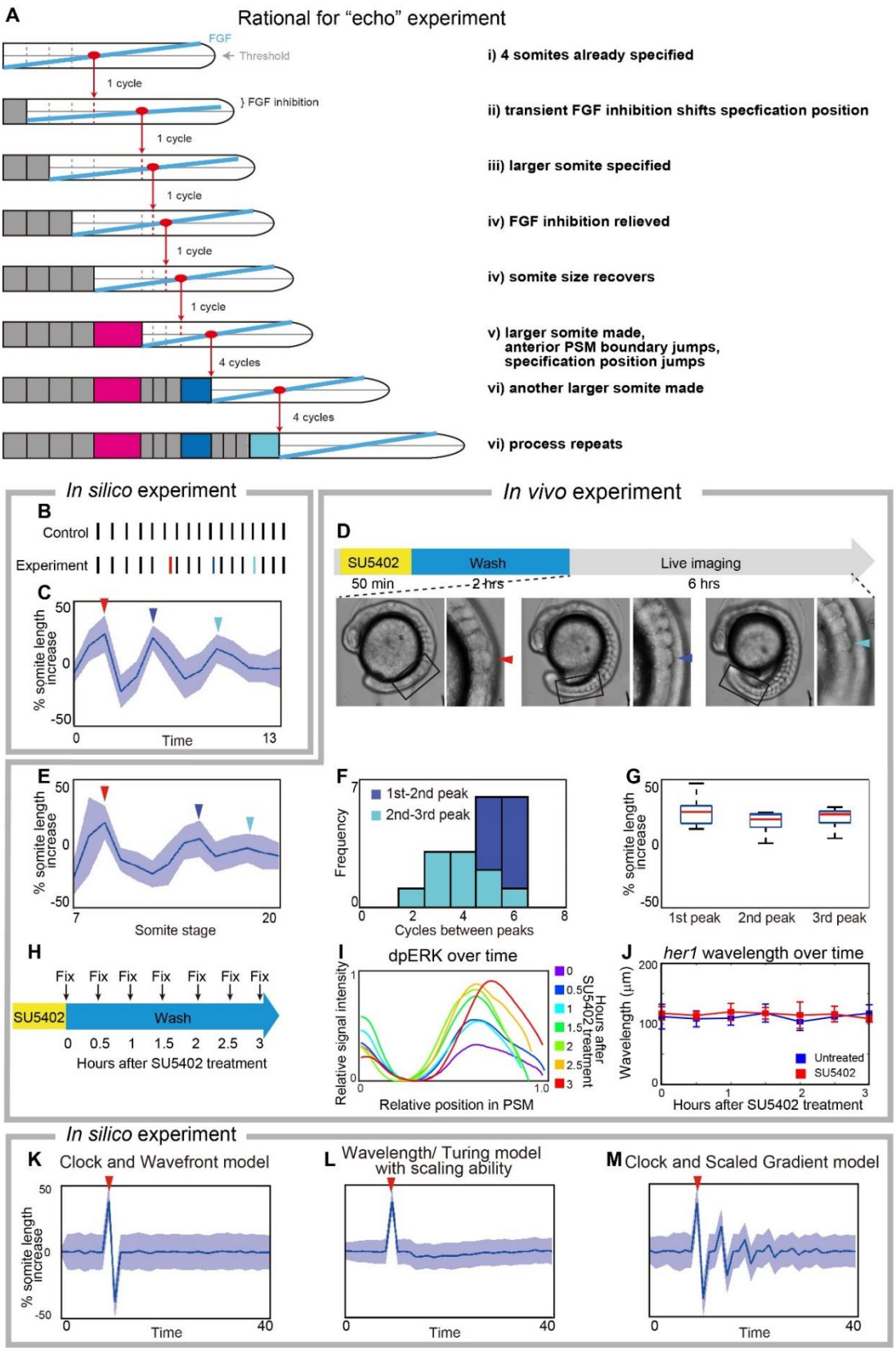


Figure 6 Experimental validation of the clock and scaled gradient model.

Figure 6 (Continued) (**A**) Schematic illustration of the outcome of the clock and scaled gradient model, following induction of one larger somite. The induced larger somite is colored in magenta, the larger somites as a result of system response are colored in blue and cyan. (**B** and **C**) Simulation results without (**B**) and with noise (**C**) for the somites size (red line in **B,** red arrow head in **C**). Blue line in **B** and blue arrowhead in **C** show the second, and cyan line in **B** and cyan arrowhead in **C** show the third large somite. (**D**) Schematic illustration of the *in vivo* experiment, and an embryo with larger somites at different time points. (**E**) Time course of percentage increase in somite length of SU5402 treated embryos, compared to those in control embryos (n=12). (**F**) Frequency distribution of somite cycles between the peaks. (**G**) Percentage increase in somite size in SU5402-treated embryos at the peaks detected in each embryo, compared to control embryos at the corresponding somite stage. In both **C** and **E**, blue lines and blue shades indicate the average somite size and the variance of one standard deviation, respectively. For **C**, **D** and **E**, red, blue and cyan arrowheads show the first, second and third larger somites. (**H-J**) Examination of ERK activity and her1 wavelength after transient SU5402 treatment. (**H**) Schematic illustration of the experiment. After fixation, the samples were subjected to dpERK and her1 in situ hybridization. (**I**) dpERK intensity curves for SU5402 treated embryos were calculated by averaging intensity curves for each time points. Relative signal intensity (y axis) was determined by scaling factors: (maximum intensity of treated embryos) / (maximum intensity of untreated embryos at the corresponding time points). Relative position (x axis) was determined by normalized positions in PSM in treated embryos by averaged PSM length of untreated embryos at the corresponding time points. (**J**) Time course analysis of her1 wavelength of untreated embryos and SU5402 treated embryos. We found no significant difference at significant level of 0.05 at any time point. Error bars denote SD. (**K-M**) Simulation results for percentage increase of somite size over time, based on different models. After induction of one larger somite (arrowheads in magenta), clock and wavefront model (when wavefront speed is associated with axis elongation only) predicts one smaller somite (**K**),

Figure 6 (Continued) wavelength/ Turing model (Cotterell et al., 2015) predicts smaller somites and the somite size eventually comes back to normal(**L**). Only the Clock and Scaled Gradient model predicts the “echo effect” that somite size dynamics shows ups and downs repeatedly every 4 cycles (**M**).

and C). Furthermore, we were encouraged by previous experiments that showed an “echo effect” of somite abnormalities following heatshock treatment (Primmett et al., 1989).

To test this prediction we transiently treated embryos with the Fgf inhibitor, SU5402, which is known to induce a larger somite (Dubrulle et al., 2001; Sawada et al., 2001), followed by extensive washes for two hours, then performed live imaging to measure the length of the newly formed somites (Figure 6D). Strikingly, we found that somite size became smaller and larger with a several-cycle period, which was uniquely predicted by the clock and scaled gradient model (Figure 6E and F, for individual data, see Figure S13). We noted that the periodicity was not always precisely 4 (Figure 6F), possibly due to internal fluctuation of the delay time or experimental variation, such as variation in washout timing of SU5402. By analyzing individual embryos (Figure S13), we confirmed that all the peaks of somite size in pulse SU5402 treated embryos are larger than those in control embryos (Figure 6G). As predicted, we observed the echo effect in long-term SU5402 treated embryos as well (Figure S14), but we chose to focus on transient treatment because the embryos were healthier. The echo effect was also seen in the embryos transiently treated with BCI (Figure S15). These results confirm that echo effect is a general phenomenon for somite formation.

We then evaluated the effect of transient SU5402 on both dpERK activity and *her1* wavelength. To perform time-course analysis, we fixed the embryos every 30 min while washing after SU5402 treatment. dpERK immunostaining confirmed quick recovery of Fgf activity after SU5402 treatment. Furthermore, as we assumed, the dpERK activity was found to scale even to the induced smaller PSM. In addition, we found no significant difference in *her1* wavelengths between control and SU5402 treated embryos.

This rebounding effect is only predicted if the “specification position” of new somites (rather than the somite itself) scales with PSM size, which is the core assumption of the clock and scaled gradient model (Figure 6K-M). Without gradient scaling, the clock and wavefront model predicts

a single smaller somite following the induced larger somite, but the size of the following somites immediately returns to normal (Figure 6K), consistent with previous theoretical work (Baker et al., 2006). Interestingly, for a class of mechanisms that assumes that the “size” of a somite is what is determined, rather than the “position” of the next somitic furrow (e.g. somite size is determined by the wavelength of traveling waves, or the wavelength of a Turing-type pattern), then the predicted result is qualitatively different (Figure 6L). In these models, somite size scales with the smaller PSM resulting from the induced larger somite, and then somite size gradually goes back to the normal size without rebounding dynamics. Together, the clock and scaled gradient model is uniquely supported by our experimental tests.

Traveling waves have a minor effect in the clock and scaled gradient model

Spatial differences in the phase of the coupled oscillators comprising the segmentation clock are known to create traveling waves of clock gene expression in the PSM from the posterior to the anterior (Ares et al., 2012; Ay et al., 2014; Giudicelli et al., 2007; Morelli et al., 2009; Uriu et al., 2009), but a mechanistic role for these waves is unclear. Thus far we have assumed synchronous oscillations throughout the PSM in our model for simplicity, as was done in the original clock and wavefront model (Cooke and Zeeman, 1976). To see how traveling waves affect the clock and scaled gradient model, we assumed a simple linear phase gradient along the AP axis (for details, see supplementary materials and methods) and repeated the simulations. As shown in Figure 7A, this results in only a minor modification to somite sizes as compared to a model without a phase gradient. Interestingly, we noticed that the somite formation period (defined as the time between successive boundaries being specified) was smaller when including a phase gradient (Figure 7B). This is consistent with the observation of the segmentation period in zebrafish being slightly faster than the intrinsic clock period (Soroldoni et al., 2014), and is reminiscent of the Doppler Effect, in which an observer moving towards a source of traveling waves measures a higher frequency than the intrinsic frequency of

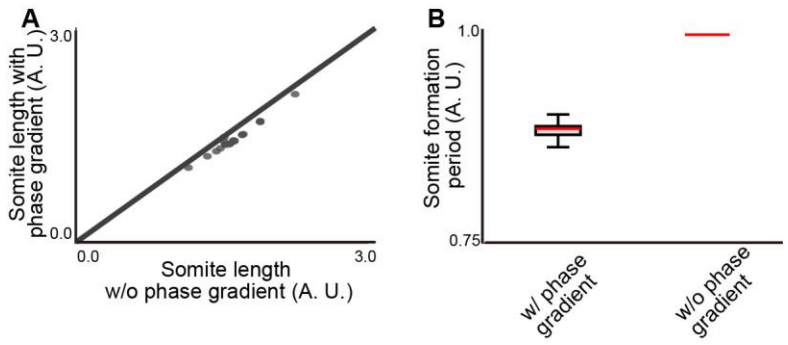


Figure 7 Traveling waves have modest effects on the clock and scaled gradient model. (**A**) Somite sizes are only slightly changed (~9%) by the presence of a phase gradient. (**B**) The phase gradient decreased the segmentation period (~11%). Error bars denote SD.

the oscillators. In Soroldoni et al., this effect is attributed to the movement of the anterior boundary of the PSM; the simulations in Figure 7 suggest it could also be due to the movement of the gradient caused by its scaling.

Discussion

Here we have proposed a novel mechanism for somite size control: the clock and scaled gradient model. This model is based on the original clock and wavefront model but the wavefront specifies new somite boundaries at a position defined at a fixed percentage along the PSM causing a somite to form with a delay of ~4 cycles. Previously, multiple models of somitogenesis have been proposed, but were difficult to experimentally distinguish since they were all consistent with existing data from wild type embryos as well as existing experimental perturbations. Here we utilized a novel perturbation—changing system size—to discriminate between existing models, and showed that only the clock and scaled gradient model can explain existing data and our new experimental data. We found that in patterning of the somites, somite length scales with PSM length *in vivo*. Importantly, we demonstrate that the delay between somite boundary specification and formation is critical to examining the relationship between

somite and PSM length. This is because the change in PSM length (and as a result, somite length) is dynamic, as a result of the changing rates of PSM production by axis elongation and consumption by somite formation (Figure 1G). Consistently, when the PSM is grown in culture conditions that do not permit axis extension, there is a monotonic decrease in PSM size and somite-PSM scaling is observable without considering delay (Lauschke et al., 2013). Considering the delay time between somite boundary specification and the appearance of a morphological somite will be essential in studying somite scaling in other situations, such as in other species, where complex dynamics of PSM length can be observed (Gomez et al., 2008; Schroter et al., 2008).

The clock and wavefront model is the classic model for somitogenesis (Cooke and Zeeman, 1976) and explains a number of previous experimental observations. In the original clock and wavefront model, what controls wavefront progression and how it is linked to axis elongation is unspecified. A simple way to specify wavefront progression is to just tie it to axis elongation (Cooke and Zeeman, 1976; Dubrulle and Pourquie, 2004; Hubaud and Pourquie, 2014; Saga, 2012). The consequence of this version of the clock and wavefront model is that somite size is equal to how far the tail moves in one clock cycle. While many of the existing perturbations can be explained by this version of the clock and wavefront model, a key feature that cannot be predicted is the phenomenon of scaling. For example, in the absence of axis elongation no somites should form, but this prediction is contradicted by *in vitro* cultured PSM which has no axis elongation yet forms a series of progressively smaller somites (Lauschke et al., 2013). Similarly, this simple clock and wavefront model does not predict the non-monotonic somite size variations following induction of a single large somite, as seen in Figure 6, since perturbations to the anterior PSM should not affect wavefront position.

An alternative class of models for explaining somite formation is based on using the wavelength of traveling waves in determining somite size (Jorg et al., 2015; Jorg et al., 2016;

Lauschke et al., 2013). However, previous studies in addition to our new results suggest that the wavelength of the traveling waves is not the primary mechanism to set somite sizes. First, the simple scenario (Lauschke et al., 2013) assumes that the phase gradient (inverse of wavelength) of the entire PSM scales with PSM length and that the scaled wavelength sets the somite size. However, (Soroldoni et al., 2014) and our results show that phase gradient does not always scale with PSM length, which argues against this simple mechanism. Second, one could still imagine some modification of the simple wavelength model would explain *in vivo* situation of somite scaling (e.g., the wavelength at B-4 locally scales with PSM length). However, this model is still hard to reconcile with the echo effect we observed after inducing one large somite (Figure 6), because regardless of the details, this class of models assumes somite "size" (not somitic furrow position) is controlled by the wavelength. In Figure 6L, we explicitly model the case where somite sizes scale with PSM size (including the 4-cycle delay) and find that it cannot explain the echo effect. In order to directly test if traveling waves are functional, one should experimentally modify the spatial pattern of the waves (for example, changing or eliminating the spatial phase gradient), without affecting the intrinsic period of the oscillators (Soza-Ried et al., 2014), and a mechanism for detecting a spatial gradient in clock gene expression level should be proposed. We suggest that traveling waves may be a byproduct of the need to synchronize oscillators locally (within the spatial scale of a somitic furrow), that while visually striking and mathematically interesting, have only a peripheral role in somite formation.

Another type of model is 'Turing-like', in which somites are formed via a combination of a periodic Turing instability (Cooke and Zeeman, 1976; Cotterell et al., 2015). There are several reasons why our data does not support Turing-like models. Firstly, a recent paper (Cotterell et al., 2015) showed how a Turing-like model of somitogenesis could, in principle, explain somite scaling, if one allowed the level of Fgf to effectively modulate the Turing-spacing of the somites. However, the change in somite size in response to PSM length is small, and is inconsistent with

our *in vivo* measurement where somite length is almost proportional to PSM length (Figure 1F and 2D). A second argument against a Turing-like model is that, unlike the clock and wavefront and clock and scaled gradient models, the 'clock' is not separable from the other components in the system. Therefore we don't necessarily expect a slower clock to increase somite size, at least not in perfect proportion as has been observed *in vivo* (Herrgen et al., 2010; Kim et al., 2011; Schroter and Oates, 2010) since a change in clock period would be associated with other parameters. Thirdly, the assumption that Fgf modulates the Turing-spacing of somites is incompatible with the results of perturbing Fgf, specifically: 1) a Turing-like model predicts consistently larger somites following sustained Fgf inhibition, which we do not see (Figure 4 L nd M); 2) a Turing-like model predicts a symmetric effect of implanting a Fgf bead (i.e. smaller somites anterior and posterior to the bead) unlike what is seen *in vivo* which shows a definite anterior-posterior bias (Dubrulle et al., 2001; Sawada et al., 2001). Finally, it is difficult to reconcile a Turing-like model with the results from Figure 6. The reason is that, like the phase-gradient model, and unlike the clock and scaled gradient model, Turing-like models fundamentally control somite size, not somite boundary position. Therefore, for exactly the same reasons as argued for the phase-based models, even with perfect somite size scaling in wildtype embryos, we do not predict the non-monotonic segment size variation following transient Fgf inhibition.

The clock and scaled gradient model presented here is a fairly simple model. We used a simple model for three reasons: 1) so that the key assumptions of the model (clock + scaling gradient) are directly supported by experimental data; 2) so that the model is at the right level of detail to make comparisons to our data; and 3) so that the model gives us a qualitative and intuitive understanding of somite size control, which may be obscured in a more complex model (Gunawardena, 2014). However, the model's simplicity does mean that it should not be viewed as a comprehensive, nor completely realistic, model of somitogenesis. Firstly, we have

assumed that somite maturation, and its effects on gradient scaling, occur instantly, whereas in reality we expect this to be a more gradual effect. Mathematically, this might mean that the 4-cycle delay should be changed from a step function to a more slowly varying function. This modification may be particularly important to understand the formation of the first four somites, and to reduce the sensitivity of somite size to initial conditions and/or perturbations. A second shortcoming of our model is that we have chosen the somite boundary to be set by a simple threshold of the gradient - an assumption that has not been directly measured, and is likely a simplification. Thirdly, we have largely focused on dpERK as readout of wavefront activity and demonstrated dpERK scaling as a proof of concept. However, the wavefront could be set by a complex function of multiple inputs such as Fgf and Wnt along with downstream signal integration (Bajard et al., 2014; Stulberg et al., 2012; Wahl et al., 2007), without affecting the core conclusions of our model. As reported, dpERK shows a steep gradient (Akiyama et al., 2014), but in our model, similar somite formation dynamics can be observed regardless of the precise shape of the gradient; even a simple linear gradient can recapitulate the *in vivo* behavior rather closely (Figure 4E and F).

One reason we chose to look at scaling of somites in size-reduced embryos is that we thought we might discover a mechanism for scaling that is *not* based on scaling of a molecular gradient (e.g. change in axis extension speed, growth rate, phase gradient, oscillation period). However, in the end we found that scaling of a molecular gradient is indeed what underlies somite scaling as has been observed in other examples of pattern scaling (Ben-Zvi et al., 2010a; Ben-Zvi et al., 2011; Inomata et al. 2013). Future research on this issue could reveal what design benefits (e.g. robustness, evolvability) systems employing gradient scaling have compared to other potential mechanisms for scaling.

Methods

Fish care.

Fish (AB) were kept at 27°C on a 14-hr-light/ 10-hr-dark cycle. Embryos were collected by natural crosses. All fish-related procedures were carried out with the approval of Institutional Animal Care and Use Committee (IACUC) at Harvard University.

Size reduction technique.

Chorions were enzymatically removed using pronase (Sigma Aldrich, 1mg/ml in egg water (Westerfield, 2000)) at ~512 cell stage. Eggs were treated with pronase until the chorions loses their tension and washed gently with egg water. Remaining chorions were removed manually using tweezers. The embryos were placed on a glass dish with 1/3 Ringer's solution (Westerfield, 2000), with 2% methylcellulose (Sigma Aldrich) in 1/3 Ringer's solution spread thinly on the bottom of the dish, to restrict movement embryos. We found using 1/3 Ringer's solution is critical for embryos to recover from the damage of chopping. Then the blastoderm was chopped at the animal pole, and the yolk was wounded, resulting in oozing out of the yolk, using either hand-pulled glass pipette or looped steel wire (30 μm in diameter) glued in the capillary glass. The chopped embryos were incubated in the 1/3 Ringer's solution for 30 min, and then moved to fresh 1/3 Ringer's solution for further incubation. The survival rate of the chopped embryos varies depending on condition of the embryos. Healthy embryos and good dissection would produce maximum ~60% of success rate; developing normally until late stages (at least several days). The ratio of remaining cells and yolk affects how normal the embryos develop; usually cutting 50% position of blastula horizontally and wounding the vegetal part of yolk produces good results.

BCI and SU5402 treatment.

Embryos were treated with BCI (Dual Specificity Protein Phosphatase 1/6 Inhibitor, Calbiochem) as described (Akiyama et al., 2014). For SU5402 treatment, embryos were treated at a low concentration (Calbiochem, 16 μM) to minimize toxicity.

Morpholino injection.

raldh1 morpholinos (Kawakami et al., 2005; Yabe et al., 2003) were injected (Nanoinject) at 1 cell stage at the concentration of 2 mM.

Imaging.

For live imaging, the embryos were mounted laterally using the dorsal mount (Megason, 2009) in egg water with 0.01% tricaine (Wentern Chemical, Inc.). Live imaging was performed using Zeiss Axio Observer Z1 and AxioCam MRm. For multiple image acquisition, we used a motorized stage, controlled by AxioVision 3.8. The temperature was maintained at 28.5 ± 0.5°C using a home-made incubator. The images were taken every 2 min, and the size of z slice varied depending on the size of embryos. The images of the *in situ* hybridization samples were also acquired using Zeiss Axio Observer Z1. The images of dpERK immunostaining samples were acquired using Leica TCS SP8. Finally, a Nikon Ti spinning disk confocal was used to acquire the images of transplanted samples.

Image processing.

Image processing was done using FIJI (Schindelin et al., 2012) and Matlab custom code. For time course measurement of axis elongation and somite size, we used the Gaussian-based stack focuser in FIJI. For axis elongation measurement, the length from 4th somite to tail tip was measured, using FIJI's LOI interpolator. For *in situ* hybridization samples and immunostaining samples, noise was first reduced using Gaussian blur (sigma = 7.0), and the signal was

extracted along AP axis, using FIJI's Plot profile function. To compare intensity profiles of BCI and SU treated embryos (Figure 6I, Figure S1 and S9), we averaged over multiple embryos. To calculate relative intensity, first, the minimum value was set to 0; and then the intensities at each position was scaled with a scaling factor of (average maximum intensity in drug treated embryo/average maximum intensity of untreated embryo).

