UNIVERSITY OF CALIFORNIA,
IRVINE

Generating robust responses to retinoic acid during development

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Biological Sciences

by

Lianna Fung

Dissertation Committee:
Professor Thomas Schilling, Chair
Associate Professor Michael Parsons
Professor Susana Cohen-Cory

2021

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGEMENTS

This work would not be possible without the support of the many individuals who accompanied and supported me along my PhD journey. I would like to first thank my thesis advisor, Dr. Tom Schilling, who was pivotal in my growth as an independent researcher, providing resources and freedom to pursue different research projects in the lab. I would also like to thank all of the current and former members of the Schilling lab I've had the fortune to work with, Arul, Danny, Ines, Olga, Pavan, David, Diego, Irene, Jessica, Lauren, Pierre, and Praveer. I want to specifically highlight the contributions of Arul and Danny for their mentorship and guidance. They were my go-to for any idea brainstorming, troubleshooting, help locating items in the lab, and just really good conversation. Time in the Schilling lab would also not be the same without all those enjoyable times discussing scientific problems and laughing over good food and coffee with Pavan, David, and Lauren. I would also like to thank my current committee members Dr. Michael Parsons and Dr. Susana Cohen-Cory, and former committee members Dr. David Gardiner and Dr. Zeba Wunderlich for their valuable advice and support over the years. I am also extremely grateful for the opportunity to work with my computational collaborator Yuchi Qiu and his adviser Qing Nie, which really allowed me to appreciate and dive into the field of computational biology and modeling. I would also not be here without the support of my family, friends made before and during my time at UCI, and scientific mentors back from my time at UCSD as an undergraduate. Most of all, I want to thank my best friend and partner, Jeff, who has stayed with me through all these years on this long, sometimes fun but often difficult scientific journey. Lastly, I would like to thank my funding sources from the NIH and NSF-Simons Center for Multiscale Cell Fate Research.

Curriculum Vitae

Lianna Fung

Education

University of California, Irvine
Irvine, California
Ph.D. Developmental and Cell Biology
Fall 2015-Summer 2022
MS, Biological Sciences
Fall 2015-June 2019

University of California, San Diego
La Jolla, California
BS, General Biology, cum laude
Fall 2010-June 2014
Phi Beta Kappa Member

Research Experience

University of California, Irvine
Irvine, California
Developmental and Cell Biology (Thomas F. Schilling Lab), PhD Candidate
Spring 2016- Summer 2022

Salk Institute for Biological Studies, San Diego
La Jolla, California
Laboratory of Genetics (Fred H. Gage Lab), Research Assistant
Fall 2013- August 2015

Publications and Presentations

Publications:

- Qiu, Y., **Fung, L**., Schilling, T., Nie, Q. (2021). Multiple morphogens and rapid elongation promote segmental patterning during development. Accepted into *PLOS Computational Biology May 2021.* https://doi.org/10.1101/2021.04.22.440966

- **Fung, L**., Guzman, H., Sevrioukov, E., Idica, A., Park, E., Bochnakian, A., Daugaard, I., Jury, D., Mortazavi, A., Zisoulis, D. G., Pedersen, I. M. (2019). miR-128 Restriction of LINE-1 (L1) Retrotransposition Is Dependent on Targeting hnRNPA1 mRNA. *International Journal of Molecular Sciences*, *20*(8), 1955. https://doi.org/10.3390/ijms20081955

- Vadodaria, K. C., Ji, Y., Skime, M., Paquola, A., Nelson, T., Hall-Flavin, D., Fredlender, Callie., Heard, K. J., Deng, Y., Le, A. T., Dave, S., **Fung, L**., Li, X., Marchetto, M. C., Weinshilboum, R., Gage, F. H. (2019). Serotonin-induced hyperactivity in SSRI-resistant major depressive disorder patient-derived neurons. *Molecular Psychiatry*, 1. https://doi.org/10.1038/s41380-019-0363-y

- Zaragoza, M. V., **Fung, L**., Jensen, E., Oh, F., Cung, K., McCarthy, L., Tran, C. K., Hoang, V., Hakim, S. A., Grosberg, A. (2016). Exome Sequencing Identifies a Novel *LMNA* Splice-site Mutation and Multigenic Heterozygosity of Potential Modifiers in a Family with Sick Sinus Syndrome, Dilated Cardiomyopathy, and Sudden Cardiac Death. *PLoS ONE.* http://dx.doi.org/10.1371/journal.pone.0155421

- Vadodaria, K. C., Mertens, J., Paquola, A., Bardy, C., Li, X., Jappelli, R., **Fung, L.**, Marchetto, M. C., Hamm, M., Gorris, M., Koch, P., and Gage, F. H. (2016). Generation of functional human serotonergic neurons from fibroblasts. *Molecular Psychiatry*, *21*(1), 49–61. http://doi.org/10.1038/mp.2015.161

Presentations and Posters:

- **Fung L**, Nie Q, Schilling T F. Visualizing Noise in Gene Expression Using MS2 in the Zebrafish Hindbrain. Poster session presented at: West Coast Society for Developmental Biology Meeting; 2019 March 21-24; Cambria, CA.

- Vadodaria K, Mertens J, Jappelli R, **Fung L**, Soltani S, Wright R, Gage FH. Directed fate conversion to human GABAergic neurons using transcriptional drivers. Poster session presented at: iPSC Cells: Directed Differentiation. 12th Annual Meeting of the International Society for Stem Cell Research; 2014 Jun 18-21; Vancouver, Canada.

- UCSF iGEM 2010 Team. Synthetic Killers. Presentation and poster session presented at: International Genetically Engineered Machine Competition; 2010 Nov 6-8; Cambridge, MA.

Academic Honors and Fellowships

- UC Irvine School of Biological Sciences Susan V Bryant Graduate Fellowship Award (2021)
- NSF-Simons Center for Multiscale Cell Fate Research Graduate Fellow (2018-2019, 2019-2020, 2020-2021)
- DTEI Graduate Fellow (Summer 2020)
- NSF GRFP Honorable Mention (2017)
- Graduate Dean's Recruitment Fellowship (University of California, Irvine 2015)

Abstract of the Dissertation

Generating robust responses to retinoic acid during development

by
Lianna Fung
Doctor of Philosophy in Biological Sciences
University of California, Irvine, 2022
Professor Thomas F. Schilling, Chair

During vertebrate development, early patterning of numerous tissues and cell types occurs consistently with great precision and accuracy. Tissues are patterned by morphogens, one major one being a derivative of dietary Vitamin A called retinoic acid (RA). Intriguingly, patterning in response to RA is both robust (i.e. less variable) and accurate, with tissues consistently forming in the same manner from embryo to embryo with sharp boundaries, despite variations in genetic background, environmental factors, and noise inherent in biological processes. Despite their importance in development, the mechanisms that regulate this robustness and accuracy are still not well understood. To unravel these regulatory mechanisms, we utilized various methods from computational modeling to dissect patterning events in hindbrain segmentation to gene knockouts in animal models to study individual roles of specific binding proteins in regulating responses to RA.

To understand the mechanisms that promote robust responses to RA, I first examine the process of hindbrain segmentation by the RA morphogen gradient. During vertebrate hindbrain development, the neural tube is subdivided along the anterior-posterior (A-P) axis into seven distinct segments called rhombomeres (r). Boundaries between rhombomeres are initially rough and jagged, with cells expressing different segment identity genes intermingling at transition zones spanning a few cell diameters wide at prospective boundaries. Mechanisms responsible for boundary sharpening have been investigated for the boundaries between r3/4 and r4/5 in the hindbrain revealing insights into interactions between RA and segment defining genes *hoxb1a*

and *krox20* in determining segmental patterning. However, these previous studies were limited in that they focused solely on those boundaries, a single morphogen gradient, and did not model the correct tissue dynamics that occurred such as convergent extension. To expand upon these studies, together with my mathematical collaborator Yuchi Qiu, we computationally modeled these processes using new parameters that included both changing tissue dimensions and more complex gene regulatory networks. Through our new model, we revealed a novel role for rapid tissue elongation in boundary sharpening and maintenance of small segment size. We also explained how two different morphogens, fibroblast growth factor (FGF) and RA which vary in range of signaling activity, synergize to specify multiple boundaries (r2-6) and consistent segment size robustly and accurately.

Then to understand at the cellular level what mechanisms regulate robust responses to RA, I investigated the roles of cellular retinoic acid binding proteins (Crabps). Vertebrates have two highly conserved Crabps, Crabp1 and Crabp2, which transport RA to its receptors in the nucleus or to RA degradation enzymes in the cytoplasm. These dual roles make Crabps excellent candidates for regulating intracellular RA levels and promoting robust cellular responses to RA. To investigate this, I generated mutants for all Crabp1s and Crabp2s in zebrafish. I discovered stark contrasts in sex ratios between Crabp null mutants, with males being overly represented in Crabp2 null mutants. Upon closer examination of mutant gonads and through exogenous RA treatments, I showed Crabp2s mediate RA signaling to promote germ cell proliferation and male sex determination as opposed to Crabp1s, which inhibit germ cell proliferation promoting female sex determination. This revealed a new role of Crabps in mediating RA responses during gonad development and maintaining appropriate sex ratios in a population. In my thesis, I provide insights into the mechanisms in development used to achieve both accurate and robust tissue patterning and cellular responses to RA.

Chapter I

Introduction and Background

Biological development is a remarkably robust yet accurate and precise process. The same cells, tissues, and structures form over and over again between different individuals within a species despite their being exposed to different environmental factors and circumstances. An extreme example of this is in organisms like Caenorhabditis elegans where every adult has the same number of somatic cells that can be traced from their earliest of cell divisions, but these same principles can be seen in development of much more complex organisms (Sulston et al., 1983). In many fish and amphibians, primary neurons such as the Mauthner cells, the largest pair of neurons in brain, are reliably specified and form interconnections with other neurons to form functional circuits amongst the sea of millions of cells in the nervous system (Eaton et al., 2001). Even looking to mammals and humans, the vast majority of individuals progress through the same stages during embryonic development and end up with identifiable shared traits. The concept that describes this phenomenon is developmental robustness.

Concept of developmental robustness

Developmental robustness is the ability of a phenotype or biological trait to arise consistently despite perturbation. These perturbations can arise genetically, through the environment, or through stochastic noise intrinsic in biological processes (Félix & Wagner, 2006). The concept of developmental robustness is synonymous with the concept of *canalization*, first coined by Conrad Hal Waddington. This concept describes that under the pressures of natural selection, most developmental processes have become canalized in that normal tissues and structures form despite minor variations in genetics or environment. (Waddington, 1942). The evidence for this is clear through observations in developmental biology where animals are formed of distinct tissues and structures and not intermediate blends,

as well as in the population where the wild-type predominates in the majority of individuals despite individual differences in genetics or environment. A more specific example of this is in the fate determination of cells, where cells can be made to differentiate into one type or another but would be hard-pressed to remain as intermediates of two distinct populations. Naturally, the concept of developmental robustness and canalization brings up questions of what buffering mechanisms exist to support these processes? How are tissues patterned so robustly under different circumstances and what processes have evolved to support them?

Morphogens and morphogen gradients

When discussing tissue wide patterning events, the most widely explanatory mechanism is via morphogen gradients. Morphogens, as first defined by Alan Turing, are chemical substances that diffuse between cells and pattern tissues in a non-uniform, concentration-dependent manner (Turing, 1952). This concept was expanded upon with the addition of the concept of positional information introduced by Lewis Wolpert, where cells were assigned, positional information based on the concentration of morphogen they were exposed to leading to the idea of diffusible morphogen gradients controlling cell fate decisions in a tissue (Wolpert, 1969). This is most famously depicted in the French flag model developed by Wolpert to explain how tissues, like the French flag, would retain a specific pattern that scaled with changes in size. Cells in this model exposed to specific thresholds of morphogens based on their position along a morphogen gradient would take on different fates. The formation of morphogen gradients was further elaborated by Francis Crick, who described the “source-sink” model for morphogen gradient formation (Crick, 1970). In this model, cells at the “source” produce and secrete the morphogen at one end of the tissue which then diffuses to the other end where it is destroyed by cells in the “sink’. This process creates a stable concentration gradient of a morphogen across a tissue. However, patterning by morphogen gradients is much more complex than the theoretical framework initially proposed by early researchers. Even for a

single morphogen alone, there is complex regulation of morphogen supply, transport, and degradation as well as different responses of cells depending on the context. These processes are only further complicated by the regular presence of multiple morphogens interacting with each other to simultaneously pattern tissues.

Retinoic acid metabolism and its role in tissue patterning

One particularly unique morphogen is retinoic acid (RA). RA is unique in that it is a dietary-derivative of Vitamin A (Kleiner-Bössaler & DeLuca, 1971). This makes it distinct from other morphogens commonly studied in development like Wnt and FGF which are secreted proteins synthesized *de novo*. Most animals are incapable of synthesizing retinoids *de novo* so they must rely on dietary intake of either carotenoids such as β-carotene from plants or Vitamin A through animal sources such as eggs and milk. This is particularly important during embryonic development as maternal vitamin A deficiency (VAD) has been shown to lead to a multitude of embryonic defects and malformations (Wilson et al., 1953). To be converted to RA, Vitamin A (retinol) must undergo a series of metabolic conversions where it is first converted to retinaldehyde (retinal) by a retinol dehydrogenase (RDH10) which then undergoes oxidation to RA by an aldehyde dehydrogenases (ALDH1A1, ALDH1A2, or ALDH1A3), (Dupé et al., 2003; Niederreither et al., 1999; Sandell et al., 2007). Of the ALDHs, ALDH1A2 has been found in vertebrates to be primarily responsible for most RA production during embryonic development with its loss resulting in phenotypes similar to those observed in VAD animal models (Niederreither & Dollé, 2008).

Equally important for proper development is the proper metabolism and clearance of RA. Early studies have shown the teratogenic effect of excess vitamin A intake during pregnancy or embryonic exposure to exogenous RA (Kesseland & Gruss, 1991; Shenefelt, 1972). In cells, RA is degraded by one of the three cytochrome P450 (CYP) enzymes (CYP26A1, CYP26A2, or CYP26A3). Loss of these enzymes, especially CYP26A1, results in lethal teratogenic defects

similar to those observed in embryos subjected to excess RA (Abu-Abed et al., 2001; Uehara et al., 2007; Yashiro et al., 2004).

FIGURE 1

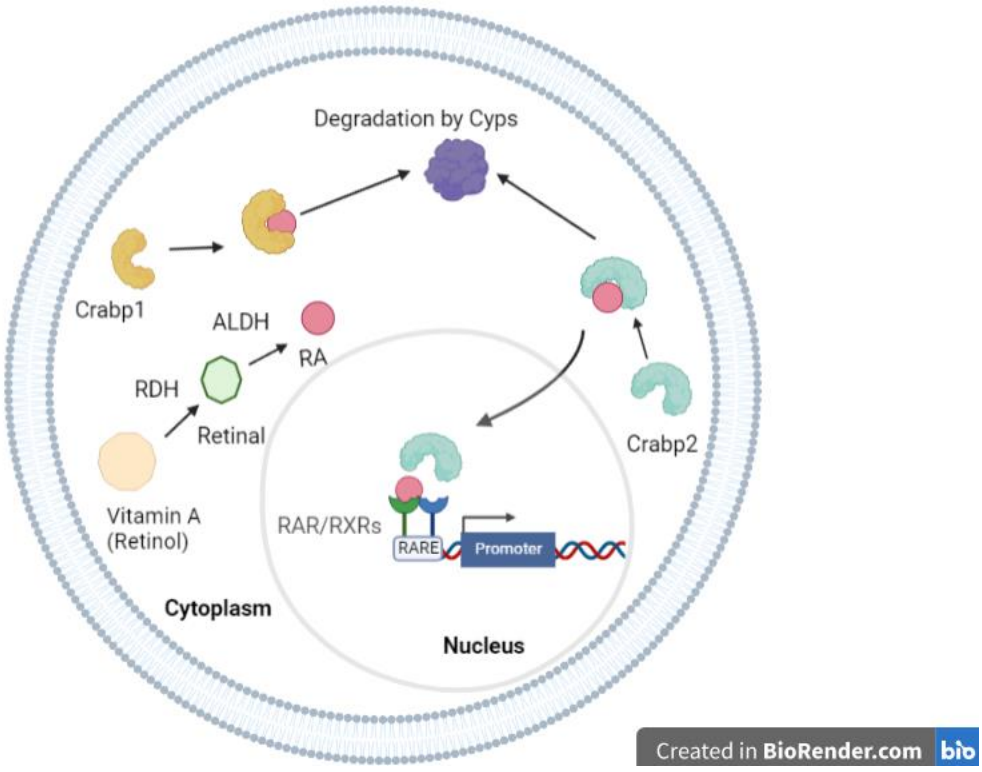

Figure 1. Schematic of RA metabolism and signaling Vitamin A (retinol) is converted to retinaldehyde (retinal) by retinol dehydrogenase (RDH) and then undergoes oxidation by aldehyde dehydrogenase (ALDH) to retinoic acid (RA). RA is transported intracellularly by cellular retinoic acid binding proteins (Crabps). Crabp1 and Crabp2 both transport RA to cytochrome P450 enzymes (Cyps) for degradation. Crabp2 also transports RA to the nucleus to bind to nuclear RA receptors (RARs) which then form heterodimers with retinoic X receptors (RXRs). RAR/RXR heterodimers bind to specific DNA elements called retinoic acid response elements (RAREs) located at gene regulatory elements of RA target genes.

Despite being a dietary derivative, RA is involved in patterning of many different tissues in animals from the earliest events in specifying embryonic tissues to patterning of adult structures in the body (Niederreither & Dolle, 2008). RA was even one of the earliest identified morphogens due to its effects on limb patterning (Thaller & Eichele, 1987), but one of RA's most essential functions in vertebrates is through its role in embryonic patterning of the anterior-

posterior (A-P) body axis. Canonical RA signaling is mediated by RA binding to nuclear RA receptors (RARs) which then form heterodimers with retinoic X receptors (RXRs) that bind to RA response elements (RAREs) located in gene regulatory elements of RA target genes (Kastner et al., 1997; Mark et al., 2006). Many of these RA targets include *Hox* genes, a family of homeobox transcription factors, which play an important role in A-P patterning of the vertebrate embryo (Balmer & Blomhoff, 2002; McGinnis & Krumlauf, 1992).

Retinoic acid patterning of the hindbrain

An excellent example of this is demonstrated by RA's role in patterning of the rhombencephalon or hindbrain in vertebrates. During vertebrate hindbrain development, the neural tube is subdivided along the A-P axis into seven distinct segments called rhombomeres (r) which are characterized by distinct domains of gene expression. In zebrafish, RA is synthesized posteriorly in the paraxial mesoderm by Aldh1a2 (source) and degraded anteriorly in the future forebrain/ midbrain by Cyp26a1 (sink). Cyp26a1 is induced by RA (White et al., 2007), and inhibited by two other posteriorizing signals, FGF and Wnt (Kudoh et al., 2002). This results in the formation of a RA gradient that patterns the A-P axis of the hindbrain (Schilling et al., 2012; White et al., 2007).

The RA gradient in the hindbrain promotes expression of more posteriorly expressed genes at progressively higher concentrations (Marshall et al., 1992). RA activates directly and indirectly two of the earliest rhombomere segmentation genes hoxb1a and krox20, which specify r4 and r3/5 respectively, (Begemann et al., 2001; Hernandez et al., 2004; Labalette et al., 2011; Niederreither et al., 1999). Subsequently, r4 becomes a local signaling center that expresses several FGF ligands (FGF3 and FGF8) (Mave et al., 2002). This local signaling center is found to be important as loss of FGF signaling results in patterning defects in the surrounding rhombomeres, particularly r5 and r6 (Marin & Charnay, 2000; Walshe et al., 2002). This results in an interesting dynamic of a local signaling center of FGF and a long-distance RA

gradient during hindbrain segmentation (White et al., 2007). Antiparallel gradients of Shh and BMP during dorsal-ventral (D-V) axis patterning of the neural tube have been shown to improve patterning accuracy by providing additional positional information (Zagorski et al., 2017). Antiparallel and antagonistic gradients of RA and FGF are seen in other contexts such as caudal body axis extension and limb bud development but in hindbrain segmentation these morphogen gradients appear to form and interact quite differently (Corral et al., 2003; Mercader et al., 2000; Niederreither & Dolle, 2008). This suggests the importance of multiple morphogens for generating robust and accurate tissue patterns, but the dynamics and roles of simultaneous morphogen gradients in different pattern formation contexts are still complex and not well understood. In Chapter II of my thesis, we examine the role of the two gradients of RA and FGF and how they synergize to promote robust and accurate patterning of the hindbrain.

Morphogenesis during embryogenesis and hindbrain development

Another component that occurs simultaneously alongside morphogen gradients during tissue patterning is morphogenesis. Morphogenesis encompasses the movements that cells go through in order for a tissue or organ to develop its shape. While morphogen gradient models and diagrams often depict patterning events with linear gradients and static tissue dimensions, the actual reality is not that simple. During development, embryonic tissues are actively growing and undergoing remodeling with cells moving and dividing in all three dimensions. These movements are certainly not random though and are instead an orchestrated set of mechanisms that reorganize the early embryo.

One of these major mechanisms is the process of convergent extension, where tissues narrow along one axis and lengthen along a perpendicular axis (Keller & Sater, 1992; Wallingford et al., 2002). In the zebrafish hindbrain, the neuroepithelium undergoes rapid convergent extension where it lengthens along the A-P axis and narrows along the left-right axis. This process is mediated by a combination of cell division and mediolateral cell

intercalations (Warga & Kimmel, 1990; Papan & Campos-Ortega, 1994). These dynamic changes in growth and cell movements not only affect cell-cell signaling but also affect the shallowness and depth of morphogen gradients. Previous studies modeling Drosophila wing disc patterning have shown gradients are able to scale with changes in tissue size due to underlying molecular mechanisms (Ben-Zvi et al., 2011; Fried & Iber., 2014; Zhu et al., 2020). However, previous modeling of the vertebrate hindbrain indicates solely elongating tissue along the A-P axis causes the RA gradient to fade, negatively impacting hindbrain patterning (Zhang et al., 2012). This brings to question how do gradients and tissue dynamics during morphogenesis cooperate in the hindbrain to generate accurate and precise patterns? The impact of the rate and timing of specific processes in morphogenesis and their interactions with other segmentation mechanisms are also not well understood. In Chapter II of my thesis, we examine how morphogenesis affects the RA morphogen gradient, cell responses to RA levels, and cell-cell signaling during hindbrain patterning.

Roles of stochastic noise in biological processes

Another aspect to consider is the role of stochastic noise, which is an innate source of biological perturbation within biological processes. During hindbrain segmentation, boundaries between gene expression domains are remarkably sharp. (Cooke & Moens, 2002). Yet, cells at boundaries experience fluctuations or noise in morphogen levels, so how are they able to accurately express the correct gene expression state of their respective segment? In spatial patterning systems, noise is largely seen to be detrimental with mechanisms having been proposed to limit noise to promote robustness of processes (Nie et al., 2020). However, for non-spatially constrained systems, noise has been shown to be a regulator of biological switching between high and low gene expression states (Hasty et al., 2000; To & Maheshri, 2010). In Drosophila segmentation, this role of noise was shown to be dependent on transcriptional and translational dynamics of target gene expression (Holloway et al., 2011).

Different roles of stochastic noise have been investigated during hindbrain boundary sharpening in zebrafish. When modeled, simulations of hindbrain segmentation with noise present solely in the RA morphogen gradient are only capable of patterning rough boundaries between segments in the hindbrain that never sharpen completely. So how do boundary cells interpret this noisy signal and form sharp boundaries? To address this, the mechanism of "noise-induced switching" was developed for sharpening the r4/5 boundary in the hindbrain (Zhang et al., 2012). In the hindbrain, *hoxb1a* (r4) and *krox20* (r5), both activated by RA initially, form a bistable switch due their self-activation and mutual inhibition of each other (Alexander et al., 2009; Barrow et al., 2000; Giudicelli et al., 2001). As a result, noise in gene expression of both *hoxb1a* and *krox20* expression was found to drive sharpening by inducing switching between either gene under different levels of RA. Direct evidence for this noise-induced switching mechanism was found by tracking krox20+ cells that had intermingled into the adjacent segment and switched cell identity (Addison et al., 2018). This reveals a positive role of gene expression noise in promoting robustness of boundary sharpening.

This brings into question what other roles gene expression noise has in tissue patterning and how are they regulated? While this topic is quite intriguing, current methods to address this question are still limited. Existing methods to examine gene expression noise such as *in situ* hybridization are largely nonquantitative and cannot be used in live samples. New developments in quantitative *in situ* hybridization methods based on the hybridization chain reaction (HCR) have allowed examination of gene expression noise spatially in fixed tissue (Choi et al., 2018), but these methods still lack the capability to examine noise temporally in live organisms. Labeling systems based on MS2 RNA stem loops address this issue and enable quantitative and dynamic imaging of gene expression in live organisms.

The MS2 RNA-labeling system functions by binary expression of a fluorescent MS2 bacteriophage coat protein (MCP) along with an RNA of interest tagged with multiple RNA hairpin MS2-binding sites (MBS) to which the MCP fusion protein binds (Hocine et al., 2013).

While the MS2 system has been used to image dynamics of one segmental stripe of *even-skipped* in the early *Drosophila* embryo (Bothma et al., 2014), this system has yet to be applied to study live transcript dynamics in vertebrate pattern formation. Recent technological developments have established a framework for MS2 imaging in the zebrafish system, though improvements are still required to study transcript dynamics in the context of noise and tissue patterning (Campbell et al., 2015). In Chapter III, I improve upon the MS2 RNA-labeling system in zebrafish, which should enable future studies into the complexities of noise inherent in gene transcription.

Cellular retinoic acid binding proteins

While there may be beneficial roles of noise, biological systems still seek ways to keep noise within ideal ranges. This is important for morphogen levels such as RA, where levels can be affected greatly by intrinsic variations in production rates and external factors like diet. One way that cells can promote robust responses to a range of morphogen levels is through non-receptor binding proteins that bind the morphogen and transport it in a more controlled manner compared to free diffusion. When morphogens are first bound to non-receptor binding proteins and then transported to enzymes for degradation or downstream signaling receptors that regulate transcriptional responses, the signaling gradient or actual cell response to a morphogen is based on availability and levels of non-receptors rather than morphogen levels alone (Lei et al., 2013). This makes cellular responses more robust to perturbations in levels of morphogen production.

Cellular retinoic acid binding proteins (Crabps) are ideal candidates for this role in regulation of RA. RA is a relatively insoluble lipophilic molecule and thus intracellular transport is mediated primarily through Crabps. Vertebrates have two classes of highly conserved Crabps, CRABP1 and CRABP2, whom are responsible for transporting RA to CYPs for degradation in the cytosol and to RARs in the nucleus (Aström et al., 1991; Budhu et al., 2001; Budhu & Noy,

2002; Delva et al., 1999). Crabps have also been found to directly influence RA metabolism. When RA is scarce and unbound Crabps are abundant, unbound Crabps prevent degradation of RA by directly binding to CYP26A1 (Yabut & Isoherranen, 2022). Despite shared functions and similar structures, CRABP1 and CRABP2 have distinct functions and expression patterns.

CRABP1's primarily roles involve regulating intracellular RA levels and transporting RA to CYP degradation enzymes for catabolism (Fiorella & Napoli, 1991, 1994). CRABP1 has been shown to have a high affinity for free RA and capable of changing the levels of free intracellular RA (Chen et al., 2003). However, in contrast to CRABP2, CRABP1 is unlikely to deliver RA to RARs. This is supported by evidence showing an inability of CRABP1 to enhance RAR interactions and promote RA-induced transcription (Dong et al., 1999; Venepally et al., 1996). In addition, non-canonical roles of CRABP1 have been described in regulating the cell cycle. When CRABP1 binds RA, it can translocate into the nucleus activating ERK1/2 independently of RARs. The CRABP1-RA complex then induces and stabilizes p27 inhibiting G1-S cell cycle progression in embryonic and neural stem cells (Lin et al., 2017; Persaud et al., 2013).

On the other hand, CRABP2 has been shown to both deliver RA to CYP enzymes for degradation as well as to RARs in the nucleus (Bastie et al., 2001; Delva et al., 1999). This dual function makes it an ideal candidate for buffering intracellular RA responses. In contrast to CRABP1 expression whose expression is quite variable across tissues with respect to RA, CRABP2 tends to be expressed in tissues of active RA biosynthesis. In cancer cells, CRABP2 was found to enhance responses to RA when overexpressed and decrease them when absent (Budhu & Noy, 2002). In zebrafish, Crabp2a was found to be uniquely RA inducible and essential for hindbrain patterning and robustness in the face of exogenous RA (Cai et al., 2012). Also in the hindbrain, Crabp2a was shown to attenuate noise in intracellular RA levels in cells along the A-P axis (Sosnik et al., 2016).

Taken as a whole, these studies tend to suggest that CRABP2 enhances and buffers while CRABP1s limits cellular responses to RA, though their actual regulation of RA responses

is likely more complex in part due to their involvement in different non-canonical activities of RA (Napoli, 2017). Additionally, prior studies are limited as most were performed in *in vitro* systems or utilized morpholinos for temporary knockdown of transcripts early in development. Early mouse knockout models for Crabps were found to display only mild digit defects but are otherwise normal (Gorry et al., 1994; Lampron et al., 1995). While this phenotype is not dramatic, it is still suggestive of robustness defects in RA- patterned tissues since RA is important for the patterning of posterior digits (Tickle et al., 1985; Niederreither et al., 2002). The individual roles of Crabps *in vivo* in different tissues are still not well understood, especially in the context of developmental robustness and during stages past embryonic development. In Chapter IV, I will elaborate on novel and discrete roles of Crabp1 and Crabp2 in the context of germ cell development and sex determination.

Overall, in my thesis, I seek to investigate the multitude of mechanisms by which organisms generate robust responses to RA in the face of various sources of perturbation whether from noise inherent in biological processes or externally from changes in environment.

Chapter II

Multiple morphogens and rapid elongation promote segmental patterning during development

Yuchi Qiu, **Lianna Fung**, Thomas F. Schilling, Qing Nie

Published in *PLOS Computational Biology* (2021) *17*(6), e1009077

Abstract

The vertebrate hindbrain is segmented into rhombomeres (r) initially defined by distinct domains of gene expression. Previous studies have shown that noise-induced gene regulation and cell sorting are critical for the sharpening of rhombomere boundaries, which start out rough in the forming neural plate (NP) and sharpen over time. However, the mechanisms controlling simultaneous formation of multiple rhombomeres and accuracy in their sizes are unclear. We have developed a stochastic multiscale cell-based model that explicitly incorporates dynamic morphogenetic changes (i.e. convergent-extension of the NP), multiple morphogens, and gene regulatory networks to investigate the formation of rhombomeres and their corresponding boundaries in the zebrafish hindbrain. During pattern initiation, the short-range signal, fibroblast growth factor (FGF), works together with the longer-range morphogen, retinoic acid (RA), to specify all of these boundaries and maintain accurately sized segments with sharp boundaries. At later stages of patterning, we show a nonlinear change in the shape of rhombomeres with rapid left-right narrowing of the NP followed by slower dynamics. Rapid initial convergence improves boundary sharpness and segment size by regulating cell sorting and cell fate both independently and coordinately. Overall, multiple morphogens and tissue dynamics synergize to regulate the sizes and boundaries of multiple segments during development.

Author Summary

In segmental pattern formation, chemical gradients control gene expression in a concentration-dependent manner to specify distinct gene expression domains. Despite the stochasticity inherent to such biological processes, precise and accurate borders form between segmental gene expression domains. Previous work has revealed synergy between gene regulation and cell sorting in sharpening borders that are initially rough. However, it is still poorly understood how size and boundary sharpness of *multiple* segments are regulated in a tissue that changes dramatically in its morphology as the embryo develops. Here we develop a stochastic multiscale cell-base model to investigate these questions. Two novel strategies synergize to promote accurate segment formation, a combination of long- and short-range morphogens plus rapid tissue convergence, with one responsible for pattern initiation and the other enabling pattern refinement.

Introduction

A fundamental question in developmental biology is how cell fate decisions are coordinated with tissue morphogenetic changes during pattern formation. During embryogenesis, cells must convert concentration-dependent positional information from diffusible chemical morphogens into coordinated cell fate decisions (Lander, 2011; Rogers & Schier, 2011; Towers & Tickle, 2009). Mathematical models have successfully integrated tissue morphogenesis and spatial signaling during patterning of embryonic segments in both flies (Jaeger, 2009; Jaeger et al., 2004) and vertebrates (Baker et al., 2008; Hubaud & Pourquié, 2014; Yoshioka-Kobayashi et al., 2020), in structures such as the wing imaginal discs in *Drosophila* (Ben-Zvi et al., 2011; Fried & Iber, 2014; Zhu et al., 2020), as well as the limb buds (Raspopovic et al., 2014), neural tube (Balaskas et al., 2012; Cohen et al., 2014; Zagorski et al., 2017), hindbrain (Bouchoucha et al., 2013; Wang et al., 2017; Zhang et al., 2012), pharyngeal arches (Meinecke et al., 2018), skin (Chou et al., 2010; Du et al., 2018) and hair follicles (Wang, et al., 2017) of vertebrates.

Understanding stochastic effects in patterning systems, particularly how precision is achieved in spite of biological noise in gene expression and spatial signals, is a major challenge in developmental biology. Noise attenuation mechanisms in gene expression have been widely explored in diverse cellular networks (Nie et al., 2020; Qiao et al., 2019). For spatial signals, binding with membrane-bound non-signaling entities (Lander et al., 2007), regulation of gradient steepness by ligand shuttling (Lander et al., 2009; Shilo et al., 2013) and self-regulated ligand uptake (Eldar et al., 2003; White et al., 2007) can reduce spatial variation in morphogen gradients. Anti-parallel morphogens (Zagorski et al., 2017) and gene regulatory networks (Exelby et al., 2021; Rackauckas et al., 2018; Sokolowski et al., 2012) that translate noisy spatial signals into cell fate decisions can also reduce patterning errors. Interestingly, noise in gene expression can counteract other stochastic effects (e.g. noise in morphogen levels) to improve pattern formation precision (Qiu et al., 2019; Zhang et al., 2012). In addition to these

molecular strategies, pattern precision can be improved through cellular strategies, such as cell sorting driven by cell-cell interactions (Du et al., 2018; Wang et al., 2017) or “community effects” of signals from adjacent cells (Schilling et al., 2001). Previous modeling studies have often neglected to take into account rapid changes in tissue morphology, and how the interaction between these and noise attenuation mechanisms impacts pattern precision remains poorly understood.

The embryonic zebrafish hindbrain is a powerful model system to study the roles of gene regulation, stochasticity, cell sorting, and tissue morphogenesis in segmental pattern formation. Neurons in the hindbrain contribute to the cranial nerves that innervate the face and neck and control many involuntary functions, such as feeding and breathing. These neurons arise in early embryonic segments, called rhombomeres (r), that progressively subdivide along the anterior-posterior (A-P) axis (Lumsden & Krumlauf, 1996). Initial gene expression domains that specify segmental cell identities in rhombomeres 1-7 (r1-7) have rough borders that subsequently sharpen (Wilkinson, 2018; Zhang et al., 2012). Several spatial signals provide positional information for the establishment of rhombomeres, such as retinoic acid (RA) (Maves & Kimmel, 2005; Rhinn & Dollé, 2012; Schilling et al., 2012; Sosnik et al., 2016; White et al., 2007) and fibroblast growth factors (FGFs) (Choe et al., 2011; Labalette et al., 2015; Maves et al., 2002; Walshe et al., 2002; Wiellette & Sive, 2003). These signals regulate numerous transcription factors, including *hox* genes, *krox20*, *val*, *vhnf1* and *irx*, with rhombomere-specific expression domains that specify rhombomere cell identity (Alexander et al., 2009; Lecaudey et al., 2004; Parker & Krumlauf, 2017). Rhombomere-specific gene regulatory networks commit cells to distinct segmental fates and can switch their identities from one segment to another by interpreting RA and FGF signals (Addison et al., 2018; Schilling et al., 2001). In addition, the complementary segmental expression of Ephrins and Eph receptors drives boundary sharpening by regulating cell sorting with differential adhesion/repulsion (Cooke et al., 2005; Hernandez et al., 2004; Pitulescu & Adams, 2010; Xu et al., 1999). Previous computational

models incorporating one morphogen, RA, and two transcription factors, *hoxb1a* and *krox20*, successfully mimic boundary sharpening in r3-5 by incorporating gene regulation and cell sorting (Wang et al., 2017; Zhang et al., 2012). During the period when these boundaries sharpen, the hindbrain grows and elongates, particularly, the hindbrain narrows along the left-right (L-R) axis and elongates along the anterior-posterior (A-P) axis, which is often termed convergent-extension (Addison et al., 2018; Warga & Kimmel, 1990). However, previous computational models have shown that tissue elongation disrupts the sharpening of rhombomere boundaries (Zhang et al., 2012). It remains unclear in any segmented tissue how multiple segments simultaneously form with accurate sizes and sharp boundaries during such morphogenetic tissue dynamics.

Here we consider hindbrain patterning across multiple stages, from pattern initiation to sharpening, across multiple segments (r2-6) and in the context of morphogenetic changes in hindbrain size and shape. We include a second morphogen in our model, FGF produced in r4, and two additional transcription factors, *vhnf1* and *irx3*. We find that FGF produced in r4 is critical to specify the r5/r6 boundary, and to achieve a robust five-segment pattern with accurate segment sizes and sharp boundaries despite variations in initial gene expression. At later stages of patterning, we show experimentally that L-R narrowing of the zebrafish hindbrain occurs rapidly at first (11-12 hours post fertilization (hpf)), but the narrowing rate drops rapidly over the following two hours (12-14 hpf). Interestingly, comparisons of hindbrain pattern formation in our model under different convergence rates suggest that such a rapid initial convergence facilitates robust patterning, both in the accuracy of segment size and boundary sharpness. This rapid initial convergence helps mediate a trade-off between boundary sharpness and segment size. Together, the cooperation between two morphogens and morphogenetic dynamics effectively regulates robust segmental patterning in the zebrafish hindbrain.

