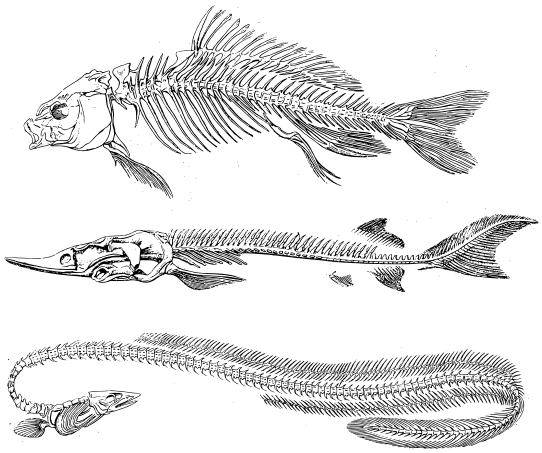
The evolvability of segment number in vertebrates



James Earwaker Hammond

Merton College, University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Hilary 2025

Acknowledgements

Every developmental biologist knows that it is the mother who most influences the life of the child, and this thesis is no exception. I would like to thank my mother for first fostering my interest in the natural world, and teaching me the names of birds, plants, and trees. Without her this thesis would never have been written. I would also like to thank my partner, Georgina Brown, for her loving and essential support throughout my studies. Thank you too to my father for his constant support and for first introducing me to the science of ichthyology on the banks of a gravel pit in Norfolk. Thank you too to my sister and my grandmother for their endless encouragement and support.

Thank you too to my supervisors, Dr Berta Verd Fernandez and Prof Ruth Baker, to whom this thesis belongs as much as it does to me. Thank you to Berta for providing me with resources, funding, training, feedback, advice, ideas, knowledge, and for first agreeing to meet with me six years ago. Thank you to Ruth for being an endless source of sage advice and ideas, and for the excellent feedback on this work. Thank you to Christine Soper, Helen Sanders, Dave Wilson, and Allex Turner, for their essential support in fish husbandry. Lastly, but by no means the least, thank you to my colleagues Callum Bucklow and Shannon Taylor for being really fantastic people to work with, and for five years of good craic.

Abstract

Understanding the developmental origins of evolvability, the capacity of a lineage to evolve novel heritable and adaptive phenotypes, is a central question in evolutionary developmental biology. Vertebrates show a high degree of evolvability in the number of vertebrae dividing their axial skeleton, which has allowed vertebrates to adopt different lifestyles and habitats across the tree of life. Vertebral number is determined by the number of somites that form in the early embryo. Somites are paired blocks of mesoderm which form sequentially in an anterior-to-posterior fashion via a process known as somitogenesis. The total number of somites formed depends on the frequency at which somites appear, and the duration of somitogenesis. The frequency of somitogenesis depends on the frequency of oscillations of a gene regulatory network known as the segmentation clock. While somitogenesis is a paradigmatic example of pattern formation in development, it is unknown which properties make somitogenesis evolvable.

In this thesis I aim to uncover the sources of evolvability in somitogenesis. I use computational modelling to predict that the frequency and duration of somitogenesis exhibit developmental modularity in Zebrafish, suggesting that vertebral number may be evolvable because large changes in vertebral number can be created by small changes in both frequency and duration. This modularity depends on properties of the segmentation clock. I compare somitogenesis in two closely related species of Cichlidae with divergent vertebral number. I find that the difference in vertebral number is due to differing duration of somitogenesis and not frequency, suggesting that duration is a more evolvable component of somitogenesis than frequency. By perturbing the segmentation clock I find that modularity of somitogenesis is conserved across most of the bony fishes, approximately 34% of extant vertebrate species. This thesis suggests that evolvability of vertebral number is underpinned by modularity of the frequency and duration of somitogenesis.

Contents

List of Figures **xi**

List of Tables **xxiii**

1 Introduction **1**

1.1 Evolvability and evo-devo 3

1.2 Evolvability of vertebral number 6

1.3 The developmental control of vertebral number — Somitogenesis 9

1.3.1 The segmentation clock 10

1.3.2 The clock and wavefront model 12

1.3.3 The wavefront 13

1.3.4 Elongation of the pre-somitic mesoderm 14

1.3.5 The initiation and termination of somitogenesis 15

1.3.6 Mathematical models of the segmentation clock 17

1.4 The evolution of somitogenesis 24

1.4.1 The respective roles of the clock and elongation in evolution 25

1.4.2 Sources of evolvability in somitogenesis 30

1.5 Thesis aims 34

2 Modularity of the segmentation clock and morphogenesis **35**

2.1 Introduction 39

2.2 Results 45

2.2.1 Varying the position of cell ingression has minimal effect on clock dynamics 45

2.2.2 The clock is robust to changes in mode of cell ingression regardless of the cell motility profile 48

2.2.3 Length and density of the PSM confer robustness to the clock . . . 51
2.2.4 Clock arrest during cell division creates asynchrony 57
2.2.5 Zebrafish clock coupling and cell mixing confer robustness to changes in cell ingression 61
2.3 Discussion 65
2.4 Methods 71
2.4.1 Computational model of the segmentation clock and PSM 71
2.4.2 Parameter values 86
2.4.3 Analysis and metrics 87

3 Navigability of a fitness landscape in somitogenesis 113
3.1 Introduction 115
3.2 Results 117
3.2.1 Modelling a general mechanism for PSM elongation 117
3.2.2 Structure of a fitness landscape for general mechanisms of PSM elongation 120
3.2.3 Computationally navigating a fitness landscape for general mechanisms of PSM elongation 126
3.3 Discussion 131
3.3.1 Future work 133
3.4 Methods 135
3.4.1 Modelling PSM morphogenesis 135
3.4.2 Random walk exploration of parameter space 138
3.4.3 Implementation 141
3.5 Supplementary material 142

4 Quantitative biology of the segmentation clock in the Lake Malawi Cichlidae *Rhamphochromis* sp. 'chilingali' and *Astatotilapia calliptera* 145
4.1 Introduction 146
4.1.1 The natural history of Lake Malawi Cichlids 148
4.1.2 Quantification of the segmentation clock using chemical inhibitors 152
4.1.3 Aims and objectives 153
4.2 Results 154

4.2.1 Dynamics of somitogenesis in *Rhamphochromis* sp. 'chilingali' and *Astatotilapia calliptera* 154
4.2.2 Disrupting clock synchrony using the γ-secretase inhibitor DAPT . 156
4.2.3 Modelling cichlid somitogenesis 169
4.3 Discussion 171
4.4 Methods 177
4.4.1 Animal husbandry and collection of embryos 177
4.4.2 Embryo culture 178
4.4.3 Quantifying somitogenesis in *Rhamphochromis* sp. 'chilingali' and *Astatotilapia calliptera* 178
4.4.4 Drug treatment of cichlid embryos 179
4.4.5 Chemical dechorionation of cichlid embryos 180
4.4.6 *in situ* visualisation of gene expression using Hybridisation Chain Reaction 180
4.4.7 Culture of PSM explants 180
4.4.8 Models and parameter fitting 181
4.5 Supplementary methods — protocols 181
4.5.1 Cichlid embryo culture 181
4.5.2 Hybridisation Chain Reaction 181
4.5.3 Pronase chemical dechorionation of cichlid embryos 184
4.5.4 PSM explant dissection and culture 186

5 Discussion **189**
5.1 Aims, approach, and findings 189
5.2 Modularity and the evolvability of vertebral number 193
5.3 The evolution of modularity 195

References **199**

List of Figures

1.1 **Vertebrate diversity.** Vertebrates exhibit diverse numbers of vertebrae, such as (left to right) 16 in the Oceanic sunfish *Mola mola*, $\geq$750 in *Nemicthys* Snipe eels, 50 in the Chicken *Gallus gallus*, and 33 vertebrae in the author (*Homo sapiens*). Image credits (left to right): I. Reijs &Hutten, J.N., CC BY 2.0; S. Raredon (Smithsonian Institution); P. Pikart CC BY-SA 4.0; J. Hammond, self-portrait. 2

1.2 **The role of the genotype-phenotype map in evolution (from Oster &Alberch, 1982 [7])**. Illustration describing how the structure of development parses genotypes into different phenotypes. The nonlinearity of development means that multiple genotypes can correspond to a given phenotype, and these genotypes can be thought of as varying the complement and magnitude of 'parameters' for development. The possible phenotypic transitions that can therefore occur in evolution are restricted to transitions between phenotypes that lie adjacent to one another in developmental parameter space. The fitness of these phenotypes is assessed by natural selection, determing thus the phenotypes that are realised in evolution and observable. © 1982 The Society for the Study of Evolution, reprinted with permission. 4

1.3 **Diversity in vertebral number across the ray-finned fishes.** Family-level phylogeny of the ray-finned fishes (Actinopterygii) after [65], with tips coloured according to the log-transformed minimum number of vertebrae reported for that family by [66]. Silhouettes were downloaded from PhyloPic (`https://www.phylopic.org/`). Image credits and licenses: Emily Troyer (CC BY 4.0), Erika Schumacher (CC BY-NC 3.0), U.S. Fish and Wildlife Service (illustration) and Timothy J. Bartley (silhouette) (CC BY-SA 3.0), Tree of Life App (CC BY 4.0), Mykle Hoban (CC BY-SA 3.0), Grupo de Ictiología de la Universidad de Antioquia (CC BY 4.0), and Richard Winterbottom and Russell Engelman (CC BY 4.0). **Methods:** Mapping of vertebral number to phylogenies was completed in R (v4.2.0) [67]. Vertebral numbers for 271 families were collated from [66]. CV Bucklow (CVB) pruned the time-calibrated phylogeny published by [65] to the families present within this vertebral number dataset using the R package, ape (v5.7.1) [68]. To reduce each family to a single branch and to account for the estimated age of each respective family, the tree was pruned by CVB to the longest (i.e. most basal) branch within each family. Log-transformed minimum total vertebral and ancestral states were mapped and inferred by CVB using the function contMap, part of the phytools (v2.1.1) package [69]. R code for generating the figure was written by CVB. 8

1.4 **Somitogenesis in Zebrafish (*Danio rerio*), after Kimmel *et al.* (1995) [74].** **A** Sketch of a 14 somite-stage Zebrafish embryo, viewed from the left-hand side (adapted from [74]). At this stage of development, the somites are visible, dividing up the developing trunk. The pre-somitic mesoderm (PSM) is at the posterior tip of the embryo. **B** Sketches of 10, 14, 18, and 21 somite-stage Zebrafish embryos, viewed from the left-hand side (adapted from [74]). Somite number increases over time, as the embryo grows posteriorly and new somites are added sequentially at the posterior end of the embryo. © 1995 Wiley-Liss, Inc. Reprinted with permission of John Wiley &Sons. 9

1.5 **The segmentation clock. A** Simplified diagram of the segmentation clock circuit. Dimers of Hes/Her transcription factors (magenta) repress their own expression and the expression of Delta protein (green), causing oscillations in the expression of Hes/Her and Delta. Delta protein transduces signals via Notch (orange), which causes cleavage of notch intracellular domain (NICD, black arrow). NICD activates expression of Hes/Her, thus communicating information about the phase of the clock to the clock in a neighbouring cell. **B** The expression of Hes/Her and Delta oscillates in space and time. Cartoon image of an idealised PSM viewed from above — the PSM is divided by the developing spinal cord and notochord and bounded at the anterior (top of figure) by the somites. 12

1.6 **The clock and wavefront model.** Travelling waves of clock gene expression (blue) emanate from the posterior of the PSM and travel towards the PSM anterior (blue arrows). At a point in space called the wavefront (magenta), the phase of clock gene expression is 'read out' by development. If cells expressing clock genes in the correct phase of the oscillatory cycle cross the wavefront, then a somite boundary is patterned. The wavefront moves slowly (relative to the frequency of the clock) towards the posterior of the PSM (magenta arrow), and thus a second somite boundary is patterned, creating a somite. Concurrent with this the PSM elongates (orange arrow), allowing the process to carry on in principle indefinitely, forming somite boundaries. 13

1.7 **Diversity in somitogenesis dynamics across the vertebrates. A Top:** the same number of somites (black rectangles) can be achieved by changes in either frequency (top) or duration (bottom) of somitogenesis. **Bottom:** Scenarios for plotting the total number of somites formed during embryogenesis against somitogenesis frequency. If diversity is due to only evolutionary changes in frequency, the datapoints should lie along a linear trendline (left, black). Conversely if diversity is due to only changes in duration, the datapoints should lie along a vertical line (middle, blue). If diversity is created by a combination of the two, then the datapoints are heterogeneously spread across the space (right, mauve). **B** The total number of somites formed during embryogenesis, plotted against the frequency at which somites are formed. A linear trendline is fitted from regression of the log-transformed data, with trendline $y = 0.023x - 1.765$ with the summary statistics $p = 0.8423$, $R^2 = 0.001$, suggesting that there is no correlation between the dependent and independent variables. **C** The time taken to reach phylotypic stage (defined as presence of four pharyngeal arches), plotted against the somitogenesis period. A linear trendline is fitted from regression of the log-transformed data, with trendline $y = 0.835x - 0.467$ with the summary statistics $p << 0.01$, $R^2 = 0.674$, suggesting a statistically significant relationship. 29

1.7 (Continued from previous page). **D** The total number of somites formed during embryogenesis, plotted against the somitogenesis period divided by the time taken to reach phylotypic stage. A linear trendline is fitted from regression of the log-transformed data, with trendline $y = 0.767x + 1.626$ with the summary statistics $p << 0.01$, $R^2 = 0.529$, suggesting a statistically significant relationship with poor explanatory power. Silhouettes were hand-drawn or downloaded from Phylopic (`https://www.phylopic.org`). Image credits and licenses: Edwin Price CC BY 4.0 and DFoidl and T. Michael Keesey CC BY-SA 3.0. 30

2.1 **Computational model of the clock and the PSM. A** Stills of a simulation of the model of [54]. Kinematic phase (θ_i) waves emerge in the posterior (right) and travel towards the tissue anterior (left, $x = x_a$), where phase is arrested. The model is parameterised to data from zebrafish, and accordingly the clock oscillates every 30 minutes. **B** Insets illustrating the key processes driving cell movements in the PSM within the model. Top: Cells advect towards the anterior of the tissue, simulating elongation of the PSM. Middle: New cells are added to replenish the loss of cellular material as cells advect towards the anterior. Bottom: Cells undergo motility-driven rearrangements. **C** Functions in the model describing (top) the intrinsic oscillation frequency, (middle) the advection velocity, (bottom) and the motility, of each cell depending on its normalised position along the anterior-posterior axis, $(x - x_a)/L$. Plots were generated using the parameters given by [54]. 44

2.2 **The effect of cell ingression position on clock frequency and synchrony. A** Diagram highlighting the position of cell addition across the three ingression scenarios tested here. Magenta shading shows where cells are added onto the tissue surface. The pale magenta and blue planes respectively correspond to the anterior limit of the PSM, $x = x_a$, and the anterior limit of cell addition, $x = x_d$. In the 'Random' condition cells are added at random positions within the PSM posterior to the plane $x = x_d$, and accordingly no magenta surface is shown. Similarly in the dorso-posterior case ('DP') cells are added at random positions in the two lateral cylinders to maintain density, and no magenta surface is shown there. In the dorso-posterior + lateral-ventral ('DP+LV') case cells are only added on the tissue surface at the positions shown by magneta shading. **B** Oscillation synchrony (r) at the PSM anterior after 1000 mins, for the three tested scenarios of cell ingression. $N = 100$ simulations. **C** Mean frequency of oscillations for cells at the PSM anterior after 1000 mins of simulation, for the three tested scenarios of cell ingression. $N = 100$ simulations. **D** Kymographs of synchrony along the x-axis on the left-hand side of the PSM for single simulations from each of the three scenarios of cell ingression tested. Black arrowheads highlight strongly asynchronous populations of cells being transported to the tissue anterior by advection. . 47

2.3 **Effect of cell motility profile on clock frequency and synchrony. A** Overview of how intrinsic cell motion is encoded in the model. Each cell is given a random direction vector $v_0(\mathbf{x}_i)\mathbf{n}_i(t)$ (black arrows) whose magnitude $v_0(\mathbf{x}_i)$ increases towards the PSM posterior. **B** Magnitude of intrinsic cell motion $v_0(\mathbf{x})$ for the parameters used by [54] (blue), and how the shape of the function can change when increasing (yellow) or decreasing (green) the inflexion point and curve steepness parameters, X_v and h, respectively. **C** Clock synchrony r (**top**) and difference from expected mean frequency $\Delta \mathrm{d}\theta/\mathrm{d}t$ (**bottom**) for different $v_0(\mathbf{x})$ specified by combinations of X_v and h. The corresponding pixels display the synchrony or frequency at the PSM anterior after 1000 mins of simulation using the specified motility profile, averaged across $N = 100$ simulations. A black + corresponds to the parameter pair used elsewhere in this paper, unless otherwise stated ($X_v = 0.4, h = 3$). Parameter ranges used are $X_v \in \{0, 0.1, 0.2, \ldots, 1\}$ and $h \in \{0, 0.5, 1, \ldots, 5\}$ 50

2.4 **Effect of tissue density and length. A** Stills from exemplar simulations illustrating the impact of changes in tissue density ρ_0. **B** Anterior synchrony after 1000 mins for changing tissue density ρ_0 and varying position of cell ingression. $N = 100$ simulations. **C** Anterior mean frequency after 1000 mins for changing tissue density ρ_0 and varying position of cell ingression. $N = 100$ simulations. **D** Stills from exemplar simulations illustrating changes in tissue length. **E** Anterior synchrony after 1000 mins for changing tissue length L and varying position of cell ingression. $N = 100$ simulations. **F** Anterior mean frequency after 1000 mins for changing tissue length L and varying position of cell ingression. $N = 100$ simulations. 53

2.5 **Effect of cell division on clock frequency and synchrony. A** Diagram showing how during cell division, clock phase θ arrests, causing a cell to fall out of phase with its neighbour. On the right hand side a still from a simulation for T_M = 15 mins is shown, illustrating how this creates asynchrony of oscillations. The effect of cell division on anterior synchrony and frequency is shown in figures **B** and **C**, respectively, for $T_M = 15$ mins after 1000 mins. $N = 100$. **D** The effect on anterior synchrony when division is restricted to only occur posterior to $x = x_{div}$, for $T_M = 15$ mins. $N = 100$. **E** The effect on anterior synchrony after 1000 mins when T_M is varied. In each case, the total length of the cell cycle is maintained at a constant length, i.e. $T_M + T_G = 187.5$ mins. $N = 100$. To rule out trivial changes in synchrony and frequency, in all analysis here we restrict measurement to non-dividing cells, i.e. cells such that $\tau \in [0, T_G)$. 60