***In situ* hybridization and immunostaining.**

In situ hybridization (Nikaido et al., 1997) was performed as previously described. dpERK immunostaining was performed basically following the protocol described in Sawada et al.(Sawada et al., 2001), except that we used Alexa Fluor 488 goat anti-mouse IgG (ThermoFisher Scientific A-11001) as the 2nd antibody. Nuclei were stained with propidium iodide (Life Technologies P1304MP).

Somite/PSM transplantation.

Transplantation was performed as described(Haines et al., 2004; Kawanishi et al., 2013), with minor modification. For making a cut on the skin, we used a mouth pipette filled with pancreatin, so the cut can be made both physically and enzymatically. Embryos for donor tissue were injected with Alexa Fluor 680 conjugated 10,000 MW Dextran, which can be detected directly after immunostaining.

Live imaging of ERK activity dynamics.

The FRET-based Erk biosensor termed Eevee-ERKnls is composed by an enhanced cyan-emitting mutant of GFP (ECFP), a WW domain (ligand domain), an EV linker, an Erk substrate (sensor domain), a yellow fluorescent protein for energy transfer (Ypet), and a nuclear localization signal (NLS) (Komatsu et al., 2011). When Erk phosphorylates the Erk substrate, the WW domain binds to the Erk substrate, leading to the induction of FRET from ECFP to Ypet.

It has been confirmed that the Erk biosensor can monitor FGF-dependent Erk activity in living zebrafish embryos (Dini Wahyu Kartika Sari, 2018). One cell stage of embryos were injected mRNA encoding a *FRET-based ERK biosensor* termed *Eevee-ERKnls* (Dini Wahyu Kartika Sari, 2018; Komatsu et al., 2011). The embryos at a certain stage were excited with a 440-nm laser, and fluorescence spectra were acquired by using a Lambda Scanning mode of a LSM710 confocal microscope (Zeiss). Using a Linear Unmixing mode, CFP and Ypet signals were separated from the original spectra data. FRET/CFP ratio images and kymographs were created with MetaMorph software (Molecular Device).

Statistical testing.

Significance was calculated by one-tailed Student's t tests, using Excel (Microsoft). Unequal variance comparison was performed for Figure 1D, Figure 2 B and C, and equal variance comparison was performed for Figure 5D and Figure 6J.

Wavelet transform.

We follow the approach of (Soroldoni et al., 2014) and use the wavelet transform to generate phase maps for *her1* along the embryo. Consider that the *her1* pattern is of the form:

$$h(x) = h_0 + A(x)\sin(\phi(x) + \Phi)$$

i.e. has a spatially varying amplitude, $A(x)$ and a spatially varying phase, $\phi(\mathrm{x})$. By performing a wavelet transform we can convert the intensity profile $h(x)$ into an effective phase profile $\phi(x)$, plotted in Figure 3E. Note, we plot the phase for positions more anterior than the first clear peak since it is only in these ranges where there is a distinct spatial pattern above noise, and, in all cases, contains the position at which the next somite boundary is specified i.e. B-4. We also measured the phase gradient manually, by identifying peaks and troughs in the intensity profile

(separated by π). This manual measurement (orange triangles in Figure 3E) was found to well match the corresponding phases as obtained from the wavelet transform, giving us confidence in our implementation. For further details of the wavelet transform, we refer the reader to (Soroldoni et al., 2014) for more discussion.

Acknowledgments

We thank O. Pourquié, A. Aulehla and A. Oates for critical discussion. Some images were taken at Nikon Imaging Center at Harvard Medical School. K.I. acknowledges a Postodoctoral Fellowship for Research Abroad (Japan Society for Promotion of Science).

Competing interest

The authors declare no competing or financial interests.

Funding

The work was supported by the PRESTO program (JPMJPR11AA) of the Japan Science and Technology. This work was partially supported by Grants-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology (MEXT), Japan (for Y.B and T.M) Agency. T.H. was supported by the Herchel Smith Graduate Fellowship.

References

Akiyama, R., Masuda, M., Tsuge, S., Bessho, Y. and Matsui, T. (2014). An anterior limit of FGF/Erk signal activity marks the earliest future somite boundary in zebrafish. *Development* 141, 1104-9.

Ares, S., Morelli, L. G., Jorg, D. J., Oates, A. C. and Julicher, F. (2012). Collective modes of coupled phase oscillators with delayed coupling. *Phys Rev Lett* 108, 204101.

Ay, A., Holland, J., Sperlea, A., Devakanmalai, G. S., Knierer, S., Sangervasi, S., Stevenson, A. and Ozbudak, E. M. (2014). Spatial gradients of protein-level time delays set the pace of the traveling segmentation clock waves. *Development* 141, 4158-67.

Bajard, L., Morelli, L. G., Ares, S., Pecreaux, J., Julicher, F. and Oates, A. C. (2014). Wnt-regulated dynamics of positional information in zebrafish somitogenesis. *Development* 141, 1381-91.

Baker, R. E., Schnell, S. and Maini, P. K. (2006). A clock and wavefront mechanism for somite formation. *Dev Biol* 293, 116-26.

Ben-Zvi, D. and Barkai, N. (2010). Scaling of morphogen gradients by an expansion-repression integral feedback control. *Proc Natl Acad Sci U S A* 107, 6924-9.

Cooke, J. (1975). Control of somite number during morphogenesis of a vertebrate, Xenopus laevis. *Nature* 254, 196-9.

Cooke, J. (1981). Scale of body pattern adjusts to available cell number in amphibian embryos. *Nature* 290, 775-8.

Cooke, J. and Zeeman, E. C. (1976). A clock and wavefront model for control of the number of repeated structures during animal morphogenesis. *J Theor Biol* 58, 455-76.

Cotterell, J., Robert-Moreno, A. and Sharpe, J. (2015). A Local, Self-Organizing Reaction-Diffusion Model Can Explain Somite Patterning in Embryos. *Cell Syst* 1, 257-69.

Delaune, E. A., Francois, P., Shih, N. P. and Amacher, S. L. (2012). Single-cell-resolution imaging of the impact of Notch signaling and mitosis on segmentation clock dynamics. *Dev Cell* 23, 995-1005.

Dequeant, M. L., Glynn, E., Gaudenz, K., Wahl, M., Chen, J., Mushegian, A. and Pourquie, O. (2006). A complex oscillating network of signaling genes underlies the mouse segmentation clock. *Science* 314, 1595-8.

Dini Wahyu Kartika Sari, R. A., Honda Naoki, Hannosuke Ishijima, Yasumasa Bessho, and Takaaki Matsui. (2018). Time-lapse observation of stepwise regression of Erk activity in zebrafish presomitic mesoderm. *Scientific Reports* 8.

Dubrulle, J., McGrew, M. J. and Pourquie, O. (2001). FGF signaling controls somite boundary position and regulates segmentation clock control of spatiotemporal Hox gene activation. *Cell* 106, 219-32.

Dubrulle, J. and Pourquie, O. (2004). fgf8 mRNA decay establishes a gradient that couples axial elongation to patterning in the vertebrate embryo. *Nature* 427, 419-22.

Elsdale, T., Pearson, M. and Whitehead, M. (1976). Abnormalities in somite segmentation following heat shock to Xenopus embryos. *J Embryol Exp Morphol* 35, 625-35.

Giudicelli, F., Ozbudak, E. M., Wright, G. J. and Lewis, J. (2007). Setting the tempo in development: an investigation of the zebrafish somite clock mechanism. *PLoS Biol* 5, e150.

Gomez, C., Ozbudak, E. M., Wunderlich, J., Baumann, D., Lewis, J. and Pourquie, O. (2008). Control of segment number in vertebrate embryos. *Nature* 454, 335-9.

Goudevenou, K., Martin, P., Yeh, Y. J., Jones, P. and Sablitzky, F. (2011). Def6 is required for convergent extension movements during zebrafish gastrulation downstream of Wnt5b signaling. *PLoS One* 6, e26548.

Gregor, T., Bialek, W., de Ruyter van Steveninck, R. R., Tank, D. W. and Wieschaus, E. F. (2005). Diffusion and scaling during early embryonic pattern formation. *Proc Natl Acad Sci U S A* 102, 18403-7.

Gregor, T., McGregor, A. P. and Wieschaus, E. F. (2008). Shape and function of the Bicoid morphogen gradient in dipteran species with different sized embryos. *Dev Biol* 316, 350-8.

Gunawardena, J. (2014). Models in biology: 'accurate descriptions of our pathetic thinking'. *BMC Biol* 12, 29.

Haines, L., Neyt, C., Gautier, P., Keenan, D. G., Bryson-Richardson, R. J., Hollway, G. E., Cole, N. J. and Currie, P. D. (2004). Met and Hgf signaling controls hypaxial muscle and lateral line development in the zebrafish. *Development* 131, 4857-69.

Herrgen, L., Ares, S., Morelli, L. G., Schroter, C., Julicher, F. and Oates, A. C. (2010). Intercellular coupling regulates the period of the segmentation clock. *Curr Biol* 20, 1244-5

Holley, S. A., Geisler, R. and Nusslein-Volhard, C. (2000). Control of her1 expression during zebrafish somitogenesis by a delta-dependent oscillator and an independent wave-front activity. *Genes Dev* 14, 1678-90.

Hubaud, A. and Pourquie, O. (2014). Signalling dynamics in vertebrate segmentation. *Nat Rev Mol Cell Biol* 15, 709-21.

Inomata, H., Shibata, T., Haraguchi, T. and Sasai, Y. (2013). Scaling of dorsal-ventral patterning by embryo size-dependent degradation of Spemann's organizer signals. *Cell* 153, 1296-311.

Jorg, D. J., Morelli, L. G., Soroldoni, D., Oates, A. C. and Julicher, F. (2015). Continuum theory of gene expression waves during vertebrate segmentation. *New J Phys* 17, 093042.

Jorg, D. J., Oates, A. C. and Julicher, F. (2016). Sequential pattern formation governed by signaling gradients. *Phys Biol* 13, 05LT03.

Kawakami, Y., Raya, A., Raya, R. M., Rodriguez-Esteban, C. and Izpisua Belmonte, J. C. (2005). Retinoic acid signalling links left-right asymmetric patterning and bilaterally symmetric somitogenesis in the zebrafish embryo. *Nature* 435, 165-71.

Kawanishi, T., Kaneko, T., Moriyama, Y., Kinoshita, M., Yokoi, H., Suzuki, T., Shimada, A. and Takeda, H. (2013). Modular development of the teleost trunk along the dorsoventral axis and zic1/zic4 as selector genes in the dorsal module. *Development* 140, 1486-96.

Kim, W., Matsui, T., Yamao, M., Ishibashi, M., Tamada, K., Takumi, T., Kohno, K., Oba, S., Ishii, S., Sakumura, Y. et al. (2011). The period of the somite segmentation clock is sensitive to Notch activity. *Mol Biol Cell* 22, 3541-9.

Komatsu, N., Aoki, K., Yamada, M., Yukinaga, H., Fujita, Y., Kamioka, Y. and Matsuda, M. (2011). Development of an optimized backbone of FRET biosensors for kinases and GTPases. *Mol Biol Cell* 22, 4647-56.

Krol, A. J., Roellig, D., Dequeant, M. L., Tassy, O., Glynn, E., Hattem, G., Mushegian, A., Oates, A. C. and Pourquie, O. (2011). Evolutionary plasticity of segmentation clock networks. *Development* 138, 2783-92.

Lander, A. D., Nie, Q., Vargas, B. and Wan, F. Y. (2011). Size-normalized Robustness of Dpp Gradient in Drosophila Wing Imaginal Disc. *J Mech Mater Struct* 6, 321-350.

Lauschke, V. M., Tsiairis, C. D., Francois, P. and Aulehla, A. (2013). Scaling of embryonic patterning based on phase-gradient encoding. *Nature* 493, 101-5.

Masamizu, Y., Ohtsuka, T., Takashima, Y., Nagahara, H., Takenaka, Y., Yoshikawa, K., Okamura, H. and Kageyama, R. (2006). Real-time imaging of the somite segmentation clock: revelation of unstable oscillators in the individual presomitic mesoderm cells. *Proc Natl Acad Sci U S A* 103, 1313-8.

McHale, P., Rappel, W. J. and Levine, H. (2006). Embryonic pattern scaling achieved by oppositely directed morphogen gradients. *Phys Biol* 3, 107-20.

Megason, S. G. (2009). In toto imaging of embryogenesis with confocal time-lapse microscopy. *Methods Mol Biol* 546, 317-32.

Morelli, L. G., Ares, S., Herrgen, L., Schroter, C., Julicher, F. and Oates, A. C. (2009). Delayed coupling theory of vertebrate segmentation. *Hfsp J* 3, 55-66.

Nikaido, M., Tada, M., Saji, T. and Ueno, N. (1997). Conservation of BMP signaling in zebrafish mesoderm patterning. *Mech Dev* 61, 75-88.

O'Connor, M. B., Umulis, D., Othmer, H. G. and Blair, S. S. (2006). Shaping BMP morphogen gradients in the Drosophila embryo and pupal wing. *Development* 133, 183-93.

Ozbudak, E. M. and Lewis, J. (2008). Notch signalling synchronizes the zebrafish segmentation clock but is not needed to create somite boundaries. *PLoS Genet* 4, e15.

Palmeirim, I., Henrique, D., IshHorowicz, D. and Pourquie, O. (1997). Avian hairy gene expression identifies a molecular clock linked to vertebrate segmentation and somitogenesis. *Cell* 91, 639-648.

Primmett, D. R., Norris, W. E., Carlson, G. J., Keynes, R. J. and Stern, C. D. (1989). Periodic segmental anomalies induced by heat shock in the chick embryo are associated with the cell cycle. *Development* 105, 119-30.

Rauch, G. J., Hammerschmidt, M., Blader, P., Schauerte, H. E., Strahle, U., Ingham, P. W., McMahon, A. P. and Haffter, P. (1997). Wnt5 is required for tail formation in the zebrafish embryo. *Cold Spring Harb Symp Quant Biol* 62, 227-34.

Roy, M. N., Prince, V. E. and Ho, R. K. (1999). Heat shock produces periodic somitic disturbances in the zebrafish embryo. *Mech Dev* 85, 27-34.

Saga, Y. (2012). The synchrony and cyclicity of developmental events. *Cold Spring Harb Perspect Biol* 4, a008201.

Sawada, A., Shinya, M., Jiang, Y. J., Kawakami, A., Kuroiwa, A. and Takeda, H. (2001). Fgf/MAPK signalling is a crucial positional cue in somite boundary formation. *Development* 128, 4873-80.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B. et al. (2012). Fiji: an open-source platform for biological-image analysis. *Nat Methods* 9, 676-82.

Schroter, C., Herrgen, L., Cardona, A., Brouhard, G. J., Feldman, B. and Oates, A. C. (2008). Dynamics of zebrafish somitogenesis. *Dev Dyn* 237, 545-53.

Schroter, C. and Oates, A. C. (2010). Segment number and axial identity in a segmentation clock period mutant. *Curr Biol* 20, 1254-8.

Shih, N. P., Francois, P., Delaune, E. A. and Amacher, S. L. (2015). Dynamics of the slowing segmentation clock reveal alternating two-segment periodicity. *Development* 142, 1785-93.

Soroldoni, D., Jorg, D. J., Morelli, L. G., Richmond, D. L., Schindelin, J., Julicher, F. and Oates, A. C. (2014). Genetic oscillations. A Doppler effect in embryonic pattern formation. *Science* 345, 222-5.

Soza-Ried, C., Ozturk, E., Ish-Horowicz, D. and Lewis, J. (2014). Pulses of Notch activation synchronise oscillating somite cells and entrain the zebrafish segmentation clock. *Development* 141, 1780-8.

Steventon, B., Duarte, F., Lagadec, R., Mazan, S., Nicolas, J. F. and Hirsinger, E. (2016). Species-specific contribution of volumetric growth and tissue convergence to posterior body elongation in vertebrates. *Development* 143, 1732-41.

Stulberg, M. J., Lin, A., Zhao, H. and Holley, S. A. (2012). Crosstalk between Fgf and Wnt signaling in the zebrafish tailbud. *Dev Biol* 369, 298-307.

Takahashi, J., Ohbayashi, A., Oginuma, M., Saito, D., Mochizuki, A., Saga, Y. and Takada, S. (2010). Analysis of Ripply1/2-deficient mouse embryos reveals a mechanism underlying the rostro-caudal patterning within a somite. *Dev Biol* 342, 134-45.

Thorpe, C. J., Weidinger, G. and Moon, R. T. (2005). Wnt/beta-catenin regulation of the Sp1-related transcription factor sp5l promotes tail development in zebrafish. *Development* 132, 1763-72.

Tsiairis, C. D. and Aulehla, A. (2016). Self-Organization of Embryonic Genetic Oscillators into Spatiotemporal Wave Patterns. *Cell* 164, 656-67.

Uriu, K., Morishita, Y. and Iwasa, Y. (2009). Traveling wave formation in vertebrate segmentation. *J Theor Biol* 257, 385-96.

Wahl, M. B., Deng, C., Lewandoski, M. and Pourquie, O. (2007). FGF signaling acts upstream of the NOTCH and WNT signaling pathways to control segmentation clock oscillations in mouse somitogenesis. *Development* 134, 4033-41.

Westerfield, M. (2000). The zebrafish book. A guide for the laboratory use of zebrafish (Danio rerio). University of Oregon Press, Eugene.

Yabe, T., Shimizu, T., Muraoka, O., Bae, Y. K., Hirata, T., Nojima, H., Kawakami, A., Hirano, T. and Hibi, M. (2003). Ogon/Secreted Frizzled functions as a negative feedback regulator of Bmp signaling. *Development* 130, 2705-16.

Supplementary Figures:

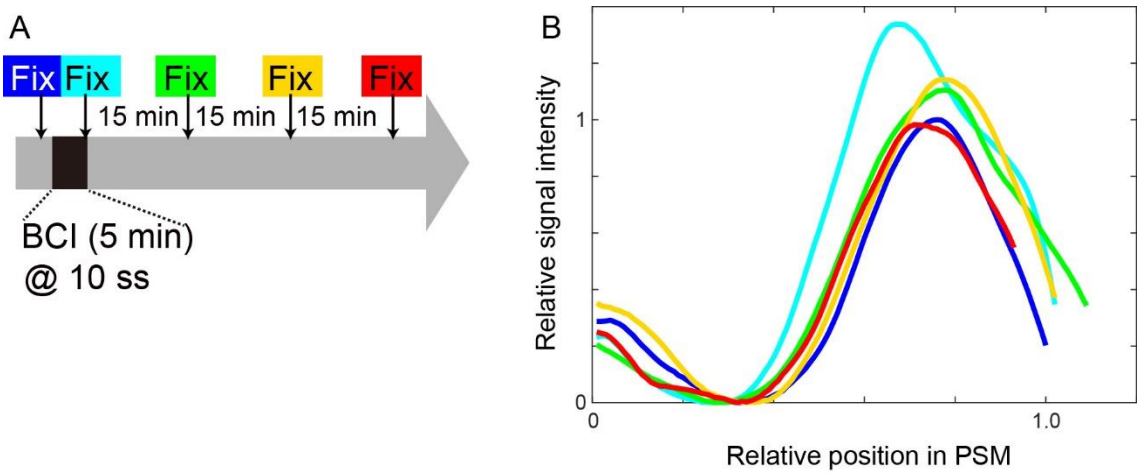


Figure S1 Time course analysis of dpERK intensity before and after transient BCI treatment. (**A**) Schematic illustration of the experiment. At each time point, control and treated embryos (seven embryos each) were fixed and analyzed. (**B**) Intensity curves were calculated by averaging intensity curves of treated embryos. Relative signal intensity (y axis) was determined by scaling factors: (maximum intensity of treated embryos) / (maximum intensity of control embryos at the corresponding time points). Relative position (x axis) was determined by normalizing positions in PSM in treated embryos by averaged PSM length of control embryos at the corresponding time points. Colors correspond to the colors in (**A**). dpERK intensity increases immediately after BCI treatment, and comes back to normal after 15 min of wash.

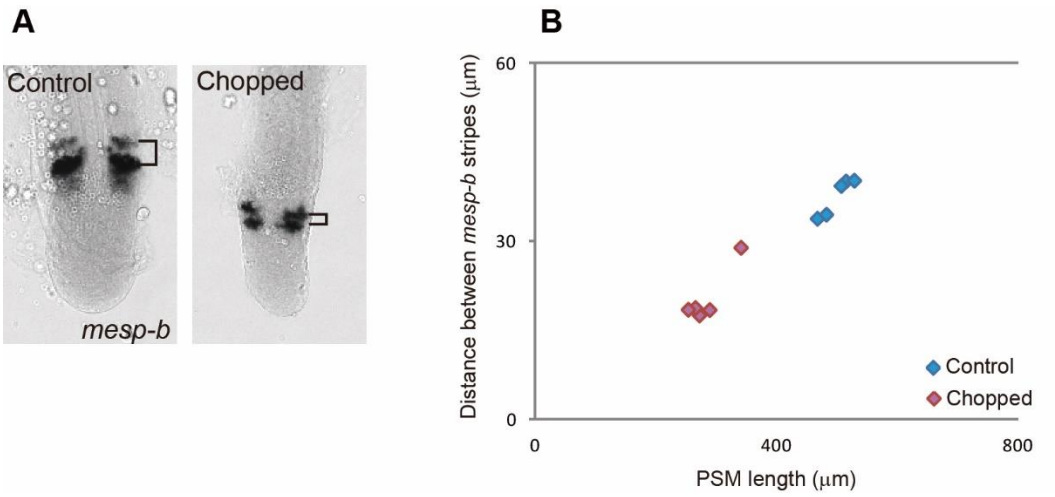


Figure S2 Scaling of *mesp-b* stripe. (**A**) *in situ* hybridization samples of *mesp-b*. (**B**) PSM length vs distance between *mesp-b* stripes.