Results

A stochastic multiscale cell-based model for hindbrain segmentation

To address how multiple morphogens and dynamics of tissue morphogenesis contribute to segmental pattern formation in the hindbrain, we developed a computational model that incorporates stochastic gene regulation, cell sorting and tissue shape changes (Fig 1). We first provide an overview of the elements, assumptions, and metrics included in our models (details see Methods and S1 Text).

FIGURE 1

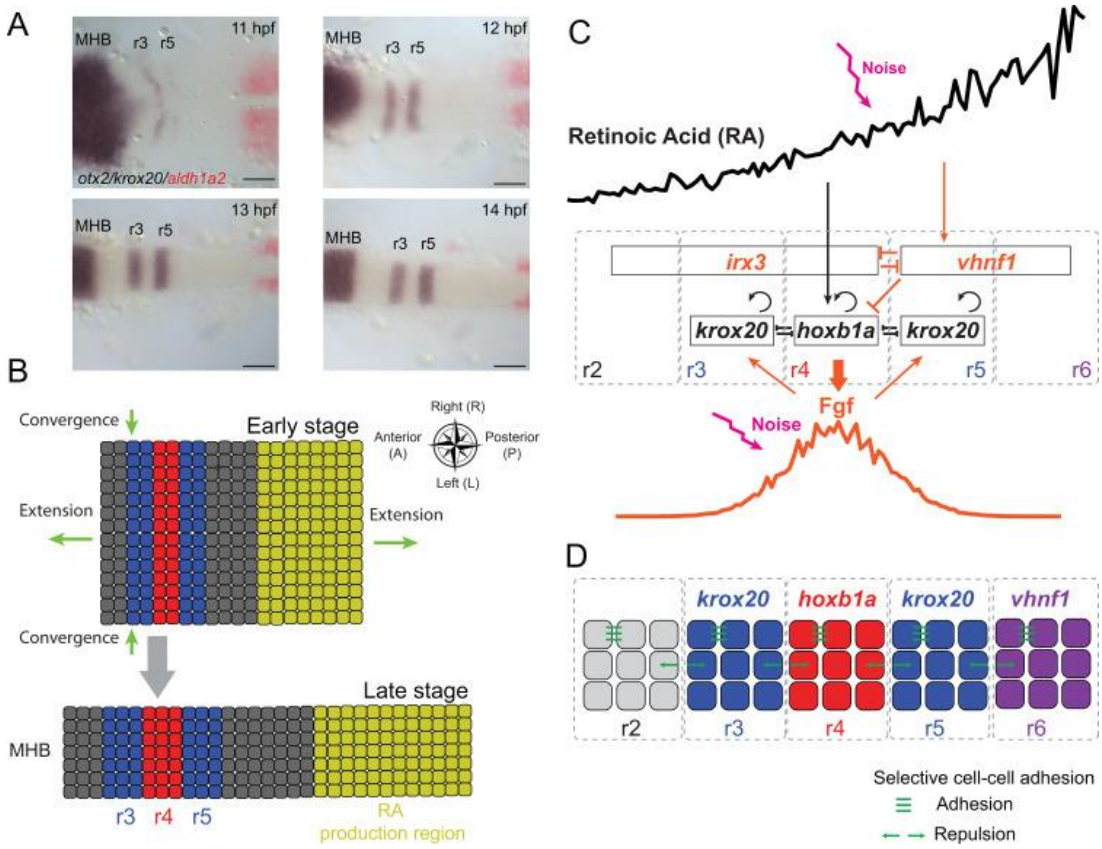

Figure 1. Model schematic and zebrafish hindbrain morphology. (A) Two-color whole mount in situ hybridization of embryonic zebrafish hindbrains for *otx2* (purple, anterior region, far left), *krox20* (purple segments, center) and *aldh1a2* (red, far right) transcripts from 11 to 14 hpf. *otx2* marks the midbrain-hindbrain boundary (MHB), *krox20* marks r3 and r5 and *aldh1a2* marks the RA production region. Embryos are flat-mounted and shown in dorsal view with anterior to the left. Scale bars: 100 μm. (B) Illustration depicting convergent-extension of the hindbrain. The entire rectangular region, including r1-7 and the RA production region, constitutes the morphogen domain. The hindbrain narrows in the L-R direction (width) and elongates in the A-P direction. (C) Gene regulatory network used to model hindbrain patterning in r2-6. Genes and morphogens with black font were previously used for modeling the r3-r5 pattern (Zhang et al., 2012), while additional factors considered in this model are shown in orange font. Pointed arrows depict up-regulation/activation and blunt arrows depict down-regulation/inhibition. Two morphogens, retinoic acid (RA) and Fibroblast Growth Factor (FGF) diffuse and form two distinct gradients to govern downstream gene expression. (D) Illustration depicting r2-6 and distinct identities (i.e. gene expression signatures) underlying selective cell sorting. Cells in r3 and r5 (blue) express *krox20* and cells in r4 (red) express *hoxb1a*, while both *krox20* and *hoxb1a* levels are low in r2 and r6. Cells in r6 (purple) have high *vhnf1* expression. Cells with the same segmental identity attract each other and cells with different identities repulse each other.

Hindbrain morphogenesis and computational domains

During embryogenesis from 11-14 hpf, the zebrafish hindbrain narrows along the L-R axis and elongates along the A-P axis (Addison et al., 2018; Warga & Kimmel, 1990). The midbrain-hindbrain boundary (MHB) (anterior to the hindbrain and adjacent to r1) and the RA production region (posterior to the hindbrain and adjacent to r7) provide A-P landmarks for the region that forms the hindbrain (Fig 1A). To quantify these changes in size and shape, we performed whole mount in situ hybridization with markers for the MHB (*otx2*), r3 and r5 (*krox20*) and RA production (*aldh1a2*). At each hour between 11-14 hpf, we measured hindbrain width at the level of r4 as well as the A-P distance between the *otx2* and *aldh1a2* domains (Fig 2A-B). Based on these experimental measurements, we modeled the hindbrain (r1-r7) along with the RA production region as a rectangle, subsequently referred to as the morphogen domain (Fig 1B). Two morphogens, RA and FGF, were modeled in the morphogen domain. Due to the expensive computational cost, instead of modeling all cells in the morphogen domain for this study, we explicitly modeled cells in a smaller region containing r2-r6, subsequently referred to as the tissue domain.

FIGURE 2

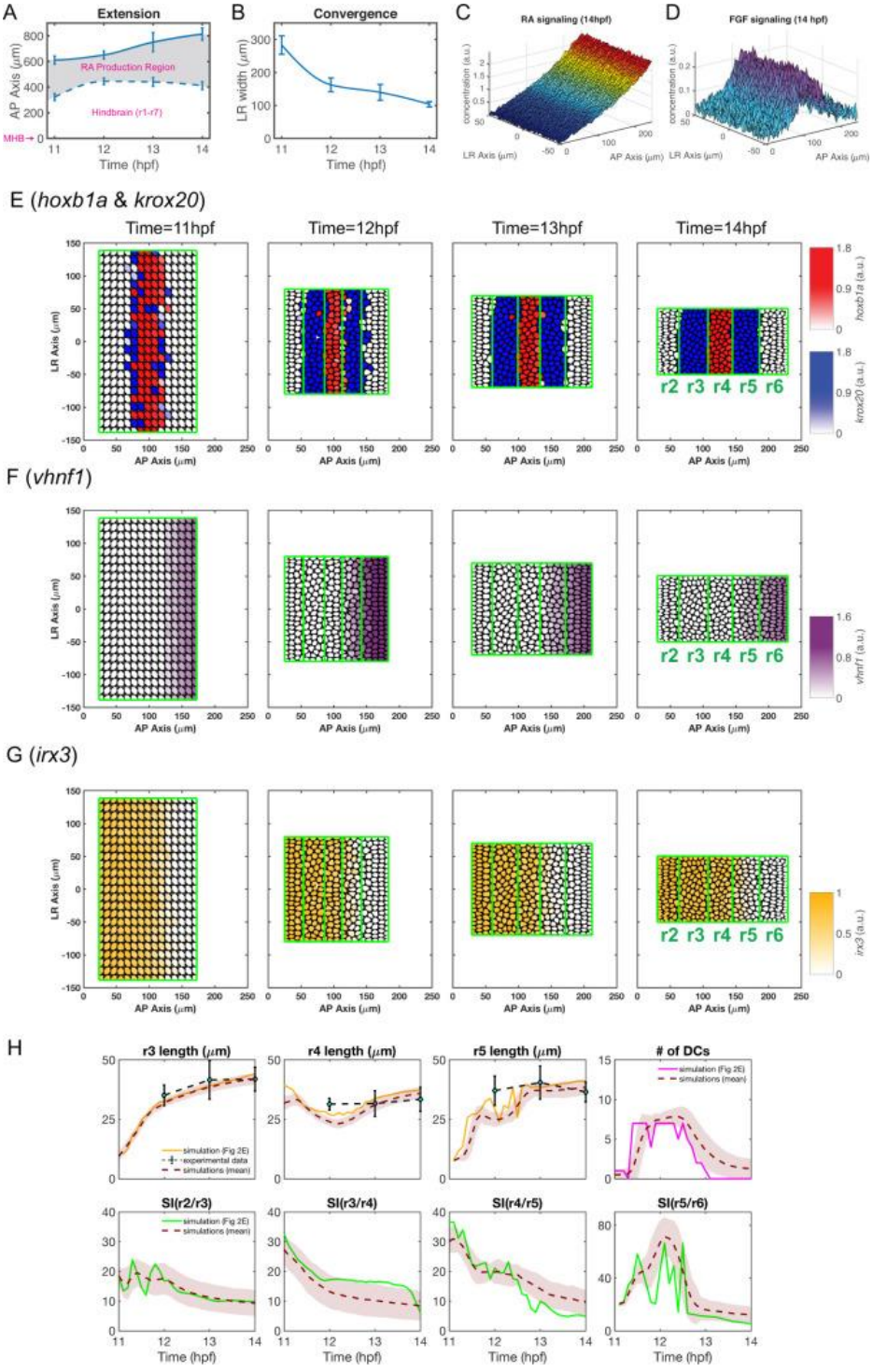
A
Extension
AP Axis (μm)
RA Production Region
Hindbrain (r1-r7)
MHB
Time (hpf)
B
Convergence
LR width (μm)
Time (hpf)
C
RA signaling (14hpf)
concentration (a.u.)
LR Axis (μm)
AP Axis (μm)
D
FGF signaling (14 hpf)
concentration (a.u.)
LR Axis (μm)
AP Axis (μm)
E (hoxb1a & krox20)
Time=11hpf
Time=12hpf
Time=13hpf
Time=14hpf
LR Axis (μm)
AP Axis (μm)
r2 r3 r4 r5 r6
hoxb1a (a.u.)
krox20 (a.u.)
F (vhnf1)
vhnf1 (a.u.)
G (irx3)
irx3 (a.u.)
H
r3 length (μm)
r4 length (μm)
r5 length (μm)
# of DCs
simulation (Fig 2E)
experimental data
simulations (mean)
SI(r2/r3)
SI(r3/r4)
SI(r4/r5)
SI(r5/r6)
Time (hpf)

Figure 2. A baseline simulation mimics rhombomere boundary sharpening. (A,B) Experimental measurements of hindbrain dimensions along the A-P (A) and L-R (B) axes at 11, 12, 13 and 14 hpf. Error bars represent standard deviation. Cubic interpolation is used to obtain the smooth curves used in the model. (A) A-P hindbrain length was measured from the posterior edge of the mid-hindbrain boundary (MHB) to the anterior edge of the RA production region. A-P length of the RA production region was based on measurements of the aldh1a2 expression domain. (B) L-R hindbrain width was measured at the A-P position of r4. (C,D) Predicted noisy distributions of morphogen signaling at 14 hpf (C) RA ([RA]in). (D) FGF ([FGF]signal). (E-G) Time series of gene expression in r2-6 (the hindbrain is represented as a rectangle for simplification): (E) hoxb1a (red) and krox20 (blue), (F) vhnf1 (purple), (G) irx3 (yellow). (H) Quantifications of rhombomere length, number of dislocated cells (DCs) and sharpness indices (SIs) versus time. Rhombomere lengths (r3-5), and SIs for four boundaries (SI(r2/r3), SI(r3/r4), SI(r4/r5), SI(r5/r6)) and DC number in multiple simulations (n = 100): 'solid line': quantities for the simulation shown in (E); 'brown dashed line' indicates the average and the width of 'brown shade" indicates standard deviation; 'black dashed line' represents rhombomere lengths from experimental measurements and the error bars represent standard deviation.

Morphogens, gene regulation and cell fate

Gene expression in the zebrafish hindbrain initially forms a rough r2-6 pattern at 11 hpf, which is refined over time into five precise segments of similar size with four sharp boundaries by 14 hpf (Fig 1C-D). RA synthesized in somites adjacent to the anterior spinal cord diffuses anteriorly and is required for proper rhombomere formation, including direct activation of *vhnf1* in r5 and r6 and *hoxb1a* in r4 (Maves & Kimmel, 2005; Rhinn & Dollé, 2012; Schilling et al., 2012). Mutual inhibition between *vhnf1* and *irx3* specifies the first pre-rhombomeric r4/r5 boundary at 9.5-10 hpf (Lecaudey et al., 2004). RA then activates *hoxb1a* and *vhnf1* which represses *hoxb1a* expression (Maves & Kimmel, 2005), restricting it to r4 where it activates FGF synthesis (Choe et al., 2011; Maves et al., 2002; Waskiewicz et al., 2002). FGF diffuses both anteriorly and posteriorly to induce *krox20* in r3 and r5 (Labalette et al., 2015; Walshe et al., 2002; Wiellette & Sive, 2003). Through auto-regulation, *krox20* has two steady-state expression levels, either zero or non-zero, depending on the FGF concentration (Bouchoucha et al., 2013; Labalette et al., 2015), and the r2/r3 and r5/r6 boundaries are specified by *krox20* levels. Auto-regulation and mutual inhibition between *hoxb1a* and *krox20* establish a toggle switch that

specifies and refines the r3/r4 and r4/r5 boundaries (Zhang et al., 2012). As a result, three distinct cell fates are specified by *hoxb1a* and *krox20* expression levels to establish the r2-r6 pattern, specifically, high *hoxb1a* and low *krox20* expression in r4, low *hoxb1a* and high *krox20* expression in r3 and r5, and low expression of both *hoxb1a* and *krox20* in r2 and r6 (Fig 1D).

In the model, morphogens (RA and FGF) were described by stochastic PDEs in a continuum fashion. Regulation of genes downstream of morphogens was modeled using stochastic ODEs for each individual cell. Interactions between morphogens were modeled at a regular rectangular mesh in the morphogen domain, and the downstream genes for each cell were modeled as being located at the center of each individual cell. Numerical interpolation was used to capture the interplay between morphogens and gene regulation modeled at different grid points (Section 1 in S1 Text).

Mechanical models for individual cells

In the cell mechanical model, we used the subcellular element method (SCEM) (Newman et al., 2017) to model individual cells and cell-cell mechanical interactions involved in cell sorting (Section 2 in S1 Text). In this computational formalism, an individual cell consists of a constant number of sub-cellular elements (i.e. nodes). Elements interact according to a prescribed force potential. This force between elements within the same cell is repulsive at short ranges and attractive at long ranges to maintain stable cell volume and circular structure. The forces between elements from different cells are repulsive at short ranges to prevent cell overlaps. At longer ranges, the intercellular forces between elements can be either repulsive or attractive depending on cell identities.

Cell sorting

In the zebrafish hindbrain, cell sorting has selectivity based on cell identities. One well-known mechanism of selectivity is cell-cell adhesion mediated by Ephrin-Eph signaling. Ephrin-

B2 ligands are expressed highly in even-numbered rhombomeres (r2, r4 and r6), while EphA4 receptors are expressed highly in odd-numbered rhombomeres (r3 and r5), and this alternating pattern controls repulsion between cells in one rhombomere and another (Hernandez et al., 2004). Depletion of EphA4 has more dramatic effects on rhombomere boundaries than EphrinB2a, but knockdown of both enhances boundary defects (Cooke et al., 2001). Krox20 directly activates transcription of *ephA4* (Theil et al., 1998), thereby regulating Ephrin-Eph mediated cell sorting. Our model mimics the selective cell-cell adhesion between two cells based on their *krox20* expression levels (Fig 1D). Specifically, cells attract each other if they have similar *krox20* levels, and repel each other if one expresses *krox20* and the other does not.

Initial conditions

We chose 11 hpf as the starting point for our modeling study soon after the initiation of gene expression in some rhombomeres. At this stage in the simulations, cells were assumed to align uniformly in the rectangular tissue domain. For initial gene expression, we first ran simulations with the stochastic gene regulation model over two simulated hours prior to 11 hpf and used those simulation results as the initial condition at 11hpf. We started with equilibrium solution for RA by solving the corresponding steady-state PDE because RA gradients are established as early as 6 hpf (Maves & Kimmel, 2005). Because *vhnf1* and *irx3* expression appear much earlier than *hoxb1a*, *krox20* and FGF (Maves & Kimmel, 2005), we ran stochastic simulations for RA, *vhnf1* and *irx3* in the first hour with zero values for *vhnf1* and *irx3*. In the second hour, we included stochastic simulations for all morphogens and genes. *krox20* and FGF start with zero expression while *hoxb1a* starts with a constant expression level because *hoxb1a* has weak expression over the hindbrain domain at early stages (Zhang et al., 2012) and shows dynamic changes in expression earlier than FGF or *krox20* (Lecaudey et al., 2004).

One-dimensional gene expression model

In the one-dimensional gene expression models, we consider steady-state solutions of genes and morphogens in a one-dimensional fixed A-P domain (Section 3 in S1 Text). The initial conditions were chosen similarly to the full model. Equilibrium solutions for RA, *vhnf1* and *irx3* were taken as the initial conditions. Both *krox20* and FGF levels were initially set at zero. We included *hoxb1a* expression at low but non-zero initial levels.

A cell sorting-only model

In the model with only cell sorting without gene regulation, cells sorted based on their pre-assigned identities, and cell identities and numbers did not change throughout the simulations. The initial "salt-and-pepper" pattern of cell identities was sampled by a mixture Gaussian probability distribution based on each cell's A-P position (Section 4 in S1 Text).

Quantification of cell fate, rhombomere boundary A-P locations and boundary sharpness

Three cell identities based on threshold levels of high or low expression of *hoxb1a* and *krox20* in the five segments (r2-6) were considered. Once these three cell identities were determined in r2-r6 we evaluated rhombomere boundaries (Section 5 in S1 Text). We defined three critical quantities in our simulations: boundary location along the A-P axis, boundary sharpness (represented as a sharpness index, SI) and the number of dislocated cells (DCs) (see Methods). Four boundaries between rhombomeres (r2/r3, r3/r4, r4/r5 and r5/r6) are all perpendicular to the A-P axis. For a single boundary determination, we selected a predefined boundary and calculated total deviations of cells located on the wrong side of this predefined boundary. The total deviations were minimized over the A-P position of the predefined boundary. SI was defined by the minima of total deviations and the A-P position of this boundary was defined based on these minima. A lower SI indicates a sharper boundary. DCs are cells located over three cell-diameters away from the rhombomere to which they belong and they are

excluded in calculating boundary location. If the number of DCs exceeds 8 cells, we consider the pattern failed.

Including multiple morphogens and tissue morphogenesis in the models recapitulates the r2-6 pattern

The stochastic multiscale cell-based model successfully recapitulated the dynamics of r2-6 formation observed in the zebrafish hindbrain. As shown for one stochastic simulation with spatial distributions of multiple genes, both RA and FGF signals have noisy distributions over the space (Fig 2C-D). The RA gradient decreases from its origins at the posterior end of the hindbrain to the anterior, while FGF levels are high in r4 where it is secreted and decreases in both anterior and posterior directions. By generating a time series of the spatial patterns of gene expression (Fig 2E-G), including *hoxb1a*, *krox20, vhnf1 and irx3*, our model recapitulates rhombomere boundary sharpening (Addison et al., 2018; Zhang et al., 2012). For example, *krox20* is weakly expressed in r3 and r5 at 11 hpf and upregulated by 12 hpf, with expression stronger in r3 than r5 (Fig 1A, 2E). At this stage, *hoxb1a*+ and *krox20*+ cells intermingle and a few cells close to the r4/r5 boundary undergo identity switching as they co-express low levels of both *krox20* and *hoxb1a* (Addison et al., 2018). By 13 hpf, cells closer to the boundaries and most of the *krox20*/*hoxb1a* co-expressing cells commit to one segment or another and boundaries become sharper. At 14 hpf, all cells segregate to their territories and the boundaries fully sharpen, producing a precise five-segment pattern.

The modeling output naturally accounts for the sharpening of other gene expression boundaries in r2-6, despite differences in their interactions. For example, the anterior edge of *vhnf1* expression and the posterior edge of *irx3* expression specify the position of the pre-rhombomeric r4/r5 boundary at 11 hpf (Fig 2F-G). At later stages, this *vhnf1*/*irx3* border shifts posteriorly to become located in r5 (Lecaudey et al., 2004). Unlike *hoxb1a* and *krox20*, *vhnf1* does not auto-induce itself to maintain its own expression without RA signals. Consequently,

vhnf1 shifts posteriorly (Maves & Kimmel, 2005) after 12 hpf as RA decreases everywhere (Fig S1C).

Stochastic simulations were repeated independently (Fig 2H). The results are consistent with experimental measurements of rhombomere A-P length at 14 hpf (r3 = 42±5 μm, r4 = 34±5 μm, r5 = 37±4 μm) as well as sharpening. For example, from 11-12 hpf, identity switching affects the sharpness of the r4/r5 boundary (Zhang et al., 2012), with a higher SI and DC number during this period. From 12-14 hpf, SI and DC number gradually decrease to the minimum as all boundaries sharpen.

Cooperation between RA and FGF improves robustness of initiation of the segmental pattern

A previous model that only considered the RA morphogen gradient without cell sorting or convergent extension successfully simulated many aspects of the formation of the r2-5 pattern (Zhang et al., 2012). In the model, *hoxb1a* and *krox20* were considered as direct downstream targets of RA, despite *krox20* being indirectly induced by RA through *hoxb1a* and FGF. Our new model incorporates both RA and FGF as well as these additional features of *krox20* regulation.

Comparing the two-morphogen (RA + FGF) to the one-morphogen (RA) model reveals many similarities and some key differences (Fig 3). In both, the borders between *hoxb1a* and *krox20* specify r3/r4 and r4/r5 boundaries, the border between *vhnf1* and *irx3* lies posterior to the r4/r5 boundary, and *krox20* has two steady-state levels induced by either RA (one-morphogen model) or FGF (two-morphogen model) depending on the morphogen levels. In the two-morphogen model, FGF has the highest expression in r4 where it is secreted and decreases in both anterior and posterior directions. By inducing *krox20* in a concentration-dependent manner, FGF can specify the r2/r3 and r5/r6 boundaries. In contrast, the one-morphogen model, with RA decreasing monotonically from posterior to anterior, can only specify r2/r3 and not the r5/r6 boundary. Overall, the two-morphogen model can specify four boundaries

(r2/r3, r3/r4, r4/r5 and r5/r6), while the one-morphogen model can specify only three boundaries (r2/r3, r3/r4 and r4/r5).

FIGURE 3

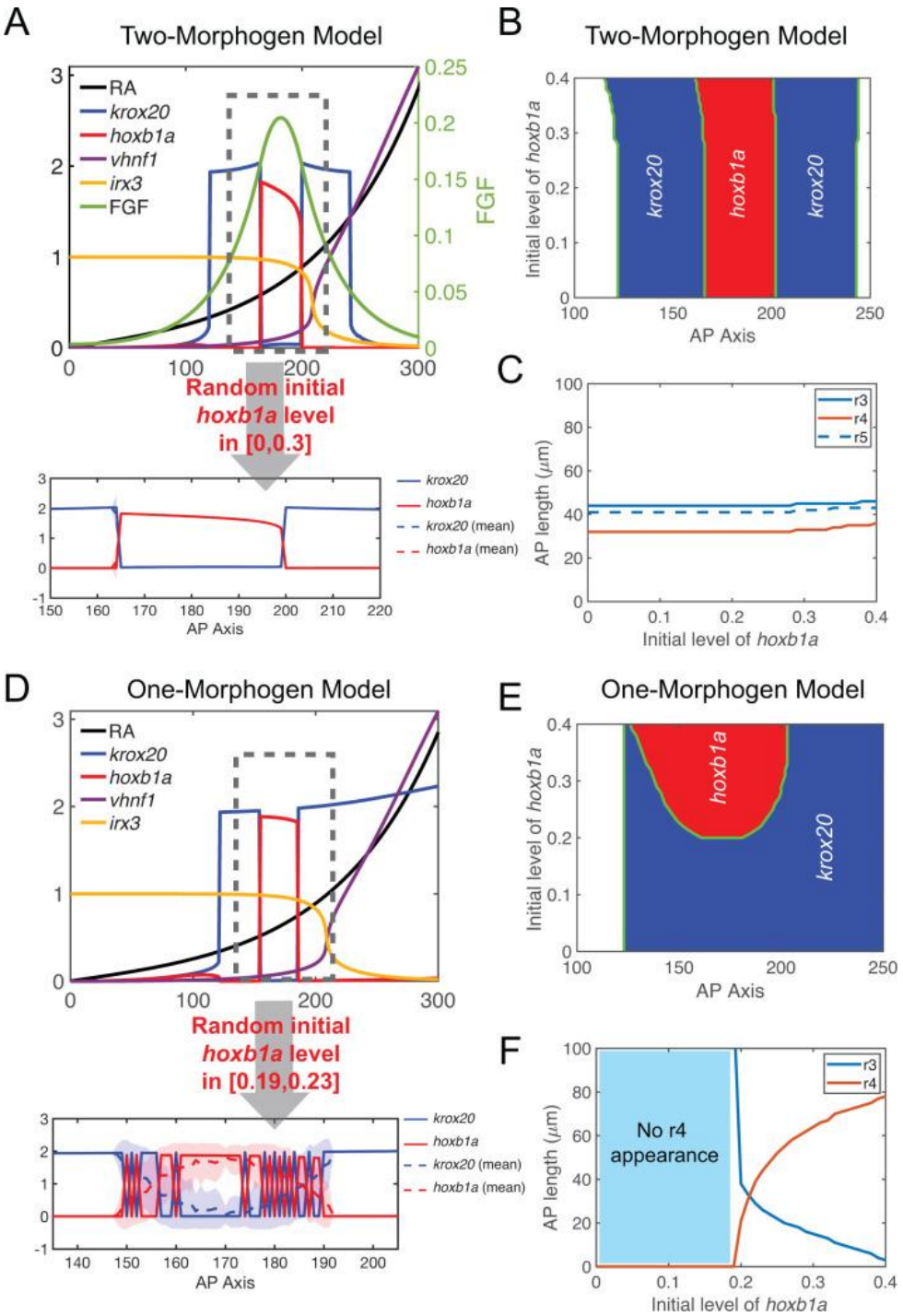

Figure 3. Comparing two-morphogen (RA, FGF) and one-morphogen (RA) models. (A-C) One-dimensional simulations for the two-morphogen model. (A) The upper panel shows spatial distributions of RA, *krox20*, *hoxb1a*, *vhnf1*, *irx3* and FGF. The initial *hoxb1a* level is modeled as a constant 0.21 over the space. In the lower panel, the initial *hoxb1a* level is randomly sampled over the space independent of the location. The value is randomly uniformly distributed at a level of [0,0.3]. Solid line represents one simulation. Dashed line represents average values and the width of the shading around each line represents the standard deviation (n = 100). Since fluctuations over multiple simulations are small, solid lines overlap with dashed lines and the small standard deviations result in shadings of negligible width around the dashed lines. X-axis, microns; Y-axis, arbitrary units. (B) Phase diagram of *hoxb1a* and *krox20* distributions with different initial *hoxb1a* levels. (C) Rhombomere lengths with different initial *hoxb1a* levels. (D-F) Similar one-dimensional simulations for the one-morphogen model. For (D), in the upper panel, the constant initial *hoxb1a* level is taken as 0.21; in the lower panel, the initial *hoxb1a* level is randomly sampled with levels in the range [0.19,0.23] with uniform distribution. Corresponding (E) phase diagram and (F) graph of rhombomere lengths with the one-morphogen model.

Additionally, we compared robustness in the two models with respect to initial *hoxb1a* expression. With constant initial *hoxb1a* level everywhere in space, we compared phase diagrams and resulting rhombomere lengths between the two-morphogen and one-morphogen models (Fig 3B, 3C, 3E, 3F). Interestingly, the inclusion of FGF makes the model relatively less sensitive to initial *hoxb1a* levels, in terms of the locations of gene expression boundaries and sizes of r3-5. When the initial *hoxb1a* level exceeds a certain level (e.g. 0.3), three rhombomeres (r3, r4 and r5) expand slightly along the A-P axis with r2/r3 and r3/r4 expanding anteriorly, and r4/r5 and r5/r6 expanding posteriorly (Fig 3B-C). In contrast, with RA alone, lengths of r3 and r4 are more sensitive to initial *hoxb1a* levels. In simulations with low initial *hoxb1a* levels (<0.2), r4 does not form (Fig 3E-F). As initial *hoxb1a* levels increase from 0.2 to 0.4, the r4 region rapidly expands at the expense of r3. A 15% increase in initial *hoxb1a* level (0.2-0.23) leads to an over two-fold expansion of r4 (21-44 μm) and r3 essentially vanishes when the initial *hoxb1a* level is close to 0.4. Thus, the two-morphogen model outperforms the one-morphogen model in robustness of rhombomere length, in that the second morphogen buffers responses to initial gene expression variation.

We also examined how the pattern reacts to perturbations of initial gene expression. We consider noisy initial *hoxb1a* levels over the space. In the two-morphogen model, such noise has negligible effects on later *hoxb1a* and *krox20* distributions resulting in clear segmental patterns and sharp r3/r4 and r4/r5 boundaries, and multiple simulations result in almost identical patterns (Fig 3A). However, in the one-morphogen model, *hoxb1a* and *krox20* distributions fluctuate more dramatically than the two-morphogen model. Despite a much smaller magnitude of perturbation (13%) in initial *hoxb1a* levels, the one-morphogen model shows fluctuating boundaries for both r3/r4 and r4/r5 (Fig 3D). Overall, the two-morphogen cooperation facilitates both accurate rhombomere lengths and sharp boundaries with perturbations of initial gene expression, providing robustness in patterning during the initial stages.

Rapid initial convergence improves boundary sharpness and segment size

Due to stochasticity, the initial pattern shows rough boundaries between rhombomeres. Later on, patterns sharpen and refine boundaries from 11-14 hpf. During these stages, the zebrafish hindbrain changes shape dramatically as it narrows in width along the L-R axis, extends in length along the A-P axis (Fig 1A) and thickens in the D-V axis (Papan & Campos-Ortega, 1994). We mainly studied A-P and L-R axes and their dimensions were experimentally measured (Fig 2A-B). Length along the A-P axis changes slowly during this period, but width along the L-R axis rapidly shrinks from 283 μm to 162 μm during the first hour, and further to 104 μm during the last two hours with approximately a 75% drop in the average rate after which patterning is largely complete. To determine how such rapid initial convergence influences hindbrain patterning we compared models incorporating medium or slow initial convergence with the rapid initial convergence rate we measured experimentally. All three types of convergence have the same initial and terminal L-R width. The curve of the medium convergence is taken as a linear function. The curves of the slow and rapid initial convergences are symmetric to the linear curve (Fig 4A).

FIGURE 4

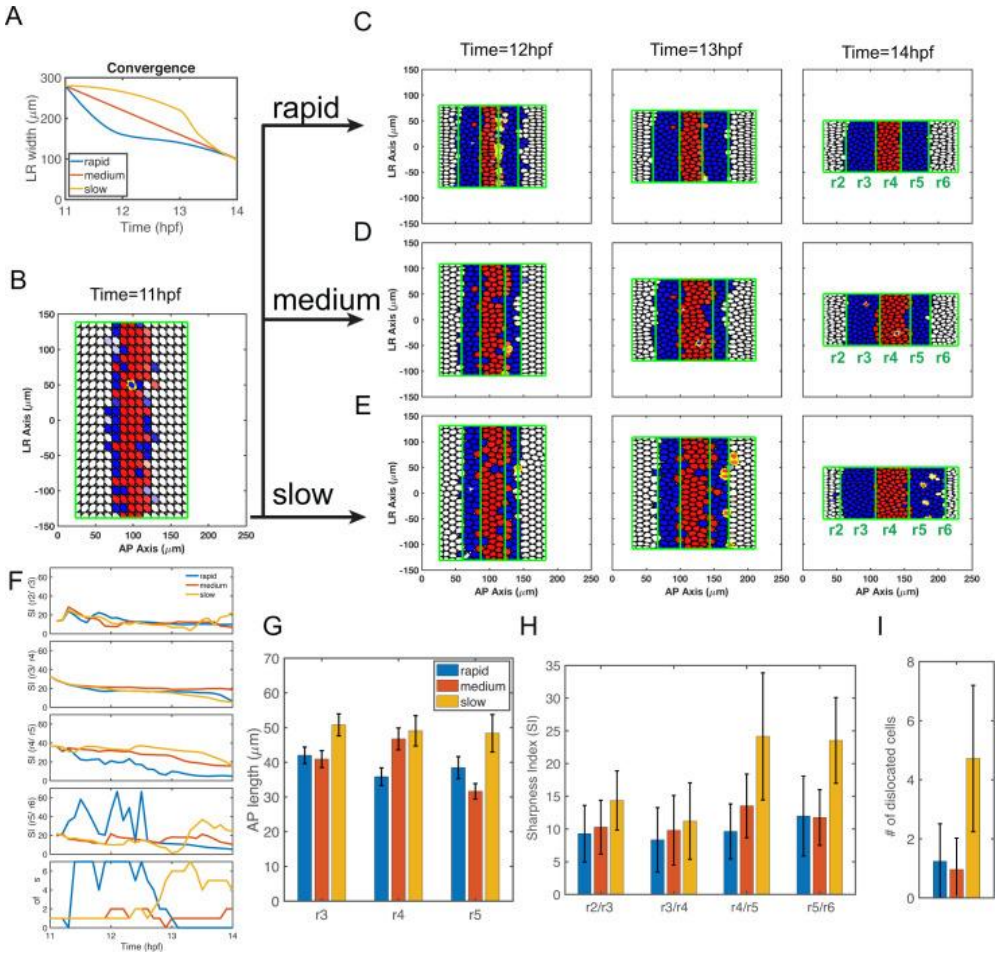

Figure 4. Simulations of full models combining gene regulation and cell sorting with different convergence rates. (A) Three convergence rates during the 11–14 hpf period are considered in the model, rapid (from experimental measurements, Fig 2B), medium and slow. All start and terminate with the same L-R width. The curve of medium convergence is depicted as a linear function. The curve of slow initial convergence is symmetric to the curve of rapid initial convergence with respect to the curve of linear function. (B-E) Time series of cell distributions with different convergence rates from 11–14 hpf. *hoxb1a* (red) and *krox20* (blue) expression. Dislocated cells (DCs) are highlighted by yellow edges. (B) Three simulations start with the same initial cell distribution (11 hpf) generated by the gene expression model (see Methods). Cell distributions with (C) rapid, (D) medium and (E) slow initial convergence rates from 12–14 hpf. (F) The boundary sharpness index (SI) for four boundaries (SI(r2/r3), SI(r3/r4), SI(r4/r5) and SI(r5/r6)) and DC number versus time. (G-I) Histograms depicting three convergence rates analyzed for (G) rhombomere lengths of r3-5, (H) SI and (I) DC number. Each represents 100 independent stochastic simulations for each convergence rate based on the same parameters. Error bars represent standard deviation.

We first performed simulations with full models including both gene regulation and cell sorting. Regardless of the convergence speed, most cells segregate to their correct territories and the final patterns display sharp boundaries between rhombomeres (Fig 4B-E). Rapid initial convergence allows the sharpest boundaries and the fewest DCs while slow initial convergence results in the roughest boundaries and more DCs (Fig 4F). SI and DC numbers at 14 hpf in multiple independent simulations confirm conclusions based on single simulations (Fig 4H-I). The exception is that rapid and medium convergence rates yield similar SI at the r5/r6 boundary and DC numbers (Fig 4H-I). Rhombomere A-P lengths also vary greatly in this model with different convergence rates (Fig 4G). Simulations with slow initial convergence result in all three rhombomeres (r3-5) elongated. Simulations with medium convergence result in a shorter r5 compared to r3 and r4. Simulations with rapid initial convergence result in r3-5 being all roughly the same length (with r4 slightly shorter) similar to experimental measurements of the hindbrain at 14 hpf (Fig 2H, 4G).

Taken together, rapid initial convergence facilitates robust patterns by optimizing boundary sharpness and rhombomere A-P length. These influences depend on multiple mechanisms, including cell sorting and gene regulation, and their coordination.

Rapid initial convergence improves boundary sharpness through cell sorting

Next, we performed simulations with models incorporating cell sorting alone and excluding gene regulation using the cell sorting-only model. Initial “salt-and-pepper” cell distributions by assigning each cell an identity generated the five-segment pattern with rough boundaries (Fig 5A). Similar to observations in full models (Fig 4), most cells segregate to their correct territories and the final patterns display sharp boundaries between rhombomeres (Fig 5B-D). Rapid initial convergence allows the sharpest boundaries and the fewest DCs, while slow initial convergence results in the roughest boundaries and more DCs (Fig 5E). SI and DC numbers at 14 hpf in multiple independent simulations confirm conclusions based on single simulations (Fig 5G-H). In

the model with cell sorting alone, while these different speeds of convergence have major effects on boundary sharpness, they have relatively minor effects on rhombomere A-P length (Fig 5F). These results suggest rapid initial convergence improves boundary sharpness by facilitating the efficiency of cell sorting.