2.6 **Synchrony for varying coupling strength κ and magnitude of intrinsic cell motion v_s.** Anterior synchrony after 1000 mins, for varying maximum magnitude of intrinsic cell motion v_s and clock phase coupling strength κ, for three different scenarios of cell ingression. Each pixel corresponds to the median value of anterior synchrony across $N = 100$ simulations. A black + marks the experimental values for zebrafish, $\kappa = 0.07$ min^{-1}, $v_s = 1$ μm · min^{-1}, derived in [49] and [140] respectively, that are used elsewhere in this paper. All other parameters are held constant at their normal values (see table **??**). 62

2.7 **Clock dynamics in the presence of a coupling delay between cells. A** Diagram illustrating phase coupling with delay. Cells couple their phase to the phase of their neighbours τ minutes ago. **B** Anterior synchrony in the presence of coupling delay ($\tau = 21$ mins) after 1000 mins, for varying maximum magnitude of intrinsic cell motion v_s and clock phase coupling strength κ, for three different scenarios of cell ingression. Each pixel corresponds to the median value of anterior synchrony across $N = 100$ simulations. A black + marks the experimental values for zebrafish, $\kappa = 0.07$ min^{-1}, $v_s = 1$ μm · min^{-1}, derived in [49] and [140] respectively, that are used elsewhere in this paper. All other parameters are held constant at their normal values (see table 2.1). **C** Anterior synchrony (**top**) r after 1000 mins, for varying phase coupling delay τ and clock phase coupling strength κ, for three different scenarios of cell ingression. Each pixel corresponds to the median value of anterior synchrony across $N = 100$ simulations. **Bottom**: Difference from expected mean frequency ($\Delta \mathrm{d}\theta/\mathrm{d}t$) at the PSM anterior after 1000 mins, for three different scenarios of cell ingression. Each pixel corresponds to the mean value across $N = 100$ simulations. A black + marks the experimental values for zebrafish, $\kappa = 0.07$ min^{-1}, $\tau = 21$ mins, derived in [49] and [139] respectively. All other parameters are held constant at their normal values (see table 2.1). 64

2.8 **Geometry of the PSM assumed in the present model. Top left:** major axes (x, y, z) used in the model. x corresponds to the anterior-posterior axis of the embryo and increases towards the tissue posterior, y corresponds to the left-right axis and increases to the right hand side of the tissue, and z corresponds to the dorsal-ventral axis of the embryo. **Top right:** schematic of the PSM in the xy plane. The PSM is comprised of two cylinders, centred at $y = r_T$ and $y = 2R + r_T$, respectively, with radius r_T. The 'tailbud' is represented as a half-torus subdomain centred at $\mathbf{x} = (X_c, Y_c, Z_c)^T$, with minor radius r_T and major radius R. **Bottom:** cross-sections of the tissue showing how a point $\mathbf{x}_i$ is assigned the polar coordinates r_i and q_i, in both the PSM cylinders ($x_i \leq X_c$) and the half-toroid tailbud ($x_i > X_c$). Adapted from [54]. . . . 73

2.9 **Exemplar dynamics of synchrony and mean anterior frequency. Top** Trace of synchrony r for a d_c-wide domain of cells at the left-hand anterior edge of the PSM over 1000 mins. Data drawn from a single simulation with random cell addition, using parameters as per [54]. **Bottom** Trace of mean frequency $\mathrm{d}\theta/\mathrm{d}t$ for a d_c-wide domain of cells at the left-hand anterior edge of the PSM over 1000 mins. Data drawn from the same simulation as above. Plotted with a black dashed line is the average intrinsic frequency $\omega(x)$ across the domain, calculated using the formula shown. 88

3.1 **Can evolution traverse parameter space between different strategies for elongating the PSM?** The PSM is elongated (magenta arrow) by at least four qualitatively distinct processes — ingression of progenitor cells (blue), spatial profiles of intrinsic cell motility (orange), cell division (purple), and compaction-extension (green). Is it possible for evolution to traverse parameter space between each of these processes in isolation without perturbing the clock? How is this affected by co-evolution of the clock? 117

3.2 **Modelling a general mechanism for PSM elongation. A** The parameter x_g can be varied continuously to transition the cell ingression regime from lateral-ventral and dorso-posterior ingression of progenitor cells (black arrowheads) to only dorso-posterior ingression (black arrowheads). **B** The parameter h can be varied to create profiles of cell motility ($v_s(x)$) with varying steepness. **C** The amount of cell division within the PSM can be adjusted by varying the total length of the G1, S, and G2 phases of the cell cycle T_G. **D** The rate at which tissue length L, radius r_T, and cell diameter d_c decrease and tissue density ρ increases over time can be co-varied by changes to the parameter α. 119

3.3 **Clock synchrony across the space of possible PMS elongation strategies.** Synchrony of the clock r at the PSM anterior at the end of a simulation, for varying anteriormost position of ventral cell ingression x_g, steepness of cell motility profile h, duration of G1, S, and G2 phases of the cell cycle T_G, and rate of tissue compaction α. Each pixel represents the median value of $N = 50$ simulations. 122

3.4 **Navigability of the fitness landscape for PSM elongation for different synchrony thresholds for somite patterning. A** SG-t-SNE embedding of the network describing the fitness landscape for $r_c = 0.7$. Blue vertices correspond to parameters with synchronous clock dynamics, and are connected by grey lines if they lie adjacent to one another in parameter space. Embedding was generated using the SG-t-SNE-π software in Julia v1.8.2 [206]. **B** Top: The proportion of parameter space with synchrony $r \geq r_c$, V, for varying r_c. Middle: The number of disconnected clusters N_s for varying threshold r_c. Bottom: The ratio of the number of elements in the second largest cluster to the number of elements in the largest cluster (C_r), for varying threshold r_c. **C** The proportion of parameter space with synchrony $r \geq r_c$, V, for varying r_c and coupling strength κ. **D** The number of disconnected clusters N_s for varying threshold r_c and coupling strength κ. **E** The ratio of the number of elements in the second largest cluster to the number of elements in the largest cluster (C_r), for varying threshold r_c and coupling strength κ. 124

3.5 **Random walks exploring parameter space. A** PCA projection of the results of $N = 400$ independent random walks ($N = 100$ from each initial condition), for instantaneous coupling ($\tau = 0$ mins) coloured according to the synchrony r at the end of the simulation. Individual parameter sets visited on a walk are shown as circles and coloured according to their clock synchrony. Lines connecting these circles show the connection of any two parameter sets as part of a walk. Locations of the four initial conditions are shown in blue, orange, green, and mauve dots (see legend, bottom). The shortest path across parameter space between the 'division' and 'motility' initial conditions is plotted in magenta. **B** Distributions of parameters visited as part of the $N = 400$ independent walks. **C** Distribution of path lengths between the 'division' and 'motlity' initial conditions. **D** PCA projection of the results of $N = 400$ independent walks ($N = 100$ from each initial condition), for delayed coupling ($\tau = 30$ mins) coloured according to the synchrony r at the end of the simulation. Individual parameter sets visited on a walk are shown as circles and coloured according to their clock synchrony. Lines connecting these circles show the connection of any two parameter sets as part of a walk. Locations of the four initial conditions are shown in blue, orange, green, and mauve dots (see legend, bottom). **E** Distributions of parameters visited as part of the $N = 400$ independent walks. 129

3.6 **Random walks across a space of PSM morphogeneses and clock coupling. A** PCA projection of the results of $N = 400$ independent heuristic searches ($N = 100$ from each initial condition), for varying coupling strength κ and coupling delay τ, coloured according to the synchrony r at the end of the simulation. Locations of the four initial conditions are shown in blue, orange, green, and mauve dots (see legend, bottom). The shortest path across parameter space between the 'division' and 'motility' initial conditions is plotted in magenta. **B** Distributions of parameters visited as part of the $N = 400$ independent walks. 130

4.1 **A cichlid model system for the rapid evolution of vertebral number, *Astatotilapia calliptera* and *Rhamphochromis* sp. 'chilingali'. A** Map of the Lake Malawi watershed, with an inset showing its position within the continent of Africa. The origins of the fish used in this study are annotated, with the stock of *Astatotilapia calliptera* being descended from fish from the Mbaka river in Tanzania at the northern end of the lake, and *Rhamphochromis* sp. 'chilingali' being endemic to the small Lake Chilingali in Malawi on the eastern coast of the lake. Photo credits: *A. calliptera*, Howard Jelks, U.S. Geological Survey. *R.* sp. 'chilingali', Aleksandra Marconi. **B** Males of each species studied here, *A. calliptera* and *R.* sp. 'chilingali', with the number of vertebrae annotated. Photo credits: *A. calliptera*, Howard Jelks, U.S. Geological Survey. *R.* sp. 'chilingali', Aleksandra Marconi. **C** Mouthbrooding female *A. calliptera*. The diagnostic distension of the lower jaw indicating mouthbrooding is highlighted with a black arrowhead. **D** Somitogenesis-stage embryos of *A. calliptera* (left) and *R.* sp. 'chilingali' (right). The outline of the embryo is shown by a white dashed line. The head anlage and tailbud are annoated with 'ha' and 'tb' respectively. White scale bar indicates 1 mm. These images of cichlid embryos were taken from [146] and reproduced under a Creative Commons CC BY license; © 2023 The Authors. Evolution &Development published by Wiley Periodicals LLC. 150

4.2 **Dynamics of somitogenesis in *Rhamphochromis* sp. 'chilingali' and *Astatotilapia calliptera*. A** Linear rate of somite formation for individual clutches of cichlid embryos. N = 9 clutches for *A. calliptera*, N = 4 clutches for *R.* sp. 'chilingali'. **B** Somite formation over time (hpf). Trendline for *A. calliptera* is $S = t - 33.547$, trendline for *R.* sp. 'chilingali' is $S = 0.901t - 33.316$. The minimum number of somites reported by Marconi *et al.* (4) is annotated with a black dashed line. The typical maximum number of somites formed in each species, 32 and 38 for *A. calliptera* and *R.* sp. 'chilingali' respectively (S. Taylor, *pers. comm.*), are annotated with dashed lines coloured according to the species. N = 178 individual embryos for *A. calliptera*, N = 135 individual embryos for *R.* sp. 'chilingali'. 155

4.3 **The effect of DAPT treatment on cichlid embryos. Left:** HCR for xirp2a (green) in 4 dpf *Astatotilapia calliptera* embryos, either treated with just DMSO (top), or 20 μM DAPT (bottom), with the chorion intact. **Right:** HCR of deltaC (yellow) in 2 dpf *Astatotilapia calliptera* embryos, either treated with just DMSO (top), or 20 μM DAPT (bottom), with the chorion intact. 157

4.4 **Effect of DAPT treatment on delta expression in dechorionated *Astatotilapia calliptera* embryos.** HCR targeting deltaC (green) and deltaD (magenta), co-stained with DAPI (grey) to show embryo morphology. Left-hand panel shows an untreated embryo, and the right-hand an embryo that was cultured under 50 μM of DAPT. DMSO-treated embryos perished during culture and as such no embryo is shown in this panel. 160

4.5 **Effect of LY411575 on segment patterning in *Astatotilapia calliptera*.** **Left** HCR of segmentation stage *Astatotilapia calliptera* embryos staining for the genes deltaC (green) and deltaD (magenta). Nuclei (DAPI staining) are shown in grey. Bottom row shows an embryo treated with LY411575, and the top row a control embryo treated with only DMSO. Maximum intensity z-projections of z-stacked images. **Right** HCR of late pharyngula (4 dpf) *Astatotilapia calliptera* embryos staining for the gene xirp2a (green). Nuclei (DAPI staining) are shown in grey. Bottom row shows an embryo treated with LY411575, and the top row a control embryo treated with only DMSO. Maximum intensity z-projections of z-stacked images. 162

4.6 **Effect of DAPT on deltaC and deltaD expression in explanted PSMs of *Astatotilapia calliptera*. A** Cartoon showing experimental protocol. **B** Max-intensity z-projection of DAPI stain (nuclei) of an explanted PSM, showing the typical 'comma' morphology of explants, with the anterior-posterior and dorsal-ventral axes annotated. **C** HCR for deltaC (green) and deltaD (magenta) in explanted PSMs of *Astatotilapia calliptera* after 4 hrs of culture. Maximum intensity z-projections of z-stacked images. 164

4.7 **Quantifying the effect of notch inhibition on segment patterning in cichlids. A** HCR for myoD (magenta) in explanted PSMs of *Astatotilapia calliptera* after 8 hrs of culture, for varying concentrations of notch inhibitor DAPT Nuclei (DAPI staining) shown in grey. Maximum intensity z-projections of z-stacked images. Posterior-most correctly patterned segment ($ALD - 1$) for each explant is annotated with a white arrowhead. **B** Anterior limit of defect (ALD) plotted against DAPT concentration for data collected from explanted PSMs of *Astatotilapia calliptera* (top) and *Rhamphochromis* sp. 'chilingali' (bottom). Solid line in each plot shows equation 4.3 fitted to data for each species. 165

4.8 **Effect of notch inhibition on somitogenesis period and somite length.** **Left:** Period of somitogenesis in PSM explants, calculated as the total number of somites formed after dissection divided by the incubation time (8 hrs), for varying concentrations of the notch signalling inhibitor DAPT. **Right:** Anterior-posterior length of the most posterior somite for varying concentrations of DAPT. Solid line represents fit of equation 4.5 to the data (see methods). 168

4.9 **Model of the cichlid segmentation clock. A** Stills of simulations of the PSM and clock dynamics for the cichlids *Astatotilapia calliptera* and *Rhamphochromis* sp. 'chilingali', after 840 mins. Coupling strength κ and delay τ for *A. calliptera* and *R.* sp. 'chilingali' are shown in orange and blue, respectively. Parameter sets used in this figure: *A. calliptera* = $\{\omega_0 = 0.105\ \text{min}^{-1}, D\theta = 0.0069\ \text{min}^{-1}, v_a = 0.76\ \mu\text{m} \cdot \text{min}^{-1}, x_q = 1, L = 400\ \mu\text{m}\}$. *R.* sp. 'chilingali' = $\{\omega_0 = 0.094\ \text{min}^{-1}, D\theta = 0.009\ \text{min}^{-1}, v_a = 0.59\ \mu\text{m} \cdot \text{min}^{-1}, x_q = 1, L = 500\ \mu\text{m}\}$. **B** Synchrony at the anterior of the PSM after 1000 mins for varying coupling strength κ and coupling delay τ, using the parameter set for *A. calliptera*. Each pixel is coloured according to the median synchrony of $N = 100$ simulations for that pair of values. The estimated coupling strength and coupling delay for *A. calliptera* ($\kappa = 0.018\ \text{min}^{-1}$, $\tau = 46$ mins) is annotated with a black +. A black dashed line marks the solutions to equation 4.4 for $\kappa \in [0.017\ \text{min}^{-1}, 0.073\ \text{min}^{-1}]$. 171

4.10 **The development-adjusted somitogenesis frequency of Lake Malawi cichlids in the context of the vertebrates.** Distribution of the frequency of somitogenesis in vertebrates, scaled by the overall pace of development (see Introduction). For raw data and references see table 1.1. Frequencies for the Lake Malawi cichlids *Astatotilapia calliptera* and *Rhamphochromis* sp. 'chilingali' are plotted with dashed lines in orange and blue, respectively. Photo credits: *A. calliptera*, Howard Jelks, U.S. Geological Survey. *R.* sp. 'chilingali', Aleksandra Marconi. 173

5.1 **Modularity across the vertebrate clade. A:** Modularity of the segmentation clock and morphogenesis of the PSM appear to be conserved across the Euteleostei, based on data from Zebrafish (*Danio rerio*) and the Lake Malawi cichlids *Astatotilapia calliptera* and *Rhamphochromis* sp. 'chilingali'. It is unclear if this property is conserved with the Eels (Anguilliformes). Photo credits: *A. calliptera*, Howard Jelks, U.S. Geological Survey. *R.* sp. 'chilingali', Aleksandra Marconi. **B:** Three scenarios for the evolution of modularity of the segmentation clock and morphogenesis, across the vertebrates. Blue: Modularity is an ancestral feature of vertebrates, Pink: modularity is ancestral only to the gnathostomes, and Green: modularity is specific to the Actinopterygii. 195

List of Tables

1.1 **Somitogenesis period at a given temperature, the time to phylotypic stage, and total somite number, across species.** For each species the periodicity of somitogenesis, the time elapsed before one somite is formed, was either taken from a previously published value ('reported'), or was inferred by linear regression of somites over time from published data ('inferred'). The time taken to reach phylotypic stage, defined as the presence of pharyngeal arches, was not always reported and so in some cases this data is missing. †A Seleit *pers. comm.* 26

2.1 Parameter values used for simulations, unless otherwise stated 86

3.1 **Initial conditions for heuristic searches.** Parameter sets corresponding to the different initial conditions used in figures 3.5 and 3.6. All other parameters, unless otherwise stated in the text, where as described in the previous chapter. 141

4.1 **The timing of teleost development.** Approximate time of appearance of various stages of early development in Lake Malawi Cichlids (based on data from *Astatotilapia calliptera*, *Rhamphochromis* sp. 'chilingali', and *Tropheops* sp. 'mauve' [146]), compared to the model teleosts Zebrafish (data from [74]) and Medaka (data from [157]) 151

4.2 **Effect of detergent Triton X-100 and culture medium on embryo survival.** Survival of *Astatotilapia calliptera* embryos after overnight incubation at 28°C with a given treatment solution, sorted by developmental stage at which treatment began. 161

4.3 **Summary of the estimated values for parameters of a model of the segmentation clock**. Parameters for Zebrafish (*Danio rerio*) derived from previous work and used in the preceding two chapters, and the values inferred for the Lake Malawi cichlids *Astatotilapia calliptera* and *Rhamphochromis* sp. 'chilingali' in this chapter. *This is the value estimated from experimental data by Riedel-Kruse *et al.* [49]. Note that as this measurement incorporates noise due to morphogenesis, Uriu *et al.* estimate the true noise of the clock to be 0.0013 min^{-1} [54]. 171

4.4 **Comparing segmentation clocks and elongation in cichlid and Zebrafish.** Clock period T_{WT} and advection velocity v_a, normalised by the rate of development (measured by the time taken to reach phylotypic stage T_P). Parameters for Zebrafish (*Danio rerio*) derived from previous work and used in the preceding two chapters, and the values inferred for the Lake Malawi cichlids *Astatotilapia calliptera* and *Rhamphochromis* sp. 'chilingali' in this chapter. 176

4.5 **Gene transcripts used to design HCR probes.** 182

1

Introduction

Author's note: *My colleague Callum V Bucklow assisted me in the creation of figure 1.3. Some of the material in this section is currently under review at Seminars in Cell and Developmental Biology for publication as a standalone literature review. All other work here is my own.*

Understanding the developmental origins of evolvability, the capacity of a biological system to evolve new heritable and adaptive phenotypes, is a major research goal within the field of evolutionary developmental biology ('evo-devo') [1–5]. While evolvability is controlled by many aspects of biology (reviewed in [6]), it is thought to be majorly determined by the structure of the embryogenic process [1, 2, 7, 8]. Therefore the distributions of number, colour, and form we observe in the natural world owe their existence, at least in part, to the development of the embryo. Simply put, development proposes phenotypes, and evolution selects them. While there is much theoretical work on the role of development in determing evolvability, it is only recently that experimental studies have begun to focus on the role of development in the

evolvability of multicellular morphology [9, 10].