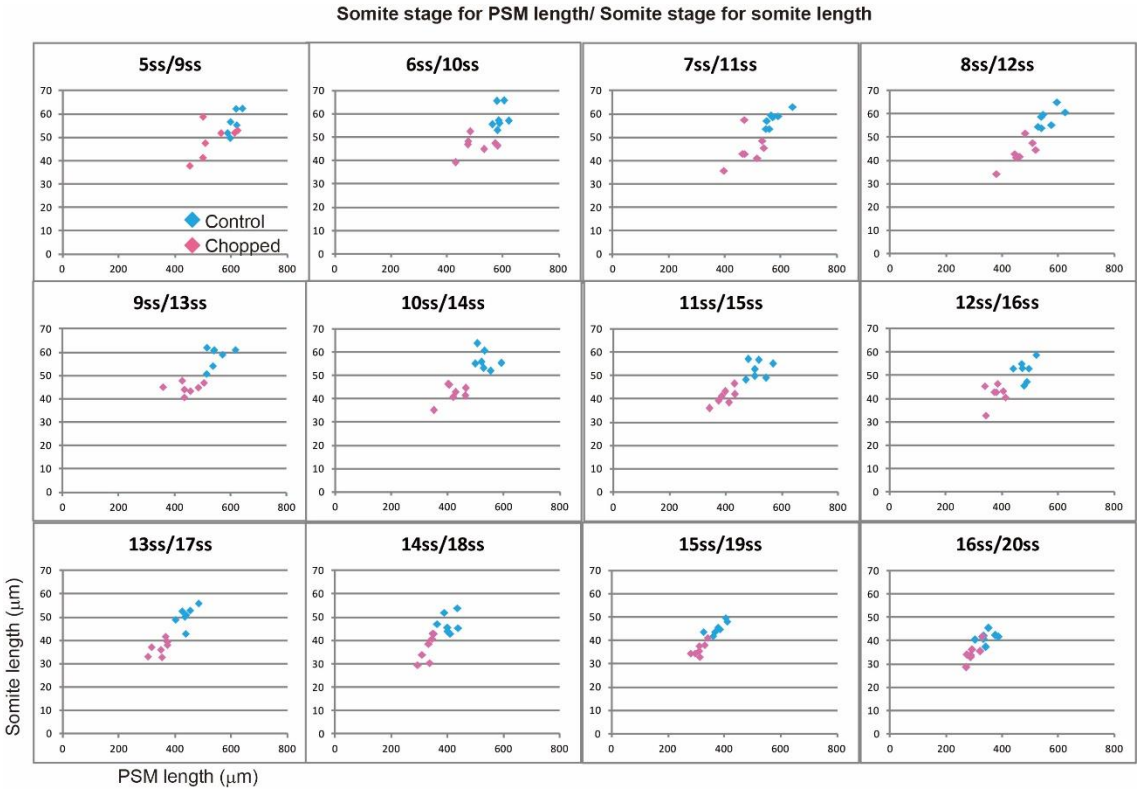


Figure S3 Somite size vs PSM size between control and chopped embryos from different somite stages.

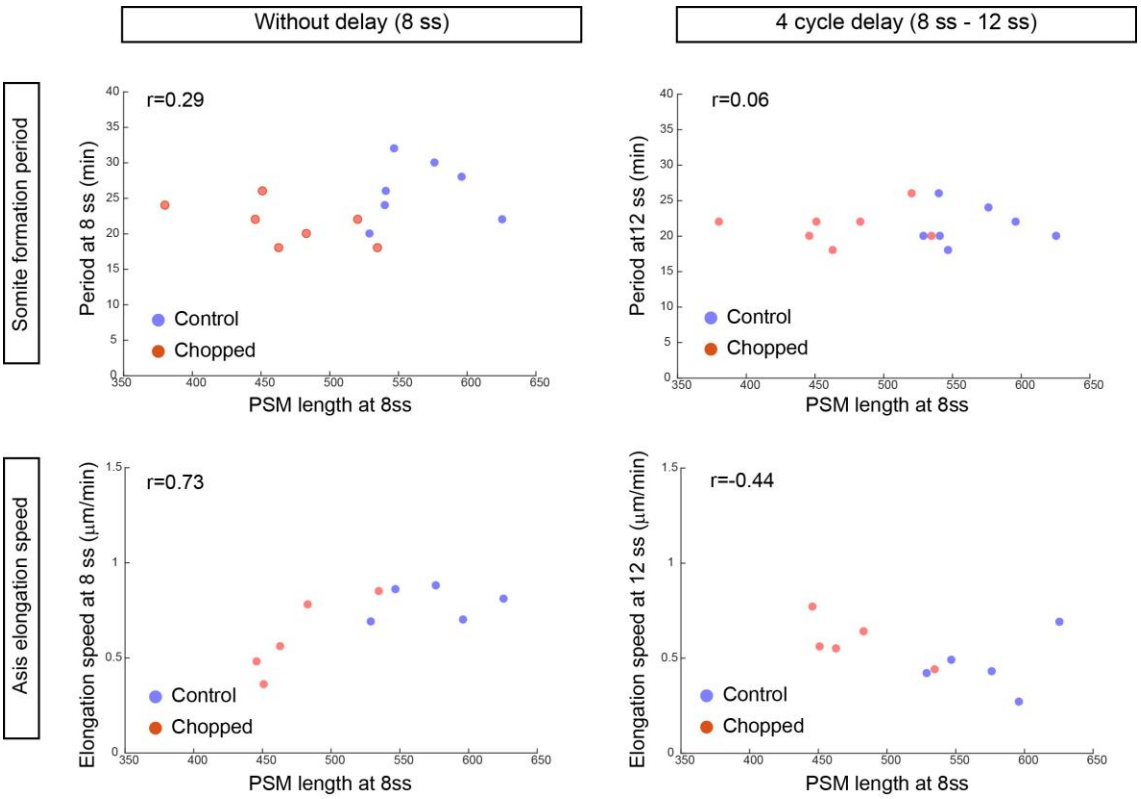


Figure S4 Clock period and axis elongation speed with and without 4 cycle delay. Neither clock period nor axis elongation speed scale with PSM length even when the time delay is taken into consideration.

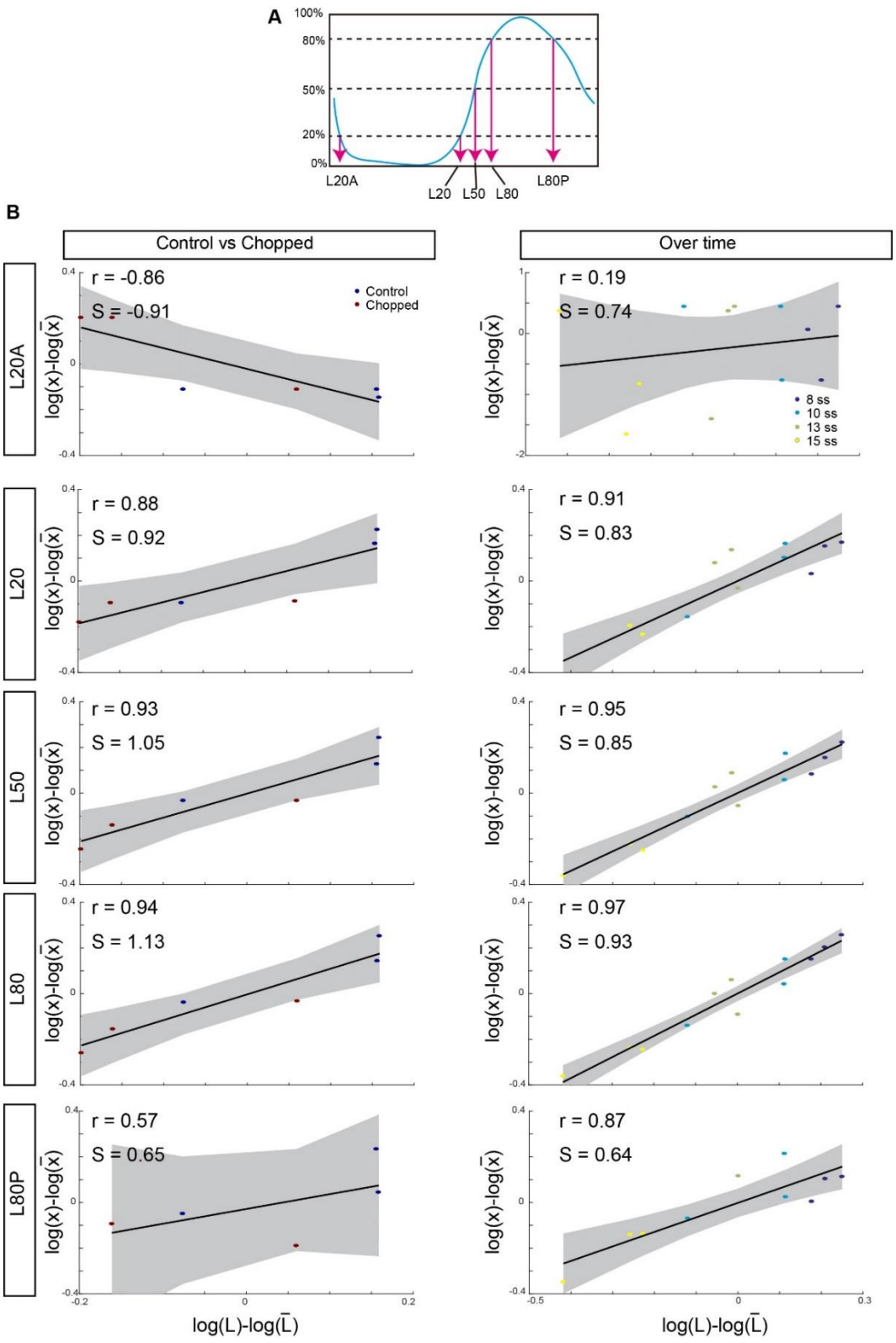


Figure S5 dpERK scaling for different thresholds.

Supplemental Figure 5 (Continued) (**A**) Schematic illustration of position names. (**B**) For each PSM, log-deviation in the position $\log(x/\bar{x}) = \log(x) - \log(\bar{x})$ are plotted against log-deviation in the PSM length $\log(L/\bar{L}) = \log(L) - \log(\bar{L})$, for different a given dpERK threshold intensities. The scaling coefficient S is obtained by linear regression (95% confidence interval on the slope is shown in gray). Correlation r is shown. In both cases (control vs chopped, and over time), L20, L50 and L80 scales with PSM length more than other positions.

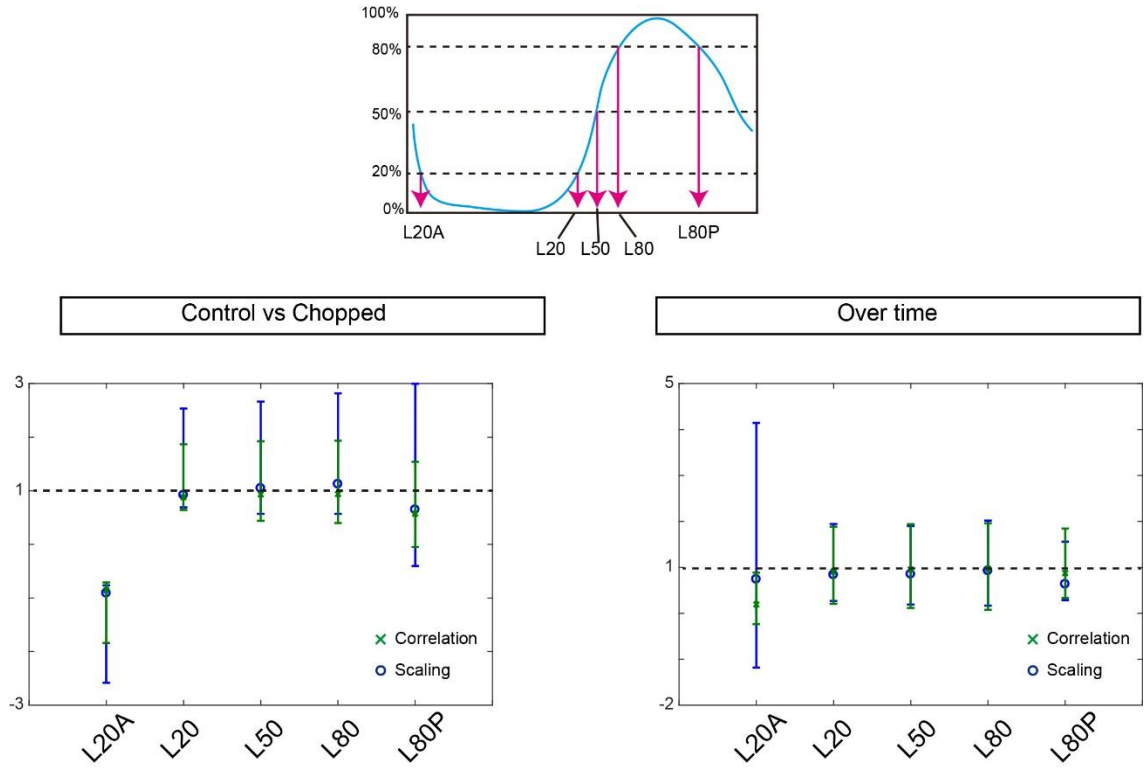


Figure S6 Correlation coefficient and scaling coefficient for dpERK gradient.

dpERK correlation (x) and scaling (o) for several threshold intensities. Error bars represent the 95% confidence intervals. At L20, L50 and L80, the correlation and scaling coefficients are closer to 1, compared to other positions, which is consistent with Figure S5.

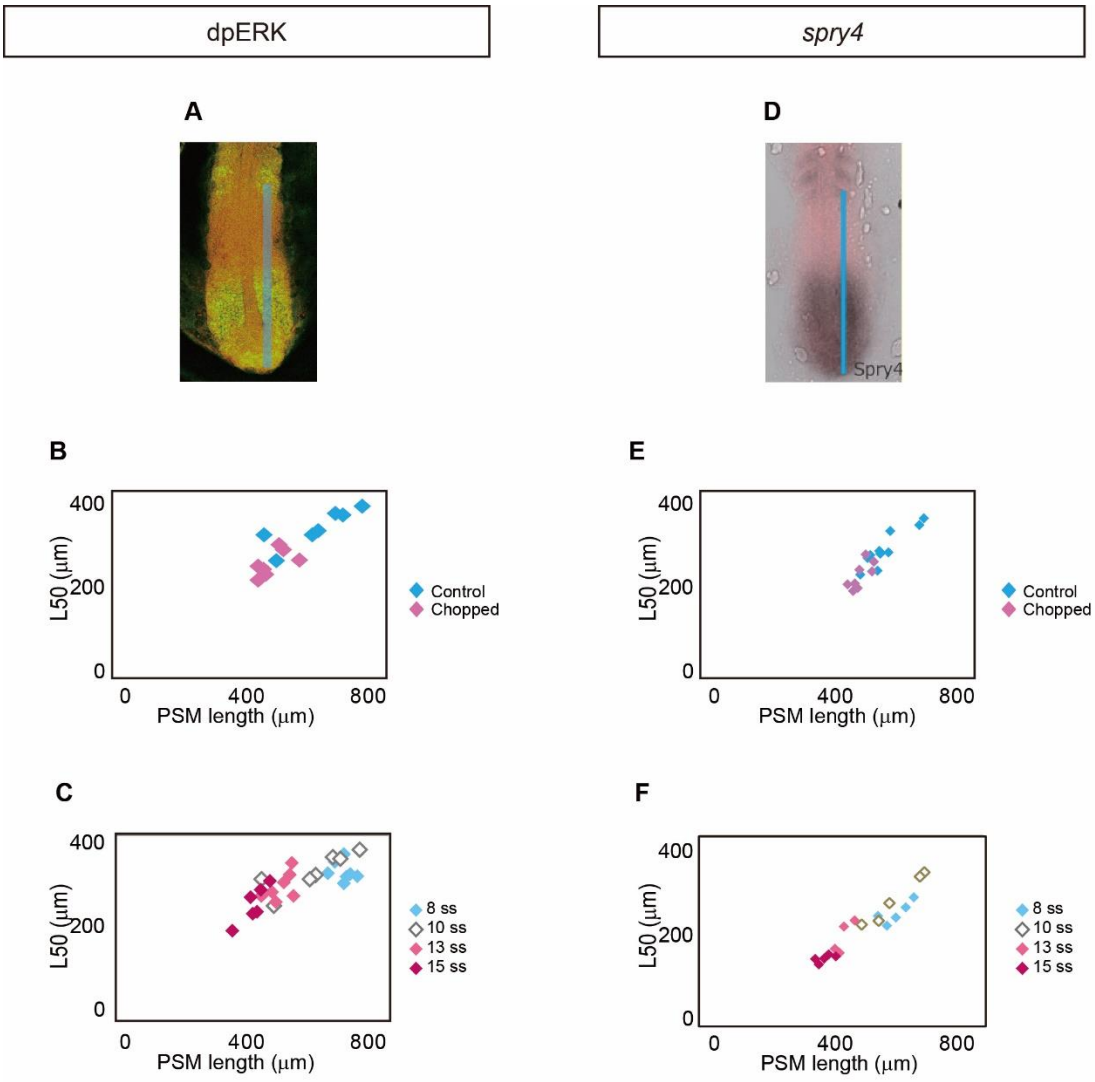


Figure S7 Fgf activity scaling detected by dpERK and *spry4.*

(**A** to **C**) L50 analysis on dpERK. (**D** to **F**) L50 analysis on *spry4* expression. (**A**) dpERK activity was detected by immunostaining. (**D**) *spry4* mRNA was detected by in situ hybridization. (**B** and **C**, **E** and **F**) L50 was calculated using *spry4* in situ hybridization samples similarly as in Figure 3**I**. For both dpERK and *spry4*, L50 was found to scale with PSM length both between control and chopped embryos (**B** and **E**) and over time (**C** and **F**)

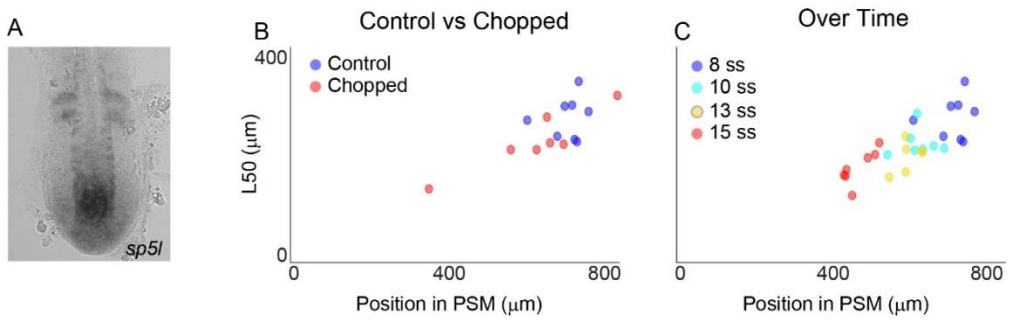


Figure S8 Wnt signal scales with PSM length.

(**A**) *sp5l* mRNA was detected by in situ hybridization. (**B** and **C**) L50 was calculated using *sp5l* in situ hybridization samples similarly as in Figure 3**I**. L50 was found to scale with PSM length both between control and chopped embryos (**B**) and over time (**C**)

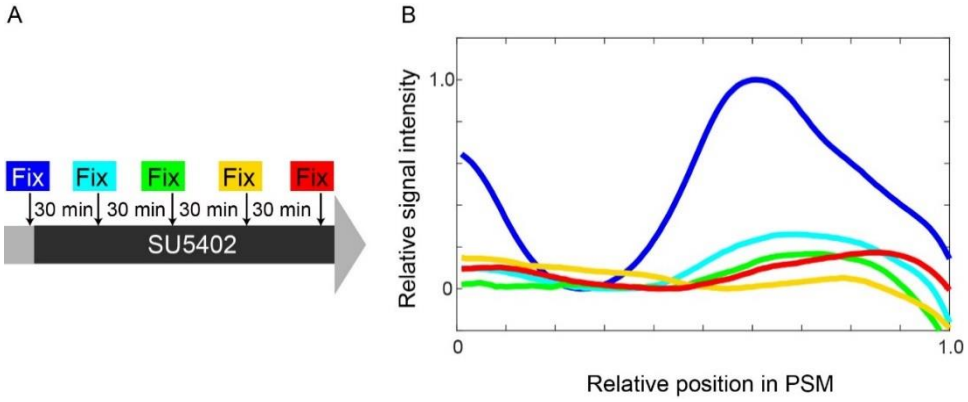


Figure S9

(**A**) Schematic illustration of the experiment. At each time point, control and treated embryos (seven embryos each) were fixed and analyzed. (**B**) Intensity curves were calculated by averaging intensity curves of treated embryos. Relative signal intensity (y axis) was determined by using scaling factors: (maximum intensity of treated embryos) / (maximum intensity of control embryos at the corresponding time points). Relative position (x axis) was determined by normalizing positions in PSM in treated embryos by averaged PSM length of control embryos at the corresponding time points. Colors correspond to the colors in (**A**). dpERK intensity drops immediately after onset of SU5402 treatment, and remain almost the same level over the course of experiment.

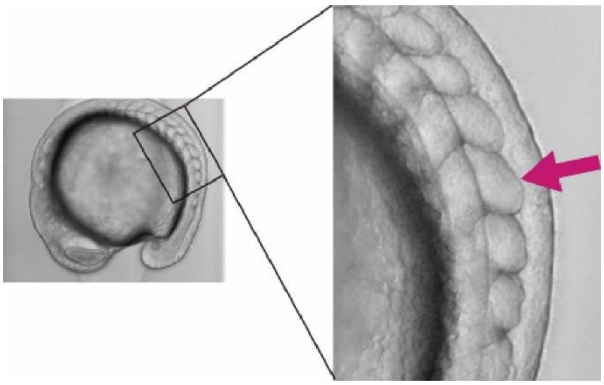


Figure S10 Long-term SU5402 treatment under constant dark condition

The embryos were treated with SU5402 at low concentration (16mM) for 4 hrs with the light completely blocked. One or two larger somites were formed (magenta arrow in the right panel) several cycles after initiation of the treatment (10 out of 11).

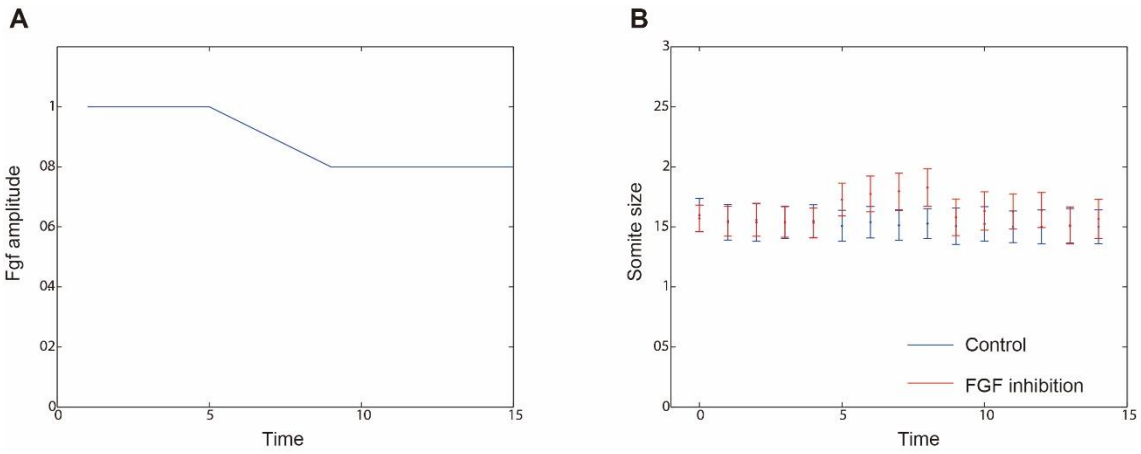


Figure S11 Increasing Fgf inhibition can cause multiple larger somites. Here using the clock and scaled gradient model, we simulated the situation in which FGF inhibition occurs increasingly, rather than in a step-wise manner (**A**). (**B**) As a result, multiple larger somites were predicted to be formed, consistent with the result in chick (Cotterell et al., 2015). Error bars denote SD.