FIGURE 5

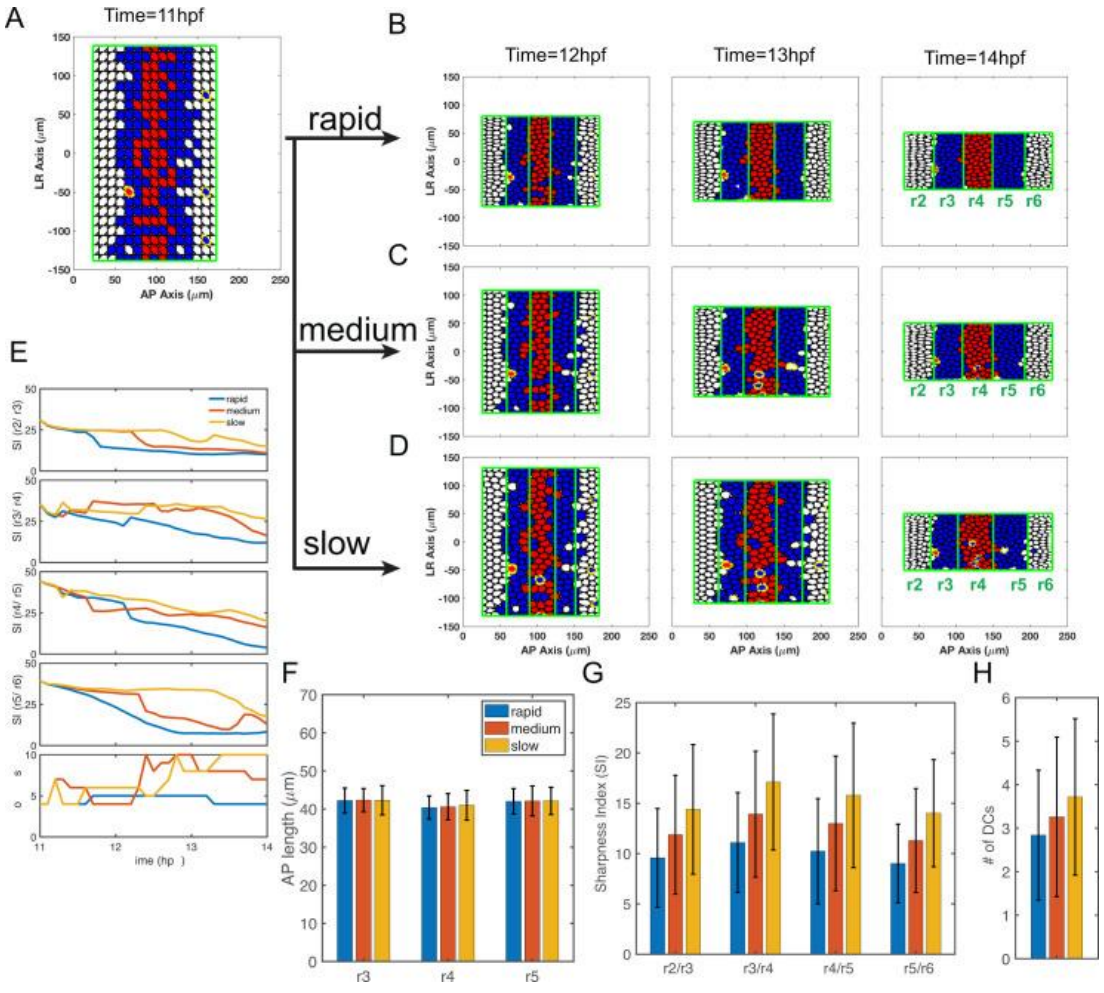

Figure 5. Simulations with selective cell-cell adhesion/sorting alone with different convergence rates. (A-D) Time series of cell distributions with different convergence rates from 11 to 14 hpf. *hoxb1a* (red) and *krox20* (blue) expression. Dislocated cells (DCs) are highlighted by yellow edges. (A) Three simulations start with the same initial cell distribution (11 hpf) generated by the Gaussian mixture distribution. Cell distributions with (B) rapid, (C) medium and (D) slow initial convergence from 12–14 hpf. (E) The boundary sharpness index (SI) for four boundaries (SI(r2/r3), SI(r3/r4), SI(r4/r5) and SI(r5/r6)) and number of DCs versus time. (F-H) Histograms depicted three convergence rates analyzed for their (F) rhombomere lengths of r3-5, (G) SI and (H) the DC number. Each represents 100 independent stochastic simulations for each convergence rate are based on the same parameters. Error bars represent standard deviation.

Rapid initial convergence helps specify correct rhombomere lengths by regulating cell fate

To investigate the effects of convergence on rhombomere lengths, we studied dynamics of gene expression under different convergence rates. We first investigated the influence of convergence rate on spatiotemporal dynamics of morphogen distribution and cell fate commitment (Fig 6). With rapid initial convergence, the RA signal and its direct target, *vhnf1*, increase quickly from 11-12 hpf, particularly in the posterior hindbrain (150 μm and 200 μm in Fig 6A-B), then decrease gradually. Near the r4/r5 boundary, *hoxb1a* is repressed by the increasing levels of *vhnf1*. FGF made in r4 then activates *krox20* leading to identity switching for cells near the r4/r5 boundary with low *hoxb1a* expression (Fig 4C). With medium or slow initial convergence, the RA signal and *vhnf1* remain relatively unchanged compared to rapid initial convergence (Fig 6A-B). Indeed, we observed that fewer cells switch from an r4 to an r5 identity and the r4/r5 boundary is located further posteriorly with medium or slow initial convergence (Fig 4C-E). Similar to RA, FGF levels increase and peak around 12 hpf with models that include rapid initial convergence (Fig 6C). With medium convergence, FGF levels remain relatively unchanged, while with slow initial convergence, FGF levels remain unchanged at the early stage then increase and peak quickly at 13-14 hpf. At the same A-P position, slow initial convergence results in higher maximum FGF levels than rapid or medium convergence rates. Since FGF induces *krox20* expression to drive identity switching from r2 identity to r3 (and r6 to

r5) identity, the higher maximum FGF levels that result from slow initial convergence led to displacement of the r2/r3 boundary anteriorly and the r5/6 boundary further posteriorly than rapid or medium convergence rates.

FIGURE 6

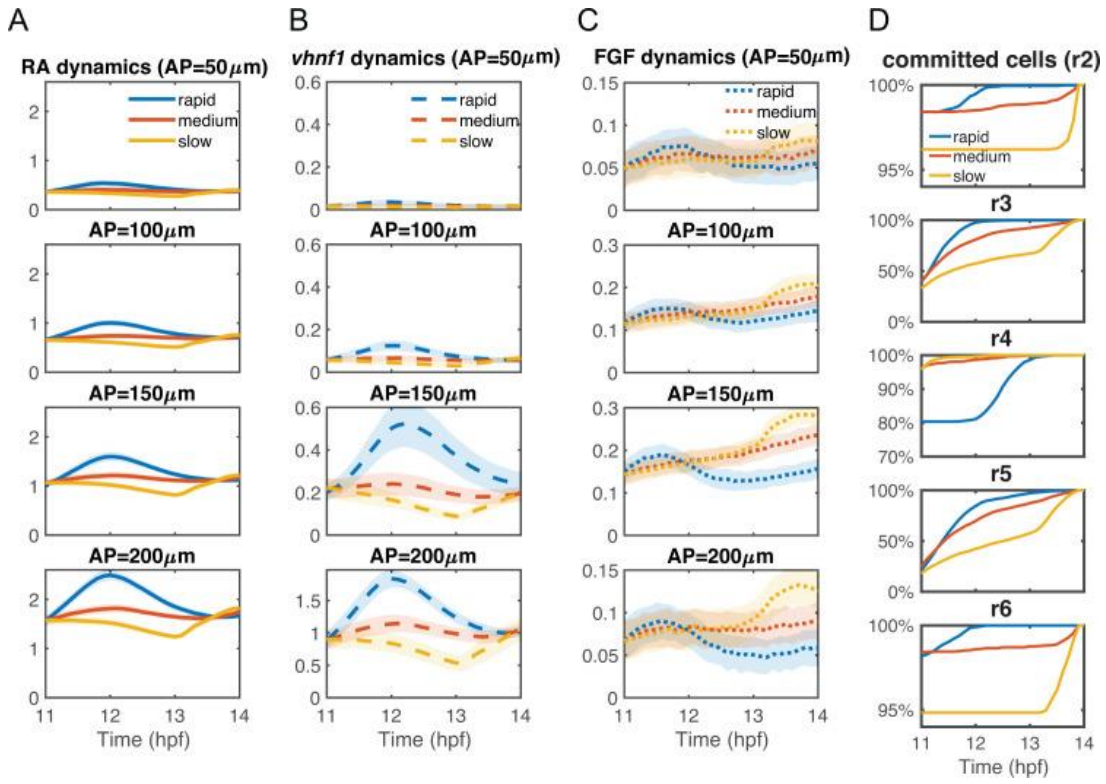

Figure 6. Dynamics of morphogens and cell commitment time with different convergence rates. The statistics of the dynamics of (A) intracellular RA $[RA]_{in}$, (B) *vhnf1*, and (C) FGF signaling $[FGF]_{signal}$ at different A-P lengths of the tissue domain: 50 μm, 100 μm, 150 μm and 200 μm. Lines represent average values and the width of the shading around each line represents the standard deviation. (D) The temporal dynamics of total percentages of cells that have committed in each rhombomere (r2-r6). Each panel shows the dynamics of the percentages of cells that have committed in each rhombomere. Data are collected from the full models, see Fig 4.

The different rhombomere A-P lengths under different convergence rates can be observed consistently either in the full model (Fig 4) or in the model excluding cell sorting (Fig S2). The dynamics of the morphogens provides an explanation for the length behaviors. Rapid initial

convergence leads to the smallest r4 because more r4 cells switch to r5 identities near the r4/r5 boundary. Slow initial convergence leads to a larger r3 and r5 because more r2 or r6 cells switch to r3 or r5 identities. Medium convergence results in the smallest r5 because fewer cells switch near the r4/r5 boundary than rapid initial convergence and fewer cells switch near the r5/r6 boundary than with slow initial convergence.

Interestingly, r3 emerges earlier and it initially has a larger A-P length than r5 since *vhnf1* expressed posteriorly represses the FGF activator, *hoxb1a*, resulting in weaker FGF signaling in r5 (Fig 4B). Through rapid convergence, r4 cells can switch to an r5 identity to compensate for the difference between r3 and r5 lengths to achieve correct A-P rhombomere lengths similar to experimental measurements.

Rapid initial convergence improves boundary sharpness through synergy between gene regulation and cell sorting

To study potential coordination between gene regulation and cell sorting in boundary formation, we restricted them individually at different patterning stages. Cells switch their fates at early stages of boundary sharpening (11-12 hpf) and this early gene regulation is critical for pattern establishment (Fig 2E). Therefore, to investigate if gene regulation also impacts pattern refinement later we extend the gene regulation period one additional hour at a time while maintaining cell sorting throughout the simulations (Fig S3). Boundary sharpness improved with longer periods of gene regulation, depending on the stage. Extension by one hour, from 11-12 hpf to 11-13 hpf, significantly changed segment size, improved boundary sharpness, and reduced the number of DCs (Fig S3B, S3C, S3F-H). Extension by an additional hour from 11-13 hpf to 11-14 hpf had negligible effects on segment size, and minor improvements in boundary sharpness and DC numbers (Fig S3C, S3D, S3F-H). Conversely, with gene regulation throughout the simulations, limiting selective cell sorting to one-hour intervals within 11-14 hpf (with cells allowed to sort uniformly at other times) yields the worst patterns when limited to 11-

12 hpf (Fig S4). These results suggest that gene regulation is more important during early patterning stages and cell sorting later for boundary sharpening.

As shown in the previous subsection, with rapid initial convergence rates morphogen levels increase and peak at around 12 hpf, while morphogen levels increase and peak much later with medium or slow initial convergence. Since *hoxb1a* and *krox20* are direct targets of RA and FGF, respectively, the timing of cell commitment as measured by expression of these genes is closely tied to the timing of peak morphogen levels. Consequently, rapid initial convergence drives the earliest cell commitment at around 12 hpf. Slow initial convergence leads to much later cell commitment (Fig 6D). One notable exception is that r4 cells commit later in the case of rapid initial convergence, due to switching near the r4/r5 boundary, but most cells still commit before 13 hpf (Fig 6D). Driven by early cell commitment, rapid initial convergence extends the effective period of cell sorting and shortens the effective period of gene regulation to improve boundary sharpness.

Rapid initial convergence mediates the trade-off between rhombomere length and boundary sharpness

To examine the sensitivity of our observations to model parameters, we performed a large number of simulations with random parameters, assaying both rhombomere A-P length and boundary sharpness. Using n=1000 independent repeats for each convergence rate, we found 513, 563 and 452 simulations that successfully produced a five-segment pattern for rapid, medium and slow initial convergence, respectively.

These results also revealed a trade-off between rhombomere length and sharpness, i.e. reduced length typically resulted in higher Sis, indicating rougher boundaries (Fig 7 and Fig S5). This can likely be explained by the fact that a shorter rhombomere has fewer cells with the same identity, thus cell sorting has less effect and cells are more susceptible to noise. As a

result, a single stray cell has more impact on the boundary sharpness index. Such trade-offs are also observed for the model without convergent extension, leading to shorter and rougher rhombomere lengths (Section 11 in S1 Text). With shorter rhombomeres, rapid initial convergence significantly reduces this trade-off, particularly in r4 and r5 (Fig 7 and Fig S5).

Figure 7

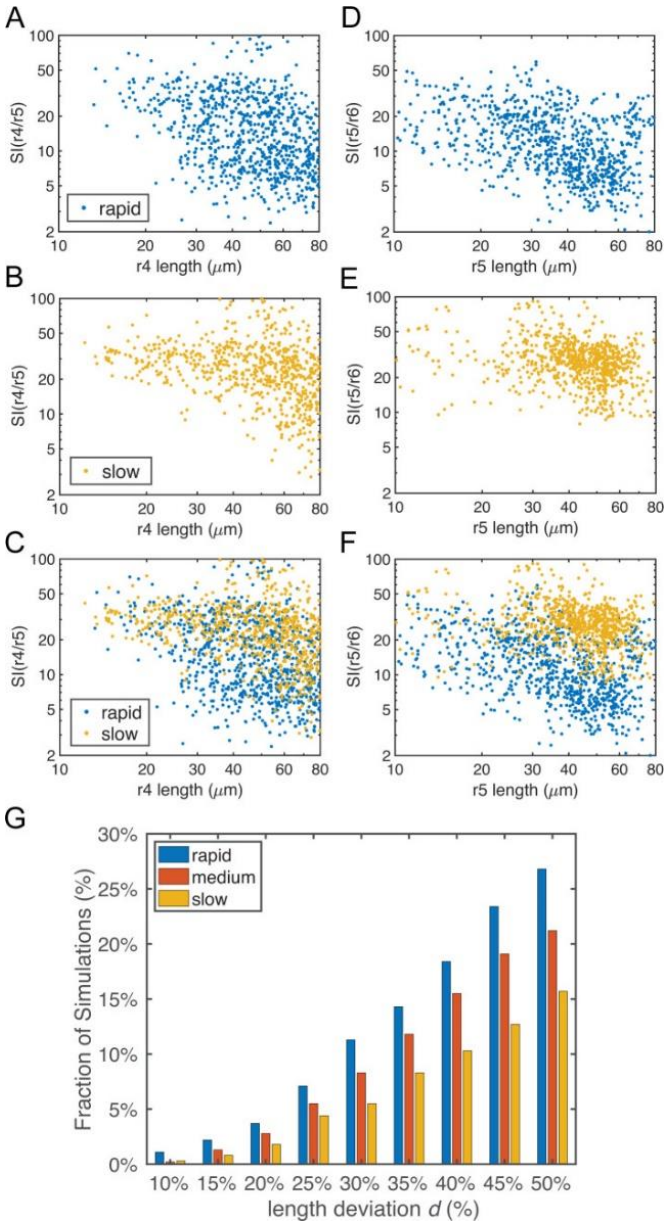

Figure 7. Boundary sharpness and rhombomere lengths based on simulations with random parameters in gene regulation. Parameters for gene regulation were randomly perturbed and a total of n = 1000 simulations are displayed for each convergence rate. There are 513, 563 and 452 simulations for rapid, medium and slow initial convergence, respectively, which successfully generate the r2-6 pattern with four boundaries. (A-F) Dot plots showing the relationship between rhombomere length and boundary sharpness. Each point represents the corresponding quantities for each simulation. (A-C) Length of r4 versus sharpness index (SI) of the r4/r5 boundary with: (A) rapid initial convergence, (B) slow initial convergence and (C) a comparison between rapid and slow initial convergence. (D-F) Length of r5 versus SI of r5/r6 boundary with: (D) rapid initial convergence, (E) slow initial convergence and (F) a comparison between rapid and slow initial convergence. (G) Fractions of simulations achieving roughly equal rhombomere lengths versus the deviation *d%*. With a deviation *d*, a simulation has roughly equal rhombomere lengths if the length of each rhombomere is deviated at most *d*% from its average experimental length (i.e. r3, r4 and r5 are in the ranges of 42*(100%±*d*%) μm, 34*(100%±*d*%) μm and 37*(100%±*d*%) μm, respectively).

Moreover, we quantified fractions of simulations achieving roughly equal segment length. We considered a simulation having equal rhombomere lengths if A-P lengths of r3, r4 and r5 were close to their experimentally measured average lengths within ranges $\left[m*\left(100\% - d\%\right),\ m*\left(100\% + d\%\right)\right]$, where *m* was the measured average length (Table G in S1 Text). With any values of deviation *d*, there are higher fractions of simulations achieving roughly equal length under rapid initial convergence compared to medium and slow initial convergence at 14 hpf (Fig 7G). Experimentally, standard deviations of r3, r4 and r5 length are within the range between 10% and 15% (Table G in S1 Text). Within this range of *d*, rapid initial convergence has at least 69% and 175% higher fractions of simulations achieving roughly equal segment length than medium and slow initial convergence, respectively (Fig 7G). These results are consistent with our findings that rapid initial convergence generates more accurate lengths of rhombomere and sharper boundaries.

Discussion

Our models suggest that a combination of two morphogens and rapid initial tissue convergence together drive robust hindbrain segmentation. Inclusion in the model of the short-range morphogen (FGF) secreted from r4, combined with the longer-range morphogen (RA) secreted posteriorly, substantially improves the robustness of segmental patterning compared with RA alone. Cooperation between morphogens is common in pattern formation in many contexts, in part because it helps maintain accuracy in size and boundary sharpness of target gene expression domains. Our previous models and experiments in the hindbrain have focused primarily on the r4/r5 boundary, where many gene regulatory interactions are known and the RA gradient is relatively steep (Addison et al., 2018; Cooke et al., 2005; Zhang et al., 2012). The current model expands upon this work to explain the formation of other rhombomere boundaries, particularly r2/r3, r3/r4 and r5/r6, with the additional positional information provided by FGF. Surprisingly, rapid initial convergence dramatically improves robustness of rhombomere patterning, both segment size and boundary sharpness. Rapid initial convergence may also be a conserved strategy for precise establishment of gene expression domains in other embryonic tissues that elongate by convergent extension (Huebner & Wallingford, 2018; Shindo, 2018; Sutherland et al., 2020) such as axial mesoderm in early vertebrate embryos (Keller & Tibbetts, 1989) or stacking of chondrocytes in developing cartilages (Ahrens et al., 2009; Kimmel et al., 1998).

Complementary roles of long- and short-range morphogens in pattern accuracy and precision

In morphogen gradient-mediated patterning, it is crucial not only that target gene expression boundaries are accurately positioned but also that they are sharp. However, there is a trade-off between accuracy and precision of boundary patterning that depends on morphogen gradient steepness (Lander, 2011; Lander et al., 2009). A steep morphogen gradient specifies

boundary locations more precisely in the face of fluctuations in signal, is less sensitive to noise and facilitates boundary sharpness. However, the trade-off is that it makes positioning boundaries more sensitive to perturbations or noise in morphogen synthesis, slight shifts in which can move the boundaries along the A-P axis. Since RA is responsible for the initial A-P patterning of the hindbrain, starting from gastrulation, it likely plays a more prominent role in accuracy, due to its shallow distribution across much of the patterning region (Sosnik et al., 2016) and self-enhanced degradation (White et al., 2007). On the other hand, FGF likely plays a prominent role in precision to help improve the sharpness of boundaries adjacent to its source, since its gradients are likely steeper due to its local effects (Choe et al., 2011; Maves et al., 2002; Waskiewicz et al., 2002). In addition, FGF synthesis most likely varies less since one of its upstream regulators, *hoxb1a*, is bi-stable and tightly controlled by a complex network (Alexander et al., 2009; Barrow et al., 2000; Parker & Krumlauf, 2017). Together, these complementary features of the long-range shallow RA gradient and the short-range steep FGF gradient help overcome the trade-offs inherent in morphogen patterning systems for achieving both accurate and precise rhombomere pattern.

During hindbrain segmentation, r4 becomes the secondary signaling center that produces FGFs (e.g. Fgf4 and Fgf8) in zebrafish (Choe et al., 2011; Maves et al., 2002; Walshe et al., 2002; Waskiewicz et al., 2002) that are regulated by the posteriorizing signal RA (White et al., 2007). The MHB is another secondary FGF (i.e. Fgf8 in zebrafish) signaling center that likely contributes to anterior rhombomere patterning (Reifers et al., 1998). In many biological contexts, two morphogens interact with each other to facilitate spatial pattern formation. Interactions between the long- and short-range morphogens induce Turing patterns, such as Sox9 and Bmp in digit patterning (Raspopovic et al., 2014), Edar and FGF in murine tooth development (Häärä et al., 2012), FGF and Shh in limb regeneration (Nacu et al., 2016), and Nodal and Lefty in early mesoderm formation and left-right patterning (Müller et al., 2012; Shen, 2007). Two long-range morphogens with anti-parallel distributions improve the precision of a single boundary, such as

Bcd and Cad in *Drosophila* embryo segmentation (Briscoe & Small, 2015), and Bmp and Shh in vertebrate neural tube patterning (Briscoe & Small, 2015; Liem et al., 2000; Zagorski et al., 2017). Unlike these examples, the novel two-morphogen mechanism presented in this work includes one long-range and one short-range morphogen that act in parallel on downstream targets. This system specifies multiple boundaries of gene expression and improves both accuracy and precision of segmental patterns.

Rapid initial convergence in tissue morphogenesis improves pattern robustness

Intuitively, elongation along the A-P axis might be expected to hinder segmental patterning and rhombomere boundary sharpening (Zhang et al., 2012), since cells quickly change neighbors and intercalate. However, we find quite the opposite (Fig 8). Rapid initial convergence facilitates boundary sharpening through two strategies. First, it induces efficient cell-cell contacts, consequently more efficient cell sorting, leading to sharper boundaries (Fig 8B). Second, it induces an early peak of morphogens that can result in early cell commitment, allowing cell sorting sufficient time for rearrangements without disrupting cell fate switching (Fig 8C). Rapid initial convergence can also regulate rhombomere length through morphogen dynamics. Initially the length of r5 is shorter than r3. Through a steeper RA distribution induced by rapid initial convergence, cells switch from an r4 to an r5 leading to similar r3 and r5 lengths (Fig 8A).

FIGURE 8

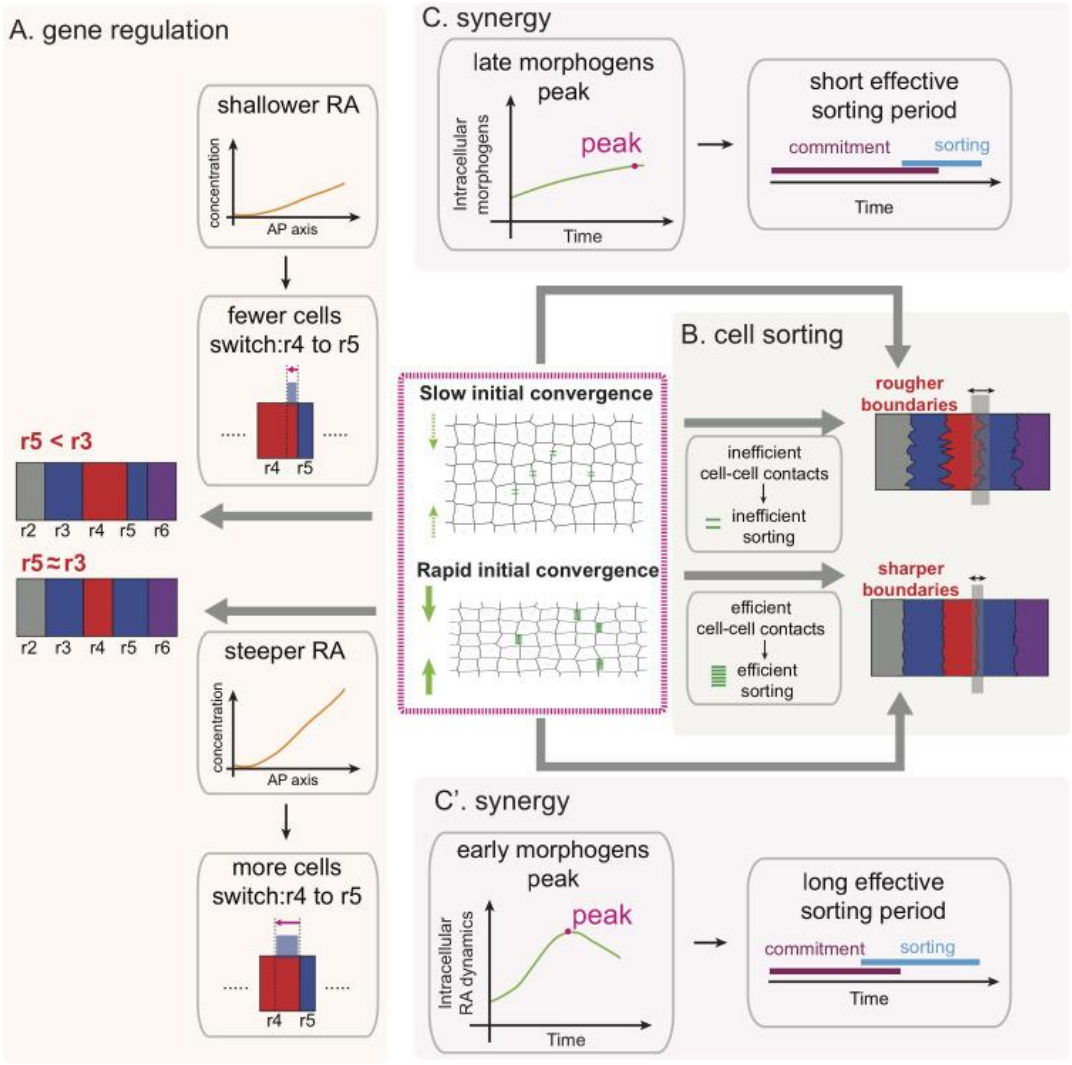

Figure 8. Schematic illustration for rapid initial convergence improves pattern robustness by comparing with slow initial convergence (A) Gene regulation/cell fate: rapid initial convergence produces a steeper RA distribution to induce more cells switching from r4 (red) to r5 (blue) identity than slow initial convergence. Consequently, r5 has similar size with r3, consistent with the experimental measurements. (B) Cell sorting: rapid initial convergence increases cell-cell contacts to enhance efficiency of cell sorting, leading to sharper boundaries comparing to slow initial convergence. Number of green lines represent efficiency of cell sorting. (C-C') Synergy between cell sorting and gene regulation: rapid initial convergence induces an early peak of morphogens for both RA and FGF, leading to an early commitment of cell fates. Cell sorting mechanisms fully function to sharpen boundaries with sufficient time without disrupting cell fate switching.

Another consequence of convergent extension is the movement of morphogen production and responding cells relative to one another. RA levels increase with time during early stages of hindbrain development due to increased synthesis and accumulation (Maves & Kimmel, 2005), then decrease due to the movement of the source of RA (*aldha1a* expression) further posteriorly as the body axis elongates (Alexander et al., 2009; Maves & Kimmel, 2005; Parker & Krumlauf, 2017). Our model successfully recapitulates these RA dynamics (Fig 6A), which are also critical for specifying rhombomeres of the correct length and boundary sharpness. Previous manipulations of convergent extension in early zebrafish or mouse embryos, which result in a shortened body axis, have shown that convergent extension is critical in establishing signaling gradients and subsequently maintaining robust segmental patterning of the hindbrain, consistent with our results (Fulton et al., 2020; García-García et al., 2008).

As tissue deforms, extracellular morphogens may have both active motion driven by the tissue dynamics as well as movements of the signals induced by morphogens within cells. In our model, both extra and intracellular morphogens were modeled in continuum context with Eulerian coordinates, where the advections are usually required to capture morphogen dynamics with moving boundaries (Čanić, 2021). We found that even if we removed intracellular advections from the models, our main results showing a positive influence of rapid initial convergence on patterning remain (Fig S11).

Tissue size, thickening and additional signals in hindbrain patterning

The embryonic zebrafish hindbrain is extremely small and composed of relatively few cells compared with most other vertebrates (Addison et al., 2018; Kulesa & Fraser, 1998; Thierion et al., 2017). The actual A-P length of each rhombomere at the stages we have examined (11-14 hpf) is approximately 3~5 cell-diameters. This small size presents a challenge for sharpening rhombomere boundaries, where a few neighbor cells with the same identity provide weak adhesion during sorting, and even more so for generating a series of rhombomeres of similar size.

A rapid initial convergence rate may be particularly important for coordinating size and boundary sharpness in such a miniaturized embryo. However, given the conserved patterns of gene expression and neuronal differentiation observed in hindbrains across species, we are confident that many of the same rules apply.

Our modeling and experimental measurements correspond in many respects, including the dynamics of RA synthesis and FGF4/8 expression as well as *hoxb1a*, *krox20*, *vhnf1* and *irx3* expression in zebrafish. However, many questions remain. In our simulations, DCs remain in r2 and r6 due to randomness in gene expression. For example, in some cases *hoxb1a*+ cells are observed in r2 because of early increases in RA, which induce *hoxb1a*. The model cannot account for how these cells switch to their correct segmental identities, but perhaps they are displaced along the D-V axis, undergo apoptosis or are extruded from the hindbrain. Such switching may also reflect a "community effect" by which cells switch identity depending on the collective influences of their neighbors, but the underlying mechanisms have not been fully identified (Addison et al., 2018; Schilling et al., 2001). Gene regulatory networks with other signals involved may also prevent cell switching or sorting by regulating *hoxb1a* or *krox20*. For example, Fgf8 is expressed at the MHB in zebrafish embryos starting at 10 hpf, and likely important for the patterning of more anterior rhombomeres, r1-3, which will be interesting to consider in future models (Parker & Krumlauf, 2017; Reifers et al., 1998). Wnt is another morphogen that controls early anterior rhombomeres and MHB formation (Buckles et al., 2004; Parker & Krumlauf, 2017).

Many other features of tissue morphogenesis also need to be considered for a comprehensive three-dimensional model of hindbrain segmentation. During the patterning period considered here, cells divide and the NP thickens along the D-V axis (Papan & Campos-Ortega, 1994). While cells divide in this period, the NP thickens and cell number in the two-dimensional plane (A-P and L-R plane) changes very little (Papan & Campos-Ortega, 1994). We also studied a two-dimensional model that incorporates cell proliferation and growth and while this led to tighter cell distributions and higher variations in rhombomere length than in other models, overall it

confirmed that rapid initial convergence improves pattern robustness (Section 10 in S1 Text). The two-dimensional nature of our models, which do not consider the complicated dynamics associated with proliferation and thickening along the D-V axis, likely explains why the computed rhombomere lengths in our model do not perfectly fit the experimental measurements (Fig 2H). More realistic, three-dimensional models that incorporate these components pose an exciting challenge and opportunity for the future.

Methods

Ethics statement

All animal work performed in this study was approved by the University of California Irvine Institutional Animal Care and Use Committee (Protocol #AUP-20-145).

Animal Husbandry

All animals used in this study were raised and handled in accordance with the guidelines of the Institutional Animal Care and Use Committee at University of California, Irvine. AB strain embryos were collected from natural crosses, raised at 28.5 °C in embryo medium (EM), and staged as previously described (Kimmel et al., 1995).

Whole-mount in situ hybridization

In situ hybridization was performed on whole embryos as previously described (Thisse et al., 1993). Digoxygenin- and fluorescein-labelled riboprobes for *aldh1a2* (Begemann et al., 2001), *krox20* (Oxtoby & Jowett, 1993), *otx2* (Mercier et al., 2004) were synthesized using an RNA labelling kit (Roche) from cDNA that had been previously cloned into PCS2+ plasmids and linearized.

Imaging and measurement of hindbrain

Embryos were flat mounted in glycerol as previously described (Cheng et al., 2014) and imaged on a Zeiss Axioplan 2 compound microscope equipped with a Micropublisher 5.0 RTV camera with Zeiss ZEN 3.1 (blue edition) software. Hindbrain measurements were performed using ImageJ/Fiji software.

Computational domains of the model

The entire hindbrain along with the RA production region, modeled as the "morphogen domain", is used to model the diffusion and distribution of morphogens (Fig 1B). In the two-dimensional model, the morphogen domain is assumed as a rectangle with the anterior-posterior (A-P) axis as its length and the left-right (L-R) axis as its width. We take the posterior end of the MHB, defined by *otx2* expression, as the anterior limit of the domain, $x^{(1)} = 0$ and the posterior end of the RA production region, defined by *aldh1a2* expression, as its posterior limit, $x^{(1)} = L_1(t)$. The L-R width of the hindbrain is $L_2(t)$. The morphogen domain has a rectangular structure with dynamic sizes:

$$M(t) = \left\{ \left(x^{(1)}, x^{(2)} \right) \in H : H = \left[0, L_1(t) \right] x \left[-\frac{1}{2} L_2(t), \frac{1}{2} L_2(t) \right] \right\} \qquad (1)$$

The RA production region is modeled as:

$$P(t) = \left\{ \left(x^{(1)}, x^{(2)} \right) : \left(x^{(1)}, x^{(2)} \right) \in [p(t), L_1(t)] x [-\frac{1}{2} L_2(t), \frac{1}{2} L_2(t)] \right\}, \qquad (2)$$

where $p(t)$ is the A-P position of the anterior boundary of the RA production region at time t. $L_1(t)$, $L_2(t)$ and $p(t)$ are obtained from experimental measurements made in zebrafish embryos at 11, 12, 13 and 14 hours postfertilization (hpf). A cubic interpolation is used to obtain the smooth curves (Fig 2A-B**)**.

Individual cells are modeled in a "tissue domain" that is contained within the morphogen domain. The tissue domain shares the same L-R axis with of the morphogen domain and its A-P range is proportional to the range of morphogen:

$$T(t) = \left\{ \left(x^{(1)}, x^{(2)} \right) : \left(x^{(1)}, x^{(2)} \right) \in [r_1 L_1(t), r_2 L_1(t)] x [-\frac{1}{2} L_2(t), \frac{1}{2} L_2(t)] \right\} \qquad (3)$$

where r_1 and r_2 are constants given in the Table A in S1 Text. At $\boldsymbol{x} = (x_1, x_2)$, the growth velocity of the tissue is given by

$$\boldsymbol{V}(\boldsymbol{x},t)=\left(L_1'(t)\frac{x^{(1)}}{L_1(t)},L_2'(t)\frac{x^{(2)}}{L_2(t)}\right). \tag{4}$$

Stochastic dynamics of morphogens

To model morphogen dynamics in the growing hindbrain, we use stochastic convection-reaction-diffusion equations. The equations for RA are given by

$$\begin{aligned}\frac{\partial[RA]_{out}}{\partial t}+\overbrace{\nabla\cdot\left(\boldsymbol{V}[RA]_{out}\right)}^{\text{convection}}&=\overbrace{D_r\Delta[RA]_{out}}^{\text{diffusion}}+\overbrace{A(\boldsymbol{x},t)}^{\text{production}}\\&\quad+\overbrace{k_r[RA]_{in}-(1+\beta_r)k_r[RA]_{out}}^{f_1}+\overbrace{\mu_{r1}\frac{dw_{r1}(t)}{dt}}^{\text{noise}},\\\frac{\partial[RA]_{in}}{\partial t}+\nabla\cdot\left(\boldsymbol{V}[RA]_{in}\right)&=\overbrace{k_r[RA]_{out}-k_r[RA]_{in}-d_r(x^{(1)})[RA]_{in}}^{f_2}\\&\quad+\mu_{r2}\frac{dw_{r2}(t)}{dt},\end{aligned} \tag{5}$$

where $[RA]_{out}$ and $[RA]_{in}$ are extracellular and intracellular forms of RA, respectively. $\mu_{r1}\frac{dw_{r1}(t)}{dt}$ and $\mu_{r2}\frac{dw_{r2}(t)}{dt}$ are additive white noise. The convection term describes the dilution and advection of RA caused by convergence extension. The production is confined to the RA production region and modeled by a Hill function of AP position $x^{(1)}$ with a large Hill coefficient:

$$A(\boldsymbol{x},t)=\frac{v_r}{1+\left(\frac{x_1}{p(t)}\right)^{-20}}. \tag{6}$$

In f_1 and f_2, k_r is the rate of exchange of morphogen between intracellular and extracellular forms. The rate of extracellular morphogen degradation is taken as a constant $\beta_r k_r$ and the degradation of intracellular morphogen rate d_r is a piecewise function with

$$d_r(x^{(1)})=\begin{cases}k_{\max}, & \text{if } x^{(1)}>p(t),\\ k_0, & \text{otherwise.}\end{cases} \tag{7}$$

The degradation rate in the RA production region has a large value k_{max}, since the RA degrading enzyme cyp26a1 is expressed in the RA production region (White et al., 2007). We use an absorbed boundary condition of $x^{(1)}=0$ since cyp26a1 is highly expressed in the forebrain and MHB, providing a sink for RA. No-flux boundary conditions are used for the other three boundaries.