Vertebrates exhibit an enormous diversity in the number of vertebrae that comprise their axial skeleton (figure 1.1). For instance, the Oceanic sunfish (*Mola mola*) exhibits merely 16 vertebrae, dwarfed by the deep-sea *Nemicthys* Snipe eels, who can posess more than 750 vertebrae [11]. In many species this trait is adaptive, and the number of vertebrate is significant for locomotion by controlling the flexibility of the body axis [12, 13]. For instance, fossorial snakes show higher numbers of vertebrae than their arboreal cousins, adapting to a lifestyle of hunting in burrows and navigating tight corners [14]. Deep-sea fish, such as *Nemicthys*, typically show elongated bodies enabling energy-efficient anguilliform swimming [12]. Similar trends have been reported from mammals, for instance Deer mice (*Peromyscus maniculatus*) appear to adapt to arboreal life styles by increasing the number of vertebrae in their tail, promoting balance when navigating across branches [15].



Figure 1.1: Vertebrate diversity. Vertebrates exhibit diverse numbers of vertebrae, such as (left to right) 16 in the Oceanic sunfish *Mola mola*, ≥750 in *Nemicthys* Snipe eels, 50 in the Chicken *Gallus gallus*, and 33 vertebrae in the author (*Homo sapiens*). Image credits (left to right): I. Reijs & Hutten, J.N., CC BY 2.0; S. Raredon (Smithsonian Institution); P. Pikart CC BY-SA 4.0; J. Hammond, self-portrait.

Changes in the vertebral column have facilitated the adaption of vertebrates to diverse niches and lifestyles [12–15], and indeed the evolvability of this trait has likely been key to the success of the vertebrates. However, little is understood about what drives this evolvability. Furthermore, despite the developmental processes determining the number of vertebrae in the early embryo being well-characterised, little is known about what aspects of development may create evolvability. In this thesis I use a combination of computational and experimental approaches to explore the developmental origins of evolvability in vertebral number. Specifically I make an argument

for this evolvability being underpinned by the modular organisation of the developmental process, and present computational and experimental data in support of this claim. In the following section I have provided an overview of the field and the developmental processes that determine vertebral number, to give the reader context for my work.

1.1 Evolvability and evo-devo

The term 'evolvability' is widely credited to Dawkins (1989), who used a simple computer program to study whether a lineage's capacity to evolve and adapt under natural selection could itself evolve [16]. However the study of how properties of development may affect the trajectory of evolution is older, with seminal work by authors such as Goodwin [1], Alberch [7, 17], and Conrad [8] laying the foundations for the field (for a contemporary review see [2]). These authors reacted to a prevailing paradigm within the modern synthesis of evolutionary biology that implicitly assumed phenotypic variation was isotropic, and that natural selection was the only significant determinant of direction in evolution [2, 18, 19]. Alberch (1980–1991) used data on morphological variation and teratologies to argue against this assumption of isotropic variation, instead highlighting that this data shows that certain morphological transitions in evolution or certain teratologies are more likely than others, or never occur at all, a result apparently incompatible with an isotropic view of phenotypic variation [7, 17, 18, 20, 21]. Alberch also noted that the nonlinearity of the developmental process was capable of creating discrete changes in morphology under small continuous changes to the genome, which could explain apparently discontinuous changes in morphology across evolution [7, 21]. Similarly Goodwin (1982) argued against the isotropic assumption, noting that the process of development itself, which involves complex physical processes such as buckling of tissues and rearrangement of cells, precludes a gene-centric view of evolution where the gene is the ultimate causal agent in determining the phenotype [1]. Goodwin also argued that biologists should seek to codify a set of rules describing which phenotypes are possible, and thus that the proper place of evo-devo within evolutionary theory is determing the rules of phenotypic variation [1].

The structure of the developmental process controls which phenotypes are possible, and thus isotropic genotypic variation does not necessarily imply isotropic variation at the phenotypic level.

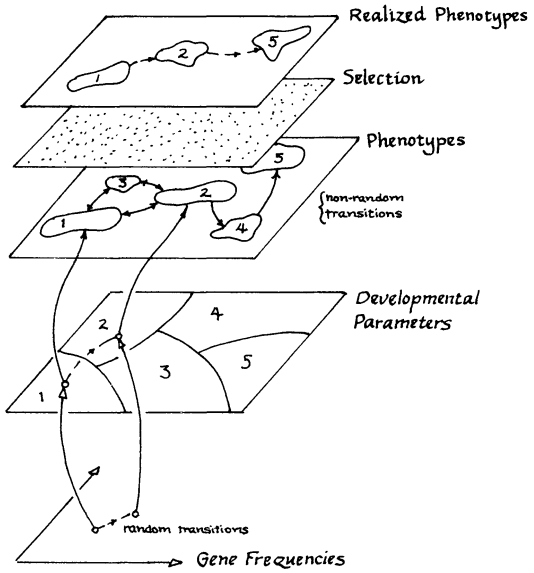


Figure 1.2: The role of the genotype-phenotype map in evolution (from Oster & Alberch, 1982 [7]). Illustration describing how the structure of development parses genotypes into different phenotypes. The nonlinearity of development means that multiple genotypes can correspond to a given phenotype, and these genotypes can be thought of as varying the complement and magnitude of 'parameters' for development. The possible phenotypic transitions that can therefore occur in evolution are restricted to transitions between phenotypes that lie adjacent to one another in developmental parameter space. The fitness of these phenotypes is assessed by natural selection, determing thus the phenotypes that are realised in evolution and observable. © 1982 The Society for the Study of Evolution, reprinted with permission.

However the role of development in evolution is not merely one of constraint [18, 19, 21]. The mapping of genotype to phenotype by development determines the different phenotypes that evolution can access from a given point in morphospace (figure 1.2), potentially directing evolution along particular trajectories in morphospace [8]. However the complexity of a developmental process can also improve acessibility of morphospace and free evolution from such biases [8]. The nonlinearity of development, the capacity for small genotypic changes to have large effects on the phenotype, allows for qualitative or rapid quantitative changes in phenotype during evolution [21]. The robustness of the developmental process to genetic change is an example of such a nonlinear behaviour, and can also faciliate such large phenotypic

changes by allowing gradual accumulation of small changes to the genotype without perturbing the phenotype, eventually allowing a final small genotypic change to effect a large change in the phenotype [22]. Analogously developmental plasticity, where the structure of development allows a phenotype to be specified by the combination of a genotype and a set of environmental conditions, can faciliate the evolution of new phenotypes by allowing rapid adaptation to a new environment followed by assimilation of the phenotype into the genome [23, 24]. Other properties of the genotype-phenotype map, such as pleiotropy [25] or stochasticity and resulting phenotypic heterogeneity [6, 26], can promote the rapid evolution of new phenotypes. Therefore, an accounting of development is essential in understanding morphological evolution.

It is worth noting that evolvability, generally defined as above, is not only dependent on properties of development or the genotype-phenotype map. For instance, changes in the rate of mutation or recombination of genetic material can have corresponding effects on the pace of evolution [27, 28]. Additionally the way in which adaptive or fit phenotypes are selected by evolution can affect the accessbility of different phenotypes to evolution [29–31]. While historically much of the work on evolvability has been theoretical, there now exists a large corpus of experimental studies investigating the causes of evolvability and their evolution (reviewed in [6]). Experiments involving techniques suited to microorganisms, such as experimental mutagenesis [32, 33], longtitudinal laboratory studies of evolution [31, 34–36] or design of synthetic gene regulatory networks [26, 37] have proven extremely powerful in furthering our understanding of evolvability. However outside of a handful of examples [9, 10, 38], there are relatively few experimental or data-driven studies of the developmental bases of evolvability in multicellular systems.

Work by Abzhanov and colleagues has revealed that the morphological evolution of the passerine beak, including within the famous adaptive radiation of 'Darwin's Finches' (*Geospiza, Camarhynchus, Certhidea, Platypsiza*, and *Pinaroloxias*) on the Galapagos and Cocos islands is faciliated by modularity of developmental pathways involved in beak development, but that the evolutionary potential is constrained by the physical mechanism by which the beak elongates

in development [10, 39–41]. Modularity in this context refers to the dissociability of two or more developmental processes by evolution, and has been suggested to be a major source of evolvability in development [42]. Additionally Jaeger and colleagues have shown that the gene regulatory networks underpinning arthropod segmentation permit neutral evolution, but can also be highly sensitive to changes in the strength of gene-gene interactions, potentially allowing for the rapid evolution of diverse modes of segment patterning [9, 43]. However such studies are few and far between, likely because they are challenging to conduct, as to identify sources of evolvability one must either work with groups of closely related organisms, for which experimental protocols and husbandry may not be established, or use data-driven models of development, which can be time-consuming to establish. With such practical constraints in mind, the evolution of vertebral number presents an excellent model for understanding of the developmental origins of evolvability. This trait is quantitative and thus it is straightforward to quantify the rate of morphological evolution across species. Additionally, the developmental mechanisms controlling vertebral number are well-characterised in a range of model species across the vertebrate clade, revealing a conserved mechanism [44, 45], thus enabling comparative studies of how this mechanism may facilitate evolutionary change [46, 47]. Finally, there also exists a rich repertoire of mathematical and computational models describing somitogenesis, which one can use to explore the evolvability of vertebral number [48–55].

1.2 Evolvability of vertebral number

The number of vertebrae in the vertebrate axial skeleton exhibits a high degree of evolvability, as is evidenced by the diversity of vertebral numbers seen in the ray-finned fish (figure 1.3). Here, we see repeated independent transitions between low and high vertebral number at a family-level resolution (figure 1.3), with the same pattern of repeated independent transitions in vertebral number being present within clades [56, 57], and at the infra-family scale, e.g. within African cichlids (Cichlidae, Pseudocrenilabrinae) [58, 59]. Vertebral number is also known to vary between populations of the same species, e.g. in japanese populations of the Medaka

(*Oryzias latipes*) which range from 30 to 31 vertebrae [60], british and irish populations of Arctic Charr (*Salvelinus alpinus*) (58 to 64 vertebrae) [61], antipodean populations of the Kōaro (*Galaxias brevipinnis*) (57 to 62 vertebrae) [62], scandinavian populations of the Northern pike (*Esox lucius*) (56 to 63 vertebrae)[63], or different ecomorphs of the Deer mouse (*Peromyscus maniculatus*) (23 to 27 caudal vertebrae) [15]. While some of this variation may be attributable to environmentally driven developmental plasticity [63, 64], in at least the Pike, Medaka, and Deer mouse such variation has been shown to be heritable [15, 60, 63]. It therefore appears that heritable changes in vertebral number can occur under presumably minor or at least evolutionarily rapid changes to the genotype, and thus that vertebral number exhibits a high degree of evolvability.

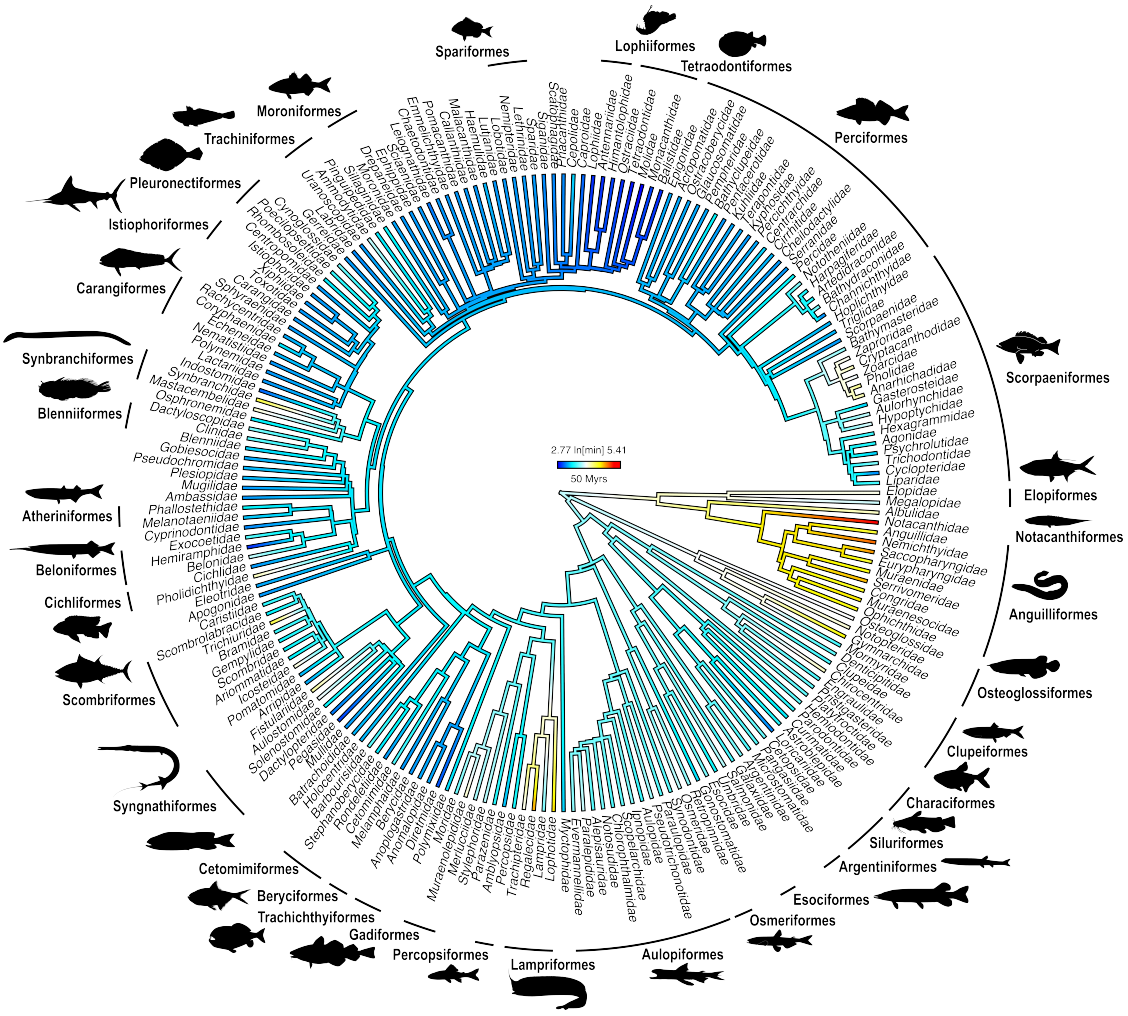


Figure 1.3: Diversity in vertebral number across the ray-finned fishes. Family-level phylogeny of the ray-finned fishes (Actinopterygii) after [65], with tips coloured according to the log-transformed minimum number of vertebrae reported for that family by [66]. Silhouettes were downloaded from PhyloPic (`https://www.phylopic.org/`). Image credits and licenses: Emily Troyer (CC BY 4.0), Erika Schumacher (CC BY-NC 3.0), U.S. Fish and Wildlife Service (illustration) and Timothy J. Bartley (silhouette) (CC BY-SA 3.0), Tree of Life App (CC BY 4.0), Mykle Hoban (CC BY-SA 3.0), Grupo de Ictiología de la Universidad de Antioquia (CC BY 4.0), and Richard Winterbottom and Russell Engelman (CC BY 4.0). **Methods:** Mapping of vertebral number to phylogenies was completed in R (v4.2.0) [67]. Vertebral numbers for 271 families were collated from [66]. CV Bucklow (CVB) pruned the time-calibrated phylogeny published by [65] to the families present within this vertebral number dataset using the R package, ape (v5.7.1) [68]. To reduce each family to a single branch and to account for the estimated age of each respective family, the tree was pruned by CVB to the longest (i.e. most basal) branch within each family. Log-transformed minimum total vertebral and ancestral states were mapped and inferred by CVB using the function contMap, part of the phytools (v2.1.1) package [69]. R code for generating the figure was written by CVB.

1.3 The developmental control of vertebral number — Somitogenesis

The number of vertebrae formed in the axial skeleton is dependent on the number of somites that form in the developing embryo [44]. Somites are paired blocks of mesoderm that segment the anterior-posterior axis of the vertebrate trunk, laying the blueprint for the segmentation of the musculature, nervous system, and skeleton [70]. The number of somites formed in an embryo generally corresponds to the number of vertebrae that will be formed in the adult organism [71, 72], with a number of the first anterior somites (two in Zebrafish (*Danio rerio*) [72]) contributing to the head mesoderm and the remaining somites contributing to the vertebrae [72]. However, in some species the number of vertebral elements can be duplicated within a somite, giving rise to extra vertebrae, such as in the caudal skeleton of the Little skate (*Leucoraja erinacea*) [73]. Such duplications are rare however in the vertebrate tree of life and in most species the number of somites is determinant of vertebral number [44].

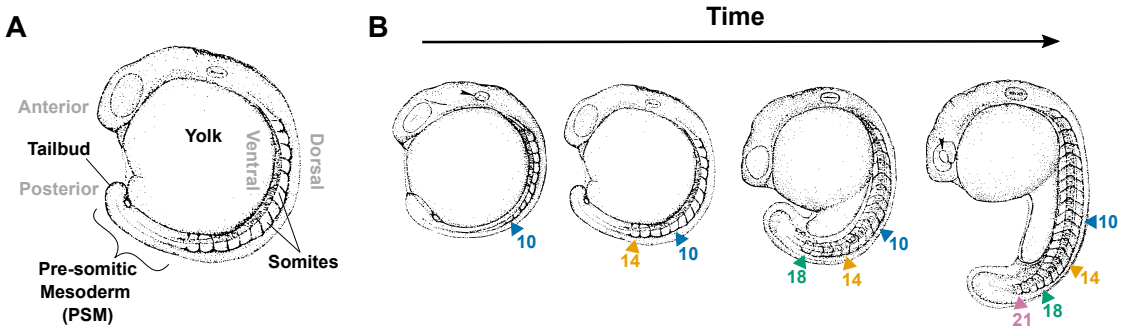


Figure 1.4: Somitogenesis in Zebrafish (*Danio rerio*), after Kimmel *et al.* (1995) [74]. A Sketch of a 14 somite-stage Zebrafish embryo, viewed from the left-hand side (adapted from [74]). At this stage of development, the somites are visible, dividing up the developing trunk. The pre-somitic mesoderm (PSM) is at the posterior tip of the embryo. **B** Sketches of 10, 14, 18, and 21 somite-stage Zebrafish embryos, viewed from the left-hand side (adapted from [74]). Somite number increases over time, as the embryo grows posteriorly and new somites are added sequentially at the posterior end of the embryo. © 1995 Wiley-Liss, Inc. Reprinted with permission of John Wiley & Sons.