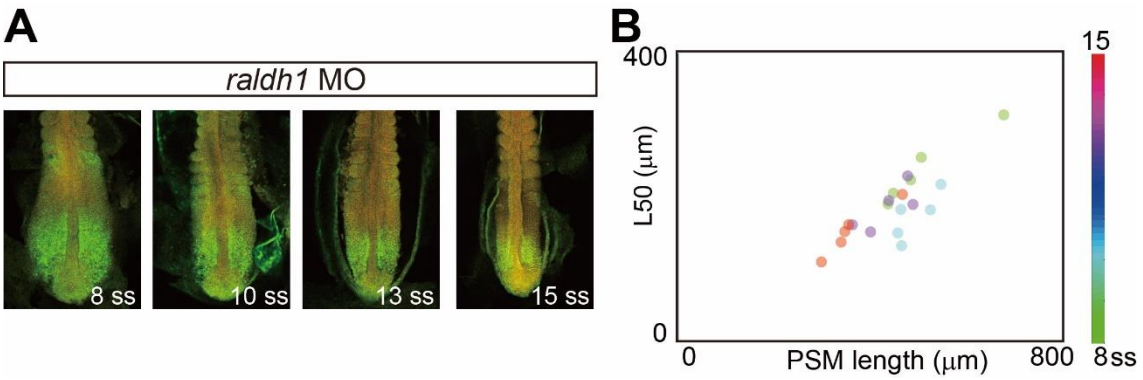


Figure S12 dpERK scaling in *raldh1* morphants. In search of the mechanism underlying Fgf activity gradient scaling, we attempted to knockdown *raldh,* a synthetase for retinoic acid. Embryos injected with *raldh* morpholino were subjected to dpERK immunostaining (**A**), and L50 was measured as described in the main text (**B**). Despite losing *raldh* expression, the Fgf gradient still scaled with PSM length.

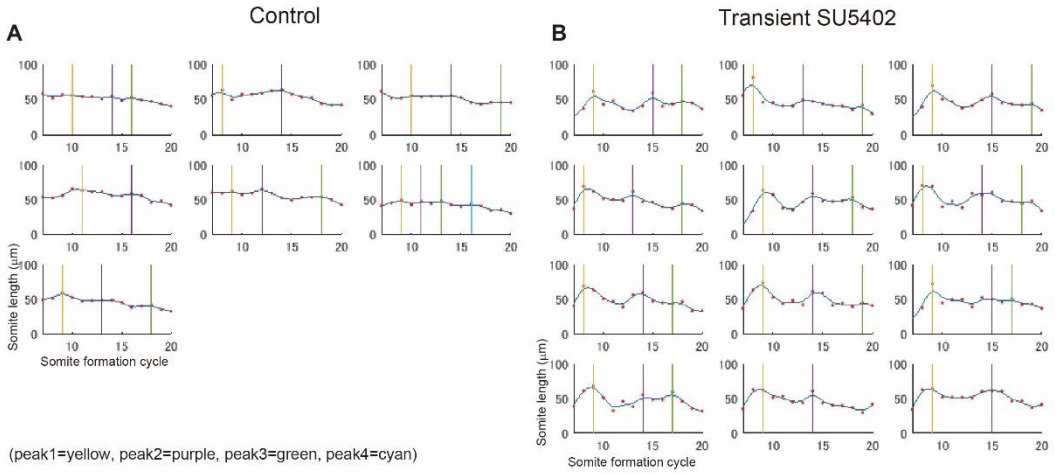


Figure S13 Somite length change in individual samples. Somite sizes were measured using time-lapse imaging both in control (**A**) and SU5402 treated embryos (**B**). The peaks are detected using matlab function. Note the periodic change in somite length is much more obvious in SU5402 treated embryos, compared to control embryos.

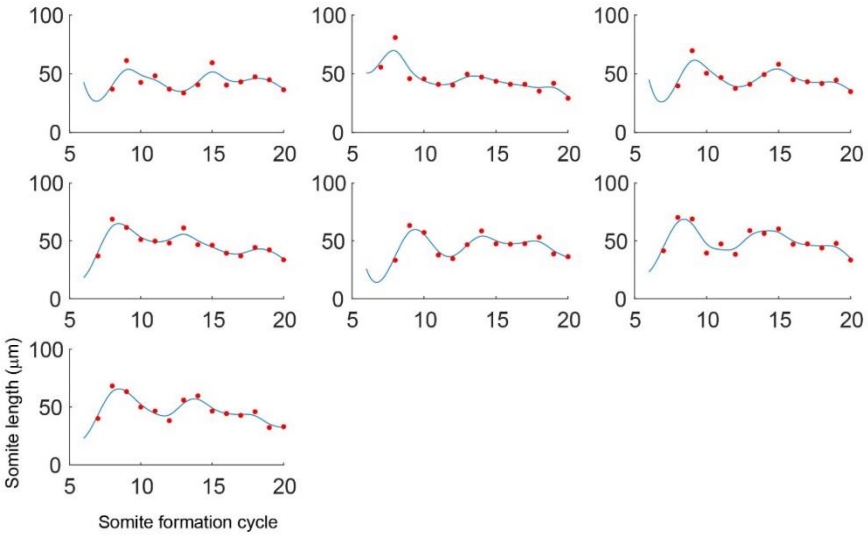


Figure S14 Somite length change in individual samples in long-term SU5402 treatment. Somite sizes were measured using time-lapse imaging. As predicted in the simulation of clock and scaled gradient model (Figure 4I), echo effect was observed.

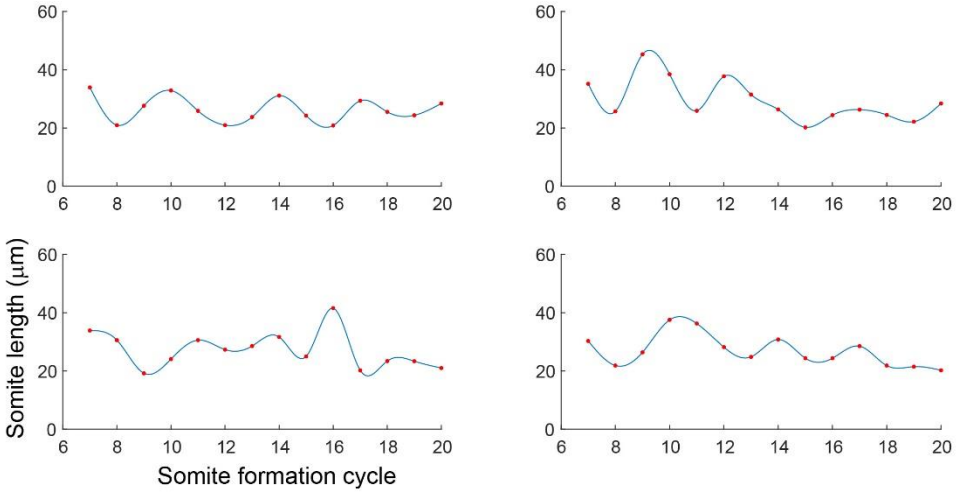


Figure S15 Somite length change in individual samples in long-term SU5402 treatment. Somite sizes were measured using time-lapse imaging. As predicted in the simulation of clock and scaled gradient model (Figure 4I), echo effect was observed.

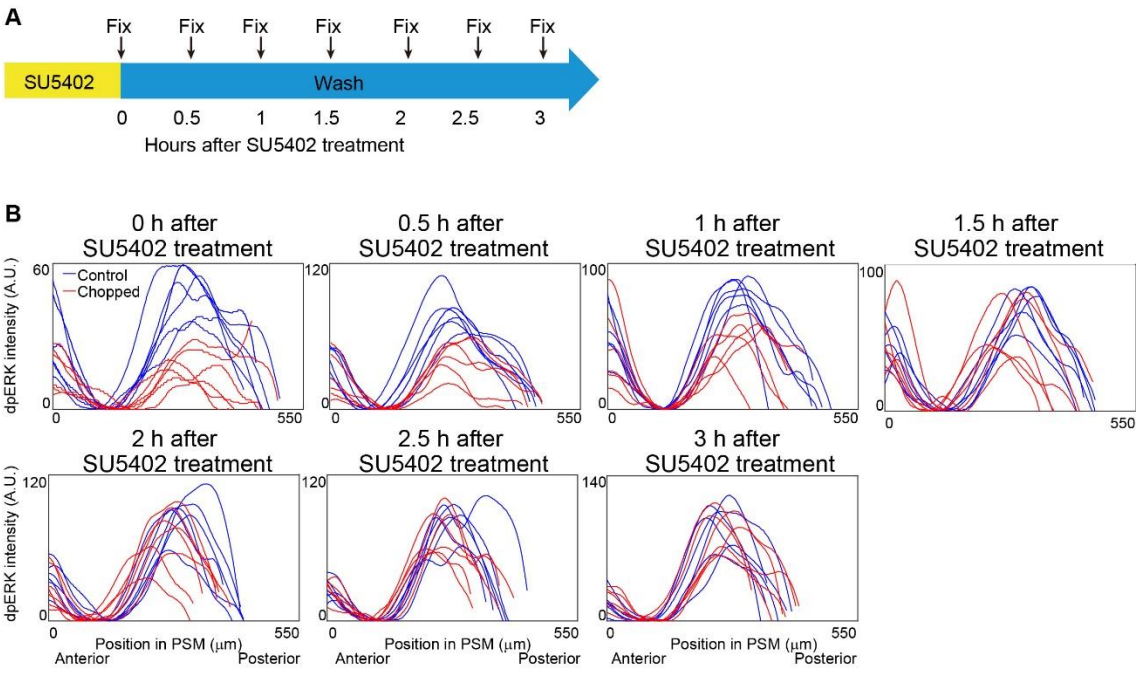


Figure S16 dpERK intensity change in individual samples for echo experiment. (**A**) Schematic illustration of the experiment. Samples were fixed every 30 min after SU5402 treatment. (**B**) Intensity profiles of dpERK immunostaining for control (blue) and treated embryos (red).

Supplemental References

Cotterell, J., Robert-Moreno, A. and Sharpe, J. (2015). A Local, Self-Organizing Reaction-Diffusion Model Can Explain Somite Patterning in Embryos. *Cell Syst* 1, 257-69.

Kawakami, Y., Raya, A., Raya, R. M., Rodriguez-Esteban, C. and Izpisua Belmonte, J. C. (2005). Retinoic acid signalling links left-right asymmetric patterning and bilaterally symmetric somitogenesis in the zebrafish embryo. *Nature* 435, 165-71.

Yabe, T., Shimizu, T., Muraoka, O., Bae, Y. K., Hirata, T., Nojima, H., Kawakami, A., Hirano, T. and Hibi, M. (2003). Ogon/Secreted Frizzled functions as a negative feedback regulator of Bmp signaling. *Development* 130, 2705-16.

Appendix 4.

Observing the Cell in Its Native State: Imaging Subcellular Dynamics in Multicellular Organisms

Authors: Tsung-Li Liu[1,9], Srigokul Upadhyayula[1,2,3,4,9], Daniel E. Milkie[1], Ved Singh[1], Kai Wang[1§], Ian A. Swinburne[5], Kishore R. Mosaliganti[5], Zachary M. Collins[5], Tom W. Hiscock[5], Jamien Shea[1], Abraham Q. Kohrman[6], Taylor N. Medwig[6], Daphne Dambournet[7], Ryan Forster[7], Brian Cunniff [2,3†], Yuan Ruan[8], Hanako Yashiro[8], Steffen Scholpp[9], Elliot M. Meyerowitz[8], Dirk Hockemeyer[7], David G. Drubin[7], Benjamin L. Martin[6], David Q. Matus[6], Minoru Koyama[1], Sean G. Megason[5], Tom Kirchhausen[1,2,3,4], Eric Betzig[1*]

Affiliations:

[1] Janelia Research Campus, Howard Hughes Medical Institute, Ashburn, VA 20147, USA. [2] Department of Cell Biology, Harvard Medical School, 200 Longwood Ave, Boston, MA 02115, USA. [3] Program in Cellular and Molecular Medicine, Boston Children's Hospital, 200 Longwood Ave, Boston, MA 02115, USA. [4] Department of Pediatrics, Harvard Medical School, 200 Longwood Ave, Boston, MA 02115, USA. [5] Department of Systems Biology, Harvard Medical School, 200 Longwood Ave, Boston, MA 02115, USA. [6] Department of Biochemistry and Cell Biology, Stony Brook University, Stony Brook, NY 11794-5215, USA. [7] Department of Molecular and Cell Biology, University of California, Berkeley, Berkeley, CA 94720, USA. [8]Howard Hughes Medical Institute and Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA 91125, USA. [9]Living System Institute, College of Life and Environmental Science, University of Exeter Exeter, UK EX4 4QD. [9]These authors contributed equally to this work.[§]Current Address: Institute of Neuroscience, Chinese Academy of Sciences, 320 Yueyang Road, Shanghai, China [†]Current Address: Department of Pathology, University of

Vermont College of Medicine, Burlington, VT, 05401, USA *Correspondence to: betzige@janelia.hhmi.org.

Authors Contributions: EB supervised the project and wrote the manuscript with input from all co-authors. TLL built the microscope with input from EB, DEM, and KW, and performed all microscope characterization experiments. DEM created the instrument control software. TLL, VS, and SU acquired all biological data with co-authors. DD, DGH, RF, and DH provided organoids and led related experiments. TK, SU, BC, and SS provided transgenic zebrafish and AP2 cells and led related clathrin experiments. ZMC, TWH, and SGM created the mCardinal-PM transgenic zebrafish. TK, SGM, SU, and IAS provided zebrafish and led related organelle dynamics and *in vivo* immune cell experiments. JS and MK created the Autobow zebrafish, and MK led related experiments. AQK, TNM, and DQM provided *C. elegans* specimens and led related experiments, BLM provided MDA-MB-231 cells and vasculature-labeled zebrafish, and led related experiments with DQM. YR, HY, and EMM provided *Arabidopsis* specimens and led related experiments. SU, KRM, TLL, and VS processed all image data. SU performed quantitative analysis of all image data. SU, TLL, and EB produced all figures and movies.

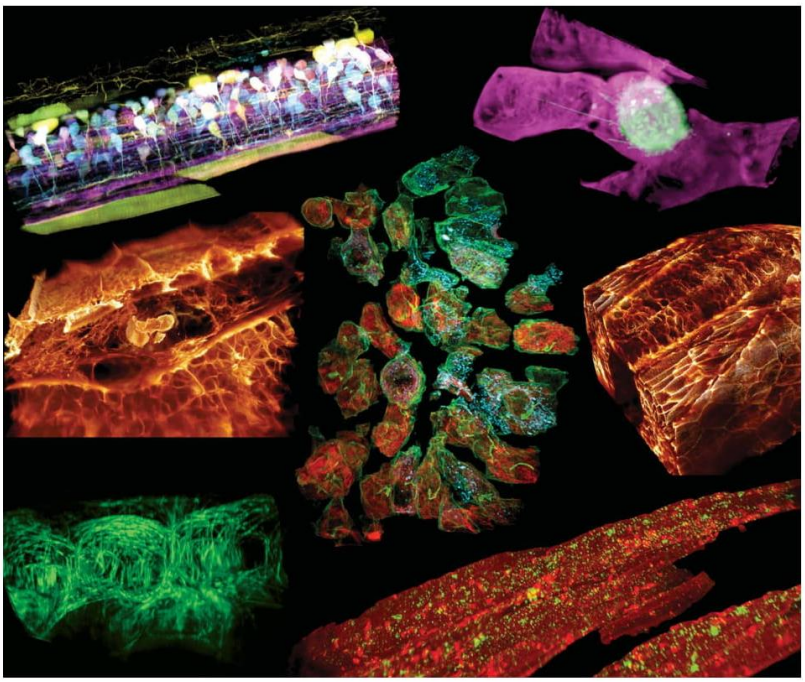
RESEARCH ARTICLE SUMMARY

ADVANCED IMAGING

Observing the cell in its native state: Imaging subcellular dynamics in multicellular organisms

Tsung-Li Liu,* Srigokul Upadhyayula,* Daniel E. Milkie, Ved Singh, Kai Wang, Ian A. Swinburne, Kishore R. Mosaliganti, Zach M. Collins, Tom W. Hiscock, Jamien Shea, Abraham Q. Kohrman, Taylor N. Medwig, Daphne Dambournet, Ryan Forster, Brian Cunniff, Yuan Ruan, Hanako Yashiro, Steffen Scholpp, Elliot M. Meyerowitz, Dirk Hockemeyer, David G. Drubin, Benjamin L. Martin, David Q. Matus, Minoru Koyama, Sean G. Megason, Tom Kirchhausen, Eric Betzig†

INTRODUCTION: Organisms live by means of the complex, dynamic, three-dimensional (3D) interplay between millions of components, from the molecular to the multicellular. Visualizing this complexity in its native form requires imaging at high resolution in space and time anywhere within the organism itself, because only there are all the environmental factors that regulate its physiology present. However, the optical heterogeneity of multicellular systems leads to aberrations that quickly compromise resolution, signal, and contrast with increasing imaging depth. Furthermore, even in the absence of aberrations, high resolution and fast imaging are usually accompanied by intense illumination, which can perturb delicate subcellular processes or even introduce permanent phototoxic effects.



High-resolution in vivo cell biology. AO-LLSM permits the study of 3D subcellular processes in their native multicellular environments at high spatiotemporal resolution, including (clockwise from upper left) growth of spinal cord axons; cancer cell metastasis; collective cellular motion; endocytosis; microtubule displacements; immune cell migration; and (center) organelle dynamics.

ON OUR WEBSITE

Read the full article at http://dx.doi.org/10.1126/science.aaq1392

RATIONALE: We combined two imaging technologies to address these problems. The first, lattice light-sheet microscopy (LLSM), rapidly and repeatedly sweeps an ultrathin sheet of light through a volume of interest while acquiring a series of images, building a high-resolution 3D movie of the dynamics within. The confinement of the illumination to a thin plane insures that regions outside the volume remain unexposed, while the parallel collection of fluorescence from across the plane permits low, less perturbative intensities to be used. The second technology, adaptive optics (AO), measures sample-induced distortions to the image of a fluorescent "guide star" created within the volume—distortions that also affect the acquired light-sheet images—and compensates for these by changing the shape of a mirror to create an equal but opposite distortion.

RESULTS: We applied AO-LLSM to study a variety of 3D subcellular processes in vivo over a broad range of length scales, from the nanoscale diffusion of clathrin-coated pits (CCPs) to axon-guided motility across 200 μm of the developing zebrafish spinal cord. Clear delineation of cell membranes allowed us to computationally isolate and individually study any desired cell within the crowded multicellular environment of the intact organism. By doing so, we could compare specific processes across different cell types, such as rates of CCP internalization in muscle fibers and brain cells, organelle remodeling during cell division in the developing brain and eye, and motility mechanisms used by immune cells and metastatic breast cancer cells. Although most examples were taken from zebrafish embryos, we also demonstrated AO-LLSM in a human stem cell–derived organoid, a *Caenorhabditis elegans* nematode, and *Arabidopsis thaliana* leaves.

CONCLUSION: AO-LLSM takes high-resolution live-cell imaging of subcellular processes from the confines of the coverslip to the more physiologically relevant 3D environment within whole transparent organisms. This creates new opportunities to study the phenotypic diversity of intracellular dynamics, extracellular communication, and collective cell behavior across different cell types, organisms, and developmental stages. ■

The list of author affiliations is available in the full article online.
*These authors contributed equally to this work.
†Corresponding author. Email: betzige@janelia.hhmi.org.
Cite this article as T.-L. Liu *et al.*, *Science* **360**, eaaq1392 (2018). DOI: 10.1126/science.aaq1392

RESEARCH ARTICLE

ADVANCED IMAGING

Observing the cell in its native state: Imaging subcellular dynamics in multicellular organisms

Tsung-Li Liu,[1]* Srigokul Upadhyayula,[1,2,3,4]* Daniel E. Milkie,[1] Ved Singh,[1] Kai Wang,[1]† Ian A. Swinburne,[5] Kishore R. Mosaliganti,[5] Zach M. Collins,[5] Tom W. Hiscock,[5] Jamien Shea,[1] Abraham Q. Kohrman,[6] Taylor N. Medwig,[6] Daphne Dambournet,[7] Ryan Forster,[7] Brian Cunniff,[2,3]‡ Yuan Ruan,[8] Hanako Yashiro,[8] Steffen Scholpp,[9,10] Elliot M. Meyerowitz,[8] Dirk Hockemeyer,[7] David G. Drubin,[7] Benjamin L. Martin,[6] David Q. Matus,[6] Minoru Koyama,[1] Sean G. Megason,[5] Tom Kirchhausen,[1,2,3,4] Eric Betzig[1]§

True physiological imaging of subcellular dynamics requires studying cells within their parent organisms, where all the environmental cues that drive gene expression, and hence the phenotypes that we actually observe, are present. A complete understanding also requires volumetric imaging of the cell and its surroundings at high spatiotemporal resolution, without inducing undue stress on either. We combined lattice light-sheet microscopy with adaptive optics to achieve, across large multicellular volumes, noninvasive aberration-free imaging of subcellular processes, including endocytosis, organelle remodeling during mitosis, and the migration of axons, immune cells, and metastatic cancer cells in vivo. The technology reveals the phenotypic diversity within cells across different organisms and developmental stages and may offer insights into how cells harness their intrinsic variability to adapt to different physiological environments.

A common tenet, oft repeated in the field of bioimaging, is "seeing is believing." But when can we believe what we see? The question becomes particularly relevant when imaging subcellular dynamics by fluorescence microscopy. Traditional imaging tools such as confocal microscopy are often too slow to study fast three-dimensional (3D) processes across cellular volumes, create out-of-focus photoinduced damage (*1*, *2*) and fluorescence photobleaching, and subject the cell at the point of measurement to peak intensities far beyond those under which life evolved. In addition, much of what fluorescence microscopy has taught us about subcellular processes has come from observing isolated adherent cells on glass. True physiological imaging requires studying cells within the organism in which they evolved, where all the environmental cues that regulate cell physiology are present (*3*). Although intravital imaging achieves this goal (*4*, *5*) and has contributed pivotally to our understanding of cellular and developmental biology, the resolution needed to study minute subcellular processes in 3D detail is compromised by the optically challenging multicellular environment.