Similarly, we model both free diffusible FGF ($[Fgf]_{free}$) and FGF ($[Fgf]_{signal}$) signaling as the following:

$$\begin{aligned}\frac{\partial[Fgf]_{free}}{\partial t}+\nabla\cdot\left(\boldsymbol{V}[Fgf]_{free}\right)&=D_f\mathrm{D}[Fgf]_{free}+A_f([H],\boldsymbol{x})+k_{rf}[Fgf]_{signal}\\&\quad-(d_{f1}+k_f)[Fgf]_{free}+\eta_1\frac{dw_{f1}(t)}{dt},\\\frac{\partial[Fgf]_{signal}}{\partial t}+\nabla\cdot\left(\boldsymbol{V}[Fgf]_{signal}\right)&=k_f[Fgf]_{free}-k_{rf}[Fgf]_{signal}\\&\quad-d_{f2}[Fgf]_{signal}+\eta_2\frac{dw_{f2}(t)}{dt}.\end{aligned}\tag{8}$$

The free diffusible FGF binds with its receptor to form a complex with rate $k_f[Fgf]_{free}$. The complex between FGF and its receptor represents the FGF signal ($[Fgf]_{signal}$) for simplification. The term $k_{rf}[Fgf]_{signal}$ describes the dissociation rate of the complex. d_{f1} and d_{f2} are degradation rates of free diffusible FGF and FGF signaling, respectively. The production of FGF is upregulated by *hoxb1a* and the production rate is modelled by a Hill function for *hoxb1a*:

$$A_f([H],\boldsymbol{x})=v_f\frac{[H](\boldsymbol{x})^2}{1+a_{hf}[H](\boldsymbol{x})^2}.\tag{9}$$

The *hoxb1a* level [*H*] is defined at the center of each cell. In Eq. (9), the term $[H](\boldsymbol{x})$, defined at arbitrary location $\boldsymbol{x}$, is obtained by interpolating [*H*] values with locations in cell centers (Section 1 in S1 Text). No-flux boundary conditions are used for FGF at all four boundaries.

Stochastic dynamics of downstream genes

We model the dynamics of gene expression with a system of stochastic differential equations based on the gene network (Fig 1C**)**. For the *i*-th cell centered at $\boldsymbol{c}_i$, the equations for the gene expression are given by

$$\begin{aligned}
\frac{d[H]_i}{dt} &= v_h \frac{a_{hh}[H]_i^2 + a_{rh}[RA]_{in}^2}{1 + a_{hh}[H]_i^2 + a_{rh}[RA]_{in}^2 + b_{kh}[K]_i^2 + b_{vh}[V]_i^2} - d_h[H]_i + \eta_h \frac{dw_h}{dt}, \\
\frac{d[K]_i}{dt} &= v_k \frac{a_{kk}[K]_i^2 + a_{fk}[Fgf]_{signal}^2}{1 + a_{kk}[K]_i^2 + a_{fk}[Fgf]_{signal}^2 + b_{hk}[H]_i^2} - d_k[K]_i + \eta_k \frac{dw_k}{dt}, \\
\frac{d[V]_i}{dt} &= v_v \frac{a_{rv}[RA]_{in}^2}{1 + a_{rv}[RA]_{in}^2 + b_{iv}[I]_i^2} - d_v[V]_i + \eta_v \frac{dw_v}{dt}, \\
\frac{d[I]_i}{dt} &= v_i \frac{1}{1 + b_{vi}[V]_i^2} - d_i[I]_i + \eta_i \frac{dw_i}{dt},
\end{aligned} \tag{10}$$

where [*H*], [*K*], [*V*] and [*I*] are gene expression of *hoxb1a*, krox20, *vhnf1* and *irx3*, respectively. $[RA]_{in} = [RA]_{in}\big|_{\mathbf{x}=\mathbf{c}_i}$ and $[Fgf]_{signal} = [Fgf]_{signal}\big|_{\mathbf{x}=\mathbf{c}_i}$ refer to the RA and FGF signaling levels at the center of i-th cell $\boldsymbol{c}_i$ to provide spatial signals for cells.

Models for individual cells and their interactions

Following our previous study (Wang et al., 2017), we use the subcellular element method (SCEM) to model individual cells (Newman). A total of N_{cell}=345 cells are modeled in each simulation, where 23 rows and 15 columns of cells align uniformly in the rectangular tissue region at the initial stage (11 hpf). Each cell consists of sub-cellular elements (nodes) and interacts according to a prescribed intercellular force potential. A cell consists of $2N_{node}$ (N_{node}=6) nodes and those nodes form two hexagonal layers (Fig S6A**)**. The radius of the outer layer is R_{out} and the radius of the inner layer is R_{in}. Initially cells are uniformly distributed in the tissue domain (Fig 4B). For a system with N_{cell} cells and $2N_{node}$ nodes per cell, the location of *i*-th node in *n*-th cell $\boldsymbol{x}_{n,i}$ is determined by the equation

$$\frac{d}{dt}\boldsymbol{x}_{n,i} = \boldsymbol{v}_{n,i}^{ext} + \boldsymbol{v}_{n,i}^{\text{inter}} + \boldsymbol{v}_{n,i}^{\text{inner}}. \tag{11}$$

On the right hand side, the first term represents cell migration due to convergent-extension (Rørth, 2012; Schumacher et al., 2016). It is given by

$$\begin{cases} \boldsymbol{v}_{n,j}^{ext} = \boldsymbol{V}(\boldsymbol{c}_n, t), \\ \boldsymbol{c}_n = \dfrac{1}{2N_{node}} \displaystyle\sum_{i=1}^{2N_{node}} \boldsymbol{x}_{n,i}. \end{cases} \tag{12}$$

The second term represents the forces between cells while the third term represents forces between nodes within the same cell to maintain stable cell morphologies (Newman) (see **Section 2 in S1 Text**).

Definition of boundary location (*m*), sharpness index (*SI*) and number of dislocated cells (*DC*)

To study boundary locations and sharpness quantitatively, we define three quantities: boundary location (m), boundary sharpness index (SI) and number of dislocated cells (DCs). For example, the A-P position of the r3/r4 boundary is denoted by $m(r3/r4)$ and its boundary SI is denoted by $SI(r3/r4)$. A cell is called a DC if: a) its identity is different from the segment in which it is located; b) its distance to the boundary of its correct segment is over three cell-diameters.

In a region with A-P coordinates in the range of (a,b), we split the index set of all cells into two sets S_L and S_R based on cell identities, where cells in S_L or S_R have segmental identities located anterior or posterior to this region. We define the distance function from the i-th cell centered at $\boldsymbol{c}_i$ to an arbitrary straight line with A-P position k (the potential location of boundary):

$$\text{dis}(\boldsymbol{c}_i, k) = \begin{cases} \text{ReLU}(c_i^{(1)} + r - k), & \text{if } i \in S_L, \\ \text{ReLU}(-c_i^{(1)} + r + k), & \text{if } i \in S_R, \end{cases} \tag{13}$$

where r is the radius of the cell, and

$$\mathrm{ReLU}(x)=\begin{cases} x, \text{ if } x>0, \\ 0, \text{ if } x\le 0, \end{cases}$$

is the rectified linear unit function. For a cell in S_L or S_R with non-zero distance, this distance function calculates the Euclidean distance between the anterior or posterior distal ends of this cell to the potential location of boundary k. This distance function is illustrated in Fig S6B.

We quantify the boundary location (m), SI and number of DCs in this region, called K, by solving an optimization problem:

$$\begin{aligned} m_K &= \arg\min_{k\in(a,b)}\left(\sum_{\mathrm{dis}(\boldsymbol{c}_i,k)\le 6r}\left(\mathrm{dis}(\boldsymbol{c}_i,k)\right)^2\right)^{\frac{1}{2}}, \\ SI_K &= \min_{k\in(a,b)}\left(\sum_{\mathrm{dis}(\boldsymbol{c}_i,k)\le 6r}\left(\mathrm{dis}(\boldsymbol{c}_i,k)\right)^2\right)^{\frac{1}{2}}, \\ DC_K &= \#\left\{k:\mathrm{dis}(\boldsymbol{c}_i,m_R)>6R_0\right\}. \end{aligned} \tag{14}$$

Particularly if the distance from a cell to the boundary is within three cell-diameters, the cell contributes to the boundary location and boundary SI. Otherwise, it is regarded as a DC and it does not contribute to the calculation of either boundary location or sharpness.

Next, we split all cells in the responding tissue domain with index set S into three sets with distinct cell types at time t. There are *hoxb1a* cells (S_h), *krox20* cells (S_k) and non-expressing cells (S_n) based on their expression level of *hoxb1a* and *krox20*.

$$\begin{aligned} S_h &= \left\{i\in S:[H]_i(t)\ge 1.2\right\}, \\ S_k &= \left\{i\in S:[K]_i(t)\ge 1.2\right\}, \\ S_n &= \left\{i\in S: i\notin S_h \cup S_k\right\}. \end{aligned} \tag{15}$$

Now, we calculate those quantities for four boundaries in the tissue domain one-by-one by utilizing Eq. (14) and Eq. (15) as shown in the flow below:

Algorithm 1: Calculate m, SI and DC for cells in domain with AP range $[r_1L_1(t), r_2L_1(t)]$

Step 1	$C = \underset{i \in S_h}{\text{mean}}(\boldsymbol{c}_i^{(1)})$	Find a point to split the tissue domain
Step 2	$K_1 = [r_1L_1(t), C]$, $S_L = \{i \in S : i \in S_k \subset S_n \text{ and } \boldsymbol{c}_i^{(1)} \in K_1\}$, $S_R = \{i \in S : i \in S_h \text{ and } \boldsymbol{c}_i^{(1)} \in K_1\}$, $m(r3/r4) = m_{K_1}$, $SI(r3/r4) = SI_{K_1}$, $DC(r3/r4) = DC_{K_1}$.	Quantify r3/r4 boundary: *hoxb1a* cells are on the right and other cells are on the left.
Step 3	$K_2 = [C, r_2L_1(t)]$, $S_L = \{i \in S : i \in S_h \text{ and } \boldsymbol{c}_i^{(1)} \in K_2\}$, $S_R = \{i \in S : i \in S_k \subset S_n \text{ and } \boldsymbol{c}_i^{(1)} \in K_2\}$, $m(r4/r5) = m_{K_2}$, $SI(r4/r5) = SI_{K_2}$, $DC(r4/r5) = DC_{K_2}$.	Quantify r4/r5 boundary
Step 4	$K_3 = [r_1L_1(t), m(r3/r4)]$, $S_L = \{i \in S : i \in S_n \text{ and } \boldsymbol{c}_i^{(1)} \in K_3\}$, $S_R = \{i \in S : i \in S_k \text{ and } \boldsymbol{c}_i^{(1)} \in K_3\}$, $m(r2/r3) = m_{K_3}$, $SI(r2/r3) = SI_{K_3}$, $DC(r2/r3) = DC_{K_3}$.	Quantify r2/r3 boundary
Step 5	$K_4 = [m(r4/r5), r_2L_1(t)]$, $S_L = \{i \in S : i \in S_k \text{ and } \boldsymbol{c}_i^{(1)} \in K_4\}$, $S_R = \{i \in S : i \in S_n \text{ and } \boldsymbol{c}_i^{(1)} \in K_4\}$,	Quantify r5/r6 boundary

	$m(r5/r6) = m_{K_4}$, $SI(r5/r6) = SI_{K_4}$, $DC(r5/r6) = DC_{K_4}$.	
Step 6	$DC = DC(r2/r3) + DC(r3/r4)$ $+ DC(r4/r5) + DC(r5/r6).$	

Chapter II

Improvement of MS2 RNA-labeling system in zebrafish for studying transcriptional noise

Introduction

Segmentation in the hindbrain (rhombencephalon) is a progressive process that divides along the anterior-posterior (A-P) axis the future hindbrain territory into seven or eight discrete segments known as rhombomeres (Lumsden, 2004; Lumsden & Krumlauf, 1996; Moens & Prince, 2002). Boundaries between rhombomeres are initially rough and jagged but eventually become razor sharp as hindbrain segmentation progresses. Rhombomere segments have distinct gene expression domains with the expression of *hoxb1a* being restricted to r4 and *krox20* to r3/5. Cell tracking of *krox20+* cells has shown cell sorting contributes to sharpening (Calzolari et al., 2014; Terriente & Pujades, 2015), and Ephrin/ Eph signaling has been found to be critical for repulsive interactions between cells in different rhombomeres that drive sorting (Cooke et al., 2001; Cooke et al., 2005). However, there is evidence from cell lineage tracing and transplantation studies in the hindbrain that a certain level of intermingling between cells of adjacent segments occurs and that these cells can acquire appropriate A-P identity in accordance with their prospective segment suggesting a role for cell identity switching mechanisms in boundary sharpening (Fraser et al., 1990; Trainor & Krumlauf., 2000; Schilling et al., 2001).

Cells near future gene expression boundaries experience stochastic fluctuations in morphogen concentration (i.e., noise), ability to respond (e.g., number of receptors), transcription/translation of target genes, and feedback (Elowitz et al., 2002; Kaern et al., 2005; Kepler & Elston, 2001). Noise in RA signaling and in its target genes are expected to compromise the ability of cells to interpret their positions within the morphogen gradient or to form sharp boundaries of gene expression. Surprisingly, when modeled computationally

simultaneous inclusion of noise in both RA and in its target genes (e.g., *krox20* and *hoxb1a*) was found to improve sharpening (Zhang et al., 2012). Based on these findings, a new mechanism of "noise-induced switching" for boundary sharpening was proposed (Schilling et al., 2012; Zhang et al., 2012). In this model, stochastic fluctuations in *hoxb1a* and *krox20* expression enable cells to transition between two steady states, from *hoxb1a*+ to *krox20*+ or vice versa, by overcoming an energetic "barrier" between states. This model is counterintuitive because it argues for a positive role for noise and suggests that the process of boundary sharpening requires noise. Aspects of this model have been validated such as: the observation of double *krox20* +/*hoxb1a*+ cells at boundaries (Zhang et al., 2012) and noise in the RA gradient (Sosnik et al., 2016). More recently, direct evidence for switching has been found with the observation that *krox20*+ cells in *egr2b*(*krox20*):H2B-citrine zebrafish transgenics were found to intermingle between hindbrain segments and switch identity due to feedback from community retinoid signaling (Addison et al., 2018). These findings reveal the complexity of the role of noise in hindbrain pattern formation and may indicate possible roles in other developmental processes.

Recently, new developments in visualizing gene expression with MS2 RNA stem loops have enabled quantitative and dynamic imaging of gene expression in live organisms. The MS2 RNA-labeling system functions by binary expression of a fluorescent MS2 bacteriophage coat protein (MCP) along with an RNA of interest tagged with multiple RNA hairpin MS2-binding sites (MBS) which the MCP fusion protein binds to (Hocine et al., 2013). This system enables dynamic live measurement of gene expression noise through time-lapsed imaging and quantitative measurements of MCP binding to MBS. While the MS2 system has been used to image dynamics of one segmental stripe of *even-skipped* in the early *Drosophila* embryo (Bothma et al., 2014), this system has yet to be applied to study transcript dynamics during pattern formation in vertebrates. The zebrafish presents itself as an ideal vertebrate model system for studying early transcriptional events in development due to its transparent embryos

which develop outside of the mother's body. Recent technological developments have established a framework for MS2 imaging in the zebrafish system (Campbell et al., 2015) though improvements are still required to study accurately transcript dynamics both spatially and temporally. In this work, I improve upon the existing tools available for the MS2-RNA labeling system in zebrafish for use in future study of transcriptional noise.

Results and Discussion

Design of new constructs to eliminate aggregates in existing tdMCP-eGFP transgenics

To study transcriptional noise, transcriptional activity has to be visualized and quantified accurately. Existing transgenic lines for tandem dimer(td) MCP-eGFP in zebrafish are unsuitable for accurate analysis of transcriptional dynamics due to the presence of fluorescent aggregates of tdMCP-eGFP that complicate identification of true fluorescent puncta that mark transcriptional activity (Campbell et al., 2015). These aggregates are likely a consequence of GFP's natural propensity to dimerize at high concentrations. To address this issue, I modified the tdMCP-eGFP construct previously generated (Campbell et al., 2015) (Fig.1). For one construct, I monomerized eGFP using a A206K mutation previously described which prevents eGFP from dimerizing (Zacharias et al., 2002). In the second construct, I deleted the nuclear localization signal (*NLS*), which disperses GFP throughout the cell reducing its concentration in the nucleus. This was the strategy used to reduce background aggregates in the *Drosophila* system (Bothma et al., 2014).

FIGURE 1

Figure 1: Schematic of construct designs for tdMCP-eGFP lines with decreased GFP aggregates Two constructs were designed to be expressed ubiquitously under a *βactin* promoter with modifications to the existing tdMCP-eGFP to delete the *NLS* tag or monomerize eGFP with the A206K mutation. Constructs have Tol2 sites to generate transgenic lines by Tol2-mediated transgenesis.

New tdMCP-eGFP transgenics have decreased GFP aggregates

After successful cloning of the new constructs, zebrafish transgenics were made using Tol2-mediated transgenesis (Kawakami, 2007; Kawakami et al., 1998) expressing the tdMCP-eGFP variants ubiquitously under a *βactin* promoter (Higashijima et al., 1997). Transgenics expressing either of the new constructs showed dramatic reduction or elimination of ectopic GFP aggregates when compared to the original *Tg(βactin:NLS-tdMCP-eGFP)* zebrafish line (Fig. 2). While the original *Tg(βactin:NLS-tdMCP-eGFP)* line had visibly bright and large aggregates throughout the cell nuclei in the absence of any MBS binding, these aggregates were largely absent in cells expressing *NLS-tdMCP-eGFP(A206K)* despite similar nuclear localization of GFP. *tdMCP-GFP* (no *NLS*) expressing cells showed an absence of aggregates, though the GFP expression was much fainter as the signal was diffused across the cell instead of localized to the nucleus. This was not seen in *NLS-tdMCP-eGFP(A206K)* expressing cells which still had strong signal localized to nuclei. The decision to utilize either line for studies would likely be based on imaging setting preferences and on the focus of the study but both provide viable options absent of background aggregates.

FIGURE 2

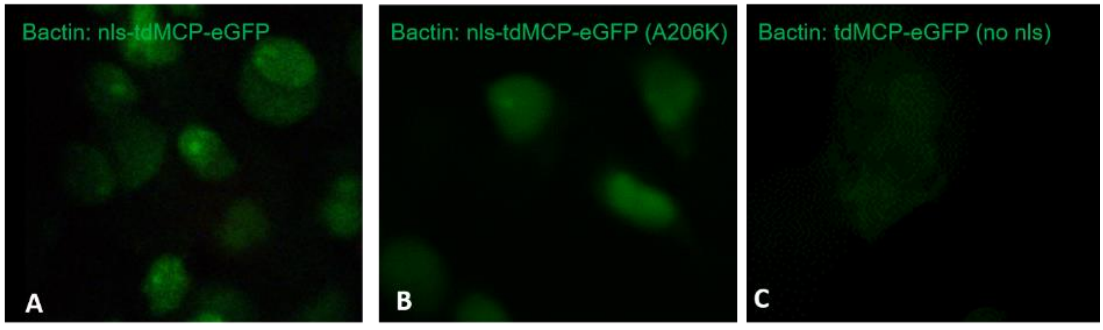

Figure 2: Introduction of A206K mutation to GFP or deletion of *NLS* eliminates aggregates in absence of MBS Imaged GFP-positive cells (green) in zebrafish transgenics lines for (A) *βactin: NLS-tdMCP-eGFP*, (B) *βactin: NLS-tdMCP-eGFP* (A206K), and (C) *βactin: tdMCP-eGFP* (no *NLS*).

Visualization of active transcription by MCP-MBS binding in new tdMCP-EGFP transgenics

After confirmation of reduction in aggregates in new constructs, the next step was to validate whether the MCP-MBS interaction is still occurring in the newly generated transgenics. To test this, transgenics were made by injecting both the newly designed tdMCP-eGFP and *mCherry-24xMBS* DNA constructs. When cells in the zebrafish embryo express either of the tdMCP-eGFP constructs along with *mCherry-24xMBS*-tagged RNA, bright green puncta representing loci of active transcription were visualized (Fig. 3A-B'). These puncta, unlike previous aggregates, were easily identifiable being smaller and brighter. This demonstrates the tdMCP-eGFP expressed by both constructs interacts with *mCherry-24xMBS*-tagged RNAs in these cells. Importantly, these puncta display dynamic expression, indicating that the interactions between tdMCP-eGFP and *mCherry-24xMBS* tagged RNA are transient (Fig. 3C) making them suitable for studying transcriptional dynamics.

FIGURE 3

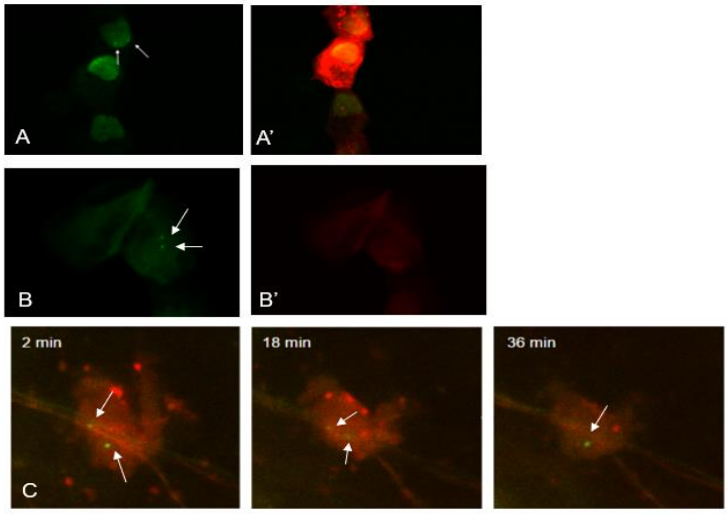

Figure 3. Dynamic tdMCP-eGFP+ puncta are observed in cells expressing MBS GFP+ puncta (arrows) in zebrafish cells co-expressing (A) *βactin: NLS-tdMCP-eGFP* (A206K) or (B) *βactin: tdMCP-eGFP* (no *NLS*) labeled in green with (A'&B') *βactin-mCherry-24xMBS* labeled in red. (C) Dynamic fluorescence intensities indicative of *βactin* -driven transcription are visible within a live zebrafish cell expressing both mCherry-24xMBS and tdMCP-GFP in a time-lapsed movie.

These results offer improvements into the viability of studying transcriptional dynamics especially noise *in vivo* in real-time in a vertebrate organism. Additionally, stable zebrafish lines of constructs described in this work have been generated reducing variability from transient expression and enabling study of later developmental timepoints. This opens up the possibility of investigating further the roles and regulation of transcriptional noise in development. A new study in Drosophila using the MS2-RNA labeling system has revealed insights into regulation of transcriptional noise by shadow enhancers in suppressing transcription factor noise (Waymack et al., 2020). With continuous technological development, there is promise that comparable studies can be performed in vertebrate systems especially in combination with site-directed transgenesis into single landing sites to reduce site integration variability (Mosimann et al., 2013).

Materials and methods:

Zebrafish husbandry and transgenic generation

All animals used in this study were raised and handled in accordance with the guidelines of the Institutional Animal Care and Use Committee at University of California, Irvine. AB strain embryos were collected from natural crosses, raised at 28.5°C in embryo medium (EM), and staged as previously described (Kimmel et al., 1995). To generate transgenics, wild-type AB embryos at 1-cell stage were injected with a 1 nl mixture of 25 ng/ul each of Tol2 Transposase RNA, either *pTol-βactin-tdMCP-eGFP* (no *NLS*) DNA or *pTol-βactin- NLS-tdMCP-eGFP*(A206K), and *pTol-βactin-mCherry-24xMBS* DNA (Campbell et al., 2015) for transcript dynamic studies.

Imaging

Zebrafish larvae were mounted live in 0.5% low melt agarose (Apex Bioresearch Products) in embryo media with tricaine. Samples were imaged on a Nikon C1 confocal with a 20x objective. Imaging processing was performed using ImageJ/Fiji software.

Molecular cloning

pTol-βactin-tdMCP-eGFP (no *NLS*) was generated from *pTol-βactin- NLS-tdMCP-eGFP* (Addgene #86248) through PCR mutagenesis using primers tdMCP_GFP_nlsdelF (5'GCG GCC GCA ATG GGC TAC CCC TAC GAC GTG CCC 3') and tdMCP_GFP_nlsdelR (5'GGG CAC GTC GTA GGG GTA GCC CAT TGC GGC CGC 3'). *pTol-βactin- NLS-tdMCP-eGFP*(A206K) was generated from *pTol-βactin- NLS-tdMCP-eGFP* (Addgene #86248) through PCR mutagenesis using primers mEGFP_AKmut_F (5' GCA CCC AGT CCA AGC TGA GCA AAG A 3') and mEGFP_AKmut_R (5' TCT TTG CTC AGC TTG GAC TGG GTG C 3'). Phusion High-Fidelity DNA Polymerase (NEB) was used for all PCR mutagenesis.

Chapter IV

Crabps and retinoic acid signaling regulate germ cell proliferation and play a role in sex determination in zebrafish

Introduction

Sex determination is a fundamental developmental process that varies greatly across organisms and involves both genetic and environmental determinants. The processes that regulate sex determination are also surprisingly plastic even within a species. In zebrafish (*Danio rerio*), wild strains have a ZZ/ZW chromosomal sex determination system with ZW animals primarily becoming female and ZZ animals becoming male (Sharma et al., 1998; Wilson et al., 2014). In contrast, laboratory strains (e.g. AB and TU) derived from domesticated animals have lost this chromosomal sex determination system (Wilson et al., 2014). Although the mechanisms of sex determination in these lab strains are still unclear, they are thought to involve contributions from both environmental determinants and multiple genetic loci.

The Vitamin-A derivative all-*trans*-retinoic acid (RA) is a cell-cell signaling molecule that plays an important role in sex determination in many vertebrate species. In mice, RA regulates initiation of meiosis in oocytes and promotes female sex differentiation through induction of Stimulated by RA gene 8 (*Stra8*), while an RA-degrading enzyme of the cytochrome p450 (Cyp) family prevents *Stra8* expression in embryonic and adult testes (Koubova et al., 2006). RA appears to function differently in zebrafish sex differentiation. Several fish species, including zebrafish, lack *Stra8* homologues in their genomes (Pasquier et al., 2016). During the bipotential stage in zebrafish (8-20 dpf), most cells throughout the gonad express aldehyde dehydrogenase 2 (*aldh1a2)*, the primary enzyme that catalyzes conversion of retinal to RA, likely generating high levels of RA throughout the tissue (Rodríguez-Marí et al., 2013). During

this time period, all zebrafish initially produce early-stage oocytes which mature and continue to be produced in females or are instead degraded in fish that will become males during sex differentiation (20-25 dpf) (Takahashi, 1977). This process is strongly influenced by the number of oocytes present at the bipotential gonad stage, since mutants with complete loss or reductions in oocytes become male (Dranow et al., 2016; Rodríguez-Marí et al., 2010; Siegfried & Nüsslein-Volhard, 2008; Slanchev et al., 2005). During sex differentiation, zebrafish express *cyp26a1* in a sexually dimorphic manner in the gonad by upregulating somatic expression in males and downregulating expression in females, resulting in low and high RA levels in male and female gonads respectively (Rodríguez-Marí et al., 2013). Prior studies have focused primarily on RA at these later stages of sex differentiation and earlier roles for RA in gonad development prior to oogenesis have not been described.

Cellular retinoic acid binding proteins (Crabps) bind RA with high affinity and transport it intracellularly. Vertebrates have two classes of highly conserved Crabps, Crabp1s and Crabp2s, which are responsible for transporting RA to Cyps for degradation in the cytosol as well as to RA-receptors (RARs) in the nucleus (Aström et al., 1991; Budhu et al., 2001; Budhu & Noy, 2002; Delva et al., 1999). Mammalian CRABP1 and CRABP2 show variable tissue-specific expression, with CRABP2 being more commonly expressed in tissues that synthesize RA. CRABP2 has also been shown to transport RA to RARs *in vitro* in addition to Cyps for degradation, while CRABP1 has been shown to primarily transport RA to Cyp26s for degradation (Boylan & Gudas, 1992; Fiorella & Napoli, 1994; Won et al., 2004). Consequently, in general, only elevating CRABP2 levels increases the transcription rate of RA responsive genes, while CRABP1 does not (Napoli, 2017). Zebrafish have four *crabp* orthologues, *crabp1a/b* and *crabp2a/b*, as a consequence of the whole genome duplication in the common ancestor of teleosts (Amores et al., 1998; Gates et al., 1999; Postlethwait et al., 1998). Our previous work studying knockdown of these genes in zebrafish morphants showed that among these orthologues, *crabp2a* is uniquely RA responsive, attenuates noise and promotes

robustness in RA levels during patterning of hindbrain rhombomeres (Cai et al., 2012; Sosnik et al., 2016). Mice lacking both Crabp1 and Crabp2 are viable but have supernumerary digits on the forelimb at low penetrance (Lampron et al., 1995), suggesting compensation for Crabp functions in RA signaling by other proteins (Romand et al., 2000). Thus, despite their extremely high conservation in all vertebrates, there is little evidence for essential functions *in vivo*.

Here, we show essential roles for zebrafish Crabp2s in gonad development and sex determination. Combined loss-of-function of both Crabp2a and Crabp2b leads to a dramatic increase in the proportion of males, which correlates with reduced gonad size, germ cell proliferation and ultimately germ cell number during early gonad development. In contrast, combined loss-of-function of both Crabp1a and Crabp1b results in an increased proportion of females and increased germ cell proliferation suggesting distinct and opposing roles of Crabp1s and Crabp2s in this context. Exogenous RA treatments at bipotential gonad stages promote germ cell proliferation, suggesting a direct role for RA in early gonad development and sex determination.

Results

Crabp2s promotes female sex determination

To investigate requirements for Crabps in zebrafish, we used CRISPR-Cas9-mediated gene editing to generate loss-of-function mutants for all four *crabp* genes (Fig. 1A). Upon generation of out-of-frame deletion mutant alleles for each gene, we did not observe any discernible phenotypes in single mutants aside from mild early embryonic and adult scoliosis defects present in *crabp1b-/-* mutants (supplementary material Fig. S1) When challenged with exogenous RA, *crabp2a-/-* mutants did not display the sensitized phenotype previously described (Cai et al., 2012). Upon analysis of RNA fold expression of *crabps* in mutants, we observed almost two-fold increase of expression in the remaining paralogue (supplementary material Fig. S2). We hypothesized that the absence of phenotypes in single mutants was due to paralogue compensation.

To address this, we produced double mutants for both Crabp1s and Crabp2s to rule out genetic compensation by the remaining paralogue. Both Crabp double mutants were viable and survived to adulthood, so we further generated maternal-zygotic homozygous double mutants for each (*crabp1a*$^{-/-}$*/crabp1b*$^{-/-}$ and *crabp2a*$^{-/-}$*/crabp2b*$^{-/-}$), which were both viable and fertile. Upon investigation of these mutants, we did not observe some of the hindbrain patterning defects found in Crabp2a morphants. This could be due to genetic compensation triggered by mutant mRNA degradation (El-Brolosy et al., 2019). Strikingly though, sex ratios in *crabp2a*$^{-/-}$*/crabp2b*$^{-/-}$ mutants were dramatically shifted such that nearly all adults appeared to be males, which was not observed in *crabp1a*$^{-/-}$*/crabp1b*$^{-/-}$ mutants or in wild-type (AB) (Fig. 1B). These data suggested that Crabp2s have essential roles in female sex determination and/or differentiation and the maintenance of proper sex ratios.

To assess possible causes of the skewed sex ratio, we examined gonads at different developmental stages. *crabp2a*$^{-/-}$*/crabp2b*$^{-/-}$ double mutants had notably smaller gonads with fewer germ cells (GCs) as marked by the GC-specific marker Ddx4/Vasa when compared to

wild-type animals at the bipotential stage at 12 dpf (Fig. 1C-D). These differences persisted through sex differentiation and into adulthood (Fig. 1E-F). Adult *crabp2a*$^{-/-}$*/crabp2b*$^{-/-}$ mutant males have testes that appeared identical to those of wild-type animals (Fig. 1G-H). This was not surprising as crabp*2a*$^{-/-}$*/crabp2b*$^{-/-}$ mutants are fertile. Based on these results, the differences in gonad size and GC number strongly suggests that Crabp2s have a role in regulating early gonad development.

FIGURE 1

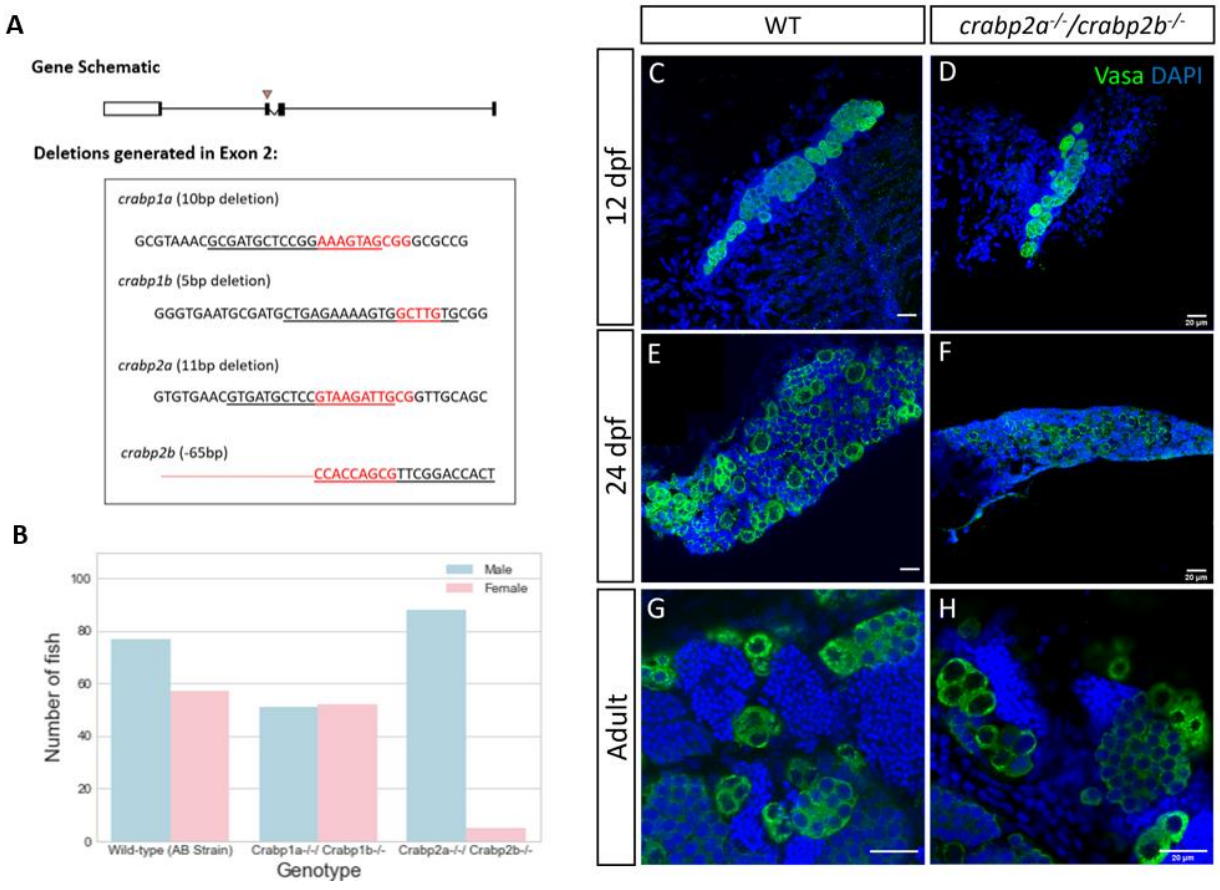

Figure 1. Crabp2s mutants are disproportionately male and have smaller gonads (A) Gene schematic of *crabp2a* with exon 2 indicated by orange arrow. Sequences of *crabp* mutant alleles depicting gRNA target sites (underlined) and deleted regions (red) in exon 2 of *crabp1a, crabp1b, crabp2a, and crabp2b.* (B) Histogram displaying sex ratios in wild-type (male n=77, female n=57), *crabp1a*$^{-/-}$*/crabp1b*$^{-/-}$ (male n=51, female n=52), and *crabp2a*$^{-/-}$*/crabp2b*$^{-/-}$ (male n=88, female n=5) animals. (C-H) Representative confocal z-projections of 12 dpf (C-D), 24 dpf (E-F), and adult (G-H) gonads in wild-type and *crabp2a*$^{-/-}$*/crabp2b*$^{-/-}$ mutant animals. Germ cells labeled by Ddx4/Vasa (green). Nuclei labeled with DAPI (blue). Scale bars = 20 μm.

Activation of RA signaling in germ cells of the bipotential gonad

As Crabps are intracellular transporters of RA and mediate various aspects of RA signaling, these results indicate that RA signaling during these early stages is important for gonad development. While we observe smaller gonads and lower number of germ cells in mutants, it is still unclear whether these are secondary effects from changes in RA signaling in somatic cells or in the GCs themselves. To investigate whether loss of GCs or skewed sex ratios are a result of GC autonomous or non-autonomous effects, we first looked at RA responses in cells of larval and juvenile zebrafish gonads using a RA response element transgenic reporter line, Tg(RARE-gata2a:NLS-YFP)^{ID1} (Perz-Edwards et al., 2001). Interestingly, we observed strong RARE-YFP expression specifically in GCs at the early bipotential gonad stage at 12 dpf (Fig. 2A-C). Also consistent with earlier reported expression (Rodríguez-Marí et al., 2013), elevated RARE:YFP expression persisted through sex differentiation at 23 dpf (Fig. 2D-F). These observations indicate a role for RA signaling in GC development.