Somites arise via a process known as somitogenesis, whereby somites are formed sequentially out of an unsegmented region of mesoderm at the posterior of the embryo known as the pre-somitic mesoderm (PSM) (figure 1.4). Together, the PSM and the somitic mesoderm comprise the

paraxial mesoderm. Somite formation is rhythmic and the time elapsed between the formation of one somite and a subsequent one is called the period of somitogenesis, or equivalently the rate at which somites appear is termed the frequency of somitogenesis. The rate of somitogenesis varies between species, and is also temperature dependent [44, 75, 76].

1.3.1 The segmentation clock

The periodicity of somite formation is underpinned by a molecular oscillator known as the segmentation clock. The architecture of the clock is complex and has been shown to vary between species, however a core circuit is conserved across all vertebrates examined thus far (summarised in figure 1.5A) [45, 77]. This circuit comprises genes in the Notch signalling pathway, specifically genes in the Hairy-and-enhancer-of-split/Hairy-and-enhancer-of-split-related (Hes/Her) family, a group of transcriptional repressors, and Notch and Delta genes. The expression of Hes/Her genes oscillates due to autorepressive regulation, and is thought to drive oscillations in the expression of Delta, causing oscillations in Notch intracellular domain in adjacent cells (figure 1.5A) [77–81]. Oscillations of Hes/Her expression are noisy [77, 82–85], but are synchronised across neighbouring cells within the PSM via this transduction of information about the phase of Hes/Her oscillations [77, 81, 82].

The existence of a segmentation clock was first posited by Cooke and Zeeman in 1976 [86], but it was not until 1997 that experimental evidence for a segmentation clock in vertebrates came to light, with the discovery of oscillations of c-hairy1 expression in the PSM of Chicken (*Gallus gallus*) embryos [86–88]. Here, Palmeirim and colleagues discovered that c-hairy1 expression oscillates, forming travelling waves of gene expression that pulse from the PSM posterior to the anterior with a periodicity equal to that of somitogenesis. Subsequent work has revealed similar expression of other Hes/Her genes in other vertebrates such as her1 and her7 in Zebrafish, her5 and her11 in Medaka, and Hes7 in Mice. The autonomous nature of Hes/Her oscillations, due to their negative autoregulation, as well as their essential role in function of the segmentation clock [79, 89–91], has lead to an often-held view in the field that the autorepressive feedback

of Hes/Her onto itself is the 'core' of the segmentation clock, and acts as the 'driver' for all downstream oscillations [77]. While this has appeared to be the case for Zebrafish [77], it was not until recently that this was considered true of amniotes, in part due to the essential role of the oscillatory gene Lunatic fringe (Lfng) in maintaining oscillations [77]. Recent evidence suggests that the role of Lfng is to mediate the temporal delay in sending signals between cells, setting it to a value that is dynamically capable of forming synchronous oscillations [92].

The frequency of clock oscillations is thought to be set by both intracellular and intercellular processes, creating an emergent collective frequency for the clock [93]. Such intracellular processes include the timescale of repressive feedback [48], and the rate at which Hes/Her transcription factors degrade [48, 94, 95]. The timescale of feedback is thought to be majorly set by delays associated with splicing of mRNA [79, 96, 97], and interspecific differences in frequency can be explained by differences in species-specific catabolic rate rather than differences in the sequence of Hes/Her genes, suggesting a major role for degradation rate in setting the tempo of the clock [95]. In Zebrafish, the effective degradation is also mediated by the promiscuous binding of Hes/Her transcription factors to other Hes/Her transcription factors [98, 99]. It is thought that a minority of these homo- and heterodimers are capable of binding DNA, and thus this sequestration of Hes/Her factors into a non-DNA-binding state effectively increases the degradation rate of DNA-binding factors and increases the frequency of oscillations [98, 99].

Across the vertebrates examined thus far, the frequency of clock oscillations does not appear to be constant along the anterior-posterior axis of the PSM, creating stripes of gene expression and a phase gradient along the PSM [46, 100–102]. In Zebrafish, this appears to be an intrinsic property of intracellular oscillations of the segmentation clock, which slow their frequency over time [85]. Additionally, this gradient self-organises in re-aggregated cultures of dissociated Mouse PSM cells [103]. The cause of this frequency gradient is unclear but is thought to be underpinned by a spatial gradient of translational delays [104], or a gradient of expression of non-oscillating Hes/Her genes such as hes6 in Zebrafish [105].

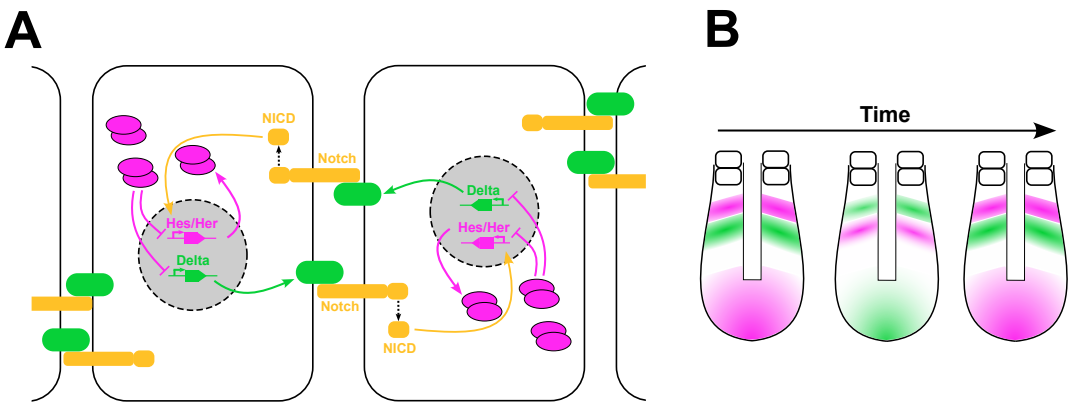


Figure 1.5: The segmentation clock. A Simplified diagram of the segmentation clock circuit. Dimers of Hes/Her transcription factors (magenta) repress their own expression and the expression of Delta protein (green), causing oscillations in the expression of Hes/Her and Delta. Delta protein transduces signals via Notch (orange), which causes cleavage of notch intracellular domain (NICD, black arrow). NICD activates expression of Hes/Her, thus communicating information about the phase of the clock to the clock in a neighbouring cell. **B** The expression of Hes/Her and Delta oscillates in space and time. Cartoon image of an idealised PSM viewed from above — the PSM is divided by the developing spinal cord and notochord and bounded at the anterior (top of figure) by the somites.

1.3.2 The clock and wavefront model

The clock and wavefront model is a conceptual framework postulated by Cooke and Zeeman in 1976 to explain the invariance of somite number in experimentally size-reduced *Xenopus* embryos [86]. The model posits the existence of a 'clock' (presciently suggested by the authors to be a molecular oscillator phase-synchronised between cells), the phase of which determines the potential of a cell in the PSM to form a somite boundary or not. The model also assumes the existence of a point in space at the PSM anterior termed the 'wavefront', at which the phase of the clock is 'read-out' by development and depending on the phase of its clock at that point, a cell either adopts somite boundary fate or somitic mesodermal fate (figure 1.6) [86]. As the embryo and PSM elongate, the wavefront also moves posteriorly, thus creating somites of length approximately equal to $v \cdot T$, where v is the velocity that the wavefront moves posteriorly, and T is the period of the segmentation clock. Cooke and Zeeman proposed that v scales with the overall size of the embryo, whereas T does not, thus explaining how size-reduced embryos maintain the typical somite number for the species while also preserving the stereotypical pace of somitogenesis [86].

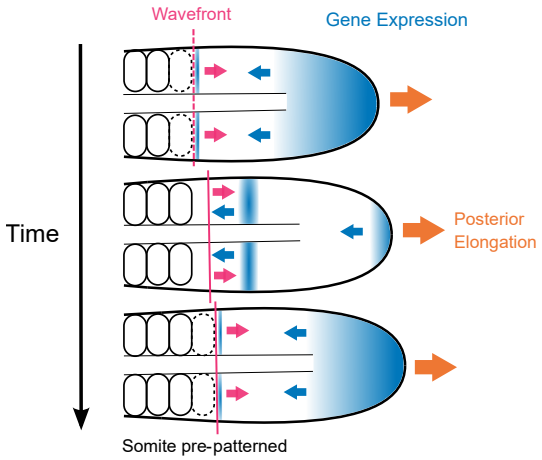


Figure 1.6: The clock and wavefront model. Travelling waves of clock gene expression (blue) emanate from the posterior of the PSM and travel towards the PSM anterior (blue arrows). At a point in space called the wavefront (magenta), the phase of clock gene expression is 'read out' by development. If cells expressing clock genes in the correct phase of the oscillatory cycle cross the wavefront, then a somite boundary is patterned. The wavefront moves slowly (relative to the frequency of the clock) towards the posterior of the PSM (magenta arrow), and thus a second somite boundary is patterned, creating a somite. Concurrent with this the PSM elongates (orange arrow), allowing the process to carry on in principle indefinitely, forming somite boundaries.

1.3.3 The wavefront

However, while the nature of the segmentation clock is well-established, it is unclear what the wavefront is, if it exists at all. The discovery of a gradient of FGF signalling in the PSM, and the fact that locally elevating FGF signalling with FGF-soaked beads in the PSM can induce the formation of atypically long somites [106], has lead to a majority opinion that the wavefront is a cue based on positional information within a gradient of FGF signalling [44, 106–109]. However, other proposals for the nature of the wavefront have been made, such as a self-organising locally diffusing repressor [53], or a cue from the phase gradient across the PSM [110]. It has also been recently demonstrated by Rohde *et al.* that differentiation of PSM cells into somites relies on a cell-intrinsic timing mechanism in Zebrafish, suggesting a possible cell-autonomous mechanism for the wavefront [85].

Recent work has established a causal link between FGF signalling and somite boundary for-

mation. Simsek *et al.* demonstrated that pulsatile inhibition of FGF signalling could rescue somite boundary formation in clock-less Zebrafish mutants, and that the transcription factors Her1 and Her7 cyclically down-regulate FGF signalling [111]. As previous work established that somite boundaries are patterned at a point in space corresponding to a critical threshold of FGF signalling [107], the authors proposed that the wavefront is this critical threshold, and that the wavefront regresses one somite length every clock cycle in discrete jumps [111]. Additionally, as the gradient in FGF signalling along the PSM is thought to be driven by gradual decay of FGF mRNAs in PSM cells [112], the wavefront slowly regresses posteriorly at a rate coupled to the elongation of the embryo. The autonomous behaviour of the wavefront proposed by Rohde *et al.*, as well as the observation that treatment with FGF sustains oscillations of dissociated PSM cells in culture [83, 85], could perhaps be explained by gradual loss of FGF mRNA molecules and thus FGF signalling within these cells, however the precise mechanism remains elusive.

1.3.4 Elongation of the pre-somitic mesoderm

Concurrent with the oscillations of the segmentation clock and patterning of the somites, the PSM undergoes elongation in the anterior-posterior axis. This elongation is necessary for the formation of the species-specific number of somites, and is achieved by diverse strategies across the vertebrate tree of life [44, 46, 113]. PSM elongation typically employs a set of various morphogenetic processes in varying combinations. One such process is the ingression of PSM progenitor cells into the posterior of the PSM. In Zebrafish and amniotes, PSM progenitor cells in the epiblast undergo an epithelial-to-mesenchymal transition and move ventrally to enter the posterior region of the PSM [114–116]. Additionally in the early stages of Zebrafish somitogenesis, progenitor cells enter the PSM laterally [113]. This addition of new cellular material allows the PSM to elongate and replenish material lost to the developing somites at the PSM anterior. Another process (and the other way by which cellular material is replenished) is cell division, although the contribution of cell division to elongation of the PSM differs between species [113]. For instance, within Zebrafish, cell division in the PSM is relatively minor

and dispensible for elongation, [113, 117], whereas it is essential for elongation in the Japanese quail (*Coturnix japonica*) [118].

Another way in which the PSM elongates is via the presence of an anterior-to-posterior increasing profile of random cell motility within the PSM [114, 119–122]. Within avian embryos, this profile causes expansion of the posterior PSM and drives elongation of the PSM, as well as elongation of the posterior body by exerting pressure on adjacent axial tissues, causing mediolateral constriction and pushing these tissues towards the posterior of the embryo [119, 121, 122]. Within Zebrafish, this profile is very abrupt and corresponds to a fluid-to-solid transition in the state of the tissue along the anterior-posterior axis of the PSM [114, 116, 120]. In Zebrafish the difference in tissue fluidity between the anterior and posterior regions of the PSM ensures that the ingression of progenitor cells into the PSM posterior creates unidirectional elongation of the PSM [116, 120].

The final mechanism by which the PSM has been thought to elongate is via compaction-extension, analogous to the mechanism of convergent extension well-characterised in amphibian embryos [113, 123, 124]. Here, migration of cells towards the PSM midline causes the PSM to shrink in volume over time and increase in density, driving elongation of the PSM [113, 124].

1.3.5 The initiation and termination of somitogenesis

It is currently unknown what causes somitogenesis to initiate or terminate. Clock gene expression first initiates during gastrulation in Zebrafish, Chicken, and Mouse [49, 125, 126]. In Chicken this occurs once PSM progenitor cells have entered the primitive streak [125]. Clock gene expression at the earliest stages of somitogenesis precedes the formation of the first somites, and shows initially synchronous oscillations across the presumptive PSM [49, 125, 126]. As oscillations continue, gradually a phase gradient forms and travelling waves appear, reminiscent of the self-organising behaviour of the phase gradient *in vitro* [103], suggesting that the anterior-posterior gradient in clock frequency is already present at the initiation of clock gene expression [49, 126]. In Chicken, at least two such oscillations precede the formation

of the first somites, versus five such oscillations in Zebrafish, and three in Mouse [49, 125, 126]. It is not clear if these differences are evolutionarily significant or if they merely reflect the differences in clock timing between these species. The initiation of clock expression appears to be independent of notch signalling, as embryos where notch signalling is disrupted from early development exhibit synchronous oscillations at the onset of clock expression, followed by a gradual decay in synchrony over time [49, 126]. This can explain the observation that mutants and morphants defective for notch signalling or Hes/Her genes typically form the first few somites correctly [49, 90]. It is unclear how or when the clock is initiated in the progenitor cells that enter the PSM during somitogenesis. In Zebrafish at least, progenitor PSM cells appear only to commence clock gene expression once within the PSM [127], however to my knowledge this question has not been explicitly addressed.

The termination of somitogenesis has been proposed to depend on morphogenesis of the PSM [44]. This idea is based on the observation that the PSM shrinks in size towards the end of somitogenesis, and that species which form larger numbers of somites often undergo more growth in their PSM. This suggests that the role of morphogenesis is to postpone shrinking of the PSM, and this shrinking is somehow causal in determining the duration of somitogenesis [44, 46, 113]. A small remnant of the PSM remains at the end of somitogenesis [128], and heterochronic grafts of cells from this tissue into the primitive streak of younger embryos shows that donor cells can contribute to the formation of somites, showing that the arrest of somitogenesis is due to some signal extrinsic to the PSM [44, 129, 130]. It has been proposed that this signal may rely on positional information from the production of retinoic acid in the somites [44, 53, 128], where clock oscillations and somite production arrest once the PSM becomes sufficiently short [44]. The overall duration of somitogenesis appears to be determined by the timing of the expression of Hox genes [131, 132], possibly due to a role in controlling proliferation [44].

1.3.6 Mathematical models of the segmentation clock

There is a long history of mathematical modelling of somitogenesis, particularly the mechanism of the segmentation clock, and this subject has developed into one of the most theory-rich areas of developmental biology [50, 55]. This area has recently been extensively reviewed in a paper by Francois and Mochulska [55], which the reader is urged to consult for further details on the subject. Here I provide a brief account of some of the models which have been particularly well-adopted or influential for our understanding of the evolution of the segmentation clock.

Models of the clock gene regulatory network

Most mathematical descriptions of the gene regulatory network underpinning the segmentation clock are related to the delay differential equation (DDE) models proposed by Monk and Lewis in 2003 [48, 133]. These models show that autorepressive feedback of transcription by Hes/Her genes creates oscillations in the levels of mRNA and protein molecules when there is a temporal delay in the production of either of these molecules. Assuming the existence of such a delay is attractive as it is more realistic than previous models of biological oscillations which assumed implausibly nonlinear autorepression [55, 134]. At their core, these models are comprised of the equations

$$\frac{\mathrm{d}p}{\mathrm{d}t} = a \cdot m(t - \tau_p) - b \cdot p(t), \tag{1.1}$$

and

$$\frac{\mathrm{d}m}{\mathrm{d}t} = f(p(t - \tau_m)) - c \cdot m(t), \tag{1.2}$$

where m and p represent the levels of intracellular mRNA and protein respectively [48, 133]. The parameters $a, b, c \in \mathbb{R}$ correspond to the rates of mRNA translation, protein degradation, and mRNA degradation respectively. The delay parameters $\tau_m, \tau_p \geq 0$ respectively correspond

to the time elapsed between the initiation of transcription and the appearance of a mature mRNA transcript at the site of translation, and the time elapsed between the initiation of mRNA translation in the ribosome and the transport of the transcription factor protein p back into the nucleus where it can effect gene regulation. Finally the function f represents the transcriptional regulation effected by the trancription factor p. In principle the identity of the oscillatory gene coding for molecules m and p is arbritrary, and such an equation could be used to model the oscillations of any autorepressive transcription factor, however within the context of the segmentation clock these are generally assumed to correspond to Hes/Her genes, which exhibit such behaviour. This model can be extended to encode intercellular coupling via notch-delta signalling [48, 135].

Models of this class have been very illuminating in furthering our understanding of the determinants of clock frequency, largely due to the relative ease of experimentally measuring the parameters for delay or degradation. For instance, one prediction of the model is that for sustained oscillations to occur, the degradation rates of m and p must be sufficiently short [48]. Hirata *et al.* verified this by showing that experimentally increasing the half-life of Hes7 protein causes a loss of Hes7 and Lfng oscillations in the Mouse PSM [94]. The frequency of oscillations is predicted to be inversely proportional to the total length of the delays $\tau_m + \tau_p$, and experimentally reducing the value of τ_m via removal of introns from the Hes7 gene has been shown to increase the frequency of Hes7 oscillations and increase the number of cervical and upper thoracic vertebrae formed in mutant mice [79]. Additionally, removal of all the introns in the Hes7 gene abolishes Hes7 oscillations, in agreement with predictions based on this model that the total delay $\tau_m + \tau_p$ must be sufficiently long to allow sustained oscillations [96]. These models have also furthered our understanding of interspecific differences in clock period, for instance the work of Matsuda *et al.* who showed that the differences in clock period between human and Mouse PSM cells can be quantitatively explained by differences in the delay and the degradation rate [95].