Two imaging tools have recently been developed to address these problems: Lattice light-sheet microscopy (LLSM) (*6*) provides a noninvasive alternative for volumetric imaging of whole living cells at high spatiotemporal resolution, often over hundreds of time points, and adaptive optics (AO) (*7*) corrects for sample-induced aberrations caused by the inhomogeneous refractive index of multicellular specimens and recovers resolution and signal-to-background ratios comparable to those attained for isolated cultured cells. The remaining challenge is to combine these technologies in a way that retains their benefits and thereby enables the in vivo study of cell biology at high resolution in conditions as close as possible to the native physiological state. Here we describe a technique based on an adaptive optical lattice light-sheet microscope designed for this purpose (AO-LLSM) and demonstrate its utility through high-speed, high-resolution, 3D in vivo imaging of a variety of dynamic subcellular processes.

Lattice light-sheet microscope with two-channel adaptive optics

Although several AO methods have been demonstrated in biological systems (*7*), including in the excitation (*8*) or detection (*9*) light paths of a light-sheet microscope, we chose an approach where the sample-induced aberrations affecting the image of a localized reference "guide star" created through two-photon excited fluorescence (TPEF) within the specimen are measured and then corrected with a phase modulation element (*10*). By scanning the guide star over the region to be imaged (*11*), an average correction is measured that is often more accurate than single-point correction—which is essential, because a poor AO correction is often worse than none at all. Scanning also greatly reduces the photon load demanded from any single point. Coupled with correction times as short as 70 ms (*11*), this AO method is compatible with the speed and noninvasiveness of LLSM.

In LLSM, light traverses different regions of the specimen for excitation and detection and therefore is subject to different aberrations. Hence, independent AO systems are needed for each. This led us to design a system (Fig. 1A, supplementary note 1, and fig. S1) where light (red) from a Ti:Sapphire ultrafast laser is ported to either the excitation or detection arm of a LLS microscope (left inset, Fig. 1A) by switching galvanometer 1. In the detection case, TPEF (green) generated within a specimen by scanning the guide star across the focal plane of the detection objective (DO) is descanned (*11*) and sent to a Shack-Hartmann wavefront sensor (DSH) via switching galvanometer 2 (SG2). We then apply the inverse of the measured aberration to a deformable mirror (DM) placed conjugate to both the DSH and the rear pupil plane of the DO (supplementary note 2). Because the signal (also green) generated by the LLS when in imaging mode travels the same path through the specimen as the guide star, and reflects from the same DM, the corrective pattern that we apply to the DM produces an AO-corrected image of the current excitation plane within the specimen on the camera when ported there by the SG2.

Similarly, for excitation correction, we send descanned TPEF generated and collected through the excitation objective (EO) to a second Shack-Hartmann sensor. However, because LLS excitation is confined to a thin annulus at the rear pupil of the DO (*6*), a DM placed conjugate to this pupil would be ineffective for AO correction over most of its surface. Instead, we apply wavefront correction at the same sample-conjugate spatial light modulator (SLM) that creates the light sheet itself, thereby enlisting thousands of independently corrective pixels. To do so, we

[1]Janelia Research Campus, Howard Hughes Medical Institute, Ashburn, VA 20147, USA. [2]Department of Cell Biology, Harvard Medical School, 200 Longwood Avenue, Boston, MA 02115, USA. [3]Program in Cellular and Molecular Medicine, Boston Children's Hospital, 200 Longwood Avenue, Boston, MA 02115, USA. [4]Department of Pediatrics, Harvard Medical School, 200 Longwood Avenue, Boston, MA 02115, USA. [5]Department of Systems Biology, Harvard Medical School, 200 Longwood Avenue, Boston, MA 02115, USA. [6]Department of Biochemistry and Cell Biology, Stony Brook University, Stony Brook, NY 11794-5215, USA. [7]Department of Molecular and Cell Biology, University of California, Berkeley, Berkeley, CA 94720, USA. [8]Howard Hughes Medical Institute and Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA 91125, USA. [9]Living Systems Institute, College of Life and Environmental Sciences, University of Exeter, Exeter EX4 4QD, UK. [10]Institute of Toxicology and Genetics, Karlsruhe Institute of Technology, 76021 Karlsruhe, Germany.
*These authors contributed equally to this work. †Present address: Institute of Neuroscience, Chinese Academy of Sciences, 320 Yueyang Road, Shanghai, China. ‡Present address: Department of Pathology, University of Vermont College of Medicine, Burlington, VT 05401, USA.
§Corresponding author. Email: betzige@janelia.hhmi.org

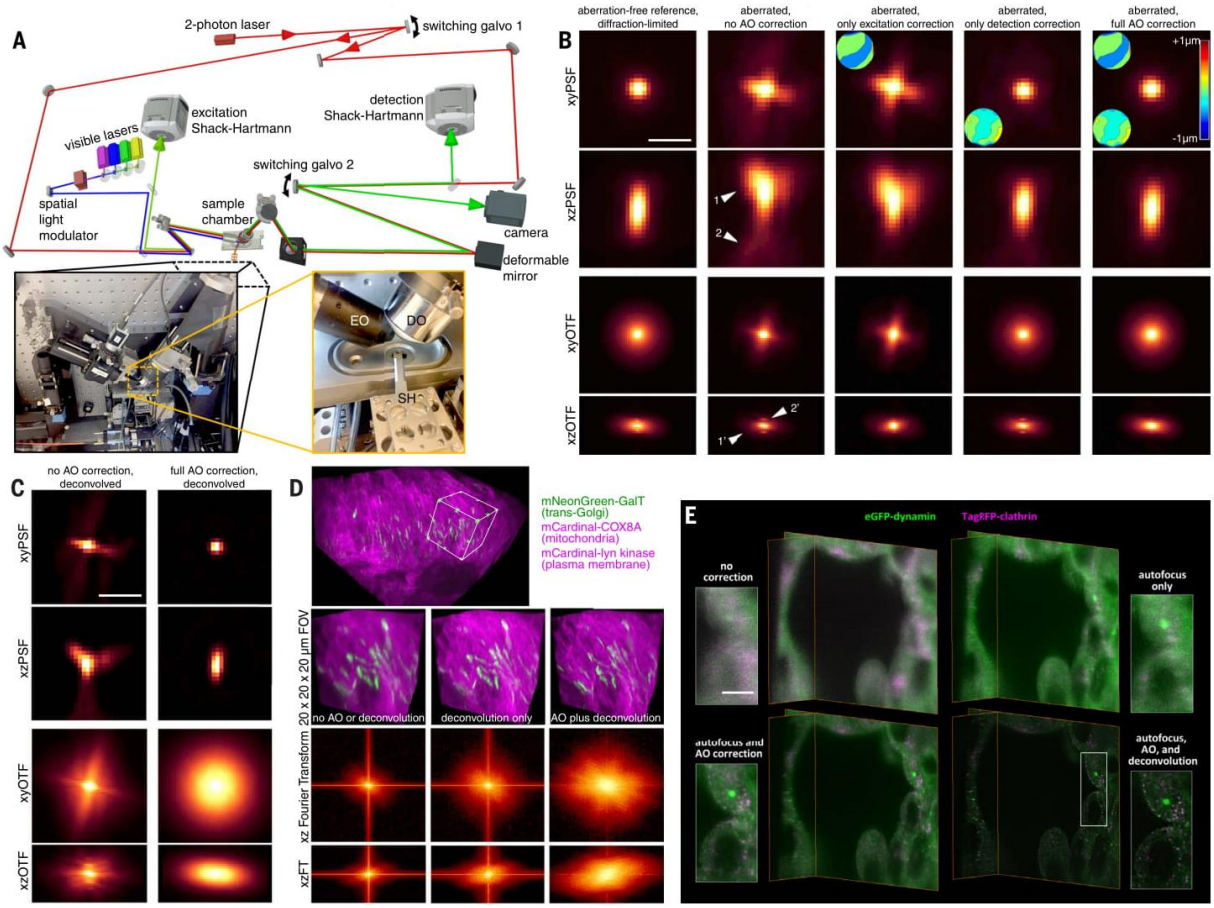
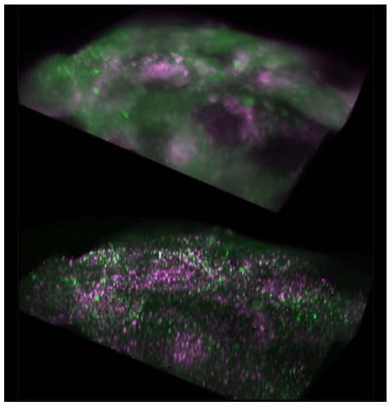


Fig. 1. Adaptive optical lattice light-sheet microscopy (AO-LLSM). (**A**) Simplified microscope schematic (fig. S1 shows a detailed version). EO, excitation objective; DO, detection objective; SH, sample holder. (**B**) *xy* and *xz* maximum intensity projections (MIPs) of the point spread function (PSF; top two rows) and corresponding optical transfer function (OTF; bottom two rows) of the microscope under five different degrees of AO correction (columns), as measured from a 200-nm fluorescent bead in an aberrating agarose gel. Insets show the corrective wavefronts applied. Arrowheads indicate lateral and axial aberrations. Scale bar, 1 μm. (**C**) MIPs and corresponding OTFs of the uncorrected (left column) and fully corrected (right column) bead images from (B), after deconvolution using the aberration-free reference PSF. Scale bar, 1 μm. (**D**) Cellular trans-Golgi, mitochondria, and plasma membranes in the spine of a live zebrafish embryo 24 hpf. Shown are unprocessed data without AO correction (left column), deconvolved data without AO correction (center), and deconvolved data after AO correction (top and right) (movie S2). MIP views (bottom two rows) of the Fourier transform (FT) of the data in all three cases indicate their respective degrees of information recovery. (**E**) Four different levels of correction, shown for orthoslices of a live human stem cell–derived organoid grown in Matrigel and gene-edited to express endogenous levels of clathrin and dynamin in coated endocytic pits (Movie 1). Scale bar, 5 μm.

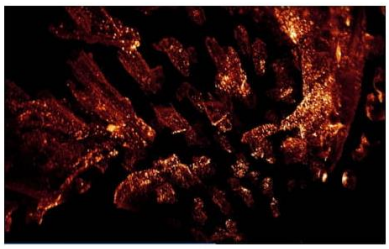
subtract the measured phase aberration from the phase of the Fourier transform (FT) of the ideal, aberration-free SLM pattern, then inverse-transform the result back to the sample-conjugate SLM plane (supplementary note 3 and fig. S2).

Lastly, optimal resolution requires the LLS to be coincident with the focal plane of the DO to less than ~0.46 μm over the entire field of view (FOV), whereas refractive index differences between the specimen and the surrounding media lead to tip-tilt alignment or axial displacement errors for the light sheet that might exceed this (*12*). Fortunately, we find that only displacement is a concern over FOVs typical of LLSM (supplementary note 4 and fig. S3) and that it can be robustly corrected by imaging, edge-on through the EO, the offset of the plane of TPEF that we generate when measuring the detection aberration relative to the plane of fluorescence generated by the LLS (supplementary note 5 and fig. S4).

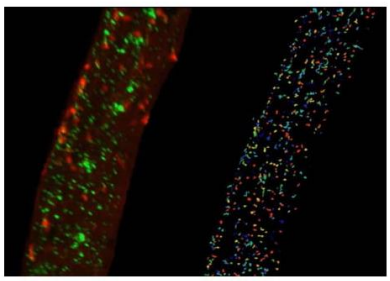
As an example (Fig. 1B), after correction for aberrations in the microscope itself, the point spread function (PSF) and corresponding optical transfer function (OTF), measured from a 200-nm bead, indicate nearly diffraction-limited performance (first column). However, when a similar bead is placed in 2% low-melt agarose, we observe substantial aberration (second column), both laterally (arrowheads 1 and 1′) and axially (arrowheads 2 and 2′). Correcting only the excitation aberration improves the axial resolution (third column) by returning the light sheet to its original width (fig. S2). Conversely, correcting only the detection aberration improves primarily the lateral performance (fourth column). However, when we combine the excitation and detection corrections, the image of the bead is returned to its diffraction-limited form (fifth column), with an eightfold recovery to its original peak intensity. Furthermore, the same



Movie 1. Endocytosis in a human stem cell–derived organoid. Gene-edited clathrin (magenta) and dynamin (green) before and after adaptive optical correction and deconvolution, showing comparative *xy* and *xz* orthoslices, volume renderings, and postcorrection tracking of the motion and lifetimes of individual CCPs and CCVs over 120 time points at 1.86-s intervals (Fig. 1E and fig. S5).



Movie 2. Clathrin-mediated endocytosis in vivo. Dynamics of CCPs and CCVs over 15 min at 10-s intervals in the dorsal tail region of a zebrafish embryo 80 hpf. Segmented cells reveal brighter clathrin puncta at the vascular endothelium than at muscle fibers (Fig. 2A).



Movie 3. Clathrin localization in muscle fibers. PM (red) and clathrin (green) in the tail of a zebrafish embryo 50 to 55 hpf, showing *xy* and *xz* orthoslices before and after AO correction and deconvolution, dynamics of individual CCPs and CCVs at and between t-tubules, large clathrin clusters and small clathrin puncta in volume-rendered and segmented cells, and tracked CCPs and CCVs in a segmented cell (Fig. 2, B to E; fig. S8; and movie S3).

correction proves valid over a 30-μm field of beads in agarose (movie S1).

One of the key advantages of complete AO correction is that it enables accurate deconvolution (Fig. 1C), giving the most truthful representation of the specimen possible within the limits of diffraction (second column). In contrast, applying deconvolution to an aberrated image gives a distorted result, because the OTF can then fall below the noise floor asymmetrically and at spatial frequencies well below the diffraction limit (first column), after which higher spatial frequencies cannot be recovered by deconvolution. These same trends can be seen in densely labeled specimens as well, such as across mitochondria, Golgi, and plasma membranes (PMs) of cells near the spinal midline in a living zebrafish embryo 24 hours postfertilization (hpf) (Fig. 1D, bottom three rows, and movie S2), where the greatest information content as seen in the FT of image volume is obtained by combined AO and deconvolution.

Next, we imaged organoids differentiated from human stem cells, gene-edited to express endogenous levels of red fluorescent protein (tagRFP)–clathrin and enhanced green fluorescent protein (EGFP)–dynamin in endocytic pits (clathrin-coated pits; CCPs). Such organoids permit the study of human cellular differentiation and tissue morphogenesis in vitro at subcellular resolution, with an experimental accessibility that is difficult to achieve in vivo. However, the extracellular matrix in which they are grown introduces considerable aberrations, and the fast dynamics and limited number of fluorophores in each CCP demand high-speed imaging with low photobleaching. The system is therefore well suited to the capabilities of AO-LLSM, with the CCPs doubling as distributed puncta of subdiffractive size to evaluate imaging performance.

Without AO or focus correction (upper left of Fig. 1E and Movie 1), no CCPs are visible, and the cell boundaries are poorly defined. Autofocus alone (upper right) reveals larger patches of clathrin and dynamin but must be combined with complete AO correction (excitation and detection, lower left) to identify individual CCPs. At that point, the imaging is diffraction-limited, so the data can be deconvolved using the system-corrected PSF to compensate for the spatial filtering properties of the microscope. Both dynamin and clathrin puncta then stand out clearly above the cytosolic background (lower right), allowing us to quantitatively measure their lifetimes and diffusion tracks (Movie 1). The increasing recovery of information as we progress through these four cases can also be seen quantitatively in their corresponding OTFs (fig. S5).

Clathrin dynamics in zebrafish

For transparent model organisms, we can apply AO-LLSM in vivo, where the complete physiological environment is preserved. The zebrafish has become the most widely used nonmammalian vertebrate model organism. We took advantage of its small size and transparency to visualize the formation of endocytic CCPs and clathrin-coated vesicles (CCVs) in the context of the developing organism. We first imaged a volume in the dorsal tail region of a fish larva 80 hpf stably expressing dsRed–clathrin light chain A (CLTA) (*13*) (Fig. 2A and Movie 2). We observed a high density of diffraction-limited clathrin spots that, after computational separation of all cells, were found to be mostly colocalized with the PM. These spots appeared and disappeared with the formation of new CCPs and their eventual uncoating. We determined that the areal density of CCPs was similar in muscle fibers (e.g., green arrowheads, Fig. 2A) and endothelial cells (magenta arrowheads) lining blood vessels, but the distribution of their intensities was broader in the latter (lower inset, Fig. 2A), which had a subpopulation of pits that were up to sixfold as bright as the median CCP intensity in the former. Because CCP size is proportional to intensity (*14*), these results indicate the presence of larger structures, possibly clusters of CCPs (*15*), in the vasculature endothelium.

To track CCPs for longer times, we turned to zebrafish embryos mRNA-injected to express the brighter and more photostable fluorescent fusion protein mNeonGreen-CLTA. These embryos also expressed mCardinal targeted to the PM, facilitating the computational separation of cells. Embryos imaged in the tail (Fig. 2B, fig. S6A, and Movie 3) and in the hindbrain (fig. S6B) displayed a variety of morphologies, trafficking behaviors (Fig. 2C), and lifetime distributions (Fig. 2D). Large, micron-scale intracellular spots of limited mobility (arrowheads, Fig. 2B) probably represent clathrin-rich vesicles clustered at the trans-Golgi network, whereas mobile diffraction-limited spots (Fig. 2C, arrowhead group 1) likely correspond to endosomal carriers similar to those seen in cultured mammalian cells (*16*). Diffraction-limited spots at the PM (Fig. 2C, arrowhead group 2) likely represent individual CCPs and CCVs. We also found CCPs at the t-tubules spanning muscle fibers (Fig. 2C, top left), in contrast to the diffuse clathrin signal observed using immunofluorescence in fixed rat muscle fibers (*17*). Most of these were pinned, but occasionally they would break free from a t-tubule and move rapidly along the fiber axis (Fig. 2C, arrowhead group 3, and movie S3), possibly by active transport along myofibrils or within the sarcoplasmic reticulum.

AO allowed us to detect more CCPs (fig. S7) and track all CCPs with higher precision (Fig. 2C, top right; fig. S8, A to C; and Movie 3). Comparing CCPs in muscle fibers and the brain, we found that, although their initiation frequencies and fluorescence intensities were similar, brain CCPs tended to internalize faster (Fig. 2E). Assuming that clathrin puncta in muscle and brain cells lasting >21 s corresponded to successful coated vesicles, each with an assumed membrane diameter of 60 nm, ~0.1% of the PM was internalized through clathrin-mediated endocytosis every minute (fig. S8D). This is similar to the values derived from measurements in cultured SUM-159 (*18*) or htertRPE-1 (*19*) mammalian cells at 37°C.

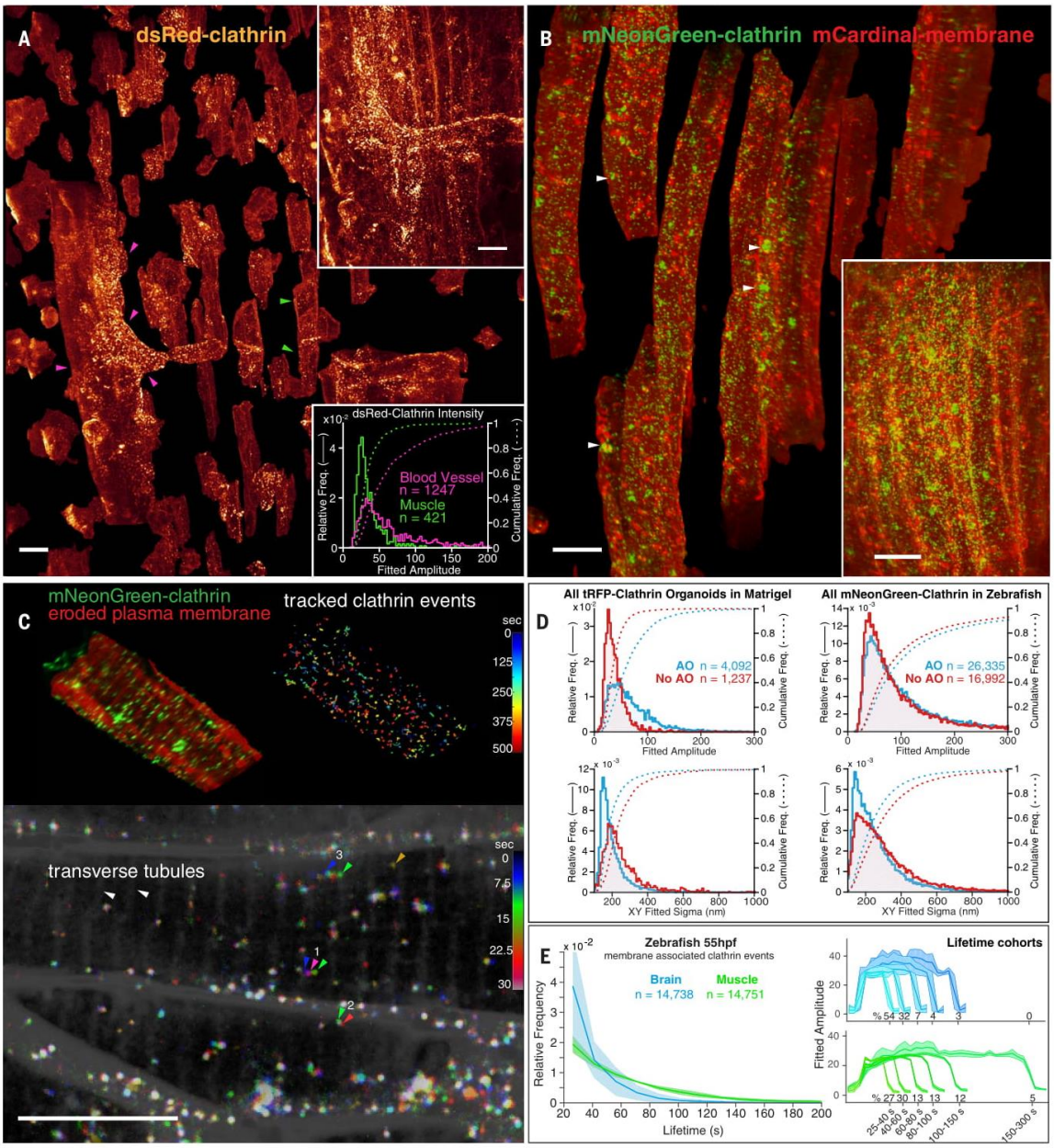


Fig. 2. Clathrin dynamics in zebrafish. (**A**) Computationally separated muscle fibers (e.g., green arrowheads) and vascular endothelial cells (e.g., magenta arrowheads), both expressing DsRed-CLTA to highlight CCPs and CCVs, from muscle tissue in a 75-μm by 99-μm by 41-μm region (upper inset) of the tail of a developing zebrafish larva 80 hpf (Movie 2). Brighter clathrin puncta were observed in the endothelial cells (lower inset). Scale bars, 10 μm. (**B**) Computationally separated muscle fibers from a region (lower inset) in the tail of a zebrafish embryo 50 hpf coexpressing an mCardinal-PM marker (red) and mNeonGreen-CLTA (green). Individual CCPs and CCVs and larger clathrin-rich vesicles (arrowheads) are visible (Movie 3). Scale bars, 10 μm. (**C**) Spatial distribution and dynamics of CCPs and CCVs tracked for 12 min at 7.5-s intervals in one muscle fiber from (B), showing CCPs localized at t-tubules (top left) and diffusion and lifetime characteristics for CCPs and CCVs across the cell (top right). A MIP through a 2-μm-thick slab at three consecutive time points (bottom) shows examples of a pinned CCV (arrowheads 2) and slowly diffusing (arrowheads 1) or rapidly shuttling CCVs (arrowheads 3) (movie S3). Scale bar, 10 μm. (**D**) Effect of AO on the measured quantity, intensity, and localization precision of CCPs and CCVs in the organoid in Fig. 1E and the zebrafish in (B). (**E**) Comparative distribution of CCP and CCV lifetimes (left) and intensity cohorts grouped by their lifetimes (right) in the brain and muscle of a developing zebrafish embryo 55 hpf.