FIGURE 2

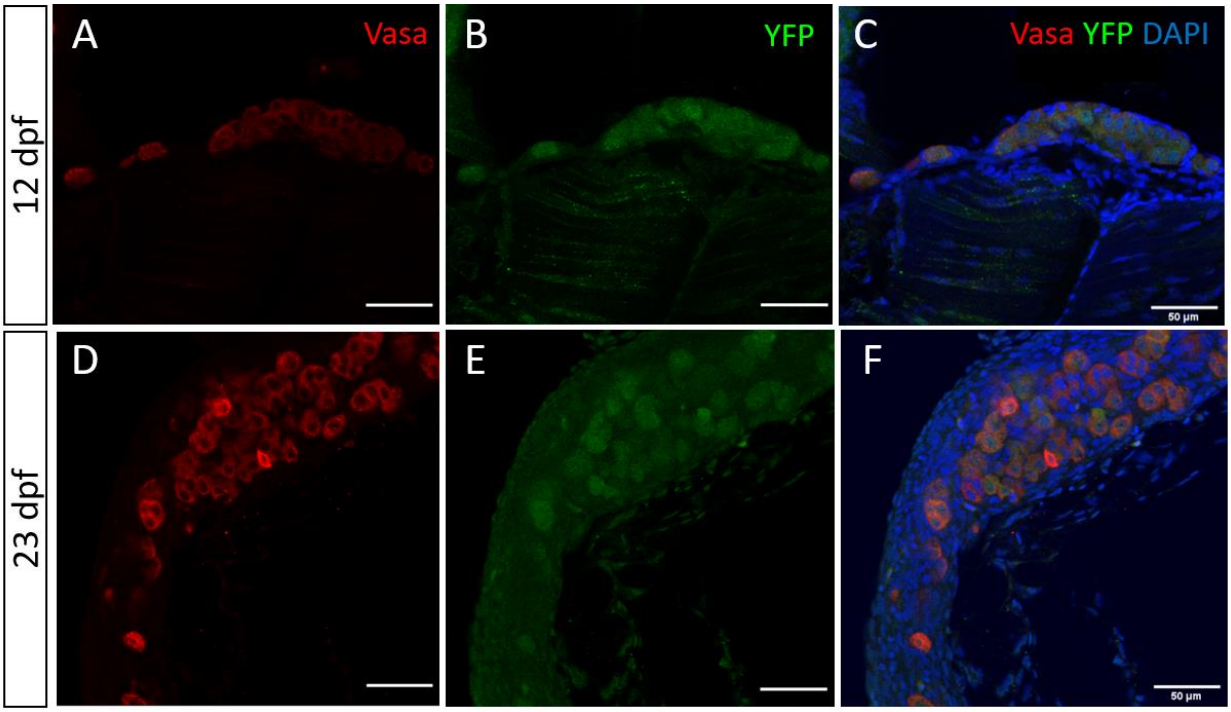

Figure 2. Germ cells are responsive to retinoic acid during early gonad development (A-F) Representative confocal z-projections of 12 dpf (A-C) and 23 dpf (D-F) gonads in *Tg(RARE:YFP)* zebrafish. Relative RA levels labeled by YFP expression (green). Germ cells labeled by Ddx4/Vasa (red). Nuclei labeled with DAPI (blue). Scale bars = 50 µm.

Crabp2s promote germ cell proliferation

Previous studies have shown that higher GC numbers during the bipotential stage in zebrafish favor female sex determination while reductions or loss of GCs favor male sex determination (Leerberg et al., 2017; Tzung et al., 2015). Notably, laboratory zebrafish larvae tend to have a unimodal distribution of GCs in the population at 7 dpf but by 14 dpf a bimodal shift can be distinguished with larvae having higher GC numbers being more likely to develop as females (Tzung et al., 2015). RA promotes proliferation of human GC-like cells and dissected chick primordial GCs in cell culture (Tan et al., 2015; Yu et al., 2012), but this has not been tested *in vivo*. Given this connection between RA and cell proliferation, we hypothesized that

Crabp2s and RA are important for regulating GC proliferation during early stages of zebrafish gonad development.

To assess this, we quantified GC numbers in *crabp2a*$^{-/-}$/*crabp2b*$^{-/-}$ mutant gonads stained with Ddx4 and compared them to wild-type. We did not observe a significant difference at 7 dpf between *crabp2a*$^{-/-}$/*crabp2b*$^{-/-}$ mutant and wild-type animals, but by 12 dpf, *crabp2a*$^{-/-}$/*crabp2b*$^{-/-}$ mutant gonads had far fewer GCs compared to wild-type gonads (Fig. 3E). To assess if this decrease is due to reduced proliferation, we incubated 11 dpf larvae for 24 hours in 10mM BrdU, dissected gonads at 12 dpf and quantified the number of BrdU and Vasa double (+) (Fig. 3A-D). We found that the proportion of BrdU(+) GCs in *crabp2a*$^{-/-}$/*crabp2b*$^{-/-}$ mutant gonads was reduced compared to wild-type gonads (Fig. 3F). This indicates that loss of Crabp2s results in decreased GC proliferation early in gonad development causing low GC number during the biopotential stage.

FIGURE 3

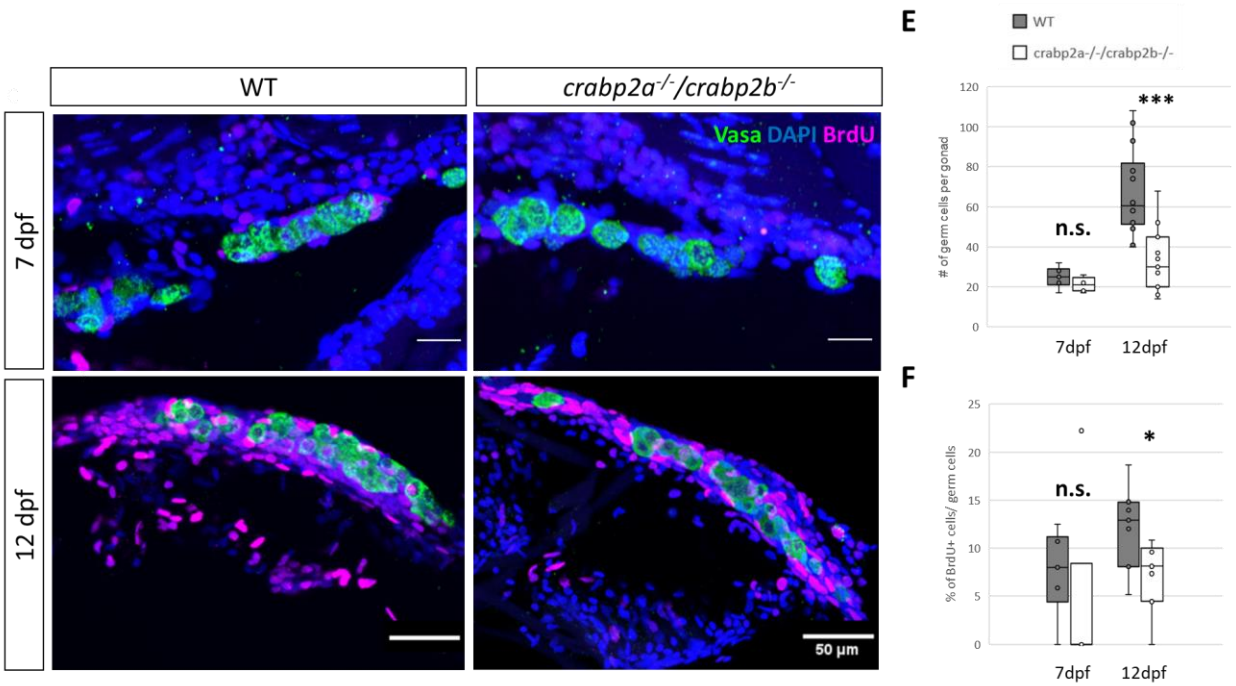

Figure 3. Crabp2 mutant gonads have fewer germ cells and decreased rate of proliferation (A-D) Representative confocal z-projections of BrdU incorporation at 7 dpf (A-B) and 12 dpf (C-D) gonads in wild-type and *crabp2a*$^{-/-}$*/crabp2b*$^{-/-}$ mutant animals. Germ cells are labeled with Ddx4/Vasa (green) and proliferating cells by BrdU (Magenta). (E) Box and whisker plots depicting the number of germ cells in a gonad at 7 dpf and 12 dpf. Each data point represents the total number of germ cells in one gonad. At 7 dpf, *crabp2a*$^{-/-}$*/crabp2b*$^{-/-}$ mutants and wild-type animals have similar germ cell numbers (mean ± SEM = 21.17 ± 1.42, 24.83 ± 2.09, respectively; P = 0.1775). At 12 dpf, *crabp2a*$^{-/-}$*/crabp2b*$^{-/-}$ mutants have a significantly lower number of germ cells compared to wild-type (mean ± SEM = 33 ± 3.85, 67.5 ± 5.79, respectively; P = 0.0001). (F) Box and whisker plots depicting the percentage of proliferating germ cells in a gonad at 7 dpf and 12 dpf. Each data point represents the number of BrdU(+) germ cells in one gonad. At 7 dpf, *crabp2a*$^{-/-}$*/crabp2b*$^{-/-}$ mutants and wild-type animals have similar proportions of BrdU(+) germ cells (mean = 4.34% and 7.52%, respectively; P = 0.4507). At 12 dpf, *crabp2a*$^{-/-}$*/crabp2b*$^{-/-}$ mutants have a significantly lower proportion of BrdU(+) germ cells compared to wild-type (mean = 7.20% and 12.22%, respectively; P = 0.0424). (Unpaired two-tailed t-test; * = P < .05; *** = P < .0001; n.s. = no significance). For box and whisker plots, genotypes are indicated for wild-type by gray bars and *crabp2a*$^{-/-}$*/crabp2b*$^{-/-}$ by white bars. Nuclei labeled with DAPI (blue). Scale bars = 50 μm.

Novel roles for RA in germ cell proliferation during early gonad development

Based on these findings, we hypothesize that reduction in RA signaling in *crabp2a*$^{-/-}$*/crabp2b*$^{-/-}$ mutants causes reduction in GC proliferation suggesting that RA positively regulates GC cell division. To test this, we treated wild-type zebrafish larvae with either DMSO vehicle alone or 0.5 μM RA starting from 8 dpf. Larvae treatment groups were also incubated in 10mM BrdU the day prior to their fixation at 10 or 12 dpf to assess proliferation. We expect to see greater GC number and proliferation in RA treated fish. (Fig. 4A-D). Both DMSO- and RA-treated larvae had reduced numbers of GCs and BrdU incorporation compared to untreated fish. This is likely due to the larvae being kept in dark conditions to prevent RA degradation and possibly less food intake, as the medium is changed daily several hours after their daily feeding of live rotifers. Despite this, we did not detect a significant difference in GC number or proliferation between either condition at 10 dpf however, we observed a significant increase in GC number and proliferation in 0.5 μM RA-treated larvae at 12 dpf (Fig. 4E-F). These results point to a novel role for RA in promoting germ cell proliferation at early stages of gonad development. These data are consistent with our hypothesis that the loss of the Crabp2s in

zebrafish causes decreased RA signaling in GCs resulting in lower levels of GC proliferation and subsequently smaller gonad size, thus promoting predominantly male development (Fig. 4G).

FIGURE 4

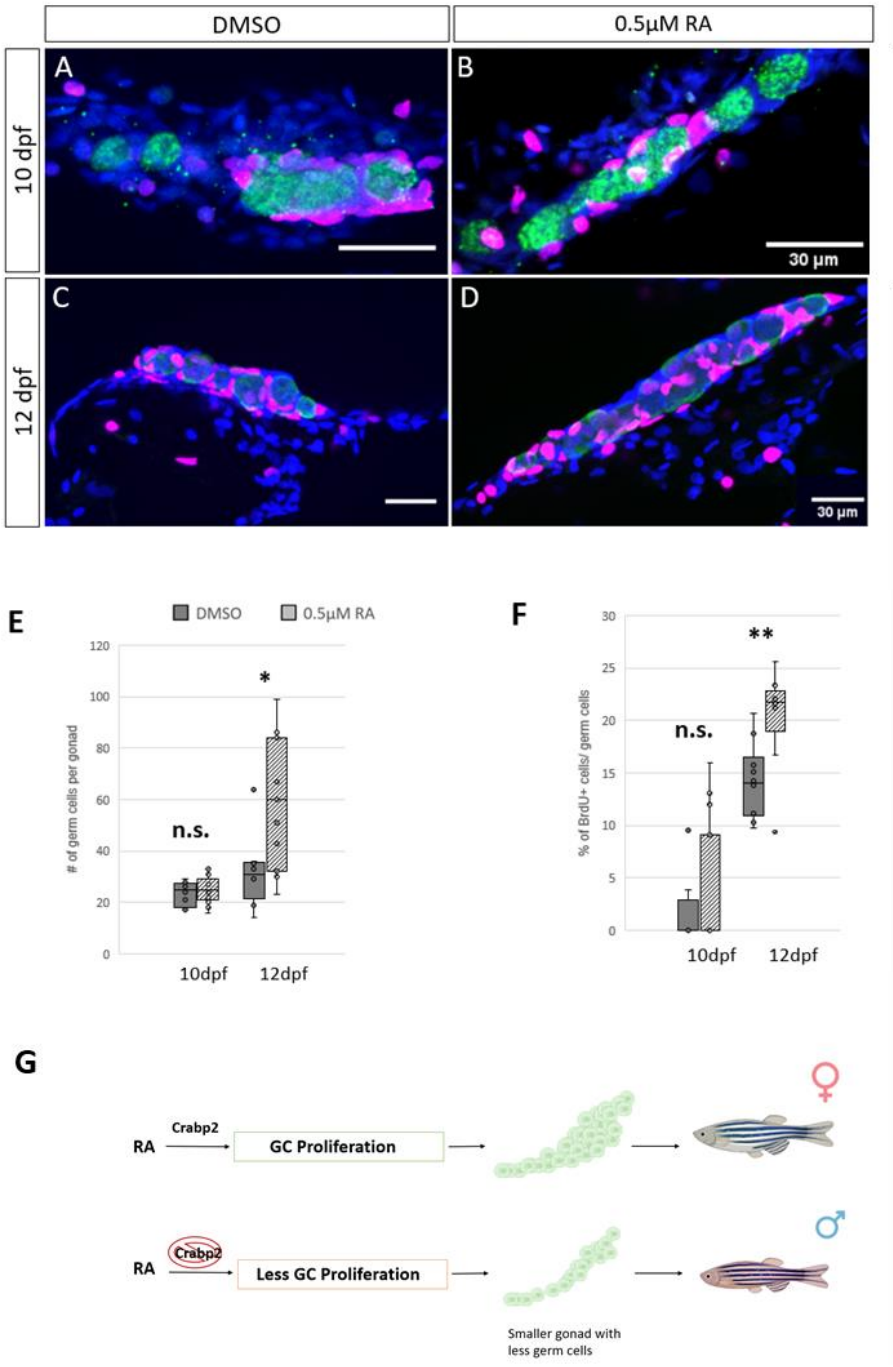

Figure 4. Retinoic acid promotes germ cell proliferation during gonad development (A-D) (A-D) Representative confocal z-projections of BrdU incorporation at 9-10 dpf (A-B) and 11-12 dpf (C-D) in gonads of DMSO vehicle (A, C) or 0.5 µM RA (B, D) treated animals. Germ cells are labeled with Ddx4/Vasa (green) and proliferating cells by BrdU (Magenta). (E) Box and whisker plots depicting the number of germ cells in a gonad at 10 dpf and 12 dpf. Each data point represents the total number of germ cells in one gonad. At 10 dpf, DMSO- and 0.5 µM RA-treated animals had similar germ cell numbers (mean ± SEM = 23.50 ± 1.66, 25.00 ± 1.25, respectively; P = 0.4930). At 12 dpf, DMSO-treated animals have a significantly lower number of germ cells compared to 0.5 µM RA treated animals (mean ± SEM = 32.38 ± 5.28, 58.45 ± 7.56, respectively; P = 0.0182). (F) Box and whisker plots depicting the percentage of proliferating germ cells in a gonad at 10 dpf and 12 dpf. Each data point represents the number of BrdU(+) germ cells in one gonad. At 10 dpf, the majority of both DMSO and 0.5 µM RA-treated animals have no BrdU(+) cells (mean = 1.67% and 3.48%, respectively; P = 0.4209). At 12 dpf, DMSO treated animals have a significantly lower proportion of BrdU(+) germ cells compared to 0.5 µM RA-treated animals (mean = 14.11% and 20.45%, respectively; P = 0.0044). (Unpaired two-tailed t-test; ** = P < .01; n.s. = no significance).(G) Model schematic for role of RA and Crabp2s on regulating germ cell development and sex differentiation (Created with BioRender.com). For box and whisker plots, treatment conditions are indicated for DMSO by solid gray bars and 0.5µM RA by diagonal striped bars. Nuclei labeled with DAPI (blue). Scale bars = 30 µm.

Crabp1s play an opposing role to Crabp2s in RA regulation and germ cell proliferation

As opposed to Crabp2s who transport RA to both receptors and degradation enzymes, Crabp1s have been primarily shown to transport RA to degradation enzymes. With absence of Crabp1s, it would be hypothesized that intracellular levels of free RA would be increased due to reduced transport to degradation enzymes. Supporting this, *crabp1b-/-* mutants have early embryonic phenotypes similar to embryos treated with exogenous RA (supplementary material Fig. S2). In the sex ratio analysis of double Crabp mutants (Fig. 1B), the proportion of females was increased in *crabp1a*$^{-/-}$*/crabp1b*$^{-/-}$ mutants compared to WT suggesting Crabp1s may promote male sex determination as opposed to female sex determination like Crabp2s.

After observing that exogenous RA promotes GC proliferation, we questioned whether Crabp1 mutants would have a similar phenotype. Since hypothetically if Crabp1s were absent and thus transport of RA to degradation enzymes decreased, it should also increase intracellular RA levels thus increasing GC proliferation. To examine this, we examined GC number and proliferation in *crabp1a*$^{-/-}$*/crabp1b*$^{-/-}$ mutant gonads at 11 dpf and compared them to both WT and *crabp2a*$^{-/-}$*/crabp2b*$^{-/-}$ mutant gonads (Fig. 5A-E). While *crabp1a*$^{-/-}$*/crabp1b*$^{-/-}$

mutants did not have significantly higher GC number or proliferation compared to WT, this could be due to low number of animals examined (Fig. 5D-E). Even so, the average level of GC proliferation tends to be higher in *crabp1a*$^{-/-}$*/crabp1b*$^{-/-}$ mutants compared to wild-type suggesting Crabps1 repress cell proliferation (Fig. 5D-E). Additionally, we found *crabp1a*$^{-/-}$ */crabp1b*$^{-/-}$ mutants had significantly higher GC number and proliferation compared to *crabp2a*$^{-/-}$ */crabp2b*$^{-/-}$ mutants suggesting distinct and opposing roles of Crabp1s and Crabp2s in GC proliferation and sex determination (Fig. 5F).

FIGURE 5

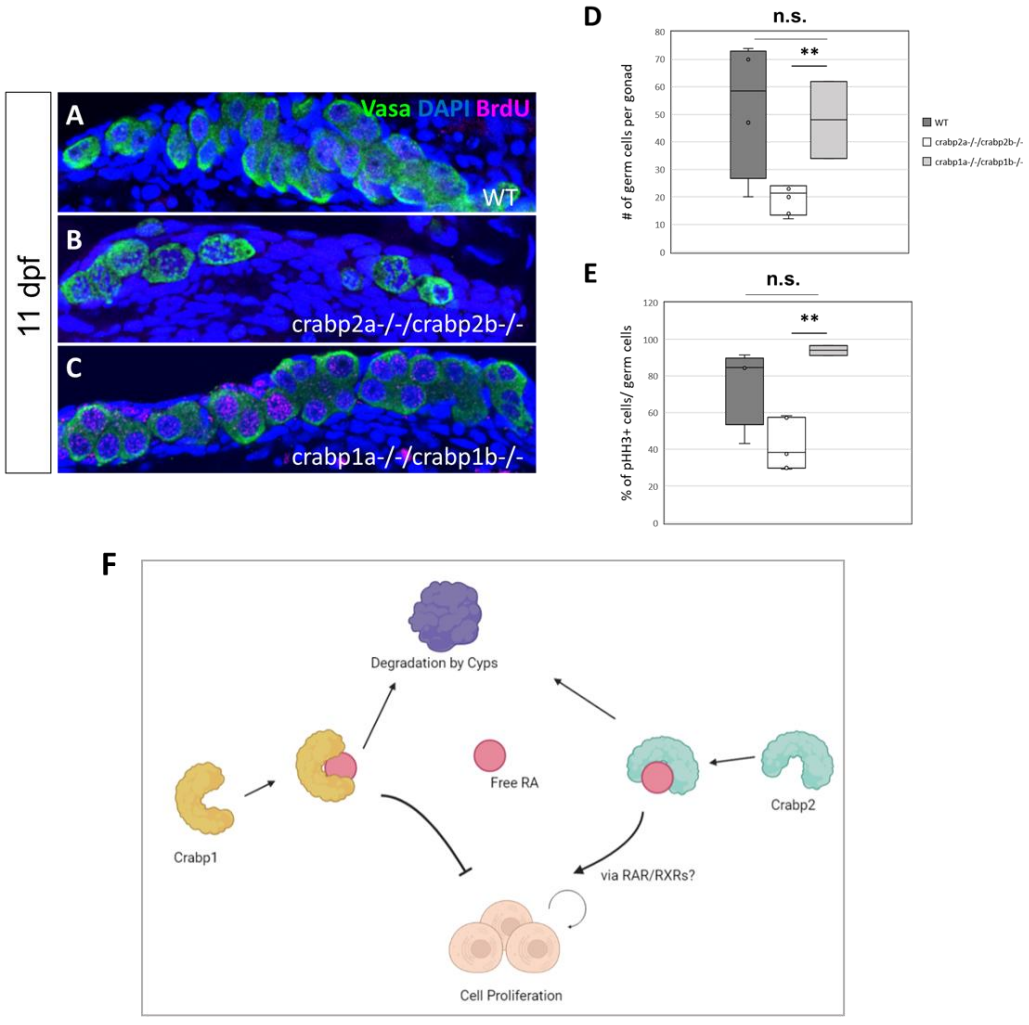

Figure 5. Crabp1s mutants have lower germ cell numbers and proliferation (A-C) Representative confocal images of antibody staining against phospho-Histone H3 at 11 dpf in gonads of (A) wild-type, (B) *crabp2a*$^{-/-}$*/crabp2b*$^{-/-}$, (C) *crabp1a*$^{-/-}$*/crabp1b*$^{-/-}$ animals. Germ cells are labeled with Ddx4/Vasa (green) and proliferating cells by pHH3 (magenta). (D) Box and whisker plots depicting the number of germ cells in a gonad at 11 dpf. Each data point represents the total number of germ cells in one gonad. *Crabp1a*$^{-/-}$*/crabp1b*$^{-/-}$ mutants and wild-type animals have similar germ cell numbers (mean ± SEM = 48.00 ± 14, 52.75 ± 12.43, respectively; P = 0.8283). *crabp1a*$^{-/-}$*/crabp1b*$^{-/-}$ animals have a significantly higher number of germ cells compared to *crabp2a*$^{-/-}$*/crabp2b*$^{-/-}$animals (mean ± SEM = 19.50 ± 2.16; P = 0.01). (E) Box and whisker plots depicting the percentage of proliferating germ cells in a gonad at 11 dpf. Each data point represents the number of pHH3(+) germ cells in one gonad. *Crabp1a*$^{-/-}$*/crabp1b*$^{-/-}$ mutants and wild-type animals have similar proportions of pHH3(+) germ cells. (mean = 93.96% and 76.00%, respectively; P = 0.3411). *Crabp1a*$^{-/-}$*/crabp1b*$^{-/-}$ animals have a significantly higher significantly lower proportion of pHH3(+) germ cells compared to *crabp2a*$^{-/-}$*/crabp2b*$^{-/-}$ animals (mean = 41.88%; P = 0.0017). (Unpaired two-tailed t-test; ** = P < .01; n.s. = no significance). (F) Model schematic of role of Crabps in regulating RA and cell proliferation (Created with BioRender.com)**.** For box and whisker plots, genotypes are indicated for wild-type by dark gray bars, *crabp2a*$^{-/-}$*/crabp2b*$^{-/-}$ by white bars, and *crabp1a*$^{-/-}$*/crabp1b*$^{-/-}$ by light gray bars. Nuclei labeled with DAPI (blue).

FIGURE 6

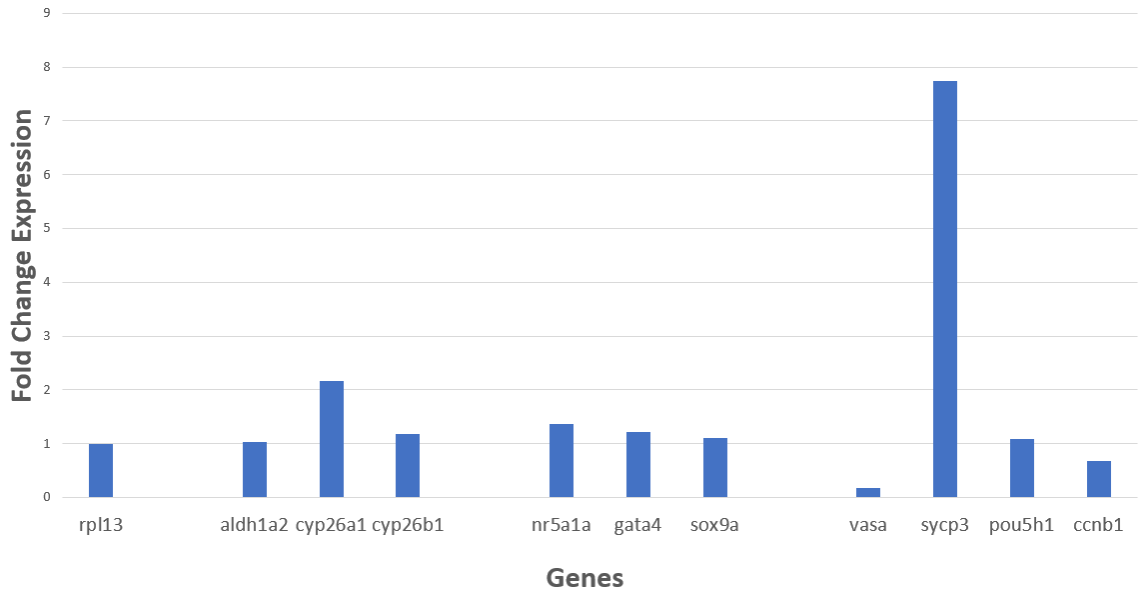

Figure 6. Gene expression changes in Crabp2 Mutants Fold change expression of select gene transcripts in Crabp2 mutants compared to wild-type. For each sample, four zebrafish larvae at 15 dpf were pooled for quantitative polymerase chain reaction (qPCR).

Lastly, to assess what RA pathways or targets Crabps may regulate during this developmental stage, we performed qPCR on whole *crabp2a*$^{-/-}$/*crabp2b*$^{-/-}$ mutant fish at 15 dpf, focusing on genes involved in RA synthesis or degradation (*aldh1a2*, *cyp26a1*, *cyp26b1*), genes implicated in early gonad development (*nr5a1a/Sf1*, *gata4*, *sox9a*) (Hu et al., 2013; Kashimada et al., 2011; Luo et al., 1994; Rodríguez-Marí et al., 2005; Sun et al., 2013), and genes that have roles in GC meiosis or proliferation (*vasa*, *sycp3*, *pou5h1/Oct4*, and *ccnb1*) (Hartung et al., 2014; Kehler et al., 2004; Strauss et al., 2018; Yuan et al., 2000) (Fig. 6). We observed slight upregulation of *cyp26a1*, an RA degradation enzyme which is induced by RA in early embryos (White et al., 2010) and upregulated in male differentiating somatic cells (Rodríguez-Marí et al., 2013). We detected only slight changes in RA target genes and the germ line stem cell marker *pou5h1*. These genes are also expressed in a variety of other tissue types, so specific effects on GC development and cell states may be obscured in whole fish gene expression analysis. We observed a significant decrease in expression of the GC marker *ddx4*/*vasa,* consistent with our data that crabp*2a*$^{-/-}$/*crabp2b*$^{-/-}$ mutants have fewer germ cells than WT. Also of note was the downregulation of cyclin b1 (*ccnb1)*, which is normally upregulated in female zebrafish and important for oocyte maturation (King et al., 2020). Surprisingly, we also observed a dramatic increase in expression of synaptonemal complex protein 3 (sycp3), which has an important role in meiosis. This might suggest that Crabp2s regulate the balance between GC proliferation and differentiation, though this requires additional studies.

Discussion

In summary, our data reveals novel and critical roles for Crabps and RA in GCs during early gonad development in zebrafish. Our work indicates Crabp2-mediated RA signaling promotes GC proliferation which in turn influences the number of oocytes produced during the bipotential stage, the number of which is a critical factor in maintaining female development and balanced sex ratios. Our findings also suggest that Crabp1s may have an opposing role in GC development, instead suppressing GC proliferation. This could be either through lowering intracellular RA levels by expediting transport of RA to degradation enzymes or by mediating non-canonical activation of ERK1/2 by RA as previously reported in stem cells of other tissues (Lin et al., 2017; Persaud et al., 2013). Future studies are needed to elucidate specific roles of individual Crabps in RA signaling in early GC development.

Additionally, our work reveals a novel role of RA in early gonad development that seems to function specifically during the period where GC number shifts from a unimodal distribution at 7 dpf to bimodal distribution at 14 dpf (Tzung et al., 2015). Also, as RA is a dietary derived metabolite from Vitamin-A and food availability is well-known to influence sex ratios in zebrafish (Lawrence et al., 2007) this could be one of the mechanisms that contribute to that environmental influence. These results emphasize the importance of dietary metabolites and the genes that regulate them in development. Overall, our work highlights yet another genetic mechanism that contributes to the complex mosaic of germ cell development and sex determination in zebrafish.

Materials and Methods

Generation of mutants

To generate mutants, we used the CRISPR/Cas9 system in zebrafish as previously described (Jao et al., 2013). Wild-type AB embryos were injected at the 1-cell stage with Cas9 mRNA and gRNAs designed using CHOPCHOPv2 web tool (Labun et al., 2019) for *crabp2a* (Gene ID: 171480) and *crabp2b* (Gene ID: 503502). gRNAs were transcribed using the MEGAshortscript T7 Transcription Kit (Invitrogen AM1354).

For primers used in study, see supplementary materials Table 1.

Retinoic acid and BrdU treatments

Wild-type or mutant larvae were treated with all-trans RA (Sigma-Aldrich R2625) or the vehicle DMSO dissolved in fish water at specified concentrations. Treatments began at 8 days postfertilization (dpf) in the dark with treatment medium being replaced daily until fixation with 4% Paraformaldehyde (PFA). For BrdU treatment, 24 hours prior to fixation larvae were transferred to a 0.5% DMSO solution of 10mM BrdU (Sigma-Aldrich B-5002) in fish water.

Immunohistochemistry

Larvae were euthanized with tricaine (Sigma-Aldrich MS-222) and fixed with 4% PFA at 4°C overnight. Samples were washed in PBS-DT (1% DMSO, 0.5% Triton x-100) and blocked with either 5% donkey or goat serum in PBS-DT. Primary antibodies and stains in this study were used at the following concentrations: Rabbit anti-Ddx4/Vasa 1:2000 (Knaut et al., 2000), Mouse anti-BrdU 1:500-1:1000 (Sigma-Aldrich B-2531), Rat anti-pHH3 1:500 (Sigma-Aldrich H9908) and DAPI nuclei stain 1:1000.

Imaging

Gonads were dissected and whole-mounted in 100% glycerol onto glass slides for imaging after dehydration in progressively higher glycerol concentrations. Confocal imaging was performed on a Leica SP8 confocal with a 40x water immersion objective. Imaging processing was performed using ImageJ/Fiji software.

Chapter V

Conclusions and Future Directions

One of the most essential aspects of development is that it is robust. From the top down, structures form from individual to individual of a species in largely the same way, morphogen gradients are reproducibly generated, and how even individual cells can be reliably specified is a wonder. These processes also show resilience despite large sources of variation from intrinsic stochasticity and external environmental influences. In my work, I describe mechanisms that drive robust responses to RA at various levels from the tissue wide patterning events by morphogen gradients to individual cells in regulating responses to morphogen levels which all culminate to big decisions that have a large effect on developmental outcomes.

Mechanisms across a tissue: morphogen gradients and morphogenesis

During tissue patterning, cells are able to determine their specific position through interactions with morphogens and cell-cell interactions. This is important in the hindbrain where segments are discrete sizes with razor sharp boundaries. To look further into this topic in Chapter II, we decided to look at hindbrain segmentation and boundary sharpening through a computational biology lens. Through computational modeling, we were able to examine and modulate specific parameters that would be almost impossible to change individually biologically to address questions of why specific events happen as they do during this biological process? This brings us back to the concept of canalization first proposed by Waddington. Natural selection is a driving factor that limits variability in decisions in development making these processes more robust.

This is the case in the zebrafish hindbrain, where we learned the rationale for induction of a second local FGF gradient by Hoxb1a in not only specifying additional segments, but in

also making the system more robust to variations in initial *hoxb1a* levels. Developmental processes have on many levels these mechanisms of positive feedback that inherently support robustness. It is fascinating how these processes tie together different aspects in both gene regulation and cell sorting to create synergistic effects. These findings can be applied to other tissues that are also patterned similarly by multiple morphogens. Additionally, in the hindbrain other morphogens are present such as Wnt, so further investigations are needed to understand more comprehensively all the morphogen gradient interactions in the tissue.

We also saw insights into how morphogenesis orchestrates tissue wide patterning events. The zebrafish hindbrain undergoes rapid convergent-extension during a specific window early in hindbrain segmentation. This specific rapid rate of elongation of the hindbrain during this period was found to be critical for improving boundary sharpening and segment size. It accomplished this in two ways: by strengthening cell-cell interactions increasing cell sorting and by creating a steeper gradient of RA for more efficient cell switching. Morphogenesis has largely been neglected in favor of studying molecular mechanisms that control patterning and cell fate specification. Yet this shows these processes are inextricably tied together, with temporal timing of morphogenesis being vital for coordinating different molecular mechanisms in hindbrain segmentation. Future studies are required to investigate roles of morphogenesis in other tissues and of different coordinated cell movements such as those along the D-V axis.

Mechanisms at the cellular level: morphogen transport and signal mediation by intracellular binding proteins

Next to shift gears and look at mechanisms that promote robust responses to RA at the cellular level**,** I examine in Chapter IV the role of individual cellular retinoic acid binding proteins in regulating responses to RA. In this work, I was able to show the essential roles of Crabp2s in mediating RA signaling and promoting germ cell proliferation. This was in contrast to Crabp1, which though structurally similar and has shared functions, has an opposite function in germ

cells by repressing cell proliferation. These specific functions of individual Crabps are shown to have larger consequences by promoting specific sex fates, influencing organisms on the population level by regulating sex ratios.

Also, while these findings may seem relevant only in this particular sex determination context, in many other organism including mammals, early onset of RA signaling and meiosis in germ cells is a fundamental event in female sex determination (Koubova et al., 2006; Bowles & Koopman, 2007). This indicates RA regulation is still important in sex determination, just that other systems gain or lose different regulatory mechanisms based on selection factors. Additionally, it is to be noted that Crabps are extremely well conserved in vertebrates and expressed in most tissues where RA signaling occurs, pointing to obvious evolutionary advantages to Crabp regulation of RA.

Future studies are still required to understand how cellular responses to RA are regulated in not only germ cells but in other cell types and tissues as well. This is only further complicated by the myriad of effects RA has on different cell types during different stages of development. In cell differentiation, RA in different contexts can lead to stem cell renewal or differentiation (Niederreither & Dollé, 2008). In cancer, RA has different effects on inhibiting or promoting growth based on the specific cell type and context (Mezquita & Mezquita, 2019). Studying Crabps can help decipher this puzzle, with specific Crabps having been already identified as biomarkers for different prognosis in specific cancers (Xiao et al., 2014; Liu et al., 2015; Celestino et al., 2018; Kim et al., 2018).

In summary, it is clear based on previous studies and my own findings that Crabp regulation of RA is complex. Differential expression and regulation of Crabps in different tissues only further amplifies this complexity. This complexity is likely one of many factors that drives the diverse yet consistent responses of cells to RA. Understanding how Crabps are regulated and how they regulate RA could potentially allow us to differentiate both the various responses of cells to RA and the amplitude of response highlighting their importance for future studies.

Mechanisms at the molecular level: transcriptional noise

Finally, previous studies have established the importance of stochastic noise and its regulation in development (Elowitz et al., 2002; Kaern et al., 2005; Kepler & Elston, 2001). In hindbrain segmentation and boundary sharpening, transcriptional noise in segmentation genes was found to have a positive role by giving cells along the boundaries a mechanism to accurately interpret noisy RA signals (Zhang et al., 2012). Studies in vertebrates on gene expression noise have been limited due to lack of existing technology suitable for this purpose. In chapter III, I focused on developing tools to look at transcriptional noise by improving the MS2 RNA-labeling system in zebrafish. This system will open up the possibilities of many future studies not only on the role of noise during patterning events, but also on the factors regulating gene transcription whether it be the basal level of transcription at different temperatures or by specific morphogens like RA. This can also be used to study roles of specific cis-regulatory elements and their spatial and temporal regulation of gene expression. Even now, existing technology is most weak in terms of ability to study events temporally, and from the sum of work in my thesis it is clear how important temporal timing is in development.