Schröter and colleagues used a similar model to describe the segmentation clock of Zebrafish,

comprised of the genes her1, her7, and the non-oscillatory gene hes6, and their regulation by the protein dimers Her1:Her1 and Hes6:Her7 [98]. This model is given by the equations

$$\frac{\mathrm{d}h_1}{\mathrm{d}s} = \frac{k_1}{1 + (h_1(s-\tau_1)^2 + h_7(s-\tau_1) \cdot h_6(s-\tau_1))^n} - h_1(s) - \delta \cdot h_1(s) \cdot (2h_1(s) + h_6(s) + h_7(s)), \tag{1.3}$$

$$\frac{\mathrm{d}h_7}{\mathrm{d}s} = \frac{k_7}{1 + (h_1(s-\tau_7)^2 + h_7(s-\tau_7) \cdot h_6(s-\tau_7))^n} - h_7(s) - \delta \cdot h_7(s) \cdot (h_1(s) + h_6(s) + 2h_7(s)), \tag{1.4}$$

and

$$\frac{\mathrm{d}h_6}{\mathrm{d}s} = k_6 - h_6(s) - \delta \cdot h_6(s) \cdot (h_1(s) + 2h_6(s) + h_7(s)), \tag{1.5}$$

where s represents time rescaled according to the global rate of protein degradation (see [98] for a full derivation). Here the production of mRNA molecules is neglected, and this model only treats the production of the proteins Her1 (h_1), Her7 (h_7), and Hes6 (h_6), and how the dimers Her1:Her1 and Hes6:Her7 repress the formation of the monomers Her1 and Her7. The parameters τ_1 and τ_7 correspond to the delays involved in the synthesis of Her1 and Her7 respectively, n describes the cooperativity of dimer binding, k_1, k_6, k_7 correspond to the production rates of Her1, Hes6, and Her7 respectively, and δ is the effective degradation of Her proteins by loss to non-DNA-binding dimers [98]. Using this model the authors were able to explain the epistatic effect of hes6 in controlling regulation of the segmentation clock by her7, as well as explain why hes6-null mutant Zebrafish have slower clock frequencies and form fewer vertebrae, as a knock-out of hes6 reduces the effective degradation rate of DNA-binding dimers [98, 136]. Though this model is of essentially the same class as the model presented above, it provides an illustrative example of how the framework can be extended to model increasingly

complex and species-specific scenarios of gene regulation and protein-protein interactions.

While a feedback delay is generally held to be necessary to generate sustained oscillations in the oscillation of Hes/Her genes, not all formulations of the segmentation clock need encode this delay as explicitly as the models above. Terry *et al*. were able to implicitly encode the delay via diffusion of her1 mRNA from the nucleus out into the cytosol, where it is then translated into Her1 protein, and diffusion of Her1 protein back into the nucleus [51]. This model is a partial differential equation (PDE) analog of the DDE model proposed by Lewis and Monk [48, 133], namely

$$\frac{\partial p}{\partial t} = D_p \nabla^2 p + H(r - A) \cdot a \cdot m - b \cdot p, \tag{1.6}$$

and

$$\frac{\partial m}{\partial t} = D_m \nabla^2 m + f(p) - c \cdot m, \tag{1.7}$$

where p and m are the concentrations of clock gene protein and mRNA respectively. The diffusion coefficients of these molecules are given by D_p and D_m respectively, as are the degradation coefficients b and c. The variable r represents the distance from the centre of the nucleus and A the minimal distance from the nucleus beyond which protein synthesis occurs (approximating the nuclear envelope). The function $H(x)$ is the heaviside step function, a is the production rate of protein p, and similar to the model of Lewis and Monk f is a function describing the transcriptional regulation of the segmentation clock gene by the transcription factor p. Using this model the authors were able to achieve oscillations of clock genes and also extend the model to achieve intercellular coupling of oscillations by modelling the notch-delta pathway similarly to Lewis [51].

Models of clock phase

Due to the complex architecture of the segmentation clock gene regulatory network [45], it is often convenient to abstract the dynamics of the clock in terms of its phase θ. This assumption is valid when the segmentation clock is a limit cycle, that is that oscillations are sustained and do not undergo damping [137]. A typical description of the clock using such assumptions would be

$$\frac{\partial \theta}{\partial t} + v \cdot \frac{\partial \theta}{\partial x} = f(x), \tag{1.8}$$

which describes the phase θ in time t along a one-dimensional PSM x. The $\partial\theta/\partial t$ term represents the frequency of oscillations, and $v \cdot \partial\theta/\partial x$ accounts for the movement of cells along the PSM during elongation [55, 138]. However other phase descriptions of the segmentation clock are more commonly used, which explicitly treat the coupling of oscillations between cells. The most commonly used formulation is a 'Kuramoto' model [49, 54, 93, 137, 139, 140]. These models are valid descriptions of the segmentation clock assuming that cell-cell coupling is weak and does not perturb the amplitude of oscillations [137]. Their general formulation is

$$\frac{\mathrm{d}\theta_i}{\mathrm{d}t} = \omega(x) + \xi_i(t) + \frac{\kappa}{N(i,t)} \sum_{j \neq i} \Gamma(\theta_j(t-\tau) - \theta_i(t)), \tag{1.9}$$

which describes the evolution of the phase θ_i in a given cell i. The function $\omega(x)$ describes the 'intrinsic' frequency of the segmentation clock oscillator, and is typically assumed to decrease towards the PSM anterior, to create a phase gradient [93]. The coupling strength between cells is given by κ (in the context of the clock, this is a proxy for the strength of delta-notch signalling), and this can also be assumed to vary in time, for instance to model the wash-out of notch signalling inhibitor molecules and the gradual recovery of synchrony [54]. The delay τ, similar to the previous DDE models of the segmentation clock, simulates the temporal delay in information exchange between cells. This coupling strength is normalised by $N(i,t)$ which is

the number of cells that couple their clocks to cell i at a given time t. For simplistic descriptions of the PSM as a one-dimensional array of cells, this number could simply be two (a cell is coupled only to the cell anterior and posterior to it), or for complex three-dimensional models of cell movements within the PSM, $N(i,t)$ could be defined by which cells are in contact with i at time t [54, 140]. The term $\xi_i(t)$ corresponds to random noise. The function Γ describes the coupling of phase between cells. In most formulations of the model [49, 54, 93, 139, 140], it is assumed that $\Gamma(x) = \sin(x)$, however recent work by Ho *et al.* suggests that a more accurate formulation for the segmentation clock may be $\Gamma(x) = \min\{0, \sin(x)\}$ [141]. It is typically assumed that Γ creates attractive coupling, i.e. that Γ will act to minimise the difference in phase between two oscillators, however this need not always be the case. Murray *et al.* introduced repulsive coupling to such a model by setting $\Gamma(x) = A \cdot \sin(x) + B \cdot (\cos(x) - 1)$ (and $\tau = \xi_i(t) = 0$), where the parameters A and B describe the magnitude of attractive and repulsive coupling respectively [142]. Interestingly, when they assumed a constant value for ω this repulsive coupling was sufficient to generate travelling phase waves along the PSM [142]. The authors showed that varying values of A and B could explain the variation in phase gradients seen across the vertebrates [142].

As Kuramoto-type models describe tissue-level phenomena, they are often relatively straightforward to parameterise from experimental data. For instance, Murray *et al.* were able to fully parameterise their model using data on somitogenesis frequency, PSM length, and *in-situ* hybridisations (ISHs) for segmentation clock genes, allowing a comparison across different vertebrate species using existing data [142]. Additionally, work by Riedel-Kruse *et al.* measured the coupling strength κ and noise $\xi_i(t)$ terms for a model of the form shown in equation 1.9 using a combination of chemical inhibitor perturbation of the segmentation clock and data from ISHs [49]. Specifically, the authors consider how rapidly the clock tends to asynchrony upon increasingly high doses of notch-signalling inhibitor [49]. By considering the synchrony of clock phase θ across N many different PSM cells numbered $1, 2, \ldots, k, \ldots, N$, given by

$$Z = \frac{1}{N}\sum_{k=1}^{N} e^{i\theta_k}, \tag{1.10}$$

it is possible to approximate the temporal evolution of clock synchrony $Z(t)$ by the function

$$Z(t) = Z_0 e^{-t(2D\theta^2-\kappa)/2}, \tag{1.11}$$

where Z_0 is the pre-perturbation synchrony of the segmentation clock, and $D\theta$ describes the magnitude of the white noise term $\xi_i(t)$ with expected value $\langle \xi_i(t) \rangle = 0$ and $\langle \xi_i(t) \cdot \xi_j(t') \rangle = 2D\theta^2\delta_{ij}\delta(t-t')$. Therefore, for $2D\theta^2 > \kappa$ the synchrony of the clock decays to 0 over time. This agrees with the experimental observation that the effects of disruption of notch signalling only manifest after several clock cycles, always leaving a number of correctly patterned somites anterior to the first defectively patterned segment [49, 90]. Using the approximation of synchrony dynamics in equation 1.11, it is therefore possible to measure the noise $D\theta$ and coupling strength κ by varying the dose of notch signalling inhibitor and using ISHs identify defectively patterned segments [49]. Similar work using varying doses of notch signalling inhibitor has also been used to measure the coupling delay τ [139].

It is worth noting that one-dimensional equations of the form in equation 1.9 are equivalent to PDE models of the form in equation 1.8 in the continuum limit of infinite cells [55, 93, 138].

Models and the evolvability of somitogenesis

However, despite these models being illuminating in understanding the mechanisms of somitogenesis, and implicitly its evolutionary potential, there has been no explicit exploration of the evolvability of somitogenesis using these models. Lewis speculated that the existence of a quick Hes/Her autorepressive loop and a slow Hes/Her — Delta — Hes/Her autorepressive loop might allow evolution to readily change between different clock frequencies [48], however this has not been explicitly studied. The large body of models for the segmentation clock is likely to be a rich area for exploring the sources of evolvability in somitogenesis.

1.4 The evolution of somitogenesis

One implication of the sequential nature of somitogenesis is that different numbers of somites can be formed by changing the period of the segmentation clock, or by changing the overall duration of somitogenesis. In a comparative study of vertebrate somitogenesis, Gomez *et al.* demonstrated that the large vertebral number of the Corn snake (*Pantherophis guttatus*) is due to a combination of an acceleration of the segmentation clock as well as a prolongation of somitogenesis relative to the whiptail lizard *Aspidoscelis uniparens*, a species that develops at a similar pace [46]. The prolongation of somitogenesis in snakes has been suggested to be due to changes in the regulation of Hox13 genes [132, 143], and a similar mechanism has been recently suggested to explain the differences in vertebral number between two populations of Deer Mice [15], suggesting a more general role for Hox13 genes in the evolution of vertebral number.

The evolution of the clock has been also examined by comparative studies, for instance the work of Krol *et al.* who showed that the oscillatory dynamics and identity of segmentation clock genes can vary across Mouse, Chicken, and Zebrafish [45], Gajewski *et al.* who performed a comparative study of Medaka and Zebrafish and showed that the role of different Hes/Her genes in segmentation can vary within the teleosts [100], or Eckalbar *et al.* who showed that the identity of oscillatory notch signalling genes varies between reptile species [102]. Additionally, recent work by Ebisuya and colleagues has explored the origins of interspecific differences in clock frequency using *in vitro* culture of PSM cells [47, 95]. This work has revealed that the 2- to 4-fold differences in the period of Hes7 oscillations between different mammalian species are not due to differences in the sequence of Hes7 gene, but rather due to biochemical parameters such as the rate of protein degradation or RNA splicing [47, 95]. Interestingly, the period of the clock does not appear to be affected by the overall metabolic rate of a species [47].

However, there have been relatively few comparative studies examining the evolution of somitogenesis, and these have typically compared highly evolutionarily divergent species [45–47, 100, 113]. While these studies have revealed conserved developmental processes, and likely

developmental constraints [45], due to the large evolutionary timescales separating these model species we have no idea which processes change most rapidly in evolution or respond quickest to evolutionary pressure, and so it is not clear which changes have been most important in general for the evolution of vertebral number. Additionally, there has not been any systematic examination of the literature to identify possible sources of evolvability in vertebral or somite number, based on the current state of knowledge regarding somitogenesis. Therefore I have dedicated the remainder of this section to a review of the literature, aiming to understand where this evolvability may arise.

1.4.1 The respective roles of the clock and elongation in evolution

One major aspect of the evolution of somitogenesis which has not been examined is the relative importance of the frequency versus the duration of somitogenesis in creating diversity in vertebral number. Comparative studies have revealed roles for both in evolution [46], but it is unclear if one process is more important for the evolution of vertebral number across the vertebrates, or if one is more evolvable. To explore this, I performed a literature survey, cataloguing the number of somites formed during embryogenesis and the rate of their appearance in different vertebrate species (see table 1.1). We can imagine three scenarios for this data (figure 1.7A). In scenario one, diversity in somite number is solely explained by evolution of clock frequency, and we would expect to see a positive correlation between somite number and frequency (figure 1.7A, left). In scenario two, where diversity in somite number can be explained by evolution of the duration of somitogenesis alone, we would expect to see the datapoints lie along a vertical line when we plot the number of somites against the clock frequency (figure 1.7A). The third scenario is that the diversity in segment number is due to co-evolution of these two processes, and we would expect to see the datapoints spread along each axis (figure 1.7A).

Species	Class	Period (mins)	Total somites	Phylotypic stage (hpf)	Temperature	Method	Source(s)
Amphiprion ocellaris	Actinopterygii	35	22	45	26°C	Reported	[144]
Anguilla japonica	Actinopterygii	37.5	114	85	22°C	Reported	[145], †
Astatotilapia calliptera	Actinopterygii	73.7	30	72	28°C	Reported	[146]
Astyanax mexicanus	Actinopterygii	20	40	—	23°C	Reported	[147]
Crassius auratus	Actinopterygii	30	35	26	24°C	Reported	[148]
Clupea harengus	Actinopterygii	52	61	—	15°C	Reported	[149]
Clupea pallasii	Actinopterygii	72.8	54	139	9.6°C	Inferred	[150]
Danio rerio	Actinopterygii	25.1	32	24	27°C	Reported	[76]
Dicentrarchus labrax	Actinopterygii	109	28	72	15°C	Inferred	[151]
Gadus morhua	Actinopterygii	207	52	250	7°C	Reported	[152, 153]
Leucopsarion petersii	Actinopterygii	187	36	—	19°C	Inferred	[154]
Luciobrama macrocephalus	Actinopterygii	22	55	60	21°C	Inferred	[155]
Misgurnus anguillicaudatus	Actinopterygii	36	50	—	20°C	Reported	[156]
Muraenesox cinereus	Actinopterygii	31.6	149	—	24°C	Reported	[145]
Oryzias latipes	Actinopterygii	57.5	35	101	27°C	Reported	[157, 158]
Rhamphochromis sp. 'chilingali'	Actinopterygii	67.7	36	72	28°C	Reported	[146]
Salmo salar	Actinopterygii	155.8	67	228.5	13°C	Reported	[75]
Anas platyrhynchos	Aves	80.4	50	96	38°C	Inferred	[159]
Columba livia	Aves	59.4	44	96	36°C	Inferred	[160]
Gallus gallus	Aves	90	52	84	38°C	Reported	[87, 161]
Lonchura striata	Aves	87.7	45	132	38°C	Inferred	[162]
Pygoscelis adeliae	Aves	187	50	240	34°C	Inferred	[163]
Chiloscyllium punctatum	Chondrichthyes	228.5	80	336	25°C	Inferred	[164]
Petromyzon marinus	Hyperoartia	241	101	384	18°C	Inferred	[165]
Bos taurus	Mammalia	310	56	626	37°C	Inferred	[47, 166]
Callithrix jacchus	Mammalia	778	56	1440	37°C	Inferred	[47, 167, 168]
Ceratotherium simum	Mammalia	236	44	—	37°C	Inferred	[47, 169]
Mus musculus	Mammalia	120	65	216	37°C	Reported	[170]
Oryctolagus cuniculus	Mammalia	143	45	252	37°C	Inferred	[47, 171, 172]
Aspidoscelis uniparens	Reptilia	240	90	—	28°C	Reported	[46]
Chelonia mydas	Reptilia	395.2	43	168	—	Inferred	[173]
Chelydra serpentina	Reptilia	1123.6	31	600	20°C	Inferred	[174]
Elaphe quadrivirgata	Reptilia	86.4	297	264	30°C	Reported	[175]
Emydura subglobosa	Reptilia	284	39	144	30°C	Inferred	[176]
Lamprophis fuliginosus	Reptilia	60	300	—	28°C	Reported	[46]
Pantherophis guttatus	Reptilia	100	315	768	28°C	Reported	[46]

Table 1.1: Somitogenesis period at a given temperature, the time to phylotypic stage, and total somite number, across species. For each species the periodicity of somitogenesis, the time elapsed before one somite is formed, was either taken from a previously published value ('reported'), or was inferred by linear regression of somites over time from published data ('inferred'). The time taken to reach phylotypic stage, defined as the presence of pharyngeal arches, was not always reported and so in some cases this data is missing. †A Seleit *pers. comm.*

Plotting the total number of somites formed in the embryo against the rate of somitogenesis (see table 1.1), we see no clear trend, and that the frequency of somitogenesis and the total number of somites is extremely diverse across different groups of vertebrates (figure 1.7B). Within this dataset, the slowest rate of somite formation is the Common snapping turtle (*Chelydra serpentina*), with one somite being formed every 19 hours [174], and the quickest is the Mexican tetra (*Astyanax mexicanus*), with one somite being formed every 20 minutes [147]. However, for appropriate comparison between species the rate of somitogenesis must be scaled against the overall pace of development, which varies across taxa [47, 95]. It should be noted that the frequency of somitogenesis is highly temperature-sensitive [75, 76], and so one must also account for differences in the temperature of development between species. To do so I assume that the frequency of somitogenesis and the overall pace of development scale together in response to changes in temperature, and so hope to implicitly normalise the data for varying

temperatures by normalising against the overall pace of development. We see that the period of somitogenesis (i.e., the time taken to form one somite) correlates well with the time taken for the embryo to reach the phylotypic stage (defined by the first appearance of the pharyngeal arches), and the slow frequency of somite formation in species like the Common snapping turtle or the Common marmoset (*Callithrix jacchus*) can largely be explained by their overall slower pace of development (figure 1.7C). This is similar to the results of [47] who observe that the period of the segmentation clock *in vitro* scales with the overall pace of development [47]. Scaling the frequency against this overall rate of development by dividing the time to reach phylotypic stage by the period of somitogenesis, we see that the total number of somites correlates with frequency, but that the data is broadly heterogeneous and is best explained by scenario three (figure 1.7D).