In vivo imaging of organelle morphology and dynamics during embryogenesis

A major focus in cell biology is the study of the structure and function of organelles within the living cell. To study the dynamics of multiple organelles simultaneously throughout the cell cycle across a population of cells in vivo, we imaged brain progenitor cells with markers for the trans-Golgi, endoplasmic reticulum (ER), mitochondria, and PM for 200 image volumes at 44-s intervals (Fig. 3, A and B, and Movie 4). In interphase, we observed multiple trans-Golgi segments in most cells, often appearing as long filaments preferentially aligned along the axis of cell polarization (Fig. 3A) that fragmented during mitosis (Fig. 3B). The ER and mitochondria largely recapitulated the forms that they commonly take in cultured cells: The ER established a reticular network in interphase and sheetlike cisternae during mitosis (*20*), whereas mitochondria formed punctate structures near the surface and longer tubules in the subset of more deeply buried interphase cells that were well labeled. Analyzing one such cell, we found that all three organelles were distributed uniformly between the PM and the nucleus in interphase (fig. S9A), but mitochondria were preferentially located nearer the PM during mitosis (fig. S9A, 109 min).

The early synchrony of cell division is lost in zebrafish at the midblastula transition (3 hpf). Nevertheless, at 14 hpf, we observed instances of cascading cell division, where adjacent cells

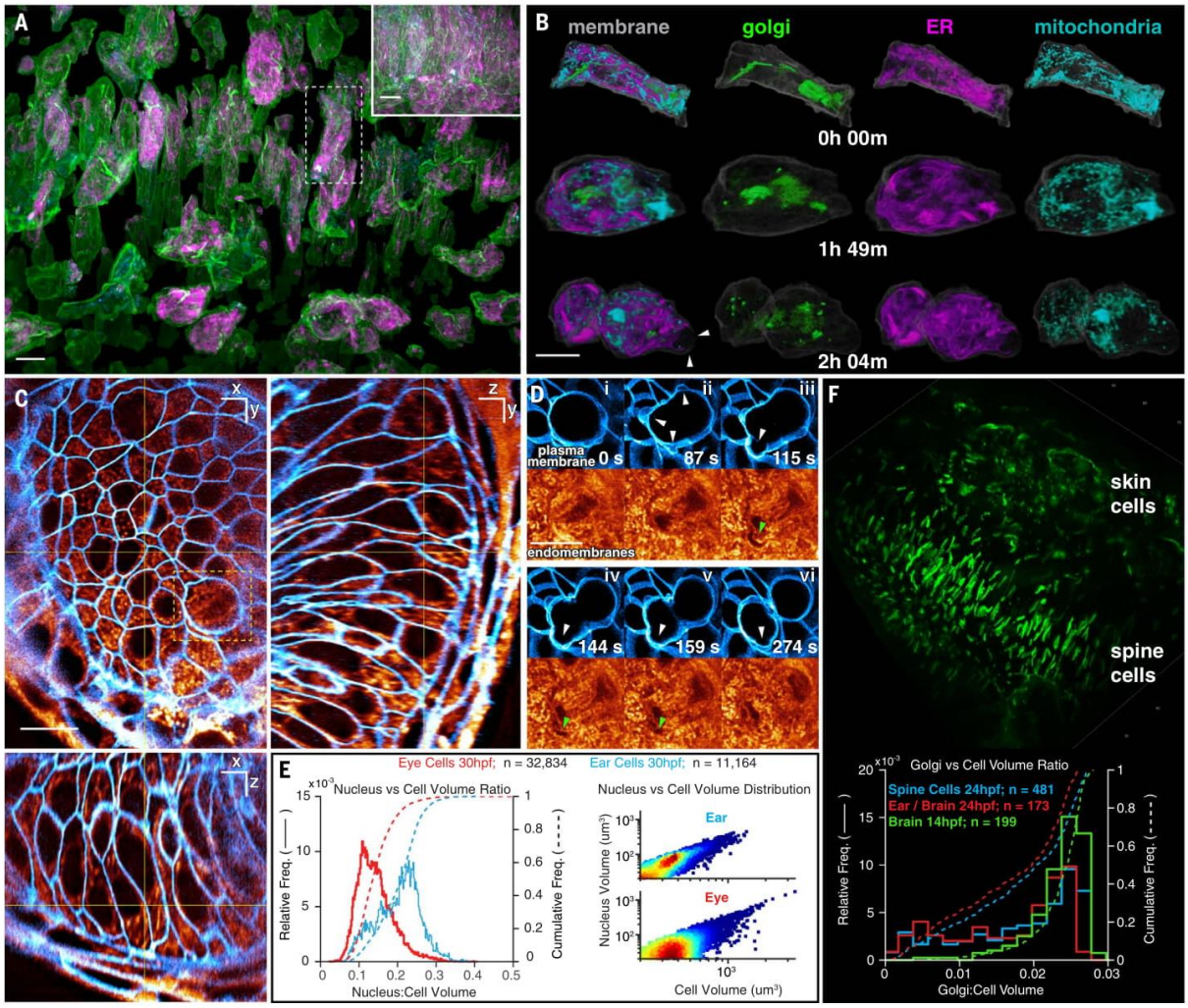


Fig. 3. Organelle morphologies and dynamics in zebrafish. (**A**) Computationally separated neural progenitor cells from a 70-μm by 35-μm by 35-μm region (inset) in the brain of a developing zebrafish embryo expressing GalT-mNeonGreen, tagRFP-Sec61β, and Citrine as markers of the trans-Golgi, ER, and PM, respectively, with additional labeling by MitoTracker Deep Red dye (Movie 4). Scale bars, 10 μm. (**B**) Changing morphologies of the organelles in the specific cell outlined in (A) at three time points through mitosis. Arrowheads indicate mitotic blebs. Scale bar, 10 μm. (**C**) MIP views from 1-μm-thick orthogonal slabs within the eye of a zebrafish embryo 30 hpf, showing PM (blue) and endomembranes (orange) (Movie 5). Scale bars, 10 μm. (**D**) Six time points from Movie 5, showing PM blebs (white arrowheads) during mitosis and the exclusion of endomembranes in early blebs (green arrowheads). (**E**) Correlation between nuclear volume and total cell volume in the eye and ear (Pearson's coefficient, 0.9 and 0.8, respectively). (**F**) Different morphologies of trans-Golgi (top) near the spine of a zebrafish embryo 24 hpf and distribution of trans-Golgi volume in different cell types and at different developmental stages (bottom).

underwent mitosis one after another (fig. S10 and Movie 4). Mitotic cells, as seen previously in cell cultures (*18*, *21*), decreased their surface area (fig. S9E) as they assumed a spheroidal shape and produced transient blebs before cytokinesis, but then recovered their initial area after division. Total cellular volume remained constant throughout mitosis (fig. S9E). We also identified instances of asymmetric cytokinesis (Fig. 3B and Movie 4), where the two daughter cells differed in surface area by ~50% (fig. S9F). When we quantified organelle intensity in one such instance, we also observed asymmetric fractional partitioning of the Golgi and mitochondria between the daughter cells during cytokinesis (fig. S9G).

Mitotic cells in the developing eye 30 hpf (white arrowheads in Fig. 3D and Movie 5) also produced transient blebs before cytokinesis. We discovered upon labeling with Bodipy–tetramethylrhodamine (TMR) that these blebs created voids (green arrowheads in Fig. 3D) that only slowly filled with endomembranes.

Lastly, we observed considerable variability in the size and morphology of specific subcellular features across different organs and developmental stages. Bodipy-TMR negative staining (e.g., Fig. 3, C and D) revealed that the nuclei of ear cells 30 hpf were nearly twice as large as those of eye cells when normalized by the total cellular volume (Fig. 3E, left) and that nuclear volume and cellular volume were positively correlated (Fig. 3E, right). Likewise, we found that the median trans-Golgi volume as a percentage of total cellular volume in brain progenitor cells 14 hpf was larger [2.46%; median absolute deviation (MAD) = 0.26%] than in ear, brain, or spine cells 24 hpf (2.0%; MAD = 0.66%) (Fig. 3F). Golgi took many forms, from the aforementioned narrow polarized filaments in the early brain (Fig. 3A) to shorter segments clustered near the midline in the spine and nuclear-wrapping filaments in skin cells (Fig. 3F).

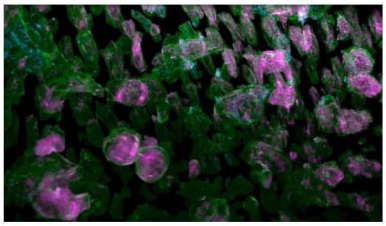
Tiled acquisition for aberrations varying in space and time

Because the refractive index profile can vary across a specimen and can also vary as the specimen develops, the AO corrections required can vary in both space and time. Unfortunately, it is difficult to estimate a priori the size or temporal stability of the isoplanatic patch (the FOV over which a single AO correction is valid). Empirically, we found in zebrafish embryos less than 72 hpf that a single excitation-detection correction pair obtained by scanning and descanning over the FOV is usually valid across 30 to 60 μm in each direction for at least 1 hour, provided that the light does not intersect the yolk. Fortuitously, these dimensions are comparable to those over which a LLS of submicron thickness does not deviate substantially in width. The examples shown above largely fall within these limits and hence, for them, a single AO correction pair at a single time point sufficed.

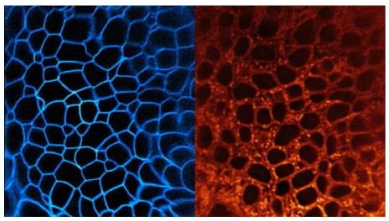
In other cases, however, we may wish to cover much larger FOVs. To do so, we must stitch together data from multiple image subvolume tiles, each with its own independent AO correction. To demonstrate the necessity of this, we imaged a 213-μm by 213-μm by 113-μm volume (Fig. 4A and movie S4) comprising 7 by 7 by 3 tiles in the tail region of a zebrafish embryo 96 hpf expressing membrane-EGFP, using three different protocols: no AO correction (Fig. 4B, left column), AO correction from the center tile applied to all tiles (middle column), and independent AO correction in each tile (right column). When viewed across 3-μm-thick slabs perpendicular (*xy*; Fig. 4B, top row) or parallel (*xz*; bottom row) to the detection axis, a small volume within the center tile (small orange boxes) showed substantial improvement by either center-tile or all-tiles AO correction, in both the *xy* lateral (upper-left orange boxes, top row) and *xz* axial (upper-left orange boxes, bottom row) planes. This is to be expected, because the site of AO correction coincided with the viewing area in these two cases. However, in a small volume at the edge of the fish (small blue boxes), only the data taken using individual AO corrections in each tile recovered optimal resolution in all directions (lower-right blue boxes, right column), because this volume was outside the isoplanatic patch over which the center-tile AO correction is valid. Applying the center correction across the larger stitched volume often results in greater wavefront errors (fig. S11) and poorer resolution (lower-right blue boxes, middle column) than applying no correction at all (lower-right blue boxes, left column), highlighting the importance of accurate and robust correction, if AO is to be applied.

Empirically, the largest aberrations that we observed in zebrafish embryos occur at regions of high curvature between the embryo and the surrounding media or at regions of rapid refractive index change, such as near the notochord (Fig. 4C and Movie 6). For example, when the LLS penetrates the embryo near-perpendicularly to its surface, the excitation aberration is initially small (Fig. 4C, red arrowhead, left column, top). However, after the light sheet passes through the notochord, it encounters substantial aberration, as seen in both the measured wavefront (green arrowhead, left column, top) and the uncorrected image (green arrowhead, middle column, bottom). On the detection side, aberrations increase with increasing depth in the embryo (left column, bottom, and middle column, top to bottom). In addition, substantial aberrations occur when the edge of the embryo is imaged tangentially (yellow arrowhead, left column, bottom), so that part of the detection light cone intersects the embryo and part does not.

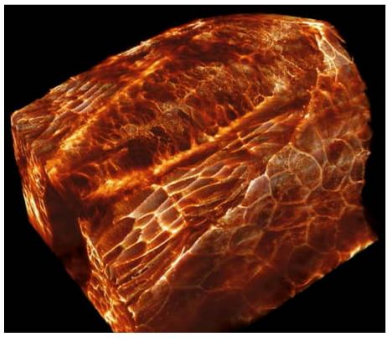
Provided that the specimen does not shift by more than a fraction of the isoplanatic patch size during imaging, a given set of tiled AO corrections can remain valid for hours (Movie 6). However, growth during development can cause an embryo to change its shape, position, or refractive index profile so that new corrections are occasionally needed. Fortunately, these changes often occur on a time scale that is slow compared with that needed to image even a large FOV by LLSM. In such cases, it is sufficient to update the



Movie 4. Subcellular imaging of organelle dynamics in the early zebrafish brain. Dynamics of PM (green or gray) and trans-Golgi (green), ER (magenta or red), and mitochondria (cyan) within neural progenitor cells over 200 time points at 44-s intervals from 14.0 hpf, showing complexity within the tissue, cross-sectional slab views through cells, sequential division of adjacent cells, segmentation and separation of all cells, and morphological changes to organelles during mitosis in one such cell (Fig. 3, A and B, and figs. S9 and S10).



Movie 5. Membrane dynamics in the zebrafish eye. PM (blue) and the endomembrane system (orange) 30 hpf viewed as *xy* orthoslices, cell divisions in a 1-μm-thick slab, and volume-rendered PM dynamics across the eye at 43.8-s intervals for 200 time points (Fig. 3, C to E).



Movie 6. Tiled AO correction for imaging large volumes. A 170-μm by 185-μm by 135-μm volume from the dorsal surface to the notochord in a PM-labeled zebrafish embryo, showing increasing aberration but continued full correction at increasing depth; corrective excitation and detection wavefronts in each of the tiled isoplanatic subvolumes of 5 by 4 by 7 tiles; and four views of PM dynamics within the complete volume from 30 to 39.5 hpf, imaged at 7.5 min intervals (Fig. 4C).

correction at only a subset of different tiles at each time point, as long as all subsets together encompass all tiles in the FOV before the previous round of corrections becomes inaccurate (Fig. 4D and movie S5). Usually, we chose subsets that broadly covered the FOV to monitor where the aberrations change the fastest.

One region that involves substantial specimen curvature, large spatial variation of the refractive index, and gradual aberration change is the eye of the developing zebrafish. Although in vivo optic cup development has been studied at subcellular 3D resolution by conventional (*22*, *23*) and multiview (*24*) light-sheet microscopy, tiled AO-LLSM permits a more detailed look at cellular morphology and organelle distributions during this process. We imaged across 4 by 4 by 3 tiles (Fig. 5A and Movie 7) spanning most of the eye of an embryo 24 to 27 hpf at 6-min intervals to study differences in the intracellular organization of various organelles (Fig. 5, B and C). Transgenic labeling of the PM allowed us to segment, isolate (Fig. 5D), and characterize each cell by type (Fig. 5E). Skin cells exhibited mitochondria clustered in the perinuclear region, like mitochondria seen in flat and thin cultured cells, to which these skin cells are morphologically similar. In contrast, mitochondria in retinal neuroepithelial (RNE) cells were generally longer, distributed across the length of the cell, and polarized along the same axis as the cell itself. The ER in RNE cells, although broadly distributed, was usually densest around the nucleus and

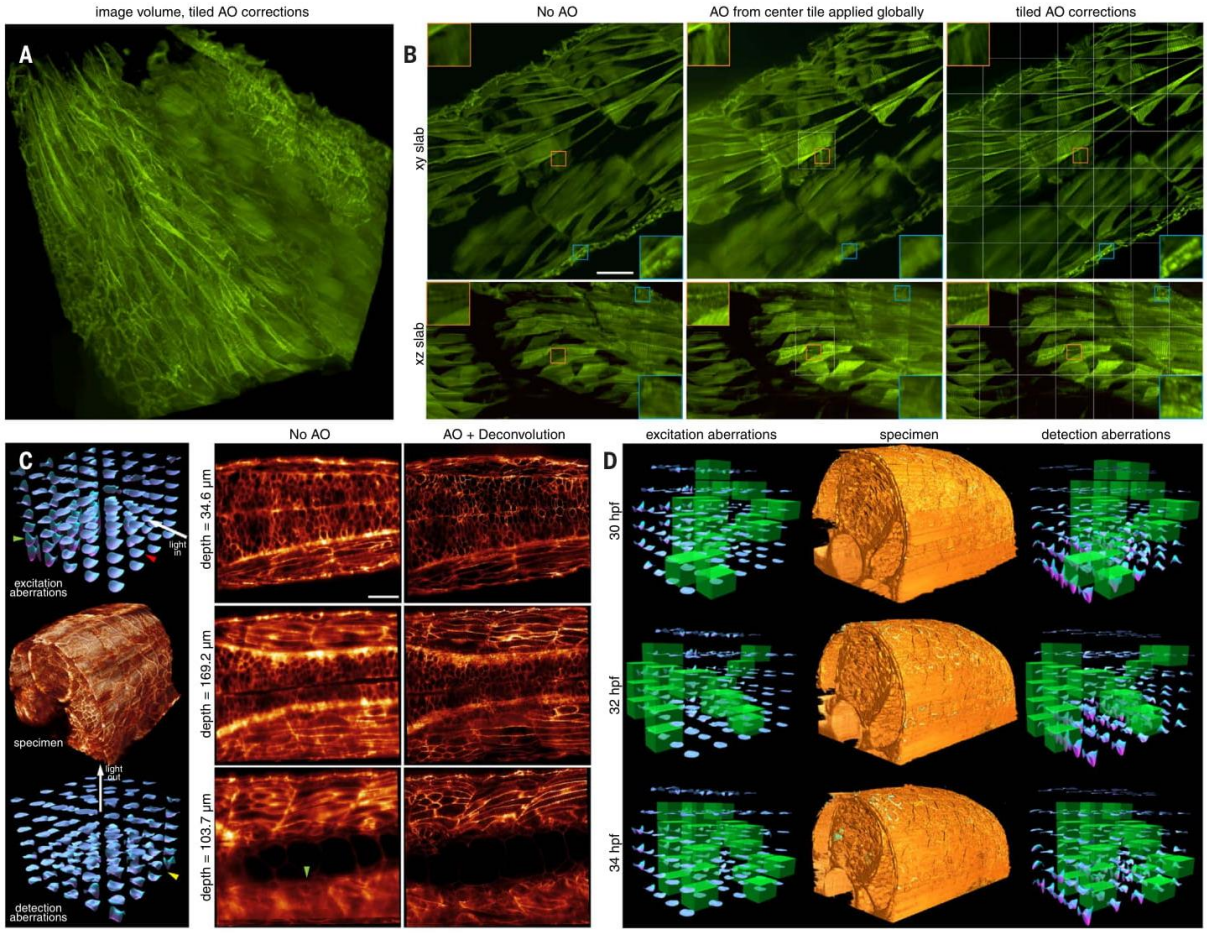


Fig. 4. AO-LLSM over large volumes. (**A**) Aberration-corrected volume rendering over 213 µm by 213 µm by 113 µm in the tail region of a live zebrafish embryo 96 hpf expressing PM-targeted EGFP, assembled from independently corrected subvolumes of 7 by 7 by 3 tiles (movie S4). (**B**) Increasing effectiveness of correction, as seen in orthogonal MIPs from 3-µm-thick slabs, under different scenarios: no AO (left column), AO correction from the center tile applied globally (middle column), and independent AO correction in each tile (right column) (fig. S11). Insets compare, at higher magnification, the quality of correction at the center tile (orange boxes) versus at the tiles at the periphery of the tail (blue boxes). Tile boundaries are shown in white. Scale bar, 30 µm. (**C**) A 5 by 4 by 7 set of measured excitation (left column, top) and detection (left column, bottom) aberrations which, after AO correction, yields diffraction-limited imaging over a 170-µm by 185-µm by 135-µm volume (left column, center) in the spine of a zebrafish embryo 30 hpf (Movie 6). Red and green arrowheads indicate excitation aberrations in specific tiles before and after passage through the notocord, respectively. The yellow arrowhead indicates a tile with a large detection aberration deep within the specimen. Orthoslices before (middle column) and after (right column) AO correction show increased aberration but continued recovery of high resolution at progressively greater depth. Scale bar, 30 µm. (**D**) Aberration-corrected volume renderings over 156 µm by 220 µm by 162 µm in the spine of a zebrafish embryo, at three points from a time series at 30-min intervals (movie S5), flanked by excitation and detection path aberrations at those points. Those tiles whose corrections were updated at a given time point are marked in green.

least dense near the polarized ends. In mitotic cells, however, we again observed (*20*) that the ER remodeled into sheetlike structures concentrated near the PM (Movie 7).