With a topic as complex and wide-ranging as developmental robustness, it is always essential to investigate the topic on various levels and from many different perspectives. In my thesis, I looked into what mechanisms drive robust responses to retinoic acid in development from the tissue to the molecular level. While the findings may seem tangentially related at first, common themes arise such as the importance of developmental timing and diversity in generating robust outcomes. These themes will form the basis of future studies that will lead us to understanding further just how development is so reliably robust.

References

Abu-Abed, S., Dollé, P., Metzger, D., Beckett, B., Chambon, P., & Petkovich, M. (2001). The retinoic acid-metabolizing enzyme, CYP26A1, is essential for normal hindbrain patterning, vertebral identity, and development of posterior structures. *Genes & Development*, *15*(2), 226–240. https://doi.org/10.1101/GAD.855001

Addison, M., Xu, Q., Cayuso, J., & Wilkinson, D. G. (2018). Cell Identity Switching Regulated by Retinoic Acid Signaling Maintains Homogeneous Segments in the Hindbrain. *Developmental Cell*, *45*(5), 606-620.e3. https://doi.org/10.1016/j.devcel.2018.04.003

Ahrens, M. J., Li, Y., Jiang, H., & Dudley, A. T. (2009). Convergent extension movements in growth plate chondrocytes require gpi-anchored cell surface proteins. *Development*, *136*(20), 3463–3474.

Alexander, T., Nolte, C., & Krumlauf, R. (2009). Hox genes and segmentation of the hindbrain and axial skeleton. *Annual Review of Cell and Developmental Biology*, *25*, 431–456. https://doi.org/10.1146/ANNUREV.CELLBIO.042308.113423

Amores, A., Force, A., Yan, Y. L., Joly, L., Amemiya, C., Fritz, A., … Postlethwait, J. H. (1998). Zebrafish hox clusters and vertebrate genome evolution. *Science (New York, N.Y.)*, *282*(5394), 1711–1714. https://doi.org/10.1126/SCIENCE.282.5394.1711

Aström, A., Tavakkol, A., Pettersson, U., Cromie, M., Elder, J. T., & Voorhees, J. J. (1991). Molecular cloning of two human cellular retinoic acid-binding proteins (CRABP). Retinoic acid-induced expression of CRABP-II but not CRABP-I in adult human skin in vivo and in skin fibroblasts in vitro. *The Journal of Biological Chemistry*, *266*(26), 17662–17666. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1654334

Baker, R. E., Schnell, S., & Maini, P. K. (2008). Mathematical models for somite formation. *Current Topics in Developmental Biology*, *81*, 183–203.

Balaskas, N., Ribeiro, A., Panovska, J., Dessaud, E., Sasai, N., Page, K. M., … Ribes, V. (2012). Gene regulatory logic for reading the Sonic Hedgehog signaling gradient in the vertebrate neural tube. *Cell*, *148*(1–2), 273–284.

Balmer, J. E., & Blomhoff, R. (2002). Gene expression regulation by retinoic acid. *Journal of Lipid Research*, *43*(11), 1773–1808. https://doi.org/10.1194/JLR.R100015-JLR200

Barrow, J. R., Stadler, H. S., & Capecchi, M. R. (2000). Roles of Hoxa1 and Hoxa2 in patterning the early hindbrain of the mouse. *Development (Cambridge, England)*, *127*(5), 933–944. https://doi.org/10.1242/DEV.127.5.933

Bastie, J. N., Despouy, G., Balitrand, N., Rochette-Egly, C., Chomienne, C., & Delva, L. (2001). The novel co-activator CRABPII binds to RARalpha and RXRalpha via two nuclear receptor interacting domains and does not require the AF-2 “core.” *FEBS Letters*, *507*(1), 67–73. https://doi.org/10.1016/S0014-5793(01)02938-6

Begemann, G., Schilling, T. F., Rauch, G. J., Geisler, R., & Ingham, P. W. (2001). The zebrafish neckless mutation reveals a requirement for raldh2 in mesodermal signals that pattern the hindbrain. *Development (Cambridge, England)*, *128*(16), 3081–3094. https://doi.org/10.1242/DEV.128.16.3081

Ben-Zvi, D., Pyrowolakis, G., Barkai, N., & Shilo, B.-Z. (2011). Expansion-repression mechanism for scaling the Dpp activation gradient in Drosophila wing imaginal discs. *Current Biology*, *21*(16), 1391–1396.

Bothma, J. P., Garcia, H. G., Esposito, E., Schlissel, G., Gregor, T., & Levine, M. (2014). Dynamic regulation of eve stripe 2 expression reveals transcriptional bursts in living Drosophila embryos. *Proceedings of the National Academy of Sciences*, *111*(29), 10598–10603. https://doi.org/10.1073/pnas.1410022111

Bouchoucha, Y. X., Reingruber, J., Labalette, C., Wassef, M. A., Thierion, E., Dinh, C. D., … Charnay, P. (2013). Dissection of a Krox20 positive feedback loop driving cell fate choices in hindbrain patterning. *Molecular Systems Biology*, *9*(1).

Bowles, J., & Koopman, P. (2007). Retinoic acid, meiosis and germ cell fate in mammals. *Development (Cambridge, England)*, *134*(19), 3401–3411. https://doi.org/10.1242/DEV.001107

Boylan, J. F., & Gudas, L. J. (1992). The level of CRABP-I expression influences the amounts and types of all-trans-retinoic acid metabolites in F9 teratocarcinoma stem cells. *The Journal of Biological Chemistry*, *267*(30), 21486–21491. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1328234

Briscoe, J., & Small, S. (2015). Morphogen rules: design principles of gradient-mediated embryo patterning. *Development*, *142*(23), 3996–4009.

Buckles, G. R., Thorpe, C. J., Ramel, M.-C., & Lekven, A. C. (2004). Combinatorial Wnt control of zebrafish midbrain–hindbrain boundary formation. *Mechanisms of Development*, *121*(5), 437–447.

Budhu, A., Gillilan, R., & Noy, N. (2001). Localization of the RAR interaction domain of cellular retinoic acid binding protein-II. *Journal of Molecular Biology*, *305*(4), 939–949. https://doi.org/10.1006/jmbi.2000.4340

Budhu, A. S., & Noy, N. (2002). Direct channeling of retinoic acid between cellular retinoic acid-binding protein II and retinoic acid receptor sensitizes mammary carcinoma cells to retinoic acid-induced growth arrest. *Molecular and Cellular Biology*, *22*(8), 2632–2641. https://doi.org/10.1128/MCB.22.8.2632-2641.2002

Cai, A. Q., Radtke, K., Linville, A., Lander, A. D., Nie, Q., & Schilling, T. F. (2012). Cellular retinoic acid-binding proteins are essential for hindbrain patterning and signal robustness in zebrafish. *Development (Cambridge)*, *139*(12), 2150–2155. https://doi.org/10.1242/DEV.077065/-/DC1

Calzolari, S., Terriente, J., & Pujades, C. (2014). Cell segregation in the vertebrate hindbrain relies on actomyosin cables located at the interhombomeric boundaries. *The EMBO Journal*, *33*(7), 686. https://doi.org/10.1002/EMBJ.201386003

Campbell, P. D., Chao, J. A., Singer, R. H., & Marlow, F. L. (2015). Dynamic visualization of transcription and RNA subcellular localization in zebrafish. *Development*, *142*(7), 1368–1374. https://doi.org/10.1242/dev.118968

Čanić, S. (2021). Moving boundary problems. *Bulletin of the American Mathematical Society*, *58*(1), 79–106.

Celestino, R., Nome, T., Pestana, A., Hoff, A. M., Gonçalves, A. P., Pereira, L., … Soares, P. (2018). CRABP1, C1QL1 and LCN2 are biomarkers of differentiated thyroid carcinoma, and predict extrathyroidal extension. *BMC Cancer*, *18*(1), 1–16. https://doi.org/10.1186/S12885-017-3948-3/TABLES/4

Chen, A. C., Yu, K., Lane, M. A., & Gudas, L. J. (2003). Homozygous deletion of the CRABPI gene in AB1 embryonic stem cells results in increased CRABPII gene expression and decreased intracellular retinoic acid concentration. *Archives of Biochemistry and Biophysics*, *411*(2), 159–173. https://doi.org/10.1016/S0003-9861(02)00732-4

Cheng, C. N., Li, Y., Marra, A. N., Verdun, V., & Wingert, R. A. (2014). Flat mount preparation for observation and analysis of zebrafish embryo specimens stained by whole mount in situ hybridization. *JoVE (Journal of Visualized Experiments)*, (89), e51604.

Choe, S.-K., Zhang, X., Hirsch, N., Straubhaar, J., & Sagerström, C. G. (2011). A screen for hoxb1-regulated genes identifies ppp1r14al as a regulator of the rhombomere 4 Fgf-signaling center. *Developmental Biology*, *358*(2), 356–367.

Choi, H. M. T., Schwarzkopf, M., Fornace, M. E., Acharya, A., Artavanis, G., Stegmaier, J., … Pierce, N. A. (2018). Third-generation in situ hybridization chain reaction: Multiplexed, quantitative, sensitive, versatile, robust. *Development (Cambridge)*, *145*(12). https://doi.org/10.1242/DEV.165753/48466

Chou, C.-S., Lo, W.-C., Gokoffski, K. K., Zhang, Y.-T., Wan, F. Y. M., Lander, A. D., … Nie, Q. (2010). Spatial dynamics of multistage cell lineages in tissue stratification. *Biophysical Journal*, *99*(10), 3145–3154.

Cohen, M., Page, K. M., Perez-Carrasco, R., Barnes, C. P., & Briscoe, J. (2014). A theoretical framework for the regulation of Shh morphogen-controlled gene expression. *Development*, *141*(20), 3868–3878.

Cooke, J. E., Moens, C. B., Roth, L. W. A., Durbin, L., Shiomi, K., Brennan, C., … Holder, N. (2001). Eph signalling functions downstream of Val to regulate cell sorting and boundary formation in the caudal hindbrain. *Development*, *128*(4), 571–580. https://doi.org/10.1242/DEV.128.4.571

Cooke, J. E., Kemp, H. A., & Moens, C. B. (2005). EphA4 Is Required for Cell Adhesion and Rhombomere-Boundary Formation in the Zebrafish. *Current Biology*, *15*(6), 536–542. https://doi.org/10.1016/J.CUB.2005.02.019

Cooke, J. E., & Moens, C. B. (2002). Boundary formation in the hindbrain: Eph only it were simple…. *Trends in Neurosciences*, *25*(5), 260–267. https://doi.org/10.1016/S0166-2236(02)02134-3

Crick, F. (1970). Diffusion in Embryogenesis. *Nature 1970 225:5231*, *225*(5231), 420–422. https://doi.org/10.1038/225420a0

Del Corral, R. D., Olivera-Martinez, I., Goriely, A., Gale, E., Maden, M., & Storey, K. (2003). Opposing FGF and Retinoid Pathways Control Ventral Neural Pattern, Neuronal Differentiation, and Segmentation during Body Axis Extension. *Neuron*, *40*(1), 65–79. https://doi.org/10.1016/S0896-6273(03)00565-8

Delva, L., Bastie, J.-N., Rochette-Egly, C., Kraïba, R., Balitrand, N., Despouy, G., … Chomienne, C. (1999). Physical and functional interactions between cellular retinoic acid binding protein II and the retinoic acid-dependent nuclear complex. *Molecular and Cellular Biology*, *19*(10), 7158–7167. https://doi.org/10.1128/MCB.19.10.7158

Dong, D., Ruuska, S. E., Levinthal, D. J., & Noy, N. (1999). Distinct roles for cellular retinoic acid-binding proteins I and II in regulating signaling by retinoic acid. *The Journal of Biological Chemistry*, *274*(34), 23695–23698. https://doi.org/10.1074/JBC.274.34.23695

Dranow, D. B., Hu, K., Bird, A. M., Lawry, S. T., Adams, M. T., Sanchez, A., … Draper, B. W. (2016). Bmp15 Is an Oocyte-Produced Signal Required for Maintenance of the Adult Female Sexual Phenotype in Zebrafish. *PLOS Genetics*, *12*(9), e1006323. https://doi.org/10.1371/JOURNAL.PGEN.1006323

Du, H., Wang, Y., Haensel, D., Lee, B., Dai, X., & Nie, Q. (2018). Multiscale modeling of layer formation in epidermis. *PLoS Computational Biology*, *14*(2), e1006006.

Dupé, V., Matt, N., Garnier, J. M., Chambon, P., Mark, M., & Ghyselinck, N. B. (2003). A newborn lethal defect due to inactivation of retinaldehyde dehydrogenase type 3 is prevented by maternal retinoic acid treatment. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(SUPPL. 2), 14036–14041. https://doi.org/10.1073/PNAS.2336223100/ASSET/2A9772E3-F062-4DE5-A0C7-7DE01B19F721/ASSETS/GRAPHIC/PQ2336223005.JPEG

Eaton, R. C., Lee, R. K. K., & Foreman, M. B. (2001). The Mauthner cell and other identified neurons of the brainstem escape network of fish. *Progress in Neurobiology*, *63*(4), 467–485. https://doi.org/10.1016/S0301-0082(00)00047-2

El-Brolosy, M. A., Kontarakis, Z., Rossi, A., Kuenne, C., Günther, S., Fukuda, N., … Stainier, D. Y. R. (2019). Genetic compensation triggered by mutant mRNA degradation. *Nature*, *568*(7751), 193–197. https://doi.org/10.1038/S41586-019-1064-Z

Eldar, A., Rosin, D., Shilo, B.-Z., & Barkai, N. (2003). Self-enhanced ligand degradation underlies robustness of morphogen gradients. *Developmental Cell*, *5*(4), 635–646.

Elowitz, M. B., Levine, A. J., Siggia, E. D., & Swain, P. S. (2002). Stochastic gene expression in a single cell. *Science*, *297*(5584), 1183–1186. https://doi.org/10.1126/SCIENCE.1070919/SUPPL_FILE/ELOWITZSOM.PDF

Exelby, K., Herrera-Delgado, E., Perez, L. G., Perez-Carrasco, R., Sagner, A., Metzis, V., … Briscoe, J. (2021). Precision of tissue patterning is controlled by dynamical properties of gene regulatory networks. *Development*.

Félix, M. A., & Wagner, A. (2006). Robustness and evolution: concepts, insights and challenges from a developmental model system. *Heredity 2008 100:2*, *100*(2), 132–140. https://doi.org/10.1038/sj.hdy.6800915

Fiorella, P. D., & Napoli, J. L. (1994). Microsomal retinoic acid metabolism. Effects of cellular retinoic acid-binding protein (type I) and C18-hydroxylation as an initial step. *Journal of Biological Chemistry*, *269*(14), 10538–10544. https://doi.org/10.1016/S0021-9258(17)34093-0

Fiorella, P. D., & Napoli, J. L. (1991). Expression of cellular retinoic acid binding protein (CRABP) in Escherichia coli. Characterization and evidence that holo-CRABP is a substrate in retinoic acid metabolism. *Journal of Biological Chemistry*, *266*(25), 16572–16579. https://doi.org/10.1016/S0021-9258(18)55339-4

Fraser, S., Keynes, R., & Lumsden, A. (1990). Segmentation in the chick embryo hindbrain is defined by cell lineage restrictions. *Nature 1990 344:6265*, *344*(6265), 431–435. https://doi.org/10.1038/344431a0

Fried, P., & Iber, D. (2014). Dynamic scaling of morphogen gradients on growing domains. *Nature Communications*, *5*(1), 1–12.

Fulton, T., Trivedi, V., Attardi, A., Anlas, K., Dingare, C., Arias, A. M., & Steventon, B. (2020). Axis Specification in Zebrafish Is Robust to Cell Mixing and Reveals a Regulation of Pattern Formation by Morphogenesis. *Current Biology*, *30*(15), 2984-2994. e3.

García-García, M. J., Shibata, M., & Anderson, K. V. (2008). Chato, a KRAB zinc-finger protein, regulates convergent extension in the mouse embryo. *Development*, *135*(18), 3053–3062.

Gates, M. A., Kim, L., Egan, E. S., Cardozo, T., Sirotkin, H. I., Dougan, S. T., … Talbot, W. S. (1999). A Genetic Linkage Map for Zebrafish: Comparative Analysis and Localization of Genes and Expressed Sequences. *Genome Research*, *9*(4), 334–347. https://doi.org/10.1101/GR.9.4.334

Giudicelli, F., Taillebourg, E., Charnay, P., & Gilardi-Hebenstreit, P. (2001). Krox-20 patterns the hindbrain through both cell-autonomous and non cell-autonomous mechanisms. *Genes & Development*, *15*(5), 567–580. https://doi.org/10.1101/GAD.189801

Gorry, P., Lufkin, T., Dierich, A., Rochette-Egly, C., Décimo, D., Dollé, P., … Chambon, P. (1994). The cellular retinoic acid binding protein I is dispensable. *Proceedings of the National Academy of Sciences of the United States of America*, *91*(19), 9032–9036. https://doi.org/10.1073/PNAS.91.19.9032

Häärä, O., Harjunmaa, E., Lindfors, P. H., Huh, S.-H., Fliniaux, I., Åberg, T., … Thesleff, I. (2012). Ectodysplasin regulates activator-inhibitor balance in murine tooth development through Fgf20 signaling. *Development*, *139*(17), 3189–3199.

Hartung, O., Forbes, M. M., & Marlow, F. L. (2014). Zebrafish vasa is required for germ-cell differentiation and maintenance. *Molecular Reproduction and Development*, *81*(10), 946. https://doi.org/10.1002/MRD.22414

Hasty, J., Pradines, J., Dolnik, M., & Collins, J. J. (2000). Noise-based switches and amplifiers forgene expression. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(5), 2075–2080. https://doi.org/10.1073/PNAS.040411297/ASSET/C408A388-45FA-44CA-AD5E-259623828722/ASSETS/GRAPHIC/PQ0404112003.JPEG

Hernandez, R. E., Rikhof, H. A., Bachmann, R., & Moens, C. B. (2004). vhnf1 integrates global RA patterning and local FGF signals to direct posterior hindbrain development in zebrafish. *Development (Cambridge, England)*, *131*(18), 4511–4520. https://doi.org/10.1242/DEV.01297

Higashijima, S. ichi, Okamoto, H., Ueno, N., Hotta, Y., & Eguchi, G. (1997). High-Frequency Generation of Transgenic Zebrafish Which Reliably Express GFP in Whole Muscles or the Whole Body by Using Promoters of Zebrafish Origin. *Developmental Biology*, *192*(2), 289–299. https://doi.org/10.1006/DBIO.1997.8779

Hiroya TAKAHASHI, B. (1977). Juvenile Hermaphroditism in the. *Bull. Fac. Fish. Hokkaido Univ*, *18*(2), 57–65.

Hocine, S., Raymond, P., Zenklusen, D., Chao, J. A., & Singer, R. H. (2013). Single-molecule analysis of gene expression using two-color RNA labeling in live yeast. *Nature Methods*, *10*(2), 119–121. https://doi.org/10.1038/nmeth.2305

Holloway, D. M., Lopes, F. J. P., da Fontoura Costa, L., Travençolo, B. A. N., Golyandina, N., Usevich, K., & Spirov, A. V. (2011). Gene expression noise in spatial patterning: hunchback promoter structure affects noise amplitude and distribution in Drosophila segmentation. *PLoS Computational Biology*, *7*(2). https://doi.org/10.1371/JOURNAL.PCBI.1001069

Hu, Y. C., Okumura, L. M., & Page, D. C. (2013). Gata4 Is Required for Formation of the Genital Ridge in Mice. *PLOS Genetics*, *9*(7), e1003629. https://doi.org/10.1371/JOURNAL.PGEN.1003629

Hubaud, A., & Pourquié, O. (2014). Signalling dynamics in vertebrate segmentation. *Nature Reviews Molecular Cell Biology*, *15*(11), 709–721.

Huebner, R. J., & Wallingford, J. B. (2018). Coming to consensus: a unifying model emerges for convergent extension. *Developmental Cell*, *46*(4), 389–396.

Jaeger, J. (2009). Modelling the Drosophila embryo. *Molecular BioSystems*, *5*(12), 1549–1568.

Jaeger, J., Surkova, S., Blagov, M., Janssens, H., Kosman, D., Kozlov, K. N., … Sharp, D. H. (2004). Dynamic control of positional information in the early Drosophila embryo. *Nature*, *430*(6997), 368–371.

Jao, L. E., Wente, S. R., & Chen, W. (2013). Efficient multiplex biallelic zebrafish genome editing using a CRISPR nuclease system. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(34), 13904–13909. https://doi.org/10.1073/PNAS.1308335110/SUPPL_FILE/PNAS.201308335SI.PDF

Kærn, M., Elston, T. C., Blake, W. J., & Collins, J. J. (2005). Stochasticity in gene expression: from theories to phenotypes. *Nature Reviews Genetics 2005 6:6*, *6*(6), 451–464. https://doi.org/10.1038/nrg1615

Kashimada, K., Svingen, T., Feng, C.-W., Pelosi, E., Bagheri-Fam, S., Harley, V. R., … Koopman, P. (2011). Antagonistic regulation of Cyp26b1 by transcription factors SOX9/SF1 and FOXL2 during gonadal development in mice. *The FASEB Journal*, *25*(10), 3561–3569. https://doi.org/10.1096/FJ.11-184333

Kastner, P., Mark, M., Ghyselinck, N., Krezel, W., Dupé, V., Grondona, J. M., & Chambon, P. (1997). Genetic evidence that the retinoid signal is transduced by heterodimeric RXR/RAR functional units during mouse development. *Development (Cambridge, England)*, *124*(2), 313–326. https://doi.org/10.1242/DEV.124.2.313

Kawakami, K. (2007). Tol2: A versatile gene transfer vector in vertebrates. *Genome Biology*, *8*(SUPPL. 1), 1–10. https://doi.org/10.1186/GB-2007-8-S1-S7/FIGURES/6

Kawakami, K., Koga, A., Hori, H., & Shima, A. (1998). Excision of the Tol2 transposable element of the medaka fish, Oryzias latipes, in zebrafish, Danio rerio. *Gene*, *225*(1–2), 17–22. https://doi.org/10.1016/S0378-1119(98)00537-X

Kehler, J., Tolkunova, E., Koschorz, B., Pesce, M., Gentile, L., Boiani, M., … Tomilin, A. (2004). Oct4 is required for primordial germ cell survival. *EMBO Reports*, *5*(11), 1078–1083. https://doi.org/10.1038/SJ.EMBOR.7400279

Keller, R., Shih, J., & Sater, A. (1992). The cellular basis of the convergence and extension of the Xenopus neural plate. *Developmental Dynamics : An Official Publication of the American Association of Anatomists*, *193*(3), 199–217. https://doi.org/10.1002/AJA.1001930302

Keller, R., & Tibbetts, P. (1989). Mediolateral cell intercalation in the dorsal, axial mesoderm of Xenopus laevis. *Developmental Biology*, *131*(2), 539–549.

Kepler, T. B., & Elston, T. C. (2001). Stochasticity in Transcriptional Regulation: Origins, Consequences, and Mathematical Representations. *Biophysical Journal*, *81*(6), 3116–3136. https://doi.org/10.1016/S0006-3495(01)75949-8

Kesseland, M., & Gruss, P. (1991). Homeotic transformations of murine vertebrae and concomitant alteration of Hox codes induced by retinoic acid. *Cell*, *67*(1), 89–104. https://doi.org/10.1016/0092-8674(91)90574-I

Kim, D. J., Kim, W. J., Lim, M., Hong, Y., Lee, S. J., Hong, S. H., … Han, S. S. (2018). Plasma CRABP2 as a Novel Biomarker in Patients with Non-Small Cell Lung Cancer. *Journal of Korean Medical Science*, *33*(26), e178. https://doi.org/10.3346/JKMS.2018.33.E178

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

Kimmel, C. B., Miller, C. T., Kruze, G., Ullmann, B., BreMiller, R. A., Larison, K. D., & Snyder, H. C. (1998). The shaping of pharyngeal cartilages during early development of the zebrafish. *Developmental Biology*, *203*(2), 245–263.

Kimmel, C. B., Warga, R. M., & Kane, D. A. (1994). Cell cycles and clonal strings during formation of the zebrafish central nervous system. *Development*, *120*(2), 265–276.

King, A. C., Gut, M., & Zenker, A. K. (2020). Shedding new light on early sex determination in zebrafish. *Archives of Toxicology*, *94*(12), 4143–4158. https://doi.org/10.1007/S00204-020-02915-Y/FIGURES/5

Kleiner-Bössaler, A., & DeLuca, H. F. (1971). Formation of retinoic acid from retinol in the kidney. *Archives of Biochemistry and Biophysics*, *142*(1), 371–377. https://doi.org/10.1016/0003-9861(71)90295-5

Knaut, H., Pelegri, F., Bohmann, K., Schwarz, H., & Nüsslein-Volhard, C. (2000). Zebrafish vasa RNA but Not Its Protein Is a Component of the Germ Plasm and Segregates Asymmetrically before Germline Specification. *The Journal of Cell Biology*, *149*(4), 875. https://doi.org/10.1083/JCB.149.4.875

Koubova, J., Menke, D. B., Zhou, Q., Cape, B., Griswold, M. D., & Page, D. C. (2006). Retinoic acid regulates sex-specific timing of meiotic initiation in mice. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(8), 2474–2479. https://doi.org/10.1073/PNAS.0510813103/SUPPL_FILE/10813FIG8.PDF

Kudoh, T., Wilson, S. W., & Dawid, I. B. (2002). Distinct roles for Fgf, Wnt and retinoic acid in posteriorizing the neural ectoderm. *Development*, *129*(18), 4335–4346. https://doi.org/10.1242/DEV.129.18.4335

Kulesa, P. M., & Fraser, S. E. (1998). Segmentation of the vertebrate hindbrain: a time-lapse analysis. *International Journal of Developmental Biology*, *42*(3), 385–392.

Labalette, C., Bouchoucha, Y. X., Wassef, M. A., Gongal, P. A., Men, J. Le, Becker, T., … Charnay, P. (2011). Hindbrain patterning requires fine-tuning of early krox20 transcription by Sprouty 4. *Development (Cambridge, England)*, *138*(2), 317–326. https://doi.org/10.1242/DEV.057299

Labalette, C., Wassef, M. A., Desmarquet-Trin Dinh, C., Bouchoucha, Y. X., Le Men, J., Charnay, P., & Gilardi-Hebenstreit, P. (2015). Molecular dissection of segment formation in the developing hindbrain. *Development (Cambridge, England)*, *142*(1), 185–195. https://doi.org/10.1242/dev.109652

Labun, K., Montague, T. G., Krause, M., Torres Cleuren, Y. N., Tjeldnes, H., & Valen, E. (2019). CHOPCHOP v3: expanding the CRISPR web toolbox beyond genome editing. *Nucleic Acids Research*, *47*(W1), W171–W174. https://doi.org/10.1093/NAR/GKZ365

Lampron, C., Rochette-Egly, C., Gorry, P., Dolle, P., Mark, M., Lufkin, T., … Chambon, P. (1995). Mice deficient in cellular retinoic acid binding protein II (CRABPII) or in both CRABPI and CRABPII are essentially normal. *Development (Cambridge, England)*, *121*(2), 539–548. https://doi.org/10.1242/DEV.121.2.539

Lander, A. D., Nie, Q., & Wan, F. Y. M. (2007). Membrane-associated non-receptors and morphogen gradients. *Bulletin of Mathematical Biology*, *69*(1), 33.

Lander, A. D. (2011). Pattern, growth, and control. *Cell*, *144*(6), 955–969. https://doi.org/10.1016/J.CELL.2011.03.009

Lander, A. D., Lo, W.-C., Nie, Q., & Wan, F. Y. M. (2009). The measure of success: constraints, objectives, and tradeoffs in morphogen-mediated patterning. *Cold Spring Harbor Perspectives in Biology*, *1*(1), a002022.

Lawrence, C., Ebersole, J. P., & Kesseli, R. V. (2007). Rapid growth and out-crossing promote female development in zebrafish (Danio rerio). *Environmental Biology of Fishes 2007 81:2*, *81*(2), 239–246. https://doi.org/10.1007/S10641-007-9195-8

Lecaudey, V., Anselme, I., Rosa, F., & Schneider-Maunoury, S. (2004). The zebrafish Iroquois gene iro7 positions the r4/r5 boundary and controls neurogenesis in the rostral hindbrain. *Development*, *131*(13), 3121–3131.

Leerberg, D. M., Sano, K., & Draper, B. W. (2017). Fibroblast growth factor signaling is required for early somatic gonad development in zebrafish. *PLOS Genetics*, *13*(9), e1006993. https://doi.org/10.1371/JOURNAL.PGEN.1006993

Lei, J., Wang, D., Song, Y., Nie, Q., & Wan, F. Y. M. (2013). ROBUSTNESS OF MORPHOGEN GRADIENTS WITH “BUCKET BRIGADE” TRANSPORT THROUGH MEMBRANE-ASSOCIATED NON-RECEPTORS. *Discrete and Continuous Dynamical Systems. Series B*, *18*(3), 721–739. https://doi.org/10.3934/DCDSB.2013.18.721

Liem, K. F., Jessell, T. M., & Briscoe, J. (2000). Regulation of the neural patterning activity of sonic hedgehog by secreted BMP inhibitors expressed by notochord and somites. *Development*, *127*(22), 4855–4866.

Lin, Y. L., Persaud, S. D., Nhieu, J., & Wei, L. N. (2017). Cellular Retinoic Acid–Binding Protein 1 Modulates Stem Cell Proliferation to Affect Learning and Memory in Male Mice. *Endocrinology*, *158*(9), 3004–3014. https://doi.org/10.1210/EN.2017-00353

Liu, R. Z., Garcia, E., Glubrecht, D. D., Poon, H. Y., Mackey, J. R., & Godbout, R. (2015). CRABP1 is associated with a poor prognosis in breast cancer: Adding to the complexity of breast cancer cell response to retinoic acid. *Molecular Cancer*, *14*(1), 1–16. https://doi.org/10.1186/S12943-015-0380-7/FIGURES/6

Lumsden, A. (2004). Segmentation and compartition in the early avian hindbrain. *Mechanisms of Development*, *121*(9), 1081–1088. https://doi.org/10.1016/J.MOD.2004.04.018

Lumsden, A., & Krumlauf, R. (1996). Patterning the vertebrate neuraxis. *Science (New York, N.Y.)*, *274*(5290), 1109–1115. https://doi.org/10.1126/SCIENCE.274.5290.1109

Luo, X., Ikeda, Y., & Parker, K. L. (1994). A cell-specific nuclear receptor is essential for adrenal and gonadal development and sexual differentiation. *Cell*, *77*(4), 481–490. https://doi.org/10.1016/0092-8674(94)90211-9

Marin, F., & Charnay, P. (2000). Hindbrain patterning: FGFs regulate Krox20 and mafB/kr expression in the otic/preotic region. *Development (Cambridge, England)*, *127*(22), 4925–4935. https://doi.org/10.1242/DEV.127.22.4925

Mark, M., Ghyselinck, N. B., & Chambon, P. (2006). FUNCTION OF RETINOID NUCLEAR RECEPTORS: Lessons from Genetic and Pharmacological Dissections of the Retinoic Acid Signaling Pathway During Mouse Embryogenesis. *Http://Dx.Doi.Org/10.1146/Annurev.Pharmtox.46.120604.141156*, *46*, 451–480. https://doi.org/10.1146/ANNUREV.PHARMTOX.46.120604.141156

Marshall, H., Nonchev, S., Sham, M. H., Muchamore, I., Lumsden, A., & Krumlauf, R. (1992). Retinoic acid alters hindbrain Hox code and induces transformation of rhombomeres 2/3 into a 4/5 identity. *Nature*, *360*(6406), 737–741. https://doi.org/10.1038/360737A0

Maves, L., Jackman, W., & Kimmel, C. B. (2002). FGF3 and FGF8 mediate a rhombomere 4 signaling activity in the zebrafish hindbrain. *Development (Cambridge, England)*, *129*(16), 3825–3837. https://doi.org/10.1242/DEV.129.16.3825

Maves, L., & Kimmel, C. B. (2005). Dynamic and sequential patterning of the zebrafish posterior hindbrain by retinoic acid. *Developmental Biology*, *285*(2), 593–605.

McGinnis, W., & Krumlauf, R. (1992). Homeobox genes and axial patterning. *Cell*, *68*(2), 283–302. https://doi.org/10.1016/0092-8674(92)90471-N

Meinecke, L., Sharma, P. P., Du, H., Zhang, L., Nie, Q., & Schilling, T. F. (2018). Modeling craniofacial development reveals spatiotemporal constraints on robust patterning of the mandibular arch. *PLoS Computational Biology*, *14*(11).

Mercader, N., Leonardo, E., Piedra, M. E., Martinez-A., C., Ros, M. A., & Torres, M. (2000). Opposing RA and FGF signals control proximodistal vertebrate limb development through regulation of Meis genes. *Development*, *127*(18), 3961–3970. https://doi.org/10.1242/DEV.127.18.3961

Mercier, P., Simeone, A., Cotelli, F., & Boncinelli, E. (2004). Expression pattern of two otx genes suggests a role in specifying anterior body structures in zebrafish. *International Journal of Developmental Biology*, *39*(4), 559–573.

Mezquita, B., & Mezquita, C. (2019). Two Opposing Faces of Retinoic Acid: Induction of Stemness or Induction of Differentiation Depending on Cell-Type. *Biomolecules*, *9*(10). https://doi.org/10.3390/BIOM9100567

Moens, C. B., & Prince, V. E. (2002). Constructing the hindbrain: insights from the zebrafish. *Developmental Dynamics : An Official Publication of the American Association of Anatomists*, *224*(1), 1–17. https://doi.org/10.1002/DVDY.10086

Mosimann, C., Puller, A. C., Lawson, K. L., Tschopp, P., Amsterdam, A., & Zon, L. I. (2013). Site-directed zebrafish transgenesis into single landing sites with the phiC31 integrase system. *Developmental Dynamics : An Official Publication of the American Association of Anatomists*, *242*(8), 949–963. https://doi.org/10.1002/DVDY.23989

Müller, P., Rogers, K. W., Jordan, B. M., Lee, J. S., Robson, D., Ramanathan, S., & Schier, A. F. (2012). Differential diffusivity of Nodal and Lefty underlies a reaction-diffusion patterning system. *Science*, *336*(6082), 721–724.

Nacu, E., Gromberg, E., Oliveira, C. R., Drechsel, D., & Tanaka, E. M. (2016). FGF8 and SHH substitute for anterior–posterior tissue interactions to induce limb regeneration. *Nature*, *533*(7603), 407–410.

Napoli, J. L. (2017). Cellular retinoid binding-proteins, CRBP, CRABP, FABP5: effects on retinoid metabolism, function and related diseases. *Pharmacology & Therapeutics*, *173*, 19. https://doi.org/10.1016/J.PHARMTHERA.2017.01.004

Newman, T. J. (n.d.). Modeling Multicellular Systems Using Subcellular Elements. *Mathematical Biosciences & Engineering*, *2*(3), 613.

Nie, Q., Qiao, L., Qiu, Y., Zhang, L., & Zhao, W. (2020). Noise control and utility: From regulatory network to spatial patterning. *Science China Mathematics 2020 63:3*, *63*(3), 425–440. https://doi.org/10.1007/S11425-019-1633-1

Niederreither, K., & Dollé, P. (2008). Retinoic acid in development: towards an integrated view. *Nature Reviews Genetics 2008 9:7*, *9*(7), 541–553. https://doi.org/10.1038/nrg2340

Niederreither, K., Vermot, J., Schuhbaur, B., Chambon, P., & Dollé, P. (2002). Embryonic retinoic acid synthesis is required for forelimb growth and anteroposterior patterning in the mouse. *Development*, *129*(15), 3563–3574. https://doi.org/10.1242/DEV.129.15.3563

Niederreither, K., Subbarayan, V., Dollé, P., & Chambon, P. (1999). Embryonic retinoic acid synthesis is essential for early mouse post-implantation development. *Nature Genetics*, *21*(4), 444–448. https://doi.org/10.1038/7788

Oxtoby, E., & Jowett, T. (1993). Cloning of the zebrafish krox-20 gene (krx-20) and its expression during hindbrain development. *Nucleic Acids Research*, *21*(5), 1087–1095.

Papan, C., & Campos-Ortega, J. A. (1994). On the formation of the neural keel and neural tube in the zebrafishDanio (Brachydanio) rerio. *Roux's Archives of Developmental Biology*, *203*(4), 178–186.

Parker, H. J., & Krumlauf, R. (2017). Segmental arithmetic: summing up the Hox gene regulatory network for hindbrain development in chordates. *Wiley Interdisciplinary Reviews: Developmental Biology*, *6*(6), e286.

Pasquier, J., Cabau, C., Nguyen, T., Jouanno, E., Severac, D., Braasch, I., … Bobe, J. (2016). Gene evolution and gene expression after whole genome duplication in fish: The PhyloFish database. *BMC Genomics*, *17*(1), 1–10. https://doi.org/10.1186/S12864-016-2709-Z/FIGURES/4

Persaud, S. D., Lin, Y. W., Wu, C. Y., Kagechika, H., & Wei, L. N. (2013). Cellular retinoic acid binding protein 1 mediates rapid non-canonical activation of ERK1/2 by all-trans retinoic acid. *Cellular Signalling*, *25*(1), 19. https://doi.org/10.1016/J.CELLSIG.2012.09.002

Perz-Edwards, A., Hardison, N. L., & Linney, E. (2001). Retinoic acid-mediated gene expression in transgenic reporter zebrafish. *Developmental Biology*, *229*(1), 89–101. https://doi.org/10.1006/DBIO.2000.9979

Pitulescu, M. E., & Adams, R. H. (2010). Eph/ephrin molecules—a hub for signaling and endocytosis. *Genes & Development*, *24*(22), 2480–2492.