Here we see that some species, such as the Japanese striped snake *Elaphe quadrivirgata* or the Corn snake, exhibit accelerated clocks and their large numbers of somites can be at least partially attributed to this fact (figure 1.7D) [46]. We also see that some active predatory fish such as the European seabass (*Dicentrarchus labrax*) form few somites and have correspondingly slow clocks (figure 1.7D), suggesting that perhaps this reflects selection for a stiffer vertebral column with fewer vertebrae, giving better propulsion when pursuing prey [177, 178]. However, some species such as the Long spiky-head carp (*Luciobrama macrocephalus*) possess rapid clocks relative to their overall pace of development, approaching the snakes' in their frequency, but do not form as many somites (figure 1.7D). This suggests that the duration of somitogenesis is shorter relative to the frequency of the clock in the Long spiky-head carp than in snakes and highlights that even in cases where the clock is dramatically accelerated, the evolution of duration cannot be discounted when considering the evolution of somite number [46].

We can also observe, as reported by Lazaro *et al.* [47], that the species of mammals analysed here have roughly equivalent somitogenesis frequencies when one accounts for differences in developmental rate (figure 1.7D), suggesting that perhaps differences in mammalian vertebral number are not driven by changes in the clock but rather differences in the duration of

somitogenesis (c.f. [15]). The evolution of the mammalian vertebral column is thought to be highly constrained by its regionalisation, with the number of pre-sacral vertebrae being highly conserved, likely reflecting mechanical constraints imposed by the running gait of most mammals [2, 13]. It is possible therefore that this invariance in scaled clock frequency reflects a constraint on the pace of the clock relative to the pace of the 'Hox timer' or the sequential activation of Hox genes in the embryo [179], thereby ensuring a constant number of pre-sacral vertebrae. However, this assumes that the tempo of this timer also scales with the overall pace of development, and it is unclear if this would be the case. While this is speculative, the accessibility of non-model mammalian species for developmental biology is rapidly improving [47], and it may soon be possible to verify if the developmentally-normalised frequency of the clock is indeed a conserved feature of mammals.

This analysis also reveals that even when the overall pace of development is accounted for, the Testudines (turtles, terrapins, and tortoises) have a very slow rate of somite formation (Figure 1.7D). The three species of Testudines in our dataset, the Common snapping turtle, the Green sea turtle (*Chelonia mydas*), and the Red-bellied short-necked turtle (*Emydura subglobosa*), all have similarly low frequencies of somite formation (Figure 1.7D), suggesting this is possibly an ancestral feature of the Testudines, as the lineages leading to *C. mydas* and the other two species diverged at the base of the testudine clade [180]. This perhaps could reflect an adaptation for decreased vertebral number and increased rigidity of the axial skeleton, which is thought to predate the evolution of the Testudine shell [181, 182]. Study of the Testudine segmentation clock may prove to be illuminating for understanding the evolvability of vertebral number, particularly in terms of understanding the clock's capacity for change in both dynamics and regulatory architecture, as well as a comparison with snakes, which appear to have taken the diametrically opposite path in evolution (Figure 1.7D) [46].

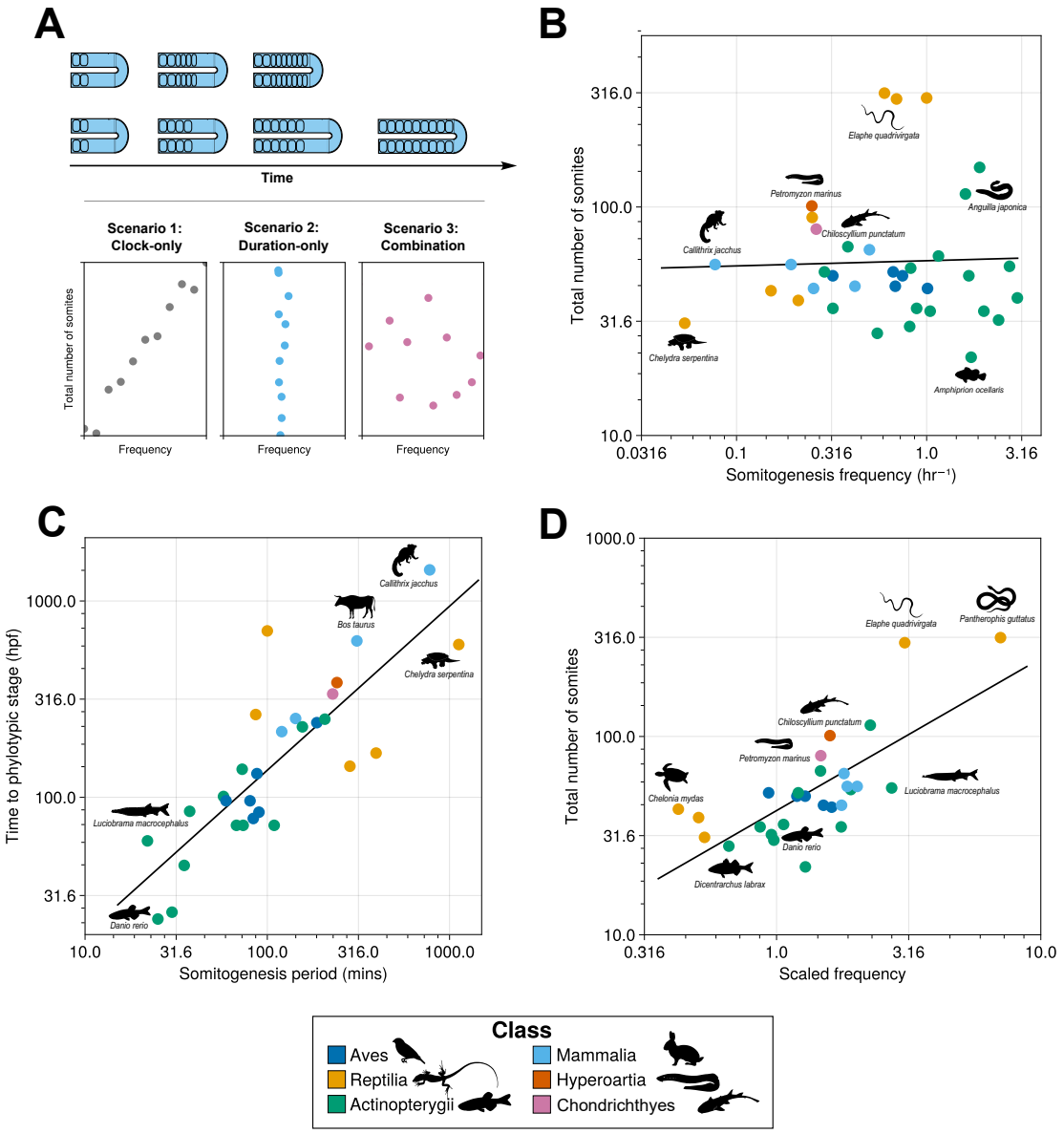


Figure 1.7: Diversity in somitogenesis dynamics across the vertebrates.
A Top: the same number of somites (black rectangles) can be achieved by changes in either frequency (top) or duration (bottom) of somitogenesis. **Bottom:** Scenarios for plotting the total number of somites formed during embryogenesis against somitogenesis frequency. If diversity is due to only evolutionary changes in frequency, the datapoints should lie along a linear trendline (left, black). Conversely if diversity is due to only changes in duration, the datapoints should lie along a vertical line (middle, blue). If diversity is created by a combination of the two, then the datapoints are heterogeneously spread across the space (right, mauve). **B** The total number of somites formed during embryogenesis, plotted against the frequency at which somites are formed. A linear trendline is fitted from regression of the log-transformed data, with trendline $y = 0.023x - 1.765$ with the summary statistics $p = 0.8423$, $R^2 = 0.001$, suggesting that there is no correlation between the dependent and independent variables. **C** The time taken to reach phylotypic stage (defined as presence of four pharyngeal arches), plotted against the somitogenesis period. A linear trendline is fitted from regression of the log-transformed data, with trendline $y = 0.835x - 0.467$ with the summary statistics $p << 0.01$, $R^2 = 0.674$, suggesting a statistically significant relationship.

Figure 1.7: (Continued from previous page).
D The total number of somites formed during embryogenesis, plotted against the somitogenesis period divided by the time taken to reach phylotypic stage. A linear trendline is fitted from regression of the log-transformed data, with trendline $y = 0.767x + 1.626$ with the summary statistics $p << 0.01$, $R^2 = 0.529$, suggesting a statistically significant relationship with poor explanatory power. Silhouettes were hand-drawn or downloaded from Phylopic (`https://www.phylopic.org`). Image credits and licenses: Edwin Price CC BY 4.0 and DFoidl and T. Michael Keesey CC BY-SA 3.0.

We can also observe that the frequency of somite formation is approximately the same in the Sea lamprey (*Petromyzon marinus*) and the Brownbanded bamboo shark (*Chiloscyllium punctatum*) (figure 1.7D). Whether this is coincidental or not remains to be seen, however I consider it of note due to the position of both species at the base of the vertebrate tree of life [183, 184]. If such a trend were to hold more generally and the developmentally-normalised rate of somitogenesis is conserved between the cyclostomes and elasmobranchs it could suggest that the extant evolvability of the clock arose only later in vertebrate evolution. Data on somitogenesis is currently extremely limited within these clades, but as more data becomes available I suggest that the field takes note of any conservation of somitogenesis frequency between these two groups.

Overall, this analysis suggests that both the morphogenesis of the PSM and the clock co-evolve to generate diversity in segment number. It cannot be identified from this dataset whether or not one process is more labile under evolution than the other, and both processes appear to have contributed to significant evolutionary changes in vertebral number. Therefore in the following section I discuss possible sources of evolvability in both processes.

1.4.2 Sources of evolvability in somitogenesis

While the segmentation clock is thought to be driven by auto-repressive feedback on Hes/Her gene expression, the clock contains a large number of oscillatory and non-oscillatory components [45, 185]. It is possible that this has been a rich substrate for evolution. The formation of non-DNA-binding Hes/Her dimers is known to accelerate clock oscillations [98, 136], and therefore duplication of Hes/Her family genes and expression of paralogs in the PSM could

potentially result in changes to the frequency of the clock. The Hes/Her family contains many paralogs within the vertebrates [186], and Hes/Her family genes have been shown to have undergone repeated duplications across the metazoa [100, 102, 187, 188], suggesting that perhaps the appearance of novel paralogs or the neo- or sub-functionalisation of paralogous genes in this rapidly-duplicating family may be a major source of evolvability for clock dynamics. Indeed, the identity of Hes/Her oscillatory genes in the clock can vary between species of teleosts [100], suggesting a role for neo-functionalisation. Similarly, altering the copy number of delta genes has been shown to increase the emergent frequency of the clock as a result of elevated clock coupling between cells [189]. Lastly, the presence of paralogs can create redundancy within the clock, such as in the case of the her1 and her7 genes in Zebrafish [98]. It is possible that such redundancy has allowed neutral exploration of genotypes in evolution, allowing the discovery of novel, highly divergent clock dynamics and expression [22]. We can expect the role of gene duplication in the evolvability of somitogenesis to become clearer as the availability of comparative data increases.

It is possible that a similar property exists for the morphogenetic processes underpinning elongation of the PSM. Across different vertebrate species elongation of the PSM is known to be driven by cell division [113], as well as ingression of nascent PSM progenitor cells from the epiblast [113, 114, 116], and spatiotemporal changes in cell motility [116, 119, 120, 122] and cell density [119, 121, 124]. The degree to which any one of these processes contributes to elongation of the PSM appears to vary across species [46, 113], which could suggest that diverse PSM elongation dynamics can be generated by the combinatorial effects of these processes. If this were the case, then the existence of a set of uncoupled morphogenetic processes such as these could be regarded as a source of evolvability for somitogenesis, however whether such processes are uncoupled is as yet deeply unclear and the development of cell-based computational models of PSM elongation [54, 118, 119, 121] will reveal whether this is the case, and with it, the scope for neutral evolutionary exploration of the mechanisms underlying PSM elongation.

The topology of the segmentation clock gene regulatory network may also be a source of

evolvability, specifically the existence of two oscillatory circuits: one being the Hes/Her auto-repressive loop, and the other the negative feedback of delta-notch signalling onto itself via intercellular signalling (figure 1.5A) [48]. Using a simple model, Lewis [48] demonstrated that the relatively long time delay created by intercellular signalling allows oscillations with much lower frequency to be created when auto-repression of Hes/Her expression is decreased, and speculated the covariation in these two processes could perhaps give rise to different clock frequencies in evolution [48]. However, simulations suggest the existence of a parameter space with irregular oscillations as the relative contribution of delta/notch to the circuit increases, and thus the evolution of this system may actually be constrained in this regard [48]. Inferring the relative contribution of each circuit in determining the clock's frequency is a hard problem as it requires experimental parameterisation of this model, and to my knowledge there have been no comparative studies examining the contributions of each circuit to clock period, or their change in evolution.

The delay in intercellular communication via delta-notch is also likely to exert constraints on the evolution of the clock gene regulatory network, and vice versa. In order for stable and synchronous oscillations to be formed the delay in intercellular signalling must satisfy a relationship with the frequency of auto-repressive intracellular oscillations of Hes/Her genes where signals are sent and received by cells in the correct phase of the clock [48, 93]. The delay in intercellular signalling is thought in part to be controlled by the number of intermediary molecules, such as lunatic fringe (Lfng), involved in the notch signalling pathway [92], and therefore the number of intermediary components in signal transduction may be constrained by the frequency of Hes/Her oscillations. We also can expect the converse to be true, and that the frequency of Hes/Her oscillations is constrained by the existing signalling delay. However, the collective frequency of oscillations is also set by the value of the delay [48, 93], and so it may be that evolutionarily neutral regions of parameter space exist where the delay and the frequency may co-vary without perturbing the overall frequency. Only with further theoretical study will it become clear to what capacity the delay and frequency of the clock mutually constrain one

another's evolution.

It is unclear how the behaviour of the oscillatory components of the FGF/Wnt pathways [45, 185, 190, 191], which are capable of changing their expression dynamics in evolution [45] might affect segmentation clock dynamics. Indeed, in Mouse at least it appears that the dynamics of Wnt and Notch signalling are coupled, with functional significance for somite polarity [191]. However, there is limited comparative data on whether this is a conserved feature of vertebrates and as such we cannot predict in what way such regulation may affect the evolution of the clock. Additionally, I note that as PSM morphogenesis is under the regulation of Wnt and FGF signalling [111, 114, 119, 120, 122], the evolution of morphogenesis and the dynamics of the clock may be prone to co-vary according to changes in Wnt or FGF signalling.

Vertebral number has been shown to depend on maternal effects [192, 193], as have the period of somitogenesis and its duration [192]. The total duration of somitogenesis appears to correlate with the amount of yolk available to the embryo [113], and the number of somites has been shown to be reduced in conjoined twin embryos of the Lake trout (*Salvelinus namaycush*), leading to a suggestion that the number of vertebrae depends on the quantity of yolk available to the embryo [194]. The maternal contribution to the offspring can vary with the age of the mother [195], suggesting this as a possible source of phenotypic heterogeneity in otherwise genetically homogenous populations, however it is unknown whether age-specific effects occur in the vertebral number of offspring.

Vertebral number is known to exhibit environmentally-induced phenotypic plasticity, with vertebral number responding to changes in salinity, light, and temperature [63, 64, 196]. Due to their external mode of development, this has been most extensively studied in the teleost fishes. The environmental effect of temperature here is perhaps counter-intuitive, as the frequency of somitogenesis increases with increasing temperature [75, 76, 149], whereas it is typically observed in nature that lower temperatures yield fish with more vertebrae [60, 63, 64, 197]. Indeed the relationship between embryo temperature and vertebral number is thought to be U-shaped, where the maximal number of vertebrae occurs at extremely low or extremely high

temperatures and at intermediary temperatures the lowest number of vertebrae are formed [196, 198]. The developmental basis of such a trend is unclear. Furthermore, the effect of temperature on somitogenesis is thought to be buffered by the invariance of somite length [76], likely due to a temperature-induced scaling in wavefront speed and clock molecular kinetics [199]. It is likely therefore that this scaling relationship, or any scaling relationship that exists between the frequency and duration of somitogenesis, breaks down at extreme temperatures.

While our understanding of somitogenesis and the segmentation clock has benefitted from a large body of theoretical work, it is still difficult to definitively identify sources of evolvability within somitogenesis. This is in large part due to a still incomplete understanding of the gene regulatory architecture underpinning the segmentation clock and the wavefront, as well as the computationally complex problem of understanding how cellular behaviours such as motility and division give rise to the elongation of the PSM, and the added difficulty of parameterising such models using experimental data. Furthermore, comparative studies of somitogenesis have largely been limited to vertebrate model species that are significantly diverged [45, 46, 100, 102], and there have been few comparative studies examining how vertebral number and the components and dynamics of somitogenesis co-vary over short evolutionary timescales [15, 158], so it is therefore difficult to suggest from observational data where the sources of evolvability may lie.

1.5 Thesis aims

As there has been little to no work studying the evolvability of vertebral number, I have focused this thesis on identifying sources of evolvability in somitogenesis. I also aim to understand their evolution. To do so, I have combined a computational approach with a comparative experimental study of somitogenesis in a closely-related lineage of fishes. It is hoped that the results presented in this thesis will improve our understanding of vertebrate evolution, as well as the developmental origins of evolvability and their evolution.

2

Modularity of the segmentation clock and morphogenesis

As discussed in the previous chapter, somitogenesis appears to be comprised of two distinct mechanisms that determine vertebral number, namely the frequency and duration of somitogenesis. If these processes were modular, i.e. if change in the function or dynamics of one process has no effect on the function or dynamics of the other, and vice versa, then this could perhaps explain the high degree of evolvability we observe in vertebral number. Large changes in the number of vertebrae would be achieveable by a combined effect of small changes in both frequency and duration, allowing for more rapid evolution. As discussed in the previous chapter, it appears that several species such as the Corn snake (*Pantherophis guttatus*) have evolved large numbers of vertebrae via changes in both duration and frequency [46], an observation that would appear to support the modularity of these two processes.

However determining whether these two processes are modular is non-trivial. While the cues determining the duration of somitogenesis are unknown, evolutionary changes in the duration of somitogenesis appear to correlate with changes in the elongation dynamics of the PSM [46, 113], suggesting that variation in the morphogenesis of the PSM drives, at least in part, the variation in the duration of somitogenesis we see in the vertebrates [44]. However there is no reason to expect that PSM morphogenesis is free to vary without disrupting the function of the segmentation clock. Cell division causes clock oscillations to arrest, creating noise [82, 200], and synchrony of the clock is promoted by random cell mixing [140, 201, 202], a process thought to be essential to elongation of the PSM [119–122]. Therefore variation in either of these processes to create diverse dynamics for the elongation of the PSM may affect synchrony of the clock, which is essential for correct patterning of the somite and development of the vertebra [91, 111], and therefore the two processes may not be modular. While a lack of modularity of frequency and duration would not explain the evolvability of vertebral number, the constraint on the evolution of PSM morphogenesis by the clock would nevertheless be interesting as it may create biases in the evolution of PSM elongation dynamics that lead to macroevolutionary trends in the evolution of vertebral number.