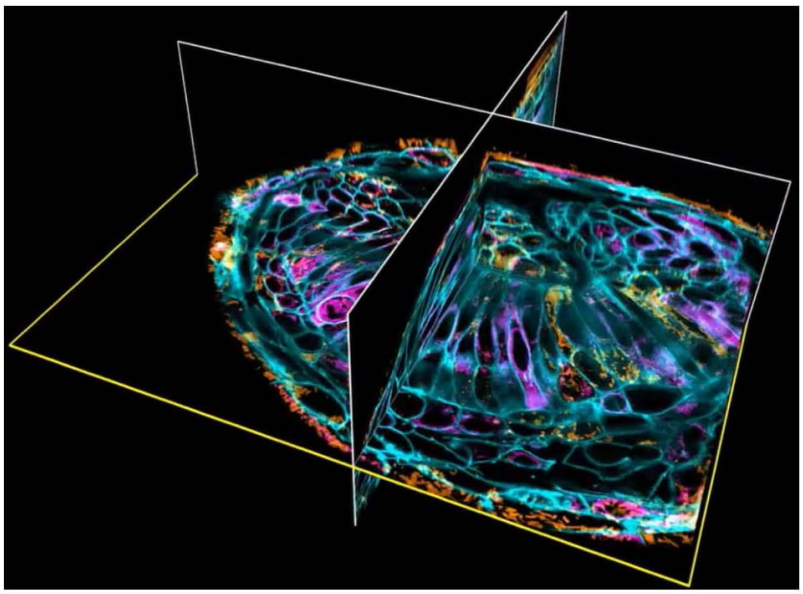
By imaging over time (Fig. 5F), we could follow the stages of RNE cell division (green arrowheads). As reported elsewhere (*25*), we found that, before mitosis, the nucleus retracts to the apical side of the retina (leftmost orange arrowheads), while the cell maintains a thin connection to the basal side (rightmost orange arrowheads). Despite its narrowness, mitochondria remain in this region. RNE cell divisions then occur at the apical surface (white arrowheads). This process has been found to be necessary to maintain the integrity of retinal tissue (*26*).

3D cell migration in vivo

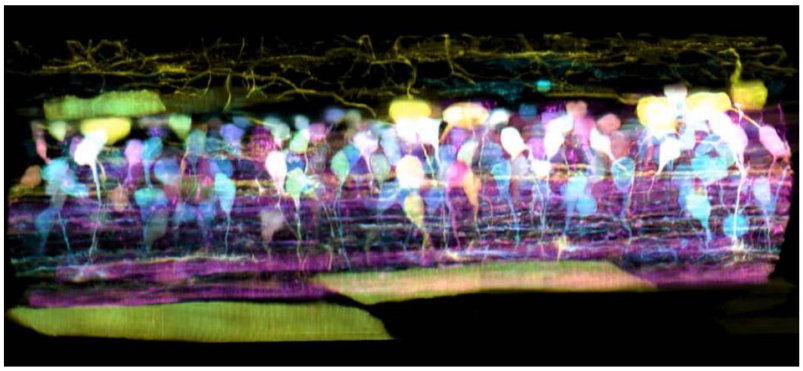
In vivo 3D migration of a cell in the densely crowded environment of living tissue involves forces, constraints, elasticity and adhesion heterogeneity, and chemical cues not found in the simple 2D environment on a cover glass. Furthermore, cell migration involves intricate and rapid remodeling of membranes, organelles, and the cytoskeleton that requires high spatiotemporal resolution to observe. It is therefore a problem well suited to AO-LLSM.

An example of this problem involves the wiring of neuronal circuits during development. To help them establish precise connections, axons are tipped with a highly complex and motile structure, the growth cone. This structure functions as both a sensor and a motor, driving the growth of neurites on the basis of environmental cues (*27*). Although its dynamics have been shown to be important for its proper function (*28*–*30*), its 3D dynamics in an intact animal have been difficult to study because of the lack of techniques for imaging the structure with sufficient resolution in vivo.

To address this, we used AO-LLSM to image growth cones in the spinal cord of a zebrafish embryo in which a subset of newly differentiated neurons expressed stochastic combinations of three different fluorophores via Autobow (*31*) (Fig. 6A and fig. S12), so that they could be spectrally distinguished from earlier differentiated neurons expressing only mCherry [e.g., those within the medial neuropil of the reticulospinal tract (magenta arrowheads, Fig. 6A)]. The Autobow-labeled neurons included Rohon-Beard sensory neurons in the dorsal spinal cord (e.g., yellow arrowheads, Fig. 6A) and interneurons with commissural axons. By imaging over more than two spinal segments at 10.4-min intervals from 58 to 70 hpf (Fig. 6B and Movie 8), we observed that the growth cones of axons migrating in the rostrocaudal direction primarily probed in the direction of their motion (Fig. 6C, top, and movie S6), whereas the growth cones of dorsoventrally aligned axons probed across a broader 2D fan (Fig. 6D, top). Transverse views (Fig. 6, C and D, bottom) revealed that most, if not all, growth cones of both types were located close to the surface of the spinal cord, with their filopodia preferentially aligned parallel to the surface,



Movie 7. Organelle dynamics across the zebrafish eye. PM (cyan), trans-Golgi (green), ER (magenta), and mitochondria (brown) across a 128-μm by 150-μm by 75-μm volume assembled from subvolumes of 4 by 4 by 3 tiles, showing orthoslices in a single tile, volume-rendered tiles before assembly into the combined volume, organelles in the combined volume, dynamics over 30 time points from 24.0 to 26.8 hpf in a 1-μm-thick slab through the combined volume, dynamics in perpendicular orthoslices, and organelle morphologies in different cell types in the computationally expanded volume (Fig. 5).



Movie 8. In vivo imaging of spinal cord neural circuit development. Autobow-labeled, newly differentiated neurons expressing stochastic combinations of three fluorophores in a zebrafish embryo, showing corrective excitation and detection wavefronts in subvolumes of 5 by 2 by 1 tiles, with scrolling updates at one tile (green box) per time point; AO-corrected orthoslices and volume-rendered views in each color channel 58 hpf; and axon pathfinding in each color channel from 58 to 70 hpf (Fig. 6, A to D; fig. S12; and movie S6).

even though the dorsoventrally projecting axons had to pass through the spinal cord to reach its surface. This is consistent with the previous observation that the neurites of late-born V2a ipsilateral projecting interneurons are located lateral to the preexisting ones, forming a layer-like organization based on the age of neurons (*32*), and extends this notion to other classes of spinal neurons. This also raises an interesting possibility that the shape of the growth cone is actively controlled in vivo to keep its exploration within a layer of its own age group.

Cell migration is also a key aspect of the innate immune system. Neutrophils, for example, migrate from the vasculature through the endothelium to reach and engulf infectious targets

(*33*). To study this process in vivo, we imaged endogenously produced immune cells moving through the perilymphatic space of the ear in a transgenic zebrafish larva ~80 hpf expressing the fluorescent protein Citrine in the PM of all cells (Fig. 6E and Movie 9). Acquiring 438 image volumes at 13-s intervals allowed us to accurately measure the 3D position and speed (fig. S13, A and B) of the cellular center of mass. Across five trials involving different embryos at 22°C, immune cells adopted a halting search pattern involving frequent changes in direction and speed (fig. S13, C to G), from nearly motionless to 10 μm/min. In contrast, neutrophilic mammalian HL-60 cells imaged in a collagen matrix at 37°C (*6*) exhibited peak speeds of ~25 μm/min (fig. S13H).

Immune cells next to the ear were remarkable for their rapidly changing and complex 3D morphologies (Fig. 6F). Their surface areas changed as much as 25% in 2 min (fig. S13I) as they remodeled themselves to exhibit a variety of protrusive features, including lamellar sheets, blebs, and short filopodia. Frequently, they also trailed a long filopodium behind them as they migrated to new regions (fig. S14). After injecting fluorescent Texas Red dextran into the heart, we also observed several immune cells containing granules of dextran up to several microns in size (light blue, Fig. 6F, and Movie 9, part 1), presumably

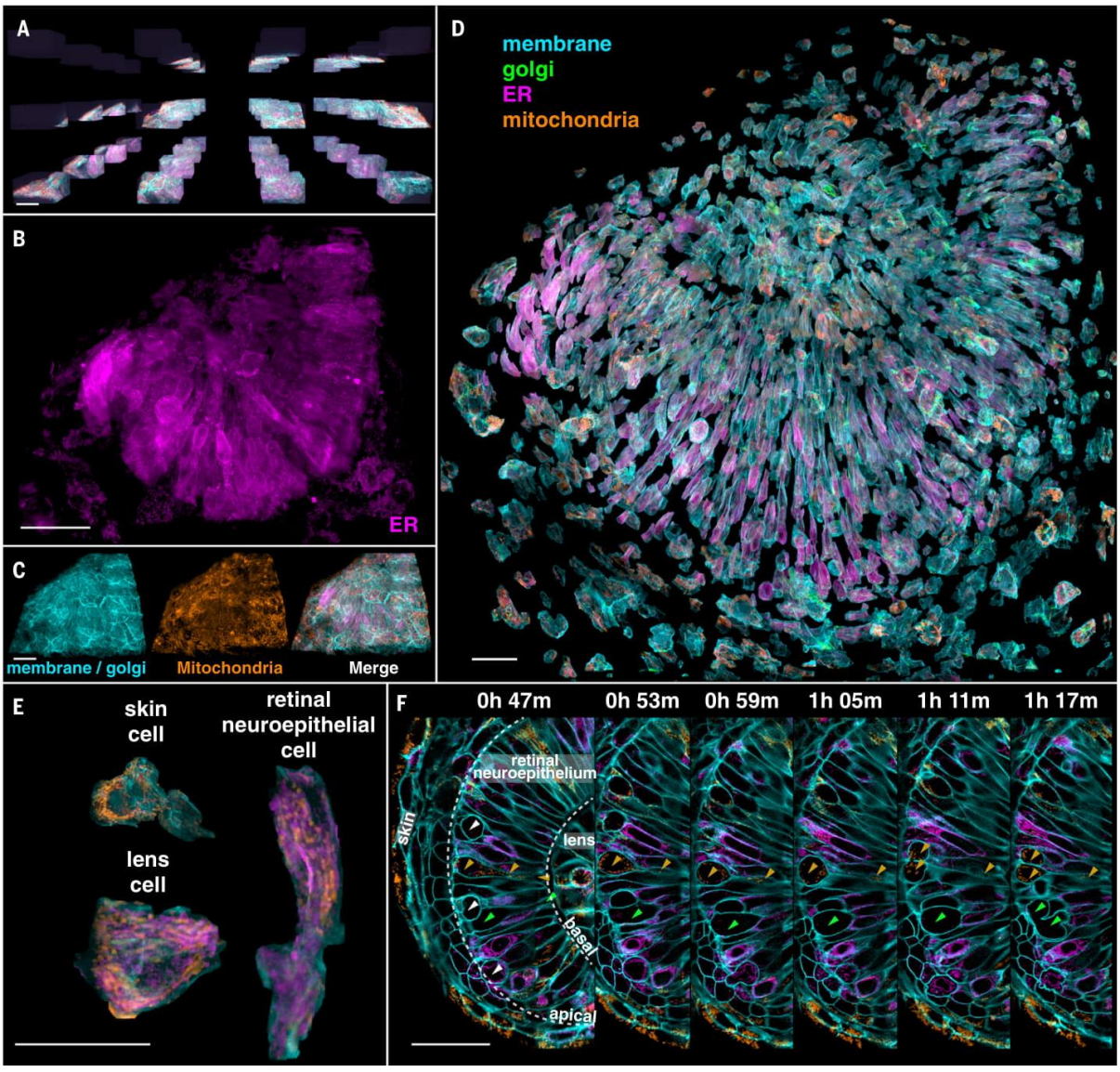


Fig. 5. Organelle diversity across the zebrafish eye. (**A**) Tiled array used to provide AO correction across the eye of a developing zebrafish embryo 24 hpf (Movie 7). Scale bar, 30 μm. (**B** and **C**) Distribution of three different types of organelles across the volume assembled from the tiles in (A). Scale bars, 30 μm. (**D**) Computationally separated cells across the eye, with the organelles colored as indicated. Scale bar, 30 μm. (**E**) Organelle morphologies in cells of three different types within the eye. Scale bar, 30 μm. (**F**) Orthoslices at six different time points highlighting cell divisions (white and green arrowheads, left panel) at the apical surface of the retinal neuroepithelium and mitochondria (orange arrowheads) present from the apical to the basal surface in one dividing cell. Scale bar, 30 μm.

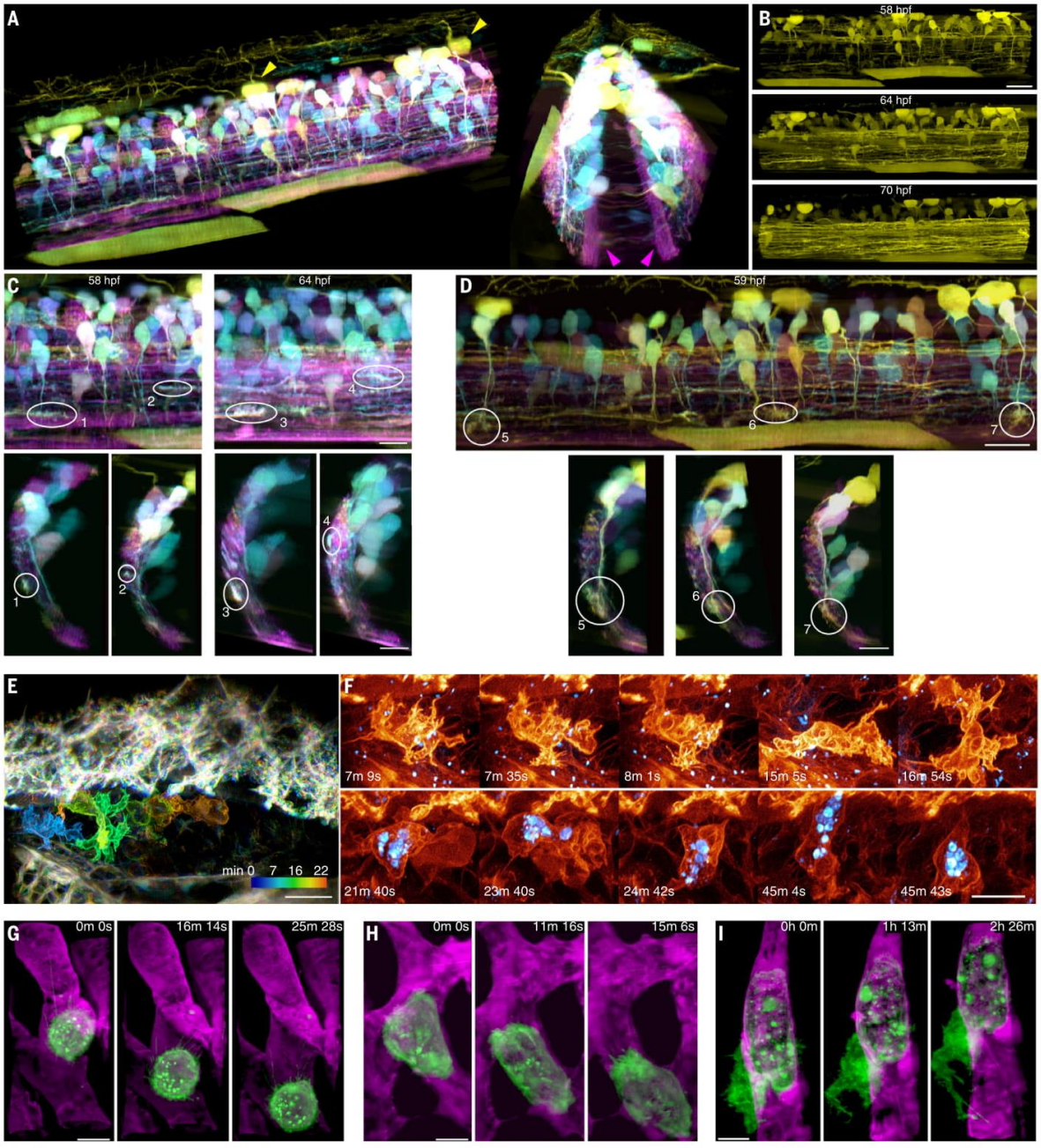


Fig. 6. 3D cell migration in vivo. (**A**) Two views of newly differentiated neurons highlighted by Autobow labeling in a 60-μm by 224-μm by 180-μm section of the spinal cord of a zebrafish embryo 58 hpf (Movie 8). Magenta and yellow arrowheads show neurons differentiated before and after Autobow expression, respectively. (**B**) Increase in the density of rostrocaudally projecting axons over time. Scale bar, 20 μm. (**C**) Sagittal (top) and transverse (bottom) views of the growth cones of four rostrocaudally projecting axons. Scale bars, 10 μm. (**D**) Sagittal (top) and transverse (bottom) views of the growth cones of three dorsoventrally projecting axons. Scale bars, 10 μm. (**E**) Time-coded color overlay of an immune cell migrating within the perilymphatic space next to the inner ear of a live transgenic zebrafish embryo 70 hpf expressing PM-targeted Citrine (Movie 9 and fig. S12). Texas Red dextran particles are shown in blue. Scale bar, 10 μm. (**F**) Changing morphologies of two different immune cells (top and bottom rows), one showing internalized dextran particles (blue) (fig. S13). Scale bar, 5 μm. (**G**) MDA-MB-231 human breast cancer cell (green) rolling in a blood vessel (magenta) in a zebrafish embryo 48 hpf. (**H**) Another MDA-MB-231 cell crawling through a blood vessel. (**I**) A partially extravasated MDA-MB-231 cell, showing an increasingly complex morphology over time (Movie 10 and fig. S1). Scale bars, 10 μm in (G) to (I).

ingested earlier by phagocytosis. However, these did not noticeably affect the motility of the cells or their ability to navigate through tight interstitial spaces.

Apart from immune cells, 3D AO-LLSM movies (e.g., Movie 9, part 2) of the developing ear region revealed a wealth of cellular morphologies and behaviors, including filopodial oscillations at the dorsal surface of skin cells, gradual inflation of the perilymphatic volume, rapid passage of cells through blood vessels, long and active filopodia on endothelial cells, and cellular motion in the hindbrain. With AO, the spatiotemporal resolution and noninvasiveness that we achieved was comparable to what we obtained when imaging cultured cells with our original LLSM approach (*6*), allowing us, for example, to follow the detailed morphological changes in a single endothelial cell lining the hindbrain over the entire course of its division (fig. S15 and Movie 9, part 2).

As a final example of 3D migration, during cancer metastasis, circulating tumor cells (CTCs) extravasate and seed new tumor formation at sites distant from the primary tumor (*34*). Extravasation has been studied extensively in vitro, but little is known about the process in vivo, owing to the highly dynamic nature of cells in circulation and the low density of CTCs in the vasculature. A long-standing hypothesis (*34*) based on in vitro studies is that CTCs co-opt a three-step process used by leukocytes to extravasate at sites of inflammation (*34*): Circulating leukocytes initially form tethers to weakly adhere to the vascular endothelium, causing them to roll slowly downstream (*35*); eventually, they attach and crawl along the endothelial wall; and finally, they penetrate the endothelium and migrate into the tissues beyond.

To determine whether CTCs follow this same pattern in vivo, we used a xenograft model (*36*), where we injected PM-labeled human breast cancer cells (MDA-MB-231) into the vasculature of 2-day-old zebrafish embryos that were transgenic for an endothelial reporter (*kdrl:gfp*). As hypothesized, we observed all three leukocyte migration behaviors in the cancer cells. First, we recorded MDA-MB-231 cells rolling through the blood vessels (Fig. 6G), trailing microvilli that adhered to the endothelium and stretched several microns before detaching as the cell continued to move downstream (Fig. 6G and Movie 10, part 1). Second, we visualized MDA-MB-231 cells crawling along the endothelium (Fig. 6H and Movie 10, part 2). Last, we observed an MDA-MB-231 cell actively engaged in transendothelial migration, with the portion of the cell outside the blood vessel projecting actin-rich extensions into the surrounding tissue (Fig. 6I and Movie 10, part 3) as the area of the cell increased by ~50% over 2 hours (fig. S16).

Discussion

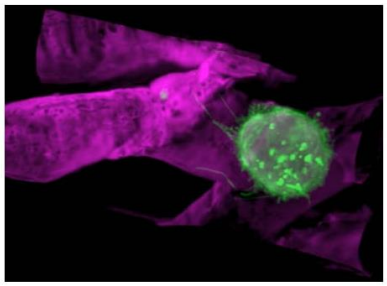
AO-LLSM enables minimally invasive high-speed 3D imaging of subcellular dynamics within optically challenging living specimens while maintaining diffraction-limited resolution, even over large FOVs. It corrects not only sample-induced aberrations, but also those introduced by mounting and immersion media (e.g., Fig. 1E) or imperfections in the optical path through the microscope. It therefore can provide practical 3D resolution exceeding that of nominally higher numerical aperture (NA) confocal or spinning disk microscopes, even in the comparatively benign optical environment encountered when imaging isolated adherent cells on cover slips.

This performance does not come without caveats, however. Because the fluorescence induced by the light sheet is captured with widefield optics, only weakly scattering specimens can be imaged. In addition, extremely sparse and/or weakly emitting fluorescent targets may require colabeling with a second, brighter color channel to provide a sufficient guide star signal for accurate wavefront measurement. Highly absorbing structures such as large blood vessels or melanin bodies within the detection light cone can block guide star light from reaching enough cells of the sensor for accurate wavefront measurement, although wavefront reconstruction algorithms might be made more robust against such missing information. Wavefront aberration can vary considerably across the specimen, and at present, this variation can only be determined empirically for each specimen type and developmental stage to determine how to subdivide the desired image volume into tiled isoplanatic subvolumes of relatively uniform aberration. Fortunately, such tile maps tend to be consistent between specimens of the same type and age, given similar mounting geometries. Lastly, specimens imaged after muscle development must be anesthetized and immobilized, or else a new correction must be measured and applied whenever sample motion exceeds the size of a given isoplanatic patch.

Another caveat is that all but one of the above described involved imaging subcellular dynamics within zebrafish embryos. Although we have shown that we can achieve substantial gains in imaging performance in both *Caenorhabditis elegans* larvae (fig. S17 and movie S7) and *Arabidopsis thaliana* leaves (fig. S18 and movie S8), *Danio rerio* represents an ideal model system in which to study cell and developmental biology in vivo, because it is a rapidly developing transparent vertebrate that is amenable to genetic manipulation. Furthermore, zebrafish embryos are small enough that most regions are optically accessible far into development, yet large enough to exhibit smoothly varying refractive index profiles that result in isoplanatic patch sizes that are comparable to imaging fields typical of LLSM. In contrast, *C. elegans* larvae and adults exhibit larger and more rapid spatial variations in refractive index, particularly near the gut, that can require a denser mesh of AO corrections, despite this organism's reputation as an optically tractable model.