Postlethwait, J. H., Yan, Y. L., Gates, M. A., Horne, S., Amores, A., Brownlie, A., … Talbot, W. S. (1998). Vertebrate genome evolution and the zebrafish gene map. *Nature Genetics*, *18*(4), 345–349. https://doi.org/10.1038/ng0498-345

Qiao, L., Zhao, W., Tang, C., Nie, Q., & Zhang, L. (2019). Network Topologies That Can Achieve Dual Function of Adaptation and Noise Attenuation. *Cell Systems*, *9*(3), 271-285. e7.

Qiu, Y., Chen, W., & Nie, Q. (2019). Stochastic dynamics of cell lineage in tissue homeostasis. *Discrete & Continuous Dynamical Systems-B*, *24*(8), 3971.

Qiu, Y., Fung, L., Schilling, T. F., & Nie, Q. (2021). Multiple morphogens and rapid elongation promote segmental patterning during development. *PLOS Computational Biology*, *17*(6), e1009077. https://doi.org/10.1371/JOURNAL.PCBI.1009077

Rackauckas, C., Schilling, T., & Nie, Q. (2018). Mean-independent noise control of cell fates via intermediate states. *IScience*, *3*, 11–20.

Raspopovic, J., Marcon, L., Russo, L., & Sharpe, J. (2014). Digit patterning is controlled by a Bmp-Sox9-Wnt Turing network modulated by morphogen gradients. *Science*, *345*(6196), 566–570.

Reifers, F., Bohli, H., Walsh, E. C., Crossley, P. H., Stainier, D. Y., & Brand, M. (1998). Fgf8 is mutated in zebrafish acerebellar (ace) mutants and is required for maintenance of midbrain-hindbrain boundary development and somitogenesis. *Development*, *125*(13), 2381–2395.

Rhinn, M., & Dollé, P. (2012). Retinoic acid signalling during development. *Development*, *139*(5), 843–858.

Rodríguez-Marí, A., Cañestro, C., BreMiller, R. A., Catchen, J. M., Yan, Y. L., & Postlethwait, J. H. (2013). Retinoic Acid Metabolic Genes, Meiosis, and Gonadal Sex Differentiation in Zebrafish. *PLoS ONE*, *8*(9), 73951. https://doi.org/10.1371/JOURNAL.PONE.0073951

Rodríguez-Marí, A., Cañestro, C., BreMiller, R. A., Nguyen-Johnson, A., Asakawa, K., Kawakami, K., & Postlethwait, J. H. (2010). Sex Reversal in Zebrafish fancl Mutants Is Caused by Tp53-Mediated Germ Cell Apoptosis. *PLOS Genetics*, *6*(7), e1001034. https://doi.org/10.1371/JOURNAL.PGEN.1001034

Rodríguez-Marí, A., Yan, Y. L., BreMiller, R. A., Wilson, C., Cañestro, C., & Postlethwait, J. H. (2005). Characterization and expression pattern of zebrafish anti-Müllerian hormone (amh) relative to sox9a, sox9b, and cyp19a1a, during gonad development. *Gene Expression Patterns*, *5*(5), 655–667. https://doi.org/10.1016/J.MODGEP.2005.02.008

Rogers, K. W., & Schier, A. F. (2011). Morphogen gradients: from generation to interpretation. *Annual Review of Cell and Developmental Biology*, *27*, 377–407. https://doi.org/10.1146/ANNUREV-CELLBIO-092910-154148

Romand, R., Sapin, V., Ghyselinck, N. B., Avan, P., Le Calvez, S., Dollé, P., … Mark, M. (2000). Spatio-temporal distribution of cellular retinoid binding protein gene transcripts in the developing and the adult cochlea. Morphological and functional consequences in CRABP- and CRBPI-null mutant mice. *The European Journal of Neuroscience*, *12*(8), 2793–2804. https://doi.org/10.1046/J.1460-9568.2000.00165.X

Rørth, P. (2012). Fellow travellers: emergent properties of collective cell migration. *EMBO Reports*, *13*(11), 984–991.

Sandell, L. L., Sanderson, B. W., Moiseyev, G., Johnson, T., Mushegian, A., Young, K., … Trainor, P. A. (2007). RDH10 is essential for synthesis of embryonic retinoic acid and is required for limb, craniofacial, and organ development. *Genes & Development*, *21*(9), 1113–1124. https://doi.org/10.1101/GAD.1533407

Schilling, T. F., Nie, Q., & Lander, A. D. (2012). Dynamics and precision in retinoic acid morphogen gradients. *Current Opinion in Genetics & Development*, *22*(6), 562–569. https://doi.org/10.1016/J.GDE.2012.11.012

Schilling, T. F., Prince, V., & Ingham, P. W. (2001). Plasticity in zebrafish hox expression in the hindbrain and cranial neural crest. *Developmental Biology*, *231*(1), 201–216.

Schumacher, L. J., Kulesa, P. M., McLennan, R., Baker, R. E., & Maini, P. K. (2016). Multidisciplinary approaches to understanding collective cell migration in developmental biology. *Open Biology*, *6*(6), 160056.

Sharma KK, Sharma OP, & Tripathi NK (1998). Female hetero-gamety in Danio rerio (Cypriniformes: Cyprinidae). Proceedings of the National Academy of Sciences India Section B, Biological Sciences, 68(B), 123–126.

Shen, M. M. (2007). Nodal signaling: developmental roles and regulation. *Development*, *134*(6), 1023–1034.

Shenefelt, R. E. (1972). Gross congenital malformations. Animal model: treatment of various species with a large dose of vitamin A at known stages in pregnancy. *The American Journal of Pathology*, *66*(3), 589. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2032719/

Shilo, B.-Z., Haskel-Ittah, M., Ben-Zvi, D., Schejter, E. D., & Barkai, N. (2013). Creating gradients by morphogen shuttling. *Trends in Genetics*, *29*(6), 339–347.

Shindo, A. (2018). Models of convergent extension during morphogenesis. *Wiley Interdisciplinary Reviews: Developmental Biology*, *7*(1), e293.

Siegfried, K. R., & Nüsslein-Volhard, C. (2008). Germ line control of female sex determination in zebrafish. *Developmental Biology*, *324*(2), 277–287. https://doi.org/10.1016/J.YDBIO.2008.09.025

Slanchev, K., Stebler, J., De La Cueva-Méndez, G., & Raz, E. (2005). Development without germ cells: The role of the germ line in zebrafish sex differentiation. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(11), 4074–4079. https://doi.org/10.1073/PNAS.0407475102/SUPPL_FILE/07475FIG6.JPG

Sokolowski, T. R., Erdmann, T., & Ten Wolde, P. R. (2012). Mutual repression enhances the steepness and precision of gene expression boundaries. *PLoS Computational Biology*, *8*(8).

Sosnik, J., Zheng, L., Rackauckas, C. V, Digman, M., Gratton, E., Nie, Q., … Nie, Q. (2016). Noise modulation in retinoic acid signaling sharpens segmental boundaries of gene expression in the embryonic zebrafish hindbrain. *ELife*, *5*, e14034. https://doi.org/10.7554/eLife.14034

Strauss, B., Harrison, A., Coelho, P. A., Yata, K., Zernicka-Goetz, M., & Pines, J. (2018). Cyclin B1 is essential for mitosis in mouse embryos, and its nuclear export sets the time for mitosis. *Journal of Cell Biology*, *217*(1), 179–193. https://doi.org/10.1083/JCB.201612147/VIDEO-10

Sulston, J. E., Schierenberg, E., White, J. G., & Thomson, J. N. (1983). The embryonic cell lineage of the nematode Caenorhabditis elegans. *Developmental Biology*, *100*(1), 64–119. https://doi.org/10.1016/0012-1606(83)90201-4

Sun, D., Zhang, Y., Wang, C., Hua, X., Zhang, X. A., & Yan, J. (2013). Sox9-related signaling controls zebrafish juvenile ovary–testis transformation. *Cell Death & Disease 2013 4:11*, *4*(11), e930–e930. https://doi.org/10.1038/cddis.2013.456

Sutherland, A., Keller, R., & Lesko, A. (2020). Convergent extension in mammalian morphogenesis. In *Seminars in cell & developmental biology* (Vol. 100, pp. 199–211). Elsevier.

Tan, H., Wang, J. J., Cheng, S. F., Ge, W., Sun, Y. C., Sun, X. F., … Shen, W. (2015). Retinoic acid promotes the proliferation of primordial germ cell-like cells differentiated from mouse skin-derived stem cells in vitro. *Theriogenology*, *85*(3), 408–418. https://doi.org/10.1016/J.THERIOGENOLOGY.2015.09.002

Terriente, J., & Pujades, C. (2015). Cell segregation in the vertebrate hindbrain: a matter of boundaries. *Cellular and Molecular Life Sciences : CMLS*, *72*(19), 3721–3730. https://doi.org/10.1007/S00018-015-1953-8

Thaller, C., & Eichele, G. (1987). Identification and spatial distribution of retinoids in the developing chick limb bud. *Nature*, *327*(6123), 625–628. https://doi.org/10.1038/327625a0

Theil, T., Frain, M., Gilardi-Hebenstreit, P., Flenniken, A., Charnay, P., & Wilkinson, D. G. (1998). Segmental expression of the EphA4 (Sek-1) receptor tyrosine kinase in the hindbrain is under direct transcriptional control of Krox-20. *Development*, *125*(3), 443–452.

Thierion, E., Le Men, J., Collombet, S., Hernandez, C., Coulpier, F., Torbey, P., … Gilardi-Hebenstreit, P. (2017). Krox20 hindbrain regulation incorporates multiple modes of cooperation between cis-acting elements. *PLoS Genetics*, *13*(7), e1006903.

Thisse, C., Thisse, B., Schilling, T. F., & Postlethwait, J. H. (1993). Structure of the zebrafish snail1 gene and its expression in wild-type, spadetail and no tail mutant embryos. *Development*, *119*(4), 1203–1215.

Tickle, C., Lee, J., & Eichele, G. (1985). A quantitative analysis of the effect of all-trans-retinoic acid on the pattern of chick wing development. *Developmental Biology*, *109*(1), 82–95. https://doi.org/10.1016/0012-1606(85)90348-3

To, T. L., & Maheshri, N. (2010). Noise can induce bimodality in positive transcriptional feedback loops without bistability. *Science*, *327*(5969), 1142–1145. https://doi.org/10.1126/SCIENCE.1178962/SUPPL_FILE/TO-SOM.PDF

Towers, M., & Tickle, C. (2009). Growing models of vertebrate limb development. *Development (Cambridge, England)*, *136*(2), 179–190. https://doi.org/10.1242/DEV.024158

Trainor, P., & Krumlauf, R. (2000). Plasticity in mouse neural crest cells reveals a new patterning role for cranial mesoderm. *Nature Cell Biology*, *2*(2), 96–102. https://doi.org/10.1038/35000051

Turing, A. M. (1952). The Chemical Basis of Morphogenesis. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *237*(641), 37–72. Retrieved from http://www.jstor.org/about/terms.html.

Tzung, K. W., Goto, R., Saju, J. M., Sreenivasan, R., Saito, T., Arai, K., … Orbán, L. (2015). Early Depletion of Primordial Germ Cells in Zebrafish Promotes Testis Formation. *Stem Cell Reports*, *4*(1), 61–73. https://doi.org/10.1016/J.STEMCR.2014.10.011

Uehara, M., Yashiro, K., Mamiya, S., Nishino, J., Chambon, P., Dolle, P., & Sakai, Y. (2007). CYP26A1 and CYP26C1 cooperatively regulate anterior-posterior patterning of the developing brain and the production of migratory cranial neural crest cells in the mouse. *Developmental Biology*, *302*(2), 399–411. https://doi.org/10.1016/J.YDBIO.2006.09.045

Venepally, P., Reddy, L. G., & Sani, B. P. (1996). Analysis of the effects of CRABP I expression on the RA-induced transcription mediated by retinoid receptors. *Biochemistry*, *35*(31), 9974–9982. https://doi.org/10.1021/BI9603276/ASSET/IMAGES/LARGE/BI9603276F00008.JPEG

Waddington, C. H. (1942). CANALIZATION OF DEVELOPMENT AND THE INHERITANCE OF ACQUIRED CHARACTERS. *Nature 1942 150:3811*, *150*(3811), 563–565. https://doi.org/10.1038/150563a0

Wallingford, J. B., Fraser, S. E., & Harland, R. M. (2002). Convergent Extension: The Molecular Control of Polarized Cell Movement during Embryonic Development. *Developmental Cell*, *2*(6), 695–706. https://doi.org/10.1016/S1534-5807(02)00197-1

Walshe, J., Maroon, H., McGonnell, I. M., Dickson, C., & Mason, I. (2002). Establishment of hindbrain segmental identity requires signaling by FGF3 and FGF8. *Current Biology : CB*, *12*(13), 1117–1123. https://doi.org/10.1016/S0960-9822(02)00899-0

Wang, Q., Holmes, W. R., Sosnik, J., Schilling, T., & Nie, Q. (2017). Cell sorting and noise-induced cell plasticity coordinate to sharpen boundaries between gene expression domains. *PLoS Computational Biology*, *13*(1).

Wang, Q., Oh, J. W., Lee, H.-L., Dhar, A., Peng, T., Ramos, R., … Cao, X. (2017). A multi-scale model for hair follicles reveals heterogeneous domains driving rapid spatiotemporal hair growth patterning. *Elife*, *6*.

Warga, R. M., & Kimmel, C. B. (1990). Cell movements during epiboly and gastrulation in zebrafish. *Development*, *108*(4), 569–580. https://doi.org/10.1242/DEV.108.4.569

Waskiewicz, A. J., Rikhof, H. A., & Moens, C. B. (2002). Eliminating zebrafish pbx proteins reveals a hindbrain ground state. *Developmental Cell*, *3*(5), 723–733.

Waymack, R., Fletcher, A., Enciso, G., & Wunderlich, Z. (2020). Shadow enhancers can suppress input transcription factor noise through distinct regulatory logic. *ELife*, *9*, 1–57. https://doi.org/10.7554/ELIFE.59351

White, R. J., Nie, Q., Lander, A. D., & Schilling, T. F. (2007). Complex regulation of cyp26a1 creates a robust retinoic acid gradient in the zebrafish embryo. *PLoS Biology*, *5*(11).

Wiellette, E. L., & Sive, H. (2003). vhnf1 and Fgf signals synergize to specify rhombomere identity in the zebrafish hindbrain. *Development*, *130*(16), 3821–3829.

Wilkinson, D. G. (2018). Establishing sharp and homogeneous segments in the hindbrain. *F1000Research*, *7*.

Wilson, C. A., High, S. K., McCluskey, B. M., Amores, A., Yan, Y. L., Titus, T. A., … Postlethwait, J. H. (2014). Wild Sex in Zebrafish: Loss of the Natural Sex Determinant in Domesticated Strains. *Genetics*, *198*(3), 1291–1308. https://doi.org/10.1534/GENETICS.114.169284

Wilson, J. G., Roth, C. B., & Warkany, J. (1953). An analysis of the syndrome of malformations induced by maternal vitamin A deficiency. Effects of restoration of vitamin A at various times during gestation. *The American Journal of Anatomy*, *92*(2), 189–217. https://doi.org/10.1002/AJA.1000920202

Wolpert, L. (1969). Positional information and the spatial pattern of cellular differentiation. *Journal of Theoretical Biology*, *25*(1), 1–47. https://doi.org/10.1016/S0022-5193(69)80016-0

Won, J. Y., Nam, E.-C., Yoo, S. J., Kwon, H. J., Um, S. J., Han, H. S., … Kim, S. Y. (2004). The effect of cellular retinoic acid binding protein-I expression on the CYP26-mediated catabolism of all-trans retinoic acid and cell proliferation in head and neck squamous cell carcinoma. *Metabolism*, *53*(8), 1007–1012. https://doi.org/10.1016/j.metabol.2003.12.015

Xiao, W., Hong, H., Awadallah, A., Yu, S., Zhou, L., & Xin, W. (2014). CRABP-II is a highly sensitive and specific diagnostic molecular marker for pancreatic ductal adenocarcinoma in distinguishing from benign pancreatic conditions. *Human Pathology*, *45*(6), 1177. https://doi.org/10.1016/J.HUMPATH.2014.01.014

Xu, Q., Mellitzer, G., Robinson, V., & Wilkinson, D. G. (1999). In vivo cell sorting in complementary segmental domains mediated by Eph receptors and ephrins. *Nature*, *399*(6733), 267–271.

Yabut, K. C. B., & Isoherranen, N. (2022). CRABPs Alter all-trans-Retinoic Acid Metabolism by CYP26A1 via Protein-Protein Interactions. *Nutrients 2022, Vol. 14, Page 1784*, *14*(9), 1784. https://doi.org/10.3390/NU14091784

Yashiro, K., Zhao, X., Uehara, M., Yamashita, K., Nishijima, M., Nishino, J., … Hamada, H. (2004). Regulation of retinoic acid distribution is required for proximodistal patterning and outgrowth of the developing mouse limb. *Developmental Cell*, *6*(3), 411–422. https://doi.org/10.1016/S1534-5807(04)00062-0

Yoshioka-Kobayashi, K., Matsumiya, M., Niino, Y., Isomura, A., Kori, H., Miyawaki, A., & Kageyama, R. (2020). Coupling delay controls synchronized oscillation in the segmentation clock. *Nature*, *580*(7801), 119–123.

Yu, M., Ge, C., Zeng, W., Mi, Y., & Zhang, C. (2012). Retinoic acid promotes proliferation of chicken primordial germ cells via activation of PI3K/Akt-mediated NF-κB signalling cascade. *Cell Biology International*, *36*(8), 705–712. https://doi.org/10.1042/CBI20110542

Yuan, L., Liu, J. G., Zhao, J., Brundell, E., Daneholt, B., & Höög, C. (2000). The Murine SCP3 Gene Is Required for Synaptonemal Complex Assembly, Chromosome Synapsis, and Male Fertility. *Molecular Cell*, *5*(1), 73–83. https://doi.org/10.1016/S1097-2765(00)80404-9

Zagorski, M., Tabata, Y., Brandenberg, N., Lutolf, M. P., Tkačik, G., Bollenbach, T., … Kicheva, A. (2017). Decoding of position in the developing neural tube from antiparallel morphogen gradients. *Science*, *356*(6345), 1379–1383.

Zhang, L., Radtke, K., Zheng, L., Cai, A. Q., Schilling, T. F., & Nie, Q. (2012). Noise drives sharpening of gene expression boundaries in the zebrafish hindbrain. *Molecular Systems Biology*, *8*, 613. https://doi.org/10.1038/MSB.2012.45

Zhu, Y., Qiu, Y., Chen, W., Nie, Q., & Lander, A. D. (2020). Scaling a Dpp Morphogen Gradient through Feedback Control of Receptors and Co-receptors. *Developmental Cell*, *53*(6), 724-739. e14.

Appendix I

Supplementary Material: Chapter I

Multiple morphogens and rapid elongation promote segmental patterning during development

Yuchi Qiu, **Lianna Fung**, Thomas F. Schilling, Qing Nie

Published in *PLOS Computational Biology* (2021) *17*(6), e1009077

S1. Signaling exchanges between continuum domain and individual cells

In the computational framework, the morphogen domain uses a regular rectangular mesh and the tissue domain uses an irregular mesh, where grids are the nodes in each moving cell. Morphogens need to transmit signals to individual cells and FGF synthesis needs *hoxb1a* expression in individual cells. We use interpolations to conduct the signaling exchanges from two domains.

From morphogen domain to tissue domain, we use a constant interpolation. For *i*-th cell, the morphogen level $[M]_i$ that is closest to the cell location:

$$\begin{cases} [M]_i = [M]\Big|_{\boldsymbol{x}=\left(x_1^{j_0}, x_2^{j_0}\right)}, \\ \left(i_0, j_0\right) = \arg\min_{i,j} \left|\left(x_1^j, x_2^j\right) - \boldsymbol{c}_i\right|, \end{cases} \tag{S1}$$

Here $[M]_i$ can represent either the intracellular RA ($[RA]_{in}$) or FGF signaling ($[Fgf]_{signal}$) in *i*-th cell. From tissue domain to morphogen domain, we use the build-in function "*griddata*" in MATLAB to interpolate the *hoxb1a* expression [*H*] defined on scatter points (i.e. centers of individual cells) to the regular mesh in the morphogen domain:

$$[H](\boldsymbol{x}) = griddata([H]_i\Big|_{i=1}^{N_{cell}}, \boldsymbol{c}_i\Big|_{i=1}^{N_{cell}}, \boldsymbol{x}) \tag{S2}$$

S2. Models for individual cells and their interactions

Following our previous work (Wang et al., 2017), we use the subcellular element method (SCEM) to model individual cells (Newman). Each cell consists of sub-cellular elements (nodes) and interacts according to a prescribed intercellular force potential. A cell consists of $2N_{node}$ (N_{node}=6) nodes and those nodes form two hexagonal layers as shown in Fig S6A. The radius of the outer layer is R_{out} and the radius of the inner layer is R_{in}. Initially cells are uniformly distributed in the tissue domain as shown in Fig 2H. For a system with N_{cell} cells and each cell has $2N_{node}$ nodes, the location of i-th node in n-th cell $\boldsymbol{x}_{n,j}$ is determined by the equation

$$\frac{d}{dt}\boldsymbol{x}_{n,i} = \boldsymbol{v}_{n,i}^{\text{ext}} + \boldsymbol{v}_{n,i}^{\text{inter}} + \boldsymbol{v}_{n,i}^{\text{inner}}. \tag{S3}$$

On the right hand side, the first term represents cell migration due to convergent extension (Rørth, 2012; Schumacher et al., 2016). It is given by

$$\begin{cases} \boldsymbol{v}_{n,j}^{ext} = \boldsymbol{V}(\boldsymbol{c}_n, t), \\ \boldsymbol{c}_n = \dfrac{1}{2N_{node}} \displaystyle\sum_{i=1}^{2N_{node}} \boldsymbol{x}_{n,i}. \end{cases} \tag{S4}$$

The second term represents the forces between cells while the third term represents forces between nodes within the same cell to maintain the stable cell morphology (Newman). For two nodes, their interactions are determined by Morse potential with a cutoff distance d:

$$\mathrm{F}(r) = \begin{cases} U\exp(-\dfrac{r}{x}) - V\exp(-\dfrac{r}{z}), & \text{if } r \le d; \\ 0, & \text{otherwise;} \end{cases} \tag{S5}$$

where r is the distance between these two nodes and d is the constant cut-off distance taken as two-fold of cell diameter ($4R_0$=16 μm). The short-range repulsion $U\exp(-\dfrac{r}{x})$ maintains a

minimum distance between nodes and the long-range attraction $-V\exp(-\frac{r}{z})$ makes nodes move closer. For nodes in the same cell, the *non-zero* intracellular forces are determined by:

$$F_{\text{intra}}(|\boldsymbol{x}_{n,i} - \boldsymbol{x}_{n,j}|) = U_{\text{intra}}\exp\left(-\frac{|\boldsymbol{x}_{n,i} - \boldsymbol{x}_{n,j}|}{x_{\text{intra}}}\right) - V_{\text{intra}}\exp\left(-\frac{|\boldsymbol{x}_{n,i} - \boldsymbol{x}_{n,j}|}{z_{\text{intra}}}\right) \quad \text{(S6)}$$

To model selective cell sorting for nodes in two cells, the *non-zero* intercellular forces depend on the identities of these two cells. In some cases, cells with the same identity sort using both long-range attractive and short-range repulsive forces, and cells with different identities sort using repulsive forces:

$$F_{\text{inter}}(r) = \begin{cases} U_{\text{inter}}^{\text{Atr}}\exp(-\frac{r}{x_{\text{inter}}}) - V_{\text{inter}}^{\text{Atr}}\exp(-\frac{r}{z_{\text{inter}}}), & \text{same identity;} \\ U_{\text{inter}}^{\text{Rep}}\exp(-\frac{r}{x_{\text{inter}}}), & \text{different identities.} \end{cases} \quad \text{(S7)}$$

In our model, selective inter-cellular forces depend on *krox20* levels in a pair of cells. We determine the inter-cellular forces using an affine combination of two types of forces in Eq. (S7). The discrepancy between *krox20* levels in these cells is used to determine the coefficients of the affine combination. For *m*-th and *n*-th cell, a similarity function is defined to describe such a discrepancy:

$$\mathcal{F}(m,n) = \frac{1}{2}(\delta^m\delta^n + 1) \in [0,1], \quad \text{(S8)}$$

where d^i is the relative level of *krox20* in *i*-th cell normalized by the maximum expression *krox20* over all cells. $d^i = 1$ indicates that the cell expresses a maximum level of *krox20* and $d^i = -1$ indicates the cell has zero *krox20* expression. The force between two cells is attractive when they have similar *krox20* levels ($\mathcal{F} = 1$ if $d_m = d_n = \pm 1$) and is repulsive when they have different *krox20* levels ($\mathcal{F} = 0$ if $d_m = -d_n = \pm 1$). The Morse potentials between a pair of nodes in two different cells with locations at $\boldsymbol{x}_{m,j}$ and $\boldsymbol{x}_{n,i}$ are defined by

$$\Phi_{\text{inter}}(|\boldsymbol{x}_{n,i}-\boldsymbol{x}_{m,j}|)=\left[\left(1-\mathcal{F}\left(m,n\right)\right)U_{\text{inter}}^{\text{Rep}}+\mathcal{F}U_{\text{inter}}^{\text{Atr}}\right]\exp\left(-\frac{|\boldsymbol{x}_{n,i}-\boldsymbol{x}_{m,j}|}{\xi_{\text{inter}}}\right) - \mathcal{F}\left(m,n\right)\mathrm{V}_{\text{inter}}^{\text{Atr}}\exp\left(-\frac{|\boldsymbol{x}_{n,i}-\boldsymbol{x}_{m,j}|}{\zeta_{\text{inter}}}\right). \quad \text{(S9)}$$

Taken together, the second term in Eq. (S3) given by

$$\boldsymbol{v}_{n,i}^{\text{inter}}=h_{n,i}-\nabla_{\boldsymbol{x}_{n,i}}\left[\sum_{j\neq i}^{2N_{node}}\mathrm{F}_{\text{intra}}(|\boldsymbol{x}_{n,i}-\boldsymbol{x}_{n,j}|)+\sum_{m\neq n}^{N_{cell}}\sum_{j}^{2N_{node}}\mathrm{F}_{\text{inter}}(|\boldsymbol{x}_{n,i}-\boldsymbol{x}_{m,j}|)\right], \quad \text{(S10)}$$

where $\eta_{n,i}\sim\mathcal{N}(0,10)$ is the Gaussian-distributed noise.

The canonical SCEM only considers Morse potentials as discussed above. To avoid squashed or exploded cells, we add other spring-type pairwise interactions to nodes within the same cell (Wang et al., 2017). In *i*-th cell, we evenly divide nodes into two layers, where each layer initially has a regular polygonal structure. These pairwise interactions act on neighboring nodes between the same layer $\boldsymbol{x}_{n,i}\sim\boldsymbol{x}_{n,i\pm1}$ and the nodes between different layers with the same index $\boldsymbol{x}_{n,i}\sim\boldsymbol{x}_{n,i\pm N_{node}}$. The force is given by

$$\mathrm{Y}(\boldsymbol{x}_{n,i},\boldsymbol{x}_{n,j})=m\frac{|\boldsymbol{x}_{n,i}-\boldsymbol{x}_{n,j}|-l_{i,j}^{n}}{|\boldsymbol{x}_{n,i}-\boldsymbol{x}_{n,j}|}(\boldsymbol{x}_{n,i}-\boldsymbol{x}_{n,j}), \quad \text{(S11)}$$

where $l_{i,j}^{n}$ is the constant length between node $\boldsymbol{x}_{n,i}$ and $\boldsymbol{x}_{n,j}$ at an initial time and it has three possible values, l_{out}, l_{in}, l_{inter}, depending on the types of two nodes (**Fig S6A**). Then for *i*-th node in the *n*-th cell, the motion from this force is

$$\boldsymbol{v}_{n,i}^{\text{inner}}=\mathrm{Y}(\boldsymbol{x}_{n,i},\boldsymbol{x}_{n,i+1})+\mathrm{Y}(\boldsymbol{x}_{n,i},\boldsymbol{x}_{n,i-1})+\mathrm{Y}(\boldsymbol{x}_{n,i},\boldsymbol{x}_{n,i\pm N_{node}}). \quad \text{(S12)}$$

To prevent cells from leaving the tissue domain, we add two layers of “ghost cells” located outside the boundary tissue domain to provide extra forces.

S3. One-dimensional gene expression model

In the main text, we compare and contrast the two-morphogen and one-morphogen models. The equations are defined for the fixed one-dimensional domain with deterministic description. The equations for the two-morphogen model are given by Eq. (10) in Methods by removing noise terms. For the one-morphogen model, RA induces both *hoxb1a* and *krox20*. Only equations for *hoxb1a* and *krox20* are modified:

$$\begin{aligned}\frac{d[H]}{dt} &= v_h \frac{a_{hh}[H]^2 + a_{rh}[RA]_{in}^2}{1 + a_{hh}[H]^2 + a_{rh}[RA]_{in}^2 + b_{kh}[K]^2 + b_{vh}[V]^2} - d_h[H] \ , \\ \frac{d[K]}{dt} &= v_k \frac{a_{kk}[K]^2 + a_{rk}[RA]_{in}^2}{1 + a_{kk}[K]^2 + a_{rk}[RA]_{in}^2 + b_{hk}[H]^2} - d_k[K] \ .\end{aligned} \tag{S13}$$

S4. Initial cell distribution for cell sorting-only model.

In Fig 5B, we use mixture Gaussian distribution to generate the "salt-and-pepper" initial cell distribution. We label the cells with r2, r3, r4, r5 and r6 identities as 1, 2, 3, 4 and 5, respectively. For a cell with A-P position x, we normalized its A-P position $\tilde{x} = \frac{x - x_{\min}}{x_{\max} - x_{\min}}$ in the range of [0,1] where $x_{\min} = r_1 L_1(0)$ and $x_{\max} = r_2 L_1(0)$ are A-P positions of boundaries of the tissue domain. The probability for this cell being *i*-th type of cell is given by

$$p_i(\tilde{x}) = \frac{\phi_i \mathcal{N}(\tilde{x} | \mu_i, \sigma_i)}{\sum_{i=1}^{5} \phi_i \mathcal{N}(\tilde{x} | \mu_i, \sigma_i)}. \tag{S14}$$

We take the weights and the variance as $f_i = 1/5$ $s_i = 0.08$ for all *i*. The means $(m_1, m_2, m_3, m_4, m_5) = \frac{1}{10}(1,3,5,7,9)$ denote the average A-P positions of each cell type. The probability distributions of different cell types are shown in Fig S6C.

S5. Determination of cell fate

In the main text, we discuss cell commitment time, defined by the latest time that a cell changes its fate. To determine cell fate numerically, we used thresholds for *hoxb1a* and *krox20* expression. There are three types of cells in the r2-r6 domain. Cells in r2 and r6 have low expression of both *hoxb1a* and *krox20*. Cells in r3 and r5 have low *hoxb1a* and high *krox20* expression. Cells in r4 have high *hoxb1a* and low *krox20* expression. “High” and “low” levels are defined by expression thresholds of 1.2, with no cell identified as having high expression of both *hoxb1a* and *krox20*. The expression thresholds are defined as the following:

$$\begin{aligned} &1)\ \text{none expressing identity: } [H]<1.2,\ [K]<1.2;\\ &2)\ \textit{hoxb1a}\ \text{identity: } [H]<1.2,\ [K]\ge 1.2;\\ &3)\ \textit{krox}20\ \text{identity: } [H]\ge 1.2,\ [K]<1.2. \end{aligned} \tag{S15}$$

S6. Numerical solvers

S6.1. Solving PDEs (morphogens) in moving boundaries

For morphogens RA and FGF, the two-dimensional spatial domains have moving boundaries. A general form of the equation is given by

$$\begin{aligned} &\frac{\partial [M]}{\partial t}+\nabla\cdot\left(\boldsymbol{V}[M]\right)=D\mathrm{D}[M]+f,\\ &\boldsymbol{x}=(x^{(1)},x^{(2)})\in[0,L_1(t)]x[-\frac{1}{2}L_2(t),\frac{1}{2}L_2(t)]. \end{aligned} \tag{S16}$$

To numerically solve these PDEs, we first transform the dynamic domain to the unit square by using a coordinate transformation:

$$\begin{cases} \boldsymbol{X}=(X^{(1)},X^{(2)})=\left(\dfrac{1}{L_1(t)}x^{(1)},\dfrac{1}{L_2(t)}x^{(2)}+\dfrac{1}{2}\right)\in[0,1]x[0,1],\\ t=t. \end{cases} \tag{S17}$$

The partial derivatives in a transformed coordinate system are given by

$$
\begin{aligned}
&\nabla_{\boldsymbol{x}} = \left(\frac{\partial}{\partial X^{(1)}}, \frac{\partial}{\partial X^{(2)}} \right) = \left(L_1(t) \frac{\partial}{\partial x^{(1)}}, L_2(t) \frac{\partial}{\partial x^{(2)}} \right), \\
&\frac{\partial^2}{\partial \left(X^{(1)} \right)^2} = \left(L_1(t) \right)^2 \frac{\partial^2}{\partial \left(x^{(1)} \right)^2}, \\
&\frac{\partial^2}{\partial \left(X^{(2)} \right)^2} = \left(L_2(t) \right)^2 \frac{\partial^2}{\partial \left(x^{(2)} \right)^2}, \\
&\frac{\partial}{\partial t} = \frac{\partial}{\partial t} + L_1^{'}(t) X^{(1)} \frac{\partial}{\partial x^{(1)}} + L_2^{'}(t) \left(X^{(2)} - \frac{1}{2} \right) \frac{\partial}{\partial x^{(2)}}.
\end{aligned}
\tag{S18}
$$

The resulting equation in the new coordinate system is given by:

$$
\frac{\P[M]}{\P t} = \frac{1}{\left(L_1(t) \right)^2} \frac{\P^2[M]}{\P \left(X^{(1)} \right)^2} + \frac{1}{\left(L_2(t) \right)^2} \frac{\P^2[M]}{\P \left(X^{(2)} \right)^2} + \frac{L_1'(t)}{L_1(t)}[M] + \frac{L_2'(t)}{L_2(t)}[M] + f \tag{S19}
$$

For spatial discretization, we use uniform rectangular grids with 151 and 71 grid points along the A-P and L-R axis, respectively. The central difference method is used for the discretization of the diffusion term. For temporal discretization, we use the Euler-Maruyama method with a fixed time step. The time step for morphogens and gene expressions model is $\Delta t_1 = 1.8$ seconds.

S6.2. Solving equations in discrete cells model

We use the Euler-Maruyama method to solve Eq. (S3) with the fixed time step $\Delta t_2 = 3.6$ seconds. For calculations of the pairwise interactions between nodes in Eq. (S10), we use GPUs algorithm by using “gpuArray()” function in MATLAB.

S7. Parameters

S7.1. Parameters value and selection

Parameter values used in our simulations are listed in Tables A, B and C. In particular, parameters used in cell mechanical models are listed in Table A, and parameters used in gene regulation models are listed in Table B. The sampling ranges of random parameters in Fig 7 are listed in Table C.

In cell-based models, the A-P range of the tissue domain, determined by ratios r_1 and r_2, is chosen to allow r2-r6 to be contained within this region, where r3-r5 A-P lengths are from experimental measurement (Table G) while r2 and r6 lengths are assumed to have similar lengths to r3 and r5 (since all rhombomeres in r2-r6 have similar A-P lengths). The cell radius, R_{out}, is around 4 μm in 11-14hpf, which is estimated from images in (Addison et al., 2018). In cell mechanical models, the parameters are taken from our previous work (Wang et al., 2017) and the time and length units have been adjusted based on new measurements. In gene regulation models, the parameters for RA (i.e. D_r, v_r, β_r, k_{max}, and k_0) are taken from previous modeling and experimental studies (White et al., 2007). Since FGF generated in r4 is likely to have a shorter decay length than RA, we use a 3-fold smaller diffusion coefficient D_f than for RA (D_r). The degradation rate chosen, d_{f2}, is also larger than the degradation rate of RA. As a result, FGF gradients are established much faster than RA gradients, Parameters for *hoxb1a* and *krox20* expression, a_{hh}, a_{kk}, a_{rh}, d_h, d_k b_{hk} and b_{kh} were initially taken from our one-morphogen model (Zhang et al., 2012) and adjusted manually to achieve values that produce a five-segment pattern according to length and time measured in this work. All other parameters in the gene regulation models were adjusted to produce a five-segment pattern with correct rhombomere A-P lengths. The noise magnitudes, (μ_* parameters) were selected manually to give effective boundary sharpening that changed with increased or decreased noise magnitudes.

Parameters	Value	Unit	Reference
r_1	0.038	--	This study
r_2	0.28	--	This study
N_{column}	15	--	This study
N_{row}	23	--	This study
N_{cell}	345	--	This study
N_{node}	6	--	This study
R_{out} (cell radius)	3.60	μm	Estimated from images in (Addison et al., 2018)
R_{in}	1.80	μm	(Wang et al., 2017)
l_{out}	3.60	μm	
l_{in}	1.80	μm	
l_{inter}	1.80	μm	
n_i	1.11	sec^{-1}	
U_{intra}	$1.33x10^{-1}$	$\mu m^2\ sec^{-1}$	
x_{intra}	2.4	μm	
V_{intra}	$5.56x10^{-2}$	$\mu m^2\ sec^{-1}$	
z_{intra}	3.6	μm	
U_{inter}^{Atr}	$1.04x10^{-1}$	$\mu m^2\ sec^{-1}$	
U_{inter}^{Rep}	$2.50x10^{-2}$	$\mu m^2\ sec^{-1}$	

x_{inter}	5.0	μm	
V_{inter}^{Atr}	$6.67x10^{-2}$	$\mu m^2\ sec^{-1}$	
z_{inter}	11.00	μm	

Table A. Parameters for the cell mechanical model.

Parameters	**Value**	**Unit**	**Reference**
D_r	2.83	$\mu m^2\ sec^{-1}$	(White et al., 2007)
v_r	$1.11x10^{-2}$	sec^{-1}	
k_r	$2.22x10^{-4}$	sec^{-1}	
b_r	1	--	
k_{max}	$5.56x10^{-1}$	sec^{-1}	
k_0	$1.11x10^{-4}$	sec^{-1}	
D_f	0.85	$\mu m^2\ sec^{-1}$	This study
v_f	$5.56x10^{-2}$	sec^{-1}	This study
a_{hf}	2	--	This study
k_f	$2.67x10^{-4}$	sec^{-1}	This study
k_{rf}	$2.67x10^{-4}$	sec^{-1}	This study
d_{f1}	$2.67x10^{-4}$	sec^{-1}	This study
d_{f2}	0.013	sec^{-1}	This study
η_{f1}	0.1	--	This study
η_{f2}	0.003	--	This study
η_{f1}	0.1	--	This study

η_{f2}	0.01	--	This study
v_h	0.056	sec^{-1}	This study
v_k	0.056	sec^{-1}	This study
v_v	0.11	sec^{-1}	This study
v_i	$3.33x10^{-4}$	sec^{-1}	This study
a_{hh}	0.85	--	This study
a_{rh}	0.13	--	This study
a_{kk}	0.9	--	This study
a_{fk}	6	--	This study
a_{rv}	0.1	--	This study
b_{kh}	40	--	This study
b_{vh}	5	--	This study
b_{hk}	20	--	This study
b_{iv}	3.5	--	This study
b_{vi}	5	--	This study
d_h	0.022	sec^{-1}	This study
d_k	0.022	sec^{-1}	This study
d_v	0.022	sec^{-1}	This study
d_i	$3.33x10^{-4}$	sec^{-1}	This study
η_h	0.01	--	This study
η_k	0.01	--	This study
η_v	0.005	--	This study
η_i	0.005	--	This study

Table B. Parameters for the equations of morphogens and intracellular genes.

S7.2 Parameters sensitivity analysis

S7.2.1 Random parameters perturbations

In Fig 7 and Fig S5, we generate the simulations with random parameters for the equations for *hoxb1a* and *krox20*. We use Latin hypercube sampling to generate high dimension random numbers and each sample is written as $\left(w_1,...,w_N\right)$. We list the parameters that are perturbed by using the random number. If not listed, those parameters are taken as the same as that in Table B.

Parameters	**Value**	**Unit**
a_{hh}	$0.85x2^{-0.4+0.8w_1}$	--
a_{kk}	$0.9x2^{-0.4+0.8w_2}$	--
a_{rh}	$0.13x2^{-1+2w_3}$	--
a_{fk}	$6x2^{-1+2w_4}$	--
b_{kh}	$40x2^{-1+2w_5}$	--
b_{vh}	$5x2^{-1+2w_6}$	--
b_{hk}	$20x2^{-1+2w_7}$	--
d_h	$0.022x2^{-1+2w_8}$	sec^{-1}
d_k	$0.022x2^{-1+2w_9}$	sec^{-1}
b_{iv}	4	--
v_h	2.5xd_h	sec^{-1}
v_k	2.5xd_k	sec^{-1}

Table C. Parameters for Fig 7 and Fig S5. If not specified, they are the same as that in Table B.

S7.2.2 Local sensitivity analysis of single parameters in gene regulation

We performed single parameter perturbations to examine sensitivity to each parameter in our morphogen equations (Eq. (5-9)) and gene regulation equations (Eq. (10)). We systematically varied each parameter up and down by 20% of its total value and measured the effect on quantities including lengths of three rhombomeres, SIs of four boundaries and DC numbers. Specifically, we calculated the ratio between quantities in perturbed cases and quantities in unperturbed cases at the end of the simulations (14 hpf) (Table D). The system is relatively insensitive to most parameters. It is relatively sensitive to v_r, k_r, d_{f1} and d_{f2} in morphogen equations. The robustness in morphogen dynamics can be improved by including many other mechanisms (Rackauckas et al., 2018), which were not included in this work for simplification. In gene regulation models, the system is sensitive to auto-activation rate (a_{hh}), as also reported in previous work (Wang et al., 2017). Since *hoxb1a* and *krox20* compete with each other, sensitivity to v_h, v_k, d_h and d_k are expected (many simulations fail to form rhombomeres) since they risk maximal values of *hoxb1a* or *krox20*. Alternatively, we can avoid these perturbations on maximal values of *hoxb1a* or *krox20* by coupling the ratio v_h/d_h or v_k/d_k. The system turns out to be robust to the perturbations on v_h or v_k with such coupling (Table D).

Parameter	Fold changed	Ratio of perturbed quantities to unperturbed quantities								# failed simulations
		r3	r4	r5	SI(r2/r3)	SI(r3/r4)	SI(r4/r5)	SI(r5/r6)	DCs	
D_r	0.8	0.96	1.05	0.85	1.23	1.46	1.11	1	1.45	0
	1.2	1.02	0.9	1.09	0.93	0.94	1.37	1.14	2.26	0
v_r	0.8	1	1.11	0.94	0.99	0.91	0.97	1.11	0.56	0
	1.2	0.95	0.92	1.07	1.44	1.83	1.39	0.98	4.39	1
b_r	0.8	0.97	0.99	1.01	0.83	0.77	0.86	0.97	0.97	0
	1.2	0.97	0.99	1.01	0.83	0.77	0.86	0.97	0.97	0

k_{max}	0.8	0.97	0.99	1.00	0.82	0.76	0.97	0.99	0.81	0
	1.2	0.97	0.98	1.02	0.82	0.76	0.90	0.97	0.97	0
k_r	0.8	--	--	--	--	--	--	--	--	10
	1.2	0.92	1.18	0.84	1.1	1.66	2.31	0.82	4.84	1
D_f	0.8	0.92	0.98	1.02	0.82	0.73	0.95	0.94	1.13	0
	1.2	1.02	1.03	0.95	0.9	0.7	1.23	1.33	0.89	0
k_f	0.8	1.04	1.03	1.03	0.93	0.66	0.89	1.04	0.48	0
	1.2	0.9	0.97	0.95	0.66	0.91	1.04	1.16	1.61	0
d_{f1}	0.8	1.12	0.95	1.18	0.85	0.75	0.64	1.07	0.4	0
	1.2	0.84	1.02	0.88	0.87	1.01	1.26	1.17	3.39	0
d_{f2}	0.8	1.27	0.89	1.3	0.99	0.69	0.78	0.91	1.05	0
	1.2	0.76	1.21	0.7	1.03	3.26	1.75	1.76	4.39	1
v_h	0.8	--	--	--	--	--	--	--	--	10
	1.2	--	--	--	--	--	--	--	--	10
v_k	0.8	--	--	--	--	--	--	--	--	10
	1.2	1.31	0.98	1.36	1.57	0.82	0.66	0.96	2.42	0
v_v	0.8	1.31	1.39	1.14	0.89	0.61	0.94	0.72	0.4	0
	1.2	0.78	0.95	0.63	1.54	4.73	1.66	1.69	6	1
v_i	0.8	--	--	--	--	--	--	--	--	10
	1.2	1.2	1.34	1.12	0.7	0.67	0.64	0.78	0.4	0
a_{hh}	0.8	0.66	0.94	0.63	1	1.57	2.68	1.87	2.26	0
	1.2	--	--	--	--	--	--	--	--	10
a_{kk}	0.8	0.75	0.97	0.83	0.7	0.93	1.2	1.4	3.39	0
	1.2	1.16	1	1.18	1.06	0.73	0.89	1.09	1.21	0
a_{rh}	0.8	0.85	0.89	0.67	1.3	4.44	1.85	1.23	4.84	5

	1.2	1.26	1.22	1.21	1.19	1.12	0.92	0.99	0.81	0
	0.8	1.29	1.36	1.19	0.95	0.59	1	0.62	0.4	0
a_{rv}	1.2	0.8	0.97	0.64	2.01	5.1	1.41	1.72	6.05	2
	0.8	0.86	1.07	0.81	0.54	0.86	1.29	1.36	3.32	1
a_{fk}	1.2	1.09	0.93	1.13	0.86	0.85	0.84	1.11	0.73	0
	0.8	0.94	0.95	1.05	0.68	0.74	0.71	1.08	0.89	0
b_{hk}	1.2	0.97	1.03	0.93	0.93	0.69	1.03	1.19	1.13	0
	0.8	0.99	1.04	0.96	0.92	0.79	1.09	0.99	1.45	0
b_{kh}	1.2	0.95	0.95	1.02	0.78	0.77	0.89	1.18	0.81	0
	0.8	1.07	1.08	1.09	0.98	0.8	0.85	0.86	0.16	0
b_{vh}	1.2	0.93	0.94	0.92	1.22	0.83	1.18	1.39	2.5	0
	0.8	1.06	1.11	1.06	0.78	0.92	1.12	0.96	0.24	0
b_{vi}	1.2	0.89	0.9	0.92	0.93	0.91	0.86	1.32	1.53	0
	0.8	0.84	0.92	0.84	1.21	2.74	1.36	1.62	4.84	1
b_{iv}	1.2	1.09	1.19	1.06	0.96	0.77	1.01	0.84	0.24	0
	0.8	--	--	--	--	--	--	--	--	10
d_h	1.2	--	--	--	--	--	--	--	--	10
	0.8	1.33	0.94	1.43	1.94	0.69	0.76	0.79	2.66	0
d_k	1.2	0.64	1.08	0.54	1.15	3.66	2.37	2.06	5.4	0
	0.8	0.82	0.82	0.62	1.51	3.58	2.03	1.38	4.64	2
d_v	1.2	1.27	1.31	1.13	0.81	0.63	0.86	0.69	0.73	0
	0.8	1.13	1.31	1	0.68	0.76	0.91	0.9	0.32	0
d_i	1.2	0.8	0.78	0.89	0.91	2.4	1.47	1.9	4.97	4
d_h and	0.8	0.96	1	0.98	1.05	1.05	1.07	1.18	0.89	0
v_h	1.2	0.96	0.99	0.99	0.91	0.94	0.91	1.3	1.05	0

d_k and v_k	0.8	0.96	0.96	0.96	0.91	0.93	0.87	1.17	0.81	0
	1.2	1.01	0.99	1.03	0.84	0.92	0.92	1.15	0.56	0
d_i and v_i	0.8	0.93	0.97	0.81	1.35	2.53	1.58	1.47	4.66	1
	1.2	1.09	1.02	1.17	0.89	0.86	0.64	0.85	0.56	0
d_v and v_v	0.8	0.96	1.01	0.99	0.82	0.83	1.02	1.07	1.05	0
	1.2	0.96	0.99	1	0.87	0.74	0.93	1.21	0.81	0

Table D. Single parameter sensitivity analysis for parameters in morphogen and gene regulation equations. Each parameter was perturbed up and down by 20% individually. In each case, 10 simulations were performed to calculate the resulting rhombomere lengths and SIs. Numbers of failed simulations missing at least one boundary or having >8 DCs are shown in the last column.

S7.2.3 Noise magnitude

Additive noise was included for both morphogen and gene regulation equations. Here we explored the effects of noise by perturbing the levels 2-fold up and down individually (Table S5). Rhombomere lengths were relatively stable to changes in noise level. The exception was noise in FGF signaling where larger η_{f2} caused expansion of r3 and r5. Our choice of noise levels was effective in disturbing boundary sharpness. Some exceptions were observed for η_{f1}, η_{f2} and η_h at a few boundaries.

Parameter	Fold change	Ratio of perturbed quantities to unperturbed quantities								# failed simulations
		r3	r4	r5	SI(r2/r3)	SI(r3/r4)	SI(r4/r5)	SI(r5/r6)	DCs	
	0.5	0.96	1.00	1.01	0.85	0.74	0.92	1.20	1.21	0

η_{f1}	2	1.00	1.00	1.01	1.06	0.95	0.88	1.15	1.45	0
η_{f2}	0.5	0.97	0.99	1.06	1.05	0.82	0.87	1.19	0.56	0
	2	1.04	1.02	1.03	1.12	1.94	0.99	1.36	2.50	0
η_{f1}	0.5	0.96	1.00	0.98	0.87	0.79	0.81	1.10	0.73	0
	2	0.97	0.99	1.00	0.91	0.75	0.83	1.19	0.81	0
η_{f2}	0.5	0.84	1.01	0.93	0.69	0.88	0.95	0.96	0.81	0
	2	1.44	0.94	1.35	4.26	0.75	0.97	1.83	5.38	1
η_{v}	0.5	0.97	0.99	1.03	0.84	0.73	0.85	1.10	0.97	0
	2	0.96	1.00	0.99	0.91	0.78	0.95	1.09	0.97	0
η_{v}	0.5	0.98	1.03	1.03	0.99	0.61	0.71	0.62	0.73	0
	2	0.95	0.88	1.11	1.03	1.15	1.91	1.52	2.15	1
η_{h}	0.5	0.95	1.00	0.96	1.16	0.69	0.96	1.18	0.73	0
	2	1.02	1.01	1.06	1.01	0.82	0.76	0.86	0.97	0
η_{k}	0.5	1.00	0.99	1.02	1.04	1.00	0.94	1.09	0.97	0
	2	0.92	0.97	0.93	1.22	1.83	1.22	1.47	3.39	0

Table E. Noise magnitude and rhombomere formation. Noise levels were perturbed 2-fold up and down and 10 simulations were performed to calculate each rhombomere length and SI. Numbers of failed simulations missing at least one boundary or having >8 DCs are shown in the last column.

S7.2.4 Cell mechanical strength and its coupling with gene regulation

The cell mechanical model relies on many parameters in Table A. Their values are based on previous work (Wang et al., 2017). We assess sensitivity with respect to the strength of the critical features in the mechanical model, such as the ratio between attraction and repulsion and the strength/timescale of the cell sorting.

We investigate how the model performs with different attraction-repulsion ratios by either perturbing attraction or repulsion strength individually (Table F). Specifically, in Eq. (S7), we perturb attraction by perturbing $U_{\text{inter}}^{\text{Atr}}$ and $V_{\text{inter}}^{\text{Atr}}$ equally or repulsion by perturbing $U_{\text{inter}}^{\text{Rep}}$ alone. We measured the effects of perturbations on lengths of three rhombomeres, SIs of four boundaries and number of DCs. We calculate the ratio between quantities in perturbed cases and quantities in unperturbed cases at the end of the simulation (14 hpf). Rhombomere lengths are relatively insensitive to changes in mechanical forces since cell fates are mainly regulated by morphogens and gene expression. The SIs are positively related to the attraction-repulsion ratio, where a smaller ratio leads to sharper boundaries. However, the valid range of the attraction-repulsion ratio to generate clear five-segment patterns is narrow (~10%, an example in Fig S7A), where a 20% reduction on attraction-repulsion ratio leads to separation and gaps of cells with different identities (Table F and Fig S7B) and a 20% increase of attraction-repulsion ratio leads to rougher boundaries (Fig S7C).

We also examined how the coupling between cell mechanics and gene regulation affects the patterns. We perturbed the strength/timescale of the cell mechanical model. The rhombomere lengths are insensitive to this perturbation since gene regulation plays a major role in controlling rhombomere lengths. Because cell sorting plays major roles in boundary refinement, increased mechanical force strength makes boundaries sharper with smaller SIs and reduced numbers of

DCs. This result supports the conclusion that the combination of global (gene regulation) and local (cell sorting) processes works synergistically in boundary sharpening.

Components	Fold change	Ratio of perturbed quantities to unperturbed quantities								# failed simulations
		r3	r4	r5	SI(r2/r3)	SI(r3/r4)	SI(r4/r5)	SI(r5/r6)	DCs	
Attractions	0.8*	1.31	1.00	1.15	0.00	0.22	0.20	0.18	0.65	0
	0.9	0.97	1.02	1.01	0.41	0.61	0.47	0.96	0.65	0
	1.1	0.94	1.00	0.98	1.23	1.28	1.61	1.30	1.45	0
	1.2	0.95	0.97	0.99	1.93	2.29	2.53	2.06	2.18	0
Repulsions	0.8	1.01	0.92	1.01	2.59	3.05	3.61	2.68	2.66	0
	0.9	0.96	0.97	1.01	1.44	1.25	1.86	1.50	1.53	0
	1.1	0.95	1.02	1.01	0.48	0.55	0.56	0.89	0.56	0
	1.2*	1.13	1.00	1.16	0.15	0.28	0.36	0.17	0.73	0
Cell mechanics strength	0.8	0.97	1.01	0.98	1.02	0.83	1.04	1.16	1.05	0
	0.9	0.95	0.99	0.98	0.85	0.84	0.85	1.08	0.89	0
	1.1	0.95	1.02	0.98	0.81	0.72	0.99	1.12	0.65	0
	1.2	0.95	0.99	1.02	0.71	0.67	0.95	1.17	1.13	0

Table F. **Attractive and repulsive forces in rhombomere formation**. Attraction or repulsion was perturbed up or down by 10% and 20%, either individually or together. In each case, 10 simulations were performed to calculate the statistics for rhombomere lengths and SIs. Numbers of failed simulations missing at least one boundary or having >8 DCs are shown in the last column. Note: *Simulations show big gaps between clones of cells with different identities.

S8. Experimental measurements of hindbrain dimensions

Here we list the measurements of hindbrain dimensions made using whole mount in situ hybridization at 11-14 hpf (Fig 1C). At 11 hpf, the *krox20* expression domain in r3 has rough, intermittent boundaries. Thus, at this stage we only measured total hindbrain length and the

RA production region. In some samples it was difficult to define the RA production region, and these were omitted in calculating average lengths.

Time	Sample NO.	LR width ($L_2(t)$; μm)	AP lengths (μm)						
			MHB to r3	r3	r3-r5	r5	r5 to RA production region	Hindbrain Total Length ($p(t)$)	RA production region length
11hpf	1	255	--	--	--	--	--	343	278
	2	237	--	--	--	--	--	338	273
	3	308	--	--	--	--	--	281	284
	4	312	--	--	--	--	--	294	--
	5	269	--	--	--	--	--	313	302
	6	303	--	--	--	--	--	355	306
	7	316	--	--	--	--	--	343	281
	8	272	--	--	--	--	--	339	--
	9	278	--	--	--	--	--	298	281
	Mean	**283**	**--**	**--**	**--**	**--**	**--**	**323**	**286**
12hpf	1	206	74	32	33	27	273	440	--
	2	152	64	34	33	40	284	455	206
	3	139	71	43	28	49	298	490	203
	4	182	74	38	35	31	283	460	209
	5	175	59	31	30	37	252	410	179
	6	154	61	28	28	37	260	414	217
	7	152	68	36	30	31	262	427	213
	8	150	67	39	32	42	277	457	206

	9	153	75	35	34	39	274	457	200
	Mean	**162**	**68**	**35**	**31**	**37**	**274**	**445**	**233**
13hpf	1	109	53	33	31	35	219	370	245
	2	131	72	40	31	44	236	423	357
	3	186	91	47	30	54	274	497	369
	4	162	81	41	37	35	266	459	308
	5	133	63	29	20	34	276	422	260
	6	123	86	43	42	41	246	458	348
	7	127	67	59	29	49	242	446	354
	8	157	88	42	31	34	256	452	271
	9	128	67	41	33	39	257	437	280
	Mean	**140**	**74**	**42**	**32**	**40**	**252**	**440**	**310**
14hpf	1	109	85	39	38	36	230	428	--
	2	110	80	48	29	33	241	431	365
	3	96	82	36	45	41	198	402	420
	4	112	71	46	29	34	234	414	375
	5	97	93	43	35	43	236	450	452
	6	99	74	46	33	36	233	423	--
	7	106	62	46	31	38	214	391	375
	8	111	92	41	31	38	223	426	379
	9	91	65	32	29	28	190	343	--
	Mean	**104**	**78**	**42**	**33**	**37**	**222**	**412**	**394**

Table G. Experimental measurement for hindbrain dimensions.

S9. Multiplicative noise in gene regulation models

Additive noise in morphogen (Eq. (5) and Eq. (8)) and gene regulation equations (Eq. (10) was used throughout most of this study. Here we investigate multiplicative noise. The equations for RA with multiplicative noise (bold font) are given as the following:

$$\begin{aligned}
&\frac{\partial [RA]_{out}}{\partial t} + \nabla \cdot \left(\boldsymbol{V} [RA]_{out} \right) = D_r \mathrm{D} [RA]_{out} + A(\boldsymbol{x}, t) \\
&\qquad + k_r [RA]_{in} - (1 + b_r) k_r [RA]_{out} + \boldsymbol{h_{r1} [RA]_{out} \frac{dm_{r1}(t)}{dt}}, \\
&\frac{\partial [RA]_{in}}{\partial t} + \nabla \cdot \left(\boldsymbol{V} [RA]_{in} \right) = k_r [RA]_{out} - k_r [RA]_{in} - d_r (x^{(1)}) [RA]_{in} \\
&\qquad + \boldsymbol{h_{r2} [RA]_{in} \frac{dm_{r2}(t)}{dt}},
\end{aligned} \tag{S20}$$

where $h_{r1}[RA]_{out}\frac{dm_{r1}(t)}{dt}$ and $h_{r2}[RA]_{in}\frac{dm_{r2}(t)}{dt}$ are the multiplicative white noise. Similarly, equations for FGF and morphogen downstream genes with multiplicative noise are given:

$$\begin{aligned}
&\frac{\partial [Fgf]_{free}}{\partial t} + \nabla \cdot \left(\boldsymbol{V} [Fgf]_{free} \right) = D_f \mathrm{D} [Fgf]_{free} + A_f([H], \boldsymbol{x}) + k_{rf} [Fgf]_{signal} \\
&\qquad - (d_{f1} + k_f) [Fgf]_{free} + \boldsymbol{h_{f1} [Fgf]_{free} \frac{dm_{f1}(t)}{dt}}, \\
&\frac{\partial [Fgf]_{signal}}{\partial t} + \nabla \cdot \left(\boldsymbol{V} [Fgf]_{signal} \right) = k_f [Fgf]_{free} - k_{rf} [Fgf]_{signal} \\
&\qquad - d_{f2} [Fgf]_{signal} + \boldsymbol{h_{f2} [Fgf]_{signal} \frac{dm_{f2}(t)}{dt}}, \\
&\frac{d[H]_i}{dt} = v_h \frac{a_{hh}[H]_i^2 + a_{rh}[RA]_{in}^2}{1 + a_{hh}[H]_i^2 + a_{rh}[RA]_{in}^2 + b_{kh}[K]_i^2 + b_{vh}[V]_i^2} - d_h [H]_i + \boldsymbol{h_h [H]_i \frac{dm_h}{dt}}, \\
&\frac{d[K]_i}{dt} = v_k \frac{a_{kk}[K]_i^2 + a_{fk}[Fgf]_{signal}^2}{1 + a_{kk}[K]_i^2 + a_{fk}[Fgf]_{signal}^2 + b_{hk}[H]_i^2} - d_k [K]_i + \boldsymbol{h_k [K]_i \frac{dm_k}{dt}}, \\
&\frac{d[V]_i}{dt} = v_v \frac{a_{rv}[RA]_{in}^2}{1 + a_{rv}[RA]_{in}^2 + b_{iv}[I]_i^2} - d_v [V]_i + \boldsymbol{h_v [V]_i \frac{dm_v}{dt}}, \\
&\frac{d[I]_i}{dt} = v_i \frac{1}{1 + b_{vi}[V]_i^2} - d_i [I]_i + \boldsymbol{h_i [I]_i \frac{dm_i}{dt}}.
\end{aligned} \tag{S21}$$

(S22)

Compared with additive noise (Fig 4), our findings suggest that the positive effects of rapid initial convergence on segmental pattern robustness also occur with multiplicative noise (Fig S8). Initial

rapid convergence results in the smallest r4 A-P length compared with medium or slow initial convergence. The initial slow convergence has largest rhombomere lengths for r3, r4 and r5. All four boundaries are also sharper with initial rapid convergence compared with slow initial convergence. In addition, rapid initial convergence results in fewer numbers of DCs compared to slow initial convergence. Medium convergence leads to sharper boundaries for r3/r4, r4/r5 and r5/r6 than rapid initial convergence, but leads to defects in rhombomere lengths, with r4 larger than r3 and r5, and r5 is smaller than r3. A difference between the effects of multiplicative versus additive noise is that with the former, the posterior boundaries (i.e. r4/r5 and r5/r6) appear much rougher than anterior boundaries (r2/r3 and r3/r4). This is likely because RA levels are higher posteriorly and RA noise levels are also higher with multiplicative noise. Noise magnitudes used in Fig S8 are provided in Table H.

Multiplicative noise magnitude	values
h_{r1}	0.002
h_{r2}	0.002
h_{f1}	0.002
h_{f2}	0.002
h_{v}	0.02
h_{i}	0.02
h_{k}	0.02
h_{h}	0.02

Table H. Multiplicative noise magnitudes used in Fig S8.

S10. Model with cell proliferation and growth

In the patterning period we have modeled, cells divide in the hindbrain. Cell proliferation and growth are coupled with cell migration along the D-V axis that thickens the neural plane. As a

result, in the two-dimensional plane consisting of A-P and L-R axes, the number of cells remains relatively unchanged (Addison et al., 2018). We therefore simplified our two-dimensional model by not considering cell proliferation and growth.

Between 11-14 hpf in zebrafish embryos, cells begin to enter the 16^{th} cell cycle and divide asynchronously with an average cell cycle length of 4 hours (Kimmel et al., 1994). Cells tend to divide and separate in the direction perpendicular to the A-P axis (Kimmel et al., 1994). We modeled cell division according to the experimental evidence on cell cycle and division orientation. In the model, we randomly assigned division times for each cell in a 4-hour window (11-15 hpf), where a cell would not divide if it had not divided by 14 hpf (Fig S9C). For each cell, division has three key stages: elongation, separation, and growth (Fig S9A). First, a cell needs to determine its division line. We modeled each cell with a hexagonal shape. The division line goes through mid-points of two opposite sides of each hexagon. With three pairs of opposite sides, there are three possible choices of division line. We selected the division line having the minimum angle relative to the A-P axis (Fig S9A). After the division line is selected, the intrinsic lengths in Eq. (S11) for the sides intersected by the line of orientation grow linearly with similar velocities. Cell elongation will continue for 30 minutes until the lengths of the growing sides increase l_{out}. Then the cell divides into two daughter cells and the intrinsic length l_{ij} in Eq. (S11) for all sides becomes 60% of their default values. The two daughter cells grow to normal size in 30 minutes with linearly growing intrinsic lengths (Fig S9A).

Similar to Fig 4, we perform simulations to compare the effects of convergence rates on pattern formation (Fig S9B and S9D-F). Our major finding, that rapid initial convergence improves pattern robustness, remains. A clear five-segment pattern is established (Fig S9B), but cells distribute more tightly than in cases without cell division. The lengths of all three rhombomeres under three different convergence patterns (Fig S9D) are similar to the cases without division (Fig

4G), but have higher variance. Rapid initial convergence yields the smallest SIs for all four boundaries with both medium and slow initial convergence rates (Fig S9E).

S11. Models without convergent extension

Without convergent extension, morphogens change less because the tissue no longer elongates to scale morphogens distributions (Fig S10). Indeed, most cells commit to fates much earlier and the five-segment pattern is established around 12 hpf (Fig S10A). From 12-14 hpf, boundaries are refined through cell sorting, though some minor refinements of boundaries are observed. Comparisons between models with and without convergent extension (Fig S10B-D) reveal that without convergent extension, much smaller r3 and r5 A-P lengths result and all boundaries are rougher with higher SIs. r3 and r5 are only one or two cells long along the A-P axis, and such a limited number of neighboring cells with similar identities provides very weak adhesion and difficulty sorting cells back to their correct segments (Fig S10A). These observations further support the idea of a trade-off between segment size and boundary sharpness.

S12. Models without advections of intracellular morphogens

As the tissue deforms, the extra-cellular morphogens may have active motion driven by the tissue dynamics while the intra-cellular signals induced by morphogens may also move due to the cell movements in Eulerian coordinates. In a continuum model using Eulerian coordinates, the advections are usually required to capture the morphogen dynamics with moving boundaries (Čanić, 2021).

We also provide a model without advection for the intracellular morphogens for both RA and FGF, using the modified equations of Eq. (5) and (9):

$$\frac{\P[RA]_{in}}{\P t} = k_r[RA]_{out} - k_r[RA]_{in} - d_r(x^{(1)})[RA]_{in} + m_{r2}\frac{dw_{r2}(t)}{dt},$$
$$\frac{\P[Fgf]_{signal}}{\P t} = k_f[Fgf]_{free} - k_{rf}[Fgf]_{signal} - d_{f2}[Fgf]_{signal} + m_{f2}\frac{dw_{f2}(t)}{dt}. \quad \text{(S23)}$$

Our main result showing positive effects of rapid initial convergence on patterning remains when such advection is removed. The rapid initial convergence induces sharper boundaries over all four boundaries and fewer DCs than medium and slow initial convergence (Fig S11C and S11D). The discrepancy in intracellular RA between rapid and slow initial convergence decreases when advection is removed but still exists (Fig S11E and S11F).

Appendix II

Supplementary Material: Chapter IV

Crabps and retinoic acid signaling regulate germ cell proliferation and play a role in sex determination in zebrafish

TABLE S1

CRISPR Primers		
crabp1a	GCA GCT AAT ACG ACT CAC TAT AGG CGA TGC TCC GGA AAG TAG GTT TTA GAG CTA GAA	
crabp1b	GCA GCT AAT ACG ACT CAC TAT AGC TGA GAA AAG TGG CTT GTG GTT TTA GAG CTA GAA	
crabp2a	GCA GCT AAT ACG ACT CAC TAT AGG TGA TGC TCC GTA AGA TTG GTT	
crabp2b	GCA GCT AAT ACG ACT CAC TAT AGG TGG TCC GAA CGC TGG TGG GTT	
Genotyping Primers		
	Forward	Reverse
crabp1a	TGG GGC AGT TTC AAA GAG TTA T	CAG TGG TGG AGG TCT TGA TGT A
crabp1b	CAG CTC AGT CAA GAG TGA CGA C	TTG ATA TAG AAC TGC TCC CCG T
crabp2a	ACA TCA CTC TCA TTC CCC CGA	TGG GTG TGT GTA TAC CGT GC
crabp2b	GCA GAC TGG ACA CGC TCA TTA	GAG AGA ATT TGT GGC GTA CTG TG

FIGURE S1

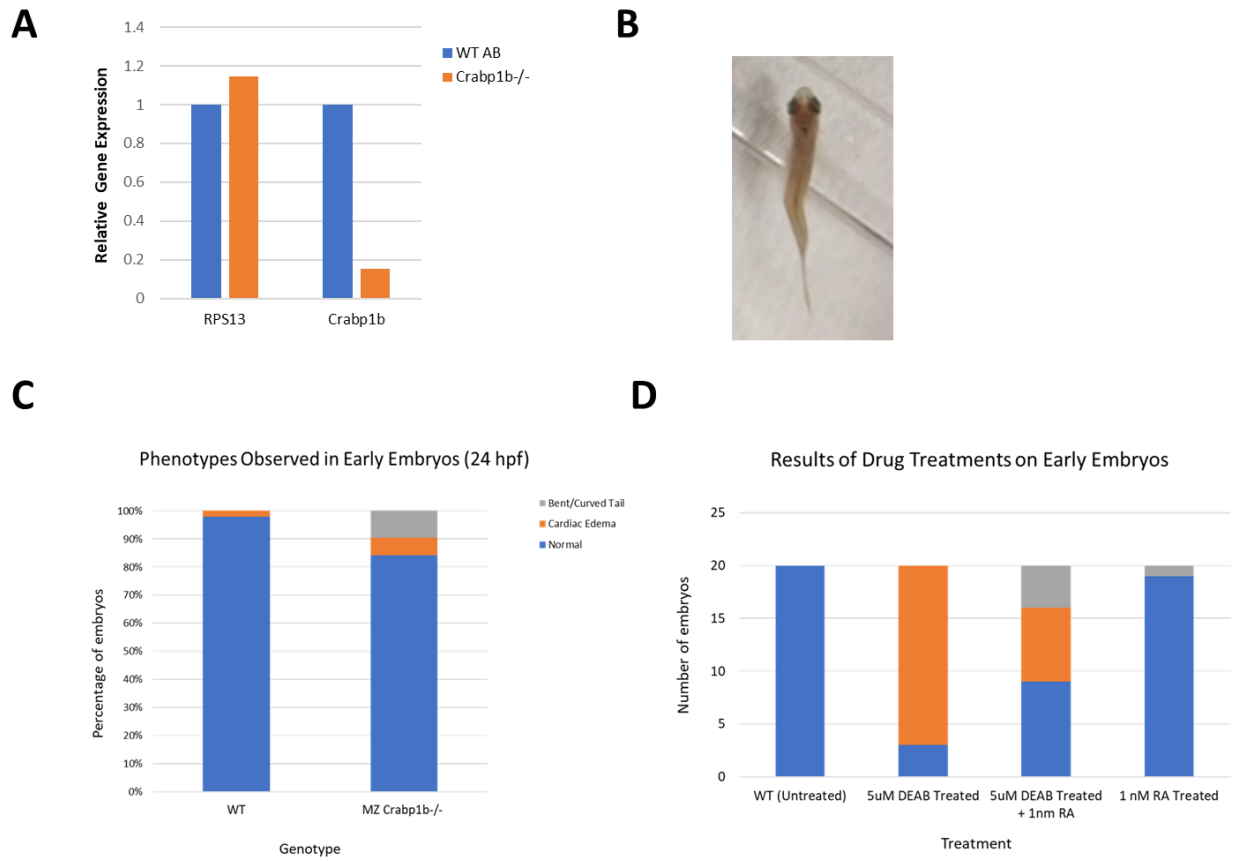

Figure S1. Phenotypic analysis of Crabp1b-/- mutants (A) qRT-PCR analysis shows relative fold-change in gene expression of *rps13* (housekeeping gene) and *crabp1b* in wild-type (WT) and Crabp1b-/- embryos at 36 hpf. Several embryos are pooled for each sample. (B) Representative image of scoliosis in Crabp1b-/- adult zebrafish (n= 7/38, 18.4%) (C) Bar graphs comparing phenotypes present in WT (cardiac edema, n= 2/96, 2.1%) and Crabp1b-/- (cardiac edema, n= 4/63, 6.3% and bent/curved tail, n=6/63, 9.52%) embryos at 24 hpf. (D) Bar graphs comparing proportion of phenotypes present in embryos treated overnight with exogenous RA and/or DEAB (inhibitor of RA biosynthesis): 5uM DEAB treated (cardiac edema, n= 3/20, 15%), 5uM DEAB treatment with 1nM RA (cardiac edema, n= 7/21, 33.33% and bent/curved tail, n=4/21, 19.05%), and 1 nM RA treated (bent/curved tail, n= 1/20, 5%). Proportion of embryos with specific phenotypes are indicated by follow colors: normal embryos (blue), bent/curved tails (grey), and cardiac edema (orange).

FIGURE S2

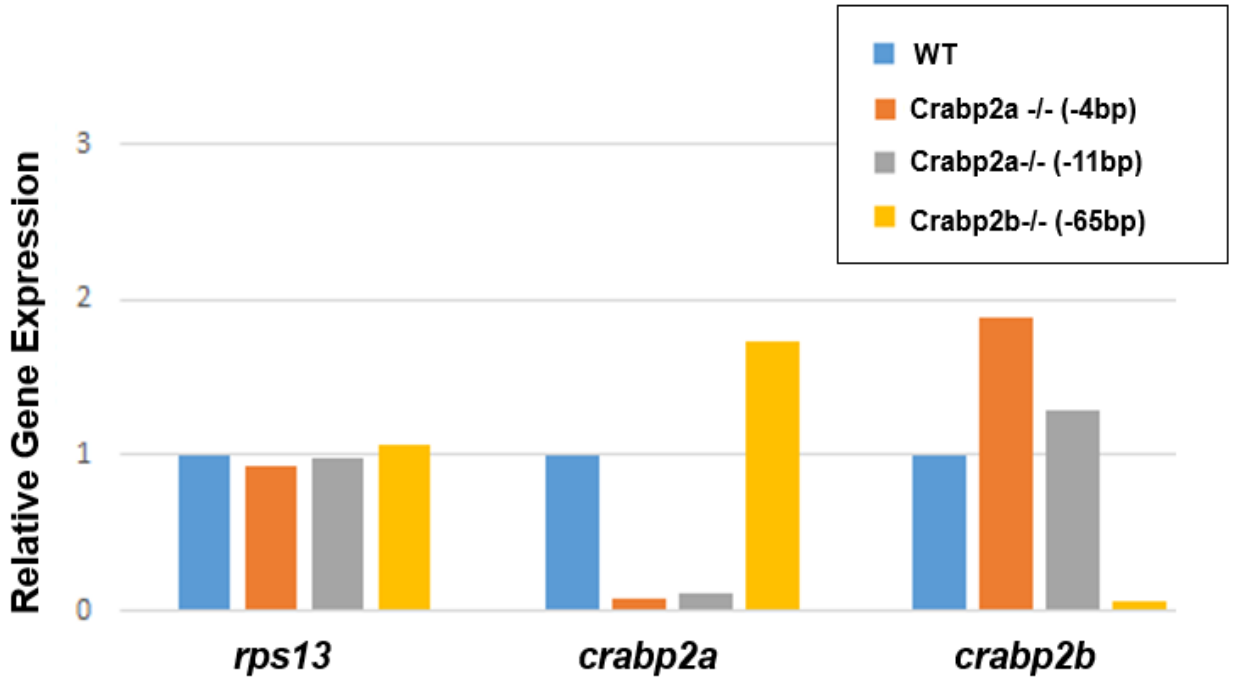

Figure S2. Upregulation of remaining *crabp2* paralogue in Crabp2 mutants qRT-PCR analysis shows relative fold-change in gene expression of *rps13*, *crabp2a*, and *crabp2b*, in WT, Crabp2a-/-, and Crabp2b-/- embryos at 24 hpf. Several embryos are pooled for each sample.