Therefore, in this chapter I have chosen to investigate whether the segmentation clock and morphogenesis of the PSM exhibit developmental modularity. To do so I first note that the disruption of clock frequency and synchrony appears to have no effect on elongation of the body [91, 136], nor on the shape of the PSM [90], suggesting that elongation of the PSM is independent of the function and dynamics of the clock. Therefore to demonstrate modularity of the clock and morphogenesis it suffices to test whether the dynamics and function of the clock are unaffected by variation in the morphogenesis of the PSM. As predictably manipulating the morphogenesis of the PSM *in vivo* is extremely challenging, I have chosen to take a computational approach, to predict how clock dynamics react to varying morphogenesis of the PSM.

Details of the computational methods, as well as the results, are presented below. The results for

this chapter have been published as a manuscript in the open-access journal eLife (Hammond, Baker, and Verd; 2025 eLife **14**, RP106316. doi:10.7554/eLife.106316.2), and consequently I have chosen to present my work for this chapter in manuscript format. This manuscript is included below in the place of a convential thesis chapter.

Modularity of the segmentation clock and morphogenesis

James E. Hammond[1], **Ruth E. Baker**[2], **Berta Verd**[1*]

[1]Biology Department, University of Oxford, Oxford, United Kingdom; [2]Mathematical Institute, University of Oxford, Oxford, United Kingdom

Abstract

Vertebrates have evolved great diversity in the number of segments dividing the trunk body, however the developmental origin of the evolvability of this trait is poorly understood. The number of segments is thought to be determined in embryogenesis as a product of morphogenesis of the pre-somitic mesoderm (PSM) and the periodicity of a molecular oscillator active within the PSM known as the segmentation clock. Here we explore whether the clock and PSM morphogenesis exhibit developmental modularity, as independent evolution of these two processes may explain the high evolvability of segment number. Using a computational model of the clock and PSM parameterised for zebrafish, we find that the clock is broadly robust to variation in morphogenetic processes such as cell ingression, motility, compaction, and cell division. We show that this robustness is in part determined by the length of the PSM and the strength of phase coupling in the clock. As previous studies report no changes to morphogenesis upon perturbing the clock, we suggest that the clock and morphogenesis of the PSM exhibit developmental modularity.

Introduction

Vertebrates exhibit great diversity in the number of segments that divide the skeleton, musculature, and nervous system of the body, along the rostral-caudal axis (***Richardson et al.*** (***1998***)). However, the developmental basis for this evolvability remains poorly understood. Segmentation of the body is established by paired blocks of mesodermal tissue, known as somites, that form periodically on either side of the developing embryo as it elongates posteriorly, in a process known as somitogenesis (***Morin-Kensicki et al.*** (***2002***); ***Gomez and Pourquié*** (***2009***); ***Oates et al.*** (***2012***)). The total number of segments corresponds to the number of somites formed in the embryo, which is thought to be an emergent property of the morphogenesis of the pre-somitic mesoderm (PSM) and the dynamics of a molecular oscillator known as the segmentation clock (***Morin-Kensicki et al.*** (***2002***); ***Cooke and Zeeman*** (***1976***); ***Gomez et al.*** (***2008***); ***Gomez and Pourquié*** (***2009***); ***Harima et al.*** (***2013***); ***Schröter and Oates*** (***2010***)). Here, we hypothesise that the evolvability of vertebrate segment number may be underpinned by developmental modularity of the segmentation clock and morphogenesis of the PSM (***Raff*** (***1996***)), and that the potential of these two processes to evolve independently from one another may explain the diversity observed in vertebrate segment number.

The segmentation clock is a complex gene regulatory network thought to be driven by cell-autonomous oscillations of transcription factors in the Hes/Her family (***Palmeirim et al.*** (***1997***); ***Lewis*** (***2003***); ***Harima et al.*** (***2013***); ***Schröter et al.*** (***2012***); ***Webb et al.*** (***2016***)). Noisy oscillations are synchronised across cells by delta-notch signalling (***Horikawa et al.*** (***2006***); ***Lewis*** (***2003***); ***Venzin and Oates*** (***2020***)), creating synchronous travelling waves of gene expression that pulse

from the posterior to the anterior of the PSM. In the anterior PSM somite boundaries are patterned by the interaction of clock oscillations with a posterior-anterior decreasing gradient of FGF signalling that is thought to act as a 'wavefront', reading-out the phase of the clock to create a spatially periodic pattern of somite boundaries (***Cooke and Zeeman*** (***1976***); ***Soroldoni et al.*** (***2014***); ***Simsek and Özbudak*** (***2018***); ***Simsek et al.*** (***2023***)). The clock synchronises the differentiation of PSM cells as they adopt somite fates at the anterior of the PSM (***Cooke and Zeeman*** (***1976***)), and so controls both the tempo at which somites are formed, and the anterior-posterior polarity of the somites (***Simsek et al.*** (***2023***)). Thus, clock synchrony controls the accuracy of somite patterning, and the frequency of clock oscillations determines the frequency of somite formation, which (together with the total duration of somitogenesis), determines the total number of somites formed.

Concurrent with somitogenesis, the PSM undergoes elongation. Comparative studies have shown that a diverse range of mechanisms are responsible for elongation of the PSM (***Gomez et al.*** (***2008***); ***Bénazéraf et al.*** (***2010***); ***Steventon et al.*** (***2016***); ***Mongera et al.*** (***2018***); ***Thomson et al.*** (***2021***); ***Michaut et al.*** (***2022***)). Importantly, the elongation dynamics of the PSM are thought to control the total duration of somitogenesis, which terminates once the PSM becomes sufficiently short in the AP direction (***Gomez et al.*** (***2008***); ***Gomez and Pourquié*** (***2009***); ***Steventon et al.*** (***2016***)). Therefore, the total number of segments formed in the developing vertebrate is thought to be the product of both the dynamics of the clock (which controls the rate of somite formation), and the morphogenesis of the PSM, via its control on the total duration of somitogenesis.

Developmental modularity is a property of two or more developmental pro-

cesses where their uncoupling in space or time permits their evolution independently of one another (***Raff*** (***1996***)). This property is thought to give rise to increased phenotypic diversity by enhancing the evolvability of the system (***Raff*** (***1996***)). In species examined thus far, such as the Corn Snake *Pantherophis guttatus*, it appears that the evolution of segment number is driven by changes in both the dynamics of the clock and the elongation of the PSM (***Gomez et al.*** (***2008***)), and it is possible that independent evolution of these two processes is responsible for the high degree of diversity observed in vertebrate segment number. However, it is not obvious whether the two processes are modular, and if so, to what degree. Perturbing the clock's periodicity and function does not appear to affect elongation of the body axis (***Schröter and Oates*** (***2010***); ***Forero et al.*** (***2018***)), so morphogenesis of the PSM is likely to be robust to changes in the segmentation clock through evolution. However, it is possible that many of the cellular- and tissue-level processes which drive PSM elongation could exert an effect on the dynamics of the clock.

For instance, cell rearrangements thought to drive elongation of the PSM in zebrafish (*Danio rerio*) (***Lawton et al.*** (***2013***); ***Mongera et al.*** (***2018***)), and in chicken (*Gallus gallus*) (***Bénazéraf et al.*** (***2010***); ***Michaut et al.*** (***2022***)), promote synchronisation of the segmentation clock (***Uriu et al.*** (***2010***, 2013); ***Uriu and Morelli*** (***2014***); ***Uriu et al.*** (***2017***)). Additionally, arrest of transcription during chromatin condensation is known to cause PSM cells to lag their clock expression relative to neighbours after cell division (***Horikawa et al.*** (***2006***); ***Delaune et al.*** (***2012***)), causing defects in clock synchrony (***Murray et al.*** (***2013***)) and suggesting clock synchrony may depend on the degree of proliferative growth in the PSM. The ingression of PSM progenitor cells from dorso-posterior and lateral donor tissues (***Kanki and Ho*** (***1997***); ***Steventon et al.*** (***2016***); ***Xiong et al.*** (***2020***)) can also

create clock asynchrony, as incoming progenitor cells do not appear to show clock gene expression (***Mara et al.*** (***2007***)) and thus may be asynchronous with their neighbours when they enter the PSM. Finally, tissue convergence movements associated with the elongation of the PSM (***Thomson et al.*** (***2021***)) have been proposed to negate the effect of random cell mixing and be deleterious for clock synchronisation (***Uriu and Morelli*** (***2014***)). It is thus non-trivial to determine whether the duration and rate of somitogenesis are modular since, in order to do so, one must examine the effect of varying morphogenesis on the dynamics of the clock.

To do this, we use a computational approach to simulate clock dynamics and cell movements within the PSM, and study how the clock responds to changes in morphogenesis. We use this approach as it is much quicker and more flexible than experimentally manipulating morphogenesis *in vivo*. As we are constrained by computational complexity and cannot simulate all possible means by which the PSM can elongate (this would require simulation of the growth and dynamics of cells in surrounding tissues which exert forces on the PSM (***Xiong et al.*** (***2020***))), we limit our study to simulating regimes of cell movement and growth within the PSM that are known to be associated with, or causal in, PSM growth and elongation across vertebrates.

We use a previously established three-dimensional model of cell movements and clock dynamics within the zebrafish PSM (see figure 1A) (***Uriu et al.*** (***2021***)), and adapt this model to simulate various regimes of cell movement and growth. Briefly (a more detailed description of the model can be found in the methods), the model assumes the PSM is in an inertial frame of reference, with elongation being encoded by the advection of cells towards the anterior PSM (the anterior

limit of which is denoted $x = x_a$, see figure 1A). Cells are modelled as point particles and are subject to random cell mixing, a cell-cell repulsion force that enforces volume exclusion of cells, and a boundary force confining the cells within a hollow horseshoe-shaped domain that reflects the geometry of the PSM. This model uses a phase-oscillator approach to describe segmentation clock dynamics and each cell is assigned a clock phase which is frozen when the cell exits the PSM at $x = x_a$, crossing the wavefront and patterning a segment (figure 1A). To replenish cells lost at the anterior, new cells are added with random phase and random position to keep the tissue at a constant density (figure 1B).

This model is appealing for several reasons, the first being that the model is three-dimensional and so can quantitatively recapture the rates of cell mixing that we observe within the PSM (***Uriu et al.*** (***2017***, 2021)). The second is that the segmentation clock is described by an abstract phase-oscillator, broadly agnostic to the molecular details which are known to vary between vertebrates (***Krol et al.*** (***2011***)). Finally, the model makes extensive use of experimentally-derived parameters for cell movements, tissue dimension, and the phase-oscillator model of the clock in zebrafish (***Uriu et al.*** (***2021***)). This allows us to make quantitative predictions as to how the zebrafish segmentation clock would react to changes in PSM morphogenesis.

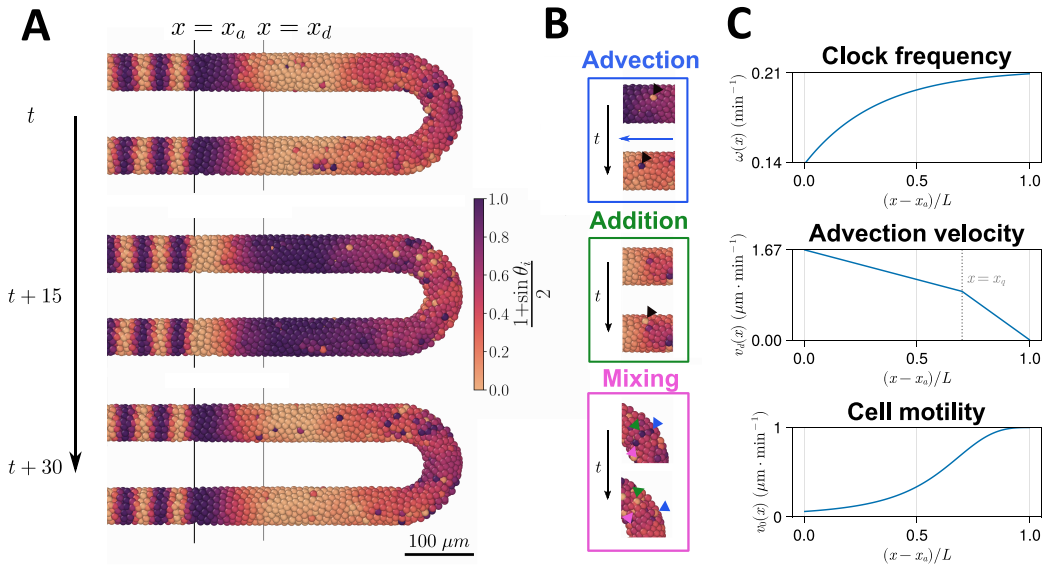


Figure 1. Computational model of the clock and the PSM. A Stills of a simulation of the model of ***Uriu et al.*** (***2021***). Kinematic phase (θ_i) waves emerge in the posterior (right) and travel towards the tissue anterior (left, $x = x_a$), where phase is arrested. The model is parameterised to data from zebrafish, and accordingly the clock oscillates every 30 minutes. **B** Insets illustrating the key processes driving cell movements in the PSM within the model. Top: Cells advect towards the anterior of the tissue, simulating elongation of the PSM. Middle: New cells are added to replenish the loss of cellular material as cells advect towards the anterior. Bottom: Cells undergo motility-driven rearrangements. **C** Functions in the model describing (top) the intrinsic oscillation frequency, (middle) the advection velocity, (bottom) and the motility, of each cell depending on its normalised position along the anterior-posterior axis, $(x - x_a)/L$. Plots were generated using the parameters given by ***Uriu et al.*** (***2021***).

Using this model we test the clock's response to changes in the position of cell ingression into the PSM, changes in the anterior-posterior profile of random cell motility, changes in the length and density of the PSM, and to mitosis. We find that clock synchrony and frequency are robust to these changes except when mitosis is introduced, where we find that clock synchrony is negatively impacted except in cases where mitosis is confined to the posterior PSM. We find that this robustness is underpinned by the tissue length and cell density in the model, as well as the rate of cell mixing and strength of clock phase coupling. Together,

these results suggest that segmentation clock dynamics and PSM morphogenesis are modular components of somitogenesis that can evolve independently from one another, conferring evolvability and helping to explain the diversity in segment number across the vertebrates.

Results

Varying the position of cell ingression has minimal effect on clock dynamics

During somitogenesis new cells enter the PSM posteriorly via ingression from dorsal tissues (***Kanki and Ho*** (***1997***); ***Banavar et al.*** (***2021***)). In the early stages of zebrafish somitogenesis, there is also a contribution of cells from lateral tissues (***Steventon et al.*** (***2016***)). As ingressing cells do not appear to express segmentation clock genes (***Mara et al.*** (***2007***)), the position at which cells ingress into the PSM can create challenges for clock patterning, as only in the 'off' phase of the clock will ingressing cells be in-phase with their neighbours. Therefore continuous ingression of PSM progenitor cells is likely to create local asynchrony of oscillations where it occurs, and so varying the position of cell ingression may have an effect on clock dynamics at the wavefront.

To test this hypothesis, we alter how cells are added to the tissue in the model. To model the ingression of cells from dorsal tissues into the posterior tailbud (***Kanki and Ho*** (***1997***); ***Banavar et al.*** (***2021***)), we define a density-dependent cell addition process where cells are added onto the dorsal surface of the posterior half-toroid subdomain (see figures 2A & 9) if the average density in this subdomain falls below ρ_0, with a constant initial phase $\theta = 0$. To maintain equal density across the PSM, we found it necessary to add cells in the two cylindrical PSM

subdomain as well, as when cells were only added to the posterior half-toroid subdomain we observed a loss in density in the two cylindrical subdomain (data not shown). To avoid bias we add cells at random positions if the average density falls below ρ_0. These cells are initialised with a random phase drawn from the uniform distribution $[0, 2\pi)$. To model the early stages of zebrafish somitogenesis where, concurrent with dorso-posterior ingression of cells, the PSM also experiences a contribution of cells from lateral tissues, we modify the above simulation so that cell density in the lateral cylindrical subdomain is maintained by addition of cells with constant phase ($\theta = 0$) on the ventral surface of these two lateral subdomain (see figures 2A & 9). These two ingression scenarios are termed 'Dorso-Posterior' ('DP') and 'Dorso-Posterior + Lateral-Ventral' ('DP+LV') ingression (figure 2A), respectively, and for brevity we refer to them as 'DP' and 'DP+LV' in the rest of the text. We compare the effect of these scenarios of cell ingression with 'Random' ingression (figure 2A), where cell density is maintained by addition of cells at random positions within the tissue with random phase ($\theta \in [0, 2\pi)$).

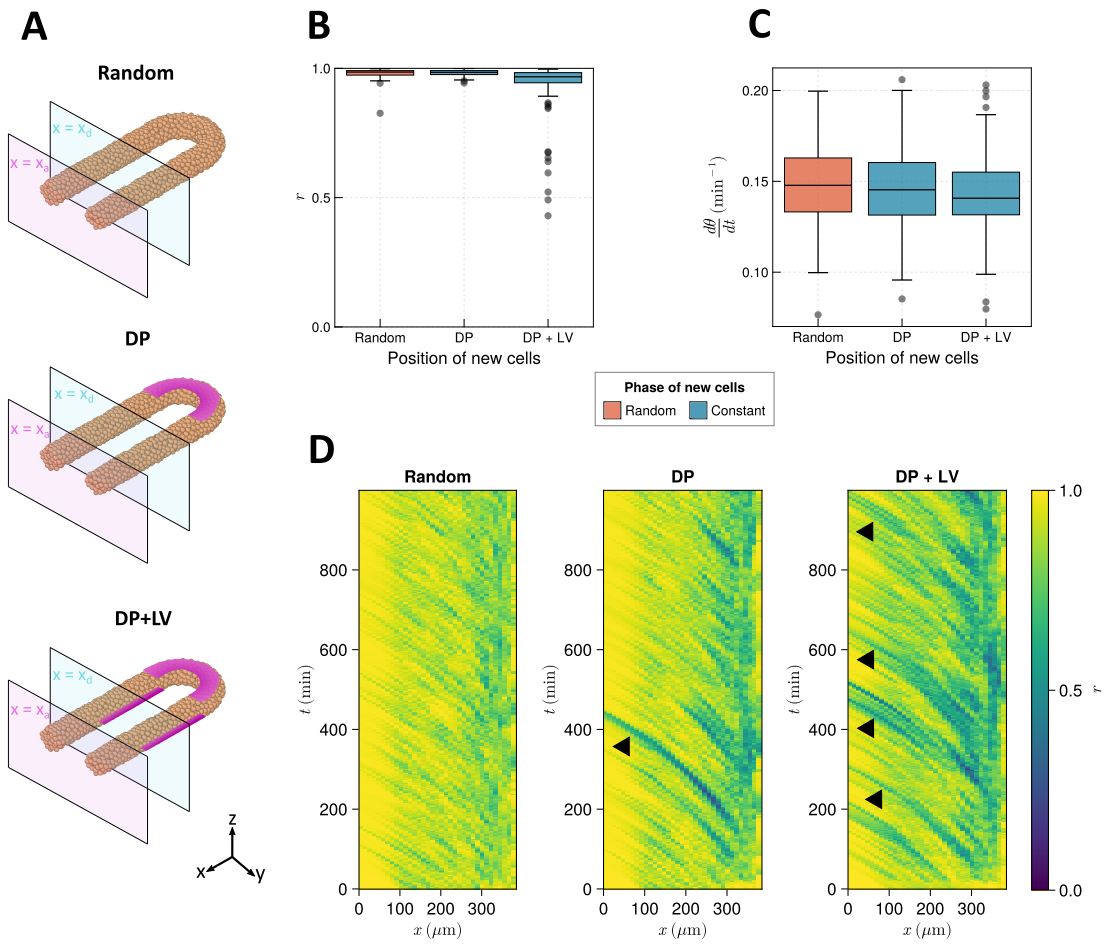


Figure 2. The effect of cell ingression position on clock frequency and synchrony. **A** Diagram highlighting the position of cell addition across the three ingression scenarios tested here. Magenta shading shows where cells are added onto the tissue surface. The pale magenta and blue planes respectively correspond to the anterior limit of the PSM, $x = x_a$, and the anterior limit of cell addition, $x = x_d$. In the 'Random' condition cells are added at random positions within the PSM posterior to the plane $x = x_d$, and accordingly no magenta surface is shown. Similarly in the dorso-posterior case ('DP') cells are added at random positions in the two lateral cylinders to maintain density, and no magenta surface is shown there. In the dorso-posterior + lateral-ventral ('DP+LV') case cells are only added on the tissue surface at the positions shown by magneta shading. **B** Oscillation synchrony (r) at the PSM anterior after 1000 mins, for the three tested scenarios of cell ingression. $N = 100$ simulations. **C** Mean frequency of oscillations for cells at the PSM anterior after 1000 mins of simulation, for the three tested scenarios of cell ingression. $N = 100$ simulations. **D** Kymographs of synchrony along the x-axis on the left-hand side of the PSM for single simulations from each of the three scenarios of cell ingression tested. Black arrowheads highlight strongly asynchronous populations of cells being transported to the tissue anterior by advection.

We test these scenarios by initialising each simulation with a synchronous initial condition for the clock and simulating cell movements and clock dynamics in the presence of each cell ingression scenario for 1000 mins (see methods for further justification). After 1000 mins we measure the synchrony (r) and mean frequency ($\mathrm{d}\theta/\mathrm{d}t$) of a cylindrical region of tissue at the left-hand PSM anterior, one cell diameter (d_c) in length (see methods). We restrict our measurement to the PSM anterior as it is here that the clock patterns somites (***Simsek and Özbudak*** (***2018***); ***Simsek et al.*** (***2023***)), and thus dynamics elsewhere in the tissue are unlikely to present a phenotype in the animal. Using this method, we find that DP ingression has no effect on synchrony at the PSM anterior after 1000 mins when compared to Random ingression, but that DP+LV ingression creates a minor decrease in anterior synchrony (figure 2B). However, in neither scenario is the anterior frequency changed when compared to Random ingression (figure 2C). We see equivalent results when ingressing cells are initiated with a random phase (figure 2 supplement 1). Kymographs of synchrony along the anterior-posterior axis reveal that cell addition onto the tissue surface creates asynchronous populations of cells that travel towards the tissue anterior with cell advection (figure 2D, black arrowheads). These are more frequent than in the DP case, likely explaining the lower anterior synchrony observed in the DP+LV simulations.

The clock is robust to changes in mode of cell ingression regardless of the cell motility profile

As cell mixing promotes clock synchrony (***Uriu et al.*** (***2010***, 2013); ***Uriu and Morelli*** (***2014***)), and an anterior-posterior profile of increasing cell mixing is present in the PSM of zebrafish and chicken (***Kanki and Ho*** (***1997***); ***Bénazéraf et al.*** (***2010***); ***Lawton et al.*** (***2013***); ***Mongera et al.*** (***2018***); ***Michaut et al.*** (***2022***)) (and presum-

ably most vertebrate taxa), we speculated that the shape of this gradient may confer robustness to clock dynamics against cell ingression. For instance, a steep profile (e.g. as reported from the early somitogenesis stages of zebrafish (***Mongera et al.*** (***2018***))) might confer robustness to lateral ingression as mixing is higher along a greater length of the PSM than a more graded profile (e.g. as reported from chicken embryos (***Bénazéraf et al.*** (***2010***); ***Michaut et al.*** (***2022***))). Therefore the shape of the cell motility profile in the PSM may constrain the evolution of cell ingression position, and vice versa.

To test this, we systematically enumerated the steepness (h) and inflexion point (X_v) of the function $v_0(x)$ (figure 3B), which controls the speed of random cell movement in the PSM (see methods, figure 3A). For each parameter pair we performed $N = 100$ simulations and recorded the median anterior synchrony and median anterior mean frequency ($\mathrm{d}\theta/\mathrm{d}t$, see methods) after 1000 mins. The results for each ingression scenario are shown in figure 3C. We find no trend in median anterior synchrony and frequency across different parameter pairs for Random ingression, DP ingression, or DP+LV ingression (figure 3C). Similar to before, results for DP ingression are indistinguishable from Random ingression. For DP+LV ingression, anterior synchrony is lower across all parameter sets than DP ingression or Random ingression (figure 3C) but the effect is still very minor in this case. We do not observe any X_v- or h-dependent trend when we examine synchrony along the anterior-posterior axis (figure 3 supplement 1). We therefore predict that, at least with respect to effects on the zebrafish segmentation clock, the evolution of the motility profile and the position of cell ingression are not constrained by one another.

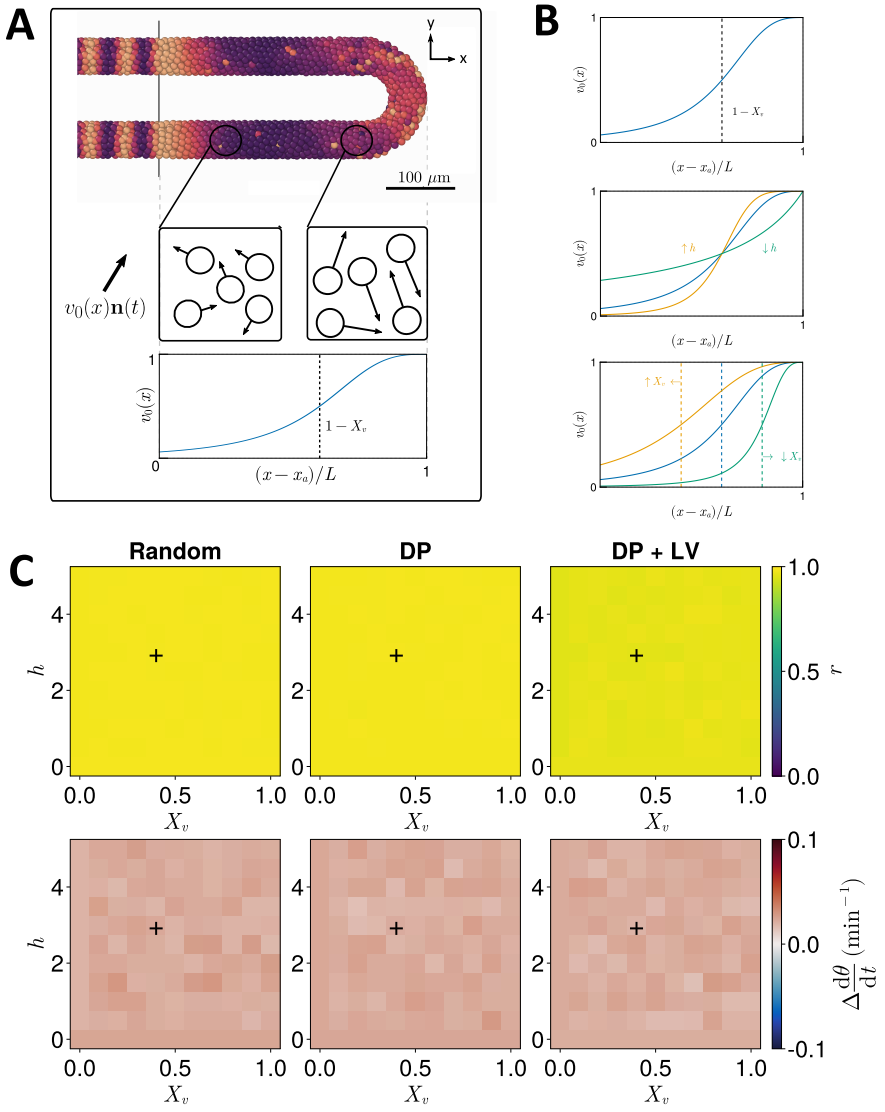


Figure 3. Effect of cell motility profile on clock frequency and synchrony. A Overview of how intrinsic cell motion is encoded in the model. Each cell is given a random direction vector $v_0(\mathbf{x}_i)\mathbf{n}_i(t)$ (black arrows) whose magnitude $v_0(\mathbf{x}_i)$ increases towards the PSM posterior. **B** Magnitude of intrinsic cell motion $v_0(\mathbf{x})$ for the parameters used by ***Uriu et al.*** (***2021***) (blue), and how the shape of the function can change when increasing (yellow) or decreasing (green) the inflexion point and curve steepness parameters, X_v and h, respectively. **C** Clock synchrony r (**top**) and difference from expected mean frequency $\Delta \mathrm{d}\theta/\mathrm{d}t$ (**bottom**) for different $v_0(\mathbf{x})$ specified by combinations of X_v and h. The corresponding pixels display the synchrony or frequency at the PSM anterior after 1000 mins of simulation using the specified motility profile, averaged across $N = 100$ simulations. A black + corresponds to the parameter pair used elsewhere in this paper, unless otherwise stated ($X_v = 0.4, h = 3$). Parameter ranges used are $X_v \in \{0, 0.1, 0.2, \ldots, 1\}$ and $h \in \{0, 0.5, 1, \ldots, 5\}$

.

Length and density of the PSM confer robustness to the clock

The PSM is a transient tissue that shrinks in length over time. While this feature is conserved across vertebrate species, the absolute length and rate of shrinkage are known to vary between taxa (***Gomez et al.*** (***2008***); ***Steventon et al.*** (***2016***); ***Thomson et al.*** (***2021***)). We postulate that PSM length is an important factor in determining the clock's response to different morphogenetic modes; for example, intuitively cells will have more time to synchronise oscillations with their neighbours before reaching the PSM anterior in a longer PSM, or if cell ingression is spread out along a longer tissue the degree of asynchrony imparted by ingression will be lessened. Similarly, we predict that the density of the tissue confers robustness to clock dynamics as increasing the average number of neighbours for each cell may act to 'correct' against clock noise. Overall, tissue density has been observed to increase over time in zebrafish (***Thomson et al.*** (***2021***)), and has been observed to vary in space in chicken, with density decreasing towards the PSM posterior (***Bénazéraf et al.*** (***2010***); ***Michaut et al.*** (***2022***)).

We thus sought to explore whether the length and density of the PSM confer robustness to changing morphogenesis. We test this by co-varying the position of cell ingression (figure 2A) with either tissue density or length. When varying the average tissue density ρ_0 and studying model dynamics after 1000 mins we see that increasing cell density marginally increases anterior synchrony (figure 4B), and decreases the variability in frequency at the PSM anterior (figure 4C). As expected, the number of neighbours for each cell is increased (figure 4 supplement 1). However the increase in density does not appear to be sufficient to 'correct' the decrease in synchrony in the DP+LV case (figure 4B). Increasing the length of the tissue does however appear to 'correct' this decrease (figure

4E). Here the PSM length L is increased but the anterior limit of cell addition x_d is held constant at $x_d = 100$ μm. Anterior synchrony after 1000 mins positively correlates with L, and decreasing L reveals a decrease in anterior synchrony for DP ingression (figure 4E). No obvious trend for mean anterior frequency is observed (figure 4F). Kymographs of synchrony along the x-axis suggest that while much of this correlation is driven by cells having more time to synchronise their neighbours with increasing L, a significant part of this may be due to cell addition occurring along a longer domain, and in so doing 'diluting' the noise-giving effect of cell ingression (figure 4 supplement 2). To control for this, we restrict cell addition to only occur along a fixed length of the tissue with $x_d = L - R - r_T - 100$ (where R represents the major radius of the half-toroid subdomain and r_T the radius of the PSM tube, see figure 9), and vary the length of the tissue L. We observe similar results as to before, with synchrony positively correlating with L and no clear trend for clock frequency (figure 4 figure supplement 3). Kymographs of synchrony along the x-axis confirm that increasing L increases the effective time cells have to synchronise oscillations with their neighbours (figure 4 figure supplement 4).

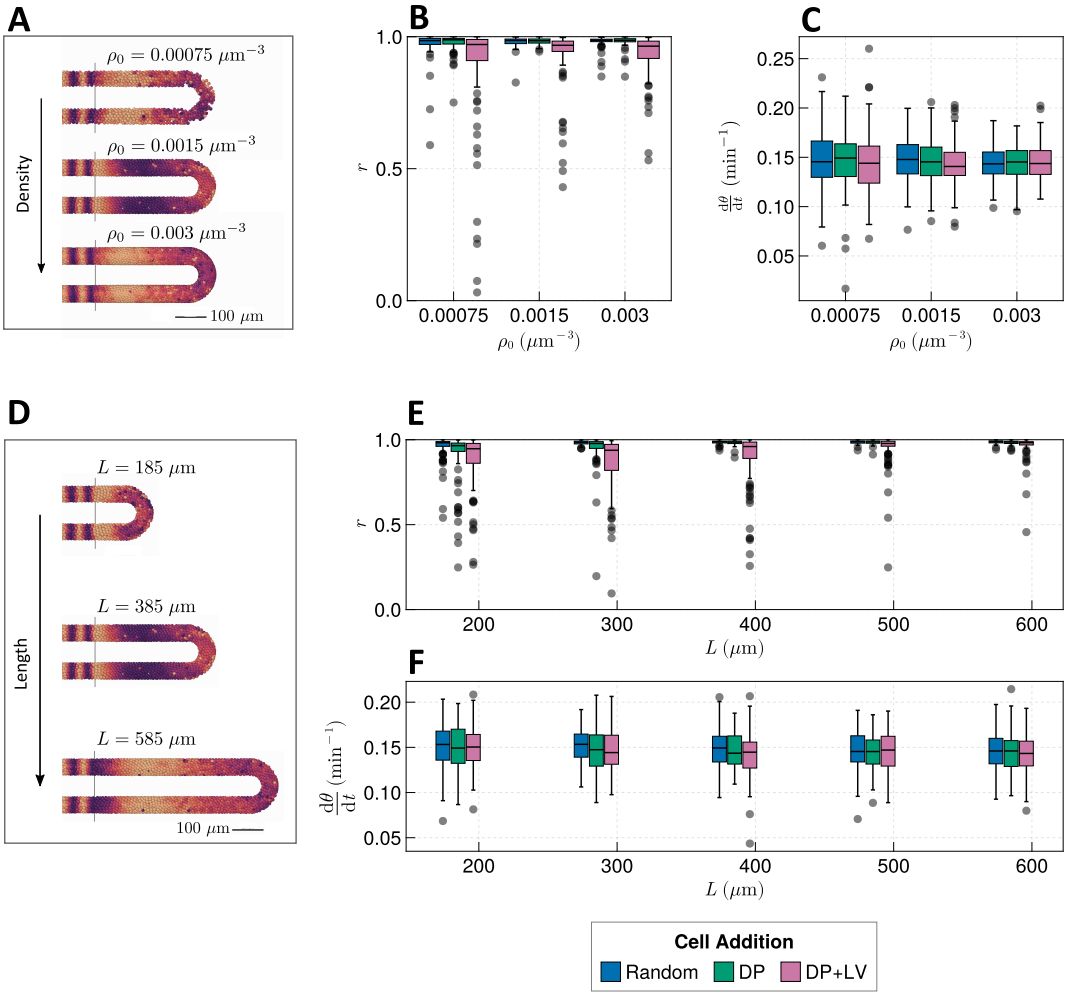


Figure 4. Effect of tissue density and length. A Stills from exemplar simulations illustrating the impact of changes in tissue density ρ_0. **B** Anterior synchrony after 1000 mins for changing tissue density ρ_0 and varying position of cell ingression. $N = 100$ simulations. **C** Anterior mean frequency after 1000 mins for changing tissue density ρ_0 and varying position of cell ingression. $N = 100$ simulations. **D** Stills from exemplar simulations illustrating changes in tissue length. **E** Anterior synchrony after 1000 mins for changing tissue length L and varying position of cell ingression. $N = 100$ simulations. **F** Anterior mean frequency after 1000 mins for changing tissue length L and varying position of cell ingression. $N = 100$ simulations.

We note that when either one of density (ρ_0) or PSM length (L) increase, the differences in synchrony between the three scenarios of cell ingression become less pronounced (figures 4B,E). This suggests that density and length are capable of independently conferring robustness to the segmentation clock against

changes in morphogenesis, however co-varying density and length suggests that the relative effect of length is stronger (figure 4 supplement 5). In zebrafish density and length are known to co-vary after the 16-somite stage (ss), with the PSM increasing in density as it shrinks in length (***Thomson et al.*** (***2021***)). We would expect decreasing the length of the PSM to decrease the clock's robustness to noise, but it is possible that the concurrent increase in density is sufficient to counteract this and preserve clock dynamics. As the dynamics of somitogenesis are consistent across individuals at these late stages (***Schröter et al.*** (***2008***)), the dynamics of the zebrafish segmentation clock are likely unaffected by compaction-extension of the PSM. It is not clear however if this can be explained by the concurrent changes in PSM length and density that occur during compaction-extension or if clock properties such as phase-coupling need to vary in order to maintain robust clock dynamics.

Here we simulate compaction-extension by shrinking the PSM radius r and length L, while increasing its density ρ over time, using published rates and initial conditions for these values (figure 5A-B, methods) (***Thomson et al.*** (***2021***)). We also decrease the cell diameter d_c, as has been observed in zebrafish (***Thomson et al.*** (***2021***)), as we expect this to be important in determining clock dynamics by changing the number of neighbouring cells that a given cell can couple its phase to.

As these changes are dynamic, we depart from previous methodology for measuring clock dynamics and plot the anterior synchrony and frequency