Conventional light-sheet microscopy using weakly focused Gaussian beams is also susceptible to aberrations and would therefore also benefit from AO correction (*8*, *9*). However, conventional systems typically cover much wider FOVs and often operate at greater depth in larger organisms, such as in applications involving functional imaging of whole neural circuits (*37*) or in toto cellular tracking during development (*38*). They therefore usually image over regions much larger than a single isoplanatic patch, making it difficult to retain even cellular-level resolution at all locations and compromising the accuracy and resolution of approaches based on multiview fusion (*39*–*41*). A single AO correction would provide at best only partial correction, and a tiled AO approach, such as we use with LLSM, would negate the high-speed, large-field advantages of the conventional light-sheet microscopy. On the other hand, simultaneous full-field AO correction would likely require

Movie 9. Immune cell migration next to the zebrafish inner ear. Immune cells within the perilymphatic space of the inner ear of transgenic zebrafish embryos expressing PM-targeted Citrine 80 hpf, showing a MIP view before and after AO correction plus deconvolution of two immune cells (orange), one of which has ingested dextran particles (blue), for 438 time points at 13-s intervals; a volume-rendered view with a migrating immune cell and a dividing endothelial cell; and tracking of the position and velocity of an immune cell (Fig. 6, E and F, and figs. S13 to S15).



Movie 10. Cancer cell migration in a zebrafish xenograft model. MDA-MB-231 human breast cancer cells (green) exhibiting three different forms of motion within the vasculature (magenta) of different zebrafish embryos: rolling within a blood vessel while extending long, adhesive microvilli; crawling while conforming to the shape of a blood vessel; and partial extravasation from a blood vessel (Fig. 6, G to I, and fig. S16).

multiconjugate adaptive optics (*42*), substantially increasing cost and complexity.

Perhaps the greatest challenge of AO-LLSM involves mining the immense and complex data that it produces to extract as much biological insight as possible. In Fig. 6, for example, panels A to D represent a minute fraction of a 0.62-terabyte raw data set which first had to be deconvolved, creating a second copy, and then imported into 3D visualization software, generating a third. We deconvolve and store data in real time, but importation and visualization can take many hours, preventing meaningful real-time feedback on whether the biological structure and dynamic process of interest are being optimally recorded. If history is any guide, problems of petabyte-scale data storage and visualization at reasonable cost will yield to continued advancements in commercial hardware, but problems of image analysis and meaningful quantification of data at this scale may prove far less tractable. Although we have demonstrated quantification on a smaller scale through single-particle (Fig. 2, D and E, and fig. S8) and single-cell (fig. S13) tracking, segmentation (Figs. 2B, 3A, and 5D), and measurement of area and volume (Fig. 3, E and F, and figs. S9 and S16), the diversity of questions that can be asked when modern genetic and pharmacological tools are combined with high-resolution 5D in vivo data spanning hundreds of cells over many hours will demand bioinformatics expertise, machine learning, and custom algorithm development on an unprecedented level. Nevertheless, such efforts promise to offer insights into how cells harness their intrinsic variability to adapt to different physiological environments and have the potential to reveal the phenotypic diversity of organelle morphologies, intracellular dynamics, extracellular communication, and collective cell behavior across different cell types, organisms, and developmental stages.

Materials and methods

Lattice light-sheet subsystem

The lattice light-sheet excitation path of the AO-LLSM was designed as described previously (*6*). Noted here are the changes introduced in the AO-LLSM. The collinear laser beams from the combiner were first expanded using a pair of cylindrical lenses and aligned such that up to three different wavelengths illuminated three vertically separated thin stripes on spatial light modulator SLM, (Holoeye, PLUTO-Vis-014 1920 × 1080 pixels; fig. S1). As a grayscale phase modulation device, SLM was introduced to not only create the light sheet but to correct sample-induced aberrations as well. The diffraction orders reflected from SLM were then filtered using annular mask MSK (Photo Sciences) as before and conjugated to galvanometer scanning mirrors G3 and G4 (3 mm mirror, Cambridge Technology, 8315H) to scan the light sheet along the x and the z axis. During imaging, different offset voltages were applied to the z galvo to sequentially realign the light sheet from each laser to the same plane within the specimen. Sample plane conjugate resonant galvanometer RG (Electro-Optical Products Corp. 7 × 8 mm, SC-30) was also added prior to the excitation objective (Special Optics, 0.65 NA, 3.74 mm WD) to wobble the light sheet in the xy plane and thereby minimize stripe artifacts due to localized absorbing or scattering objects in the specimen. The fluorescence generated in the excitation plane was collected with detection objective DO (Nikon, CFI Apo LWD 25XW, 1.1 NA, 2 mm WD) and reflected off deformable mirror DM (ALPAO 97-15) conjugate to the rear focal plane of DO before being imaged at sCMOS camera CAM 1 (Hamamatsu Orca Flash 4.0 v2). Complete details of the optical design are given in supplementary note 1.

Adaptive optics subsystems

In principle, independent AO corrective systems are needed for excitation and detection in LLSM, since light traverses different regions of the specimen in each case and hence is subject to different aberrations. However, given that: (i) aberrations decrease quickly with decreasing numerical aperture (NA) (*7*); (ii) we use at most 0.6 NA for excitation, versus 1.1 NA for detection in LLSM; and (iii) only aberrations within a narrow annulus at the rear pupil of the excitation objective will affect a lattice light sheet, it is not obvious that AO correction of the light sheet itself is necessary. To check, we simulated the effect of aberrations consisting of random combinations of the 55 lowest-order Zernike modes up to a root mean square (RMS) amplitude of two wavelengths (λ). We found (fig. S19) that aberrations at this level could expand a 0.7-μm-thick lattice light sheet to as much as 20 μm, and displace it perpendicular to its plane by up to ±8 μm, indicating that correction of excitation as well as detection is essential.

Hence, during aberration measurement, light from Ti:Sapphire ultrafast pulsed laser 2PL (Coherent Cameleon Ultra II) was ported to either the excitation or detection arm by switching galvanometer SG1 (fig. S1). In either case, TPEF generated within the specimen by scanning the guide star focused by EO or DO was collected by the same objective, descanned (*11*) and sent to homebuilt Shack-Hartmann wavefront sensor ESH or DSH, each consisting of a square microlens array (Edmund Optics) focused onto an EMCCD camera (Andor iXon). Corrective wavefronts were then applied to SLM or DM as described in supplementary notes 3 or 2, respectively. Further hardware details are given in supplementary note 1.

Autofocus measurement was achieved by viewing, side-on through EO, both the light-sheet fluorescence and the plane of fluorescence generated by guide star TPE excitation through DO on camera CAM4, and correcting for any displacement between them as outlined in supplementary note 5.

Zebrafish immobilization, mounting, and imaging conditions

Zebrafish embryos were paralyzed with ~1 ng of α-bungarotoxin protein injected prior to imaging (*43*) or anesthetized using tricaine (0.16 mg/ml) for 15 min. 12-mm-diameter glass coverslips were precleaned as follows: ~20 coverslips were placed in a 50 ml Falcon tube containing 0.1M NaOH and the tube placed in a sonicator for 15 min, followed by at least 5 consecutives washes with Mili-Q water and then immobilized in the sample holder using superglue. An agarose holder containing narrow groves for mounting the embryos was created by solidifying a few drops of 0.5 to 2% (wt/wt) low-melting agarose between the coverslip and a mold containing ridges. For the immune cell experiments, larvae were placed in 3D-printed volcano-shaped mounts (https://www.shapeways.com/shops/megason-lab). A home-made hair-loop was used to position the embryo in the mold, which was then stabilized with a thin layer of agarose made by applying on top of the immobilized embryo ~10 to 20 μl 1% low-melting agarose at 37° to 40°C and then wicking the excess. After solidification, the sample holder was bolted onto a three-axis set of sample stages (Attocube, ECS3030 for x and y, ECS3050 for z) and submerged in a sample bath containing ~8 ml of 1x Danieau buffer. This assembly was then raised by a motorized actuator (Newport, LTA-HS Actuator, integrated with CONEX-CC Controller, CONEX-LTA-HS) until EO and DO were immersed in the media. The sample stages then positioned the desired FOV to the mutual focal point of the objectives. Detailed imaging conditions for each experiment discussed in the paper, including excitation power, imaging time, image, tile and voxel sizes, fluorophores, and proteins, are in table S1. Additional preparation conditions are discussed in supplementary note 6.

REFERENCES AND NOTES

1. J. Icha, M. Weber, J. C. Waters, C. Norden, Phototoxicity in live fluorescence microscopy, and how to avoid it. *BioEssays* **39**, 1700003 (2017). doi: 10.1002/bies.201700003; pmid: 28749075
2. P. P. Laissue, R. A. Alghamdi, P. Tomancak, E. G. Reynaud, H. Shroff, Assessing phototoxicity in live fluorescence imaging. *Nat. Methods* **14**, 657–661 (2017). doi: 10.1038/nmeth.4344; pmid: 28661494
3. F. Pampaloni, E. G. Reynaud, E. H. Stelzer, The third dimension bridges the gap between cell culture and live tissue. *Nat. Rev. Mol. Cell Biol.* **8**, 839–845 (2007). doi: 10.1038/nrm2236; pmid: 17684528
4. M. J. Pittet, R. Weissleder, Intravital imaging. *Cell* **147**, 983–991 (2011). doi: 10.1016/j.cell.2011.11.004; pmid: 22118457
5. R. Weigert, M. Sramkova, L. Parente, P. Amornphimoltham, A. Masedunskas, Intravital microscopy: A novel tool to study cell biology in living animals. *Histochem. Cell Biol.* **133**, 481–491 (2010). doi: 10.1007/s00418-010-0692-z; pmid: 20372919
6. B. C. Chen *et al.*, Lattice light-sheet microscopy: Imaging molecules to embryos at high spatiotemporal resolution. *Science* **346**, 1257998 (2014). doi: 10.1126/science.1257998; pmid: 25342811
7. N. Ji, Adaptive optical fluorescence microscopy. *Nat. Methods* **14**, 374–380 (2017). doi: 10.1038/nmeth.4218; pmid: 28362438
8. D. Wilding, P. Pozzi, O. Soloviev, G. Vdovin, M. Verhaegen, Adaptive illumination based on direct wavefront sensing in a light-sheet fluorescence microscope. *Opt. Express* **24**, 24896–24906 (2016). doi: 10.1364/OE.24.024896; pmid: 27828430
9. C. Bourgenot, C. D. Saunter, J. M. Taylor, J. M. Girkin, G. D. Love, 3D adaptive optics in a light sheet microscope. *Opt. Express* **20**, 13252–13261 (2012). doi: 10.1364/OE.20.013252; pmid: 22714353

10. R. Aviles-Espinosa *et al*., Measurement and correction of in vivo sample aberrations employing a nonlinear guide-star in two-photon excited fluorescence microscopy. *Biomed. Opt. Express* **2**, 3135–3149 (2011). doi: 10.1364/BOE.2.003135; pmid: 22076274
11. K. Wang *et al*., Rapid adaptive optical recovery of optimal resolution over large volumes. *Nat. Methods* **11**, 625–628 (2014). doi: 10.1038/nmeth.2925; pmid: 24727653
12. L. A. Royer *et al*., Adaptive light-sheet microscopy for long-term, high-resolution imaging in living organisms. *Nat. Biotechnol.* **34**, 1267–1278 (2016). doi: 10.1038/nbt.3708; pmid: 27798562
13. F. Arbizzani, M. Mayrhofer, M. Mione, Novel transgenic lines to fluorescently label clathrin and caveolin endosomes in live zebrafish. *Zebrafish* **12**, 202–203 (2015). doi: 10.1089/zeb.2015.1501; pmid: 25748563
14. M. Ehrlich *et al*., Endocytosis by random initiation and stabilization of clathrin-coated pits. *Cell* **118**, 591–605 (2004). doi: 10.1016/j.cell.2004.08.017; pmid: 15339664
15. D. Li *et al*., Extended-resolution structured illumination imaging of endocytic and cytoskeletal dynamics. *Science* **349**, aab3500 (2015). doi: 10.1126/science.aab3500; pmid: 26315442
16. C. Kural *et al*., Dynamics of intracellular clathrin/AP1- and clathrin/AP3-containing carriers. *Cell Rep.* **2**, 1111–1119 (2012). doi: 10.1016/j.celrep.2012.09.025; pmid: 23103167
17. S. Vassilopoulos *et al*., Actin scaffolding by clathrin heavy chain is required for skeletal muscle sarcomere organization. *J. Cell Biol.* **205**, 377–393 (2014). doi: 10.1083/jcb.201309096; pmid: 24798732
18. F. Aguet *et al*., Membrane dynamics of dividing cells imaged by lattice light-sheet microscopy. *Mol. Biol. Cell* **27**, 3418–3435 (2016). doi: 10.1091/mbc.E16-03-0164; pmid: 27535432
19. Z. Kadlecova *et al*., Regulation of clathrin-mediated endocytosis by hierarchical allosteric activation of AP2. *J. Cell Biol.* **216**, 167–179 (2017). doi: 10.1083/jcb.201608071; pmid: 28003333
20. L. Lu, M. S. Ladinsky, T. Kirchhausen, Cisternal organization of the endoplasmic reticulum during mitosis. *Mol. Biol. Cell* **20**, 3471–3480 (2009). doi: 10.1091/mbc.E09-04-0327; pmid: 19494040
21. E. Boucrot, T. Kirchhausen, Endosomal recycling controls plasma membrane area during mitosis. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 7939–7944 (2007). doi: 10.1073/pnas.0702511104; pmid: 17483462
22. J. Icha, C. Kunath, M. Rocha-Martins, C. Norden, Independent modes of ganglion cell translocation ensure correct lamination of the zebrafish retina. *J. Cell Biol.* **215**, 259–275 (2016). doi: 10.1083/jcb.201604095; pmid: 27810916
23. J. Sidhaye, C. Norden, Concerted action of neuroepithelial basal shrinkage and active epithelial migration ensures efficient optic cup morphogenesis. *eLife* **6**, e22689 (2017). doi: 10.7554/eLife.22689; pmid: 28372636
24. J. Icha *et al*., Using light sheet fluorescence microscopy to image zebrafish eye development. *J. Vis. Exp.* **110**, e53966 (2016).
25. F. C. Sauer, Mitosis in the neural tube. *J. Comp. Neurol.* **62**, 377–405 (1935). doi: 10.1002/cne.900620207
26. P. J. Strzyz *et al*., Interkinetic nuclear migration is centrosome independent and ensures apical cell division to maintain tissue integrity. *Dev. Cell* **32**, 203–219 (2015). doi: 10.1016/j.devcel.2014.12.001; pmid: 25600237
27. L. A. Lowery, D. Van Vactor, The trip of the tip: Understanding the growth cone machinery. *Nat. Rev. Mol. Cell Biol.* **10**, 332–343 (2009). doi: 10.1038/nrm2679; pmid: 19373241
28. D. Bentley, A. Toroian-Raymond, Disoriented pathfinding by pioneer neurone growth cones deprived of filopodia by cytochalasin treatment. *Nature* **323**, 712–715 (1986). doi: 10.1038/323712a0; pmid: 3773996
29. C. B. Chien, D. E. Rosenthal, W. A. Harris, C. E. Holt, Navigational errors made by growth cones without filopodia in the embryonic *Xenopus* brain. *Neuron* **11**, 237–251 (1993). doi: 10.1016/0896-6273(93)90181-P; pmid: 8352941
30. J. Q. Zheng, J. J. Wan, M. M. Poo, Essential role of filopodia in chemotropic turning of nerve growth cone induced by a glutamate gradient. *J. Neurosci.* **16**, 1140–1149 (1996). pmid: 8558243
31. D. Cai, K. B. Cohen, T. Luo, J. W. Lichtman, J. R. Sanes, Improved tools for the Brainbow toolbox. *Nat. Methods* **10**, 540–547 (2013). doi: 10.1038/nmeth.2450
32. A. Kinkhabwala *et al*., A structural and functional ground plan for neurons in the hindbrain of zebrafish. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 1164–1169 (2011). doi: 10.1073/pnas.1012185108; pmid: 21199947
33. S. de Oliveira, E. E. Rosowski, A. Huttenlocher, Neutrophil migration in infection and wound repair: Going forward in reverse. *Nat. Rev. Immunol.* **16**, 378–391 (2016). doi: 10.1038/nri.2016.49; pmid: 27231052
34. N. Reymond, B. B. d'Água, A. J. Ridley, Crossing the endothelial barrier during metastasis. *Nat. Rev. Cancer* **13**, 858–870 (2013). doi: 10.1038/nrc3628; pmid: 24263189
35. A. Marki, K. Buscher, Z. Mikulski, A. Pries, K. Ley, Rolling neutrophils form tethers and slings under physiologic conditions in vivo. *J. Leukoc. Biol.* **103**, 67–70 (2017).
36. K. Stoletov *et al*., Visualizing extravasation dynamics of metastatic tumor cells. *J. Cell Sci.* **123**, 2332–2341 (2010). doi: 10.1242/jcs.069443; pmid: 20530574
37. M. B. Ahrens, M. B. Orger, D. N. Robson, J. M. Li, P. J. Keller, Whole-brain functional imaging at cellular resolution using light-sheet microscopy. *Nat. Methods* **10**, 413–420 (2013). doi: 10.1038/nmeth.2434; pmid: 23524393
38. R. Tomer, K. Khairy, F. Amat, P. J. Keller, Quantitative high-speed imaging of entire developing embryos with simultaneous multiview light-sheet microscopy. *Nat. Methods* **9**, 755–763 (2012). doi: 10.1038/nmeth.2062; pmid: 22660741
39. R. K. Chhetri *et al*., Whole-animal functional and developmental imaging with isotropic spatial resolution. *Nat. Methods* **12**, 1171–1178 (2015). doi: 10.1038/nmeth.3632; pmid: 26501515
40. A. Kumar *et al*., Dual-view plane illumination microscopy for rapid and spatially isotropic imaging. *Nat. Protoc.* **9**, 2555–2573 (2014). doi: 10.1038/nprot.2014.172; pmid: 25299154
41. J. Swoger, P. Verveer, K. Greger, J. Huisken, E. H. Stelzer, Multi-view image fusion improves resolution in three-dimensional microscopy. *Opt. Express* **15**, 8029–8042 (2007). doi: 10.1364/OE.15.008029; pmid: 19547131
42. R. D. Simmonds, M. J. Booth, Modelling of multi-conjugate adaptive optics for spatially variant aberrations in microscopy. *J. Opt.* **15**, 094010 (2013). doi: 10.1088/2040-8978/15/9/094010
43. I. A. Swinburne, K. R. Mosaliganti, A. A. Green, S. G. Megason, Improved long-term imaging of embryos with genetically encoded α-bungarotoxin. *PLOS ONE* **10**, e0134005 (2015). doi: 10.1371/journal.pone.0134005; pmid: 26244658

ACKNOWLEDGMENTS

We thank the Shared Resource teams at the Janelia Research Campus for their skill and dedication in specimen handling and preparation, K. Schaefer for the preparation of cancer cells for the xenograph experiments, and the Instrument Design and Fabrication team for their manufacturing expertise. We also gratefully acknowledge the support of the Janelia Visitor Program. S.U. thanks H. Elliott, D. Richmond, and R. Gao for discussions and acknowledges the MATLAB code repository received from the Computational Image Analysis Workshop supported by NIH grant GM103792. **Funding:** T.-L.L., D.E.M., V.S., J.S., M.K., E.M.M., and E.B. are funded by the Howard Hughes Medical Institute (HHMI). T.K. and S.U. are funded by grants from Biogen, Ionis Pharmaceuticals, and NIH grant R01GM075252 (to T.K.). S.U. is a Fellow at the Image and Data Analysis core at Harvard Medical School. S.S. is funded by a Living Systems Institute start-up grant, University of Exeter. K.R.M., I.A.S., Z.M.C., T.W.H., and S.G.M. were supported by NIH grant R01DC015478. D.Q.M. is funded by the NIH (5R00CA154870-05 and 1R01GM121597-01). D.Q.M. and B.L.M. are funded by the Carol M. Baldwin Foundation and are Damon Runyon-Rachleff Innovators supported (in part) by the Damon Runyon Cancer Research Foundation (DRR-47-17). B.L.M. is also funded by the NSF (IOS1452928). D.H. is a Pew-Stewart Scholar for Cancer Research supported by the Pew Charitable Trusts and NIH grant R01CA196884. D.D. was supported by a Human Frontier Science Program Fellowship. D.G.D. was supported by NIH grant R35GM118149. **Author contributions:** E.B. supervised the project and wrote the manuscript with input from all coauthors. T.-L.L. built the microscope with input from E.B., D.E.M., and K.W. and performed all microscope characterization experiments. D.E.M. created the instrument control software. T.-L.L., V.S., and S.U. acquired all biological data with coauthors. D.D., D.G.D., R.F., and D.H. provided organoids and led related experiments. T.K., S.U., B.C., and S.S. provided transgenic zebrafish and AP2 cells and led related clathrin experiments. Z.M.C., T.W.H., and S.G.M. created the mCardinal-PM transgenic zebrafish. T.K., S.G.M., S.U., and I.A.S. provided zebrafish and led related organelle dynamics and in vivo immune cell experiments. J.S. and M.K. created the Autobow zebrafish, and M.K. led related experiments. A.Q.K., T.N.M., and D.Q.M. provided *C. elegans* specimens and led related experiments. B.L.M. provided MDA-MB-231 cells and vasculature-labeled zebrafish and led related experiments with D.Q.M. Y.R., H.Y., and E.M.M. provided *Arabidopsis* specimens and led related experiments. S.U., K.R.M., T.-L.L., and V.S. processed all image data. S.U. performed quantitative analysis of all image data. S.U., T.-L.L., and E.B. produced all figures and movies. **Competing interests:** Portions of the technology described herein are covered by U.S. Patent 7,894,136 issued to E.B., assigned to Lattice Light of Ashburn, VA, and licensed to Carl Zeiss Microscopy; U.S. Patents 8,711,211 and 9,477,074 issued to E.B., assigned to HHMI, and licensed to Carl Zeiss Microscopy; U.S. Patent application 13/844,405 filed by E.B. and K.W. and assigned to HHMI; and U.S. Patent 9,500,846 issued to E.B. and K.W. and assigned to HHMI. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper or the supplementary materials. Documentation for construction of a lattice light-sheet microscope can be obtained by execution of a research license agreement with HHMI.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/360/6386/eaaq1392/suppl/DC1
Supplementary Text
Figs. S1 to S20
Table S1
References (*44*–*75*)
Movies S1 to S8

5 October 2017; accepted 19 February 2018
10.1126/science.aaq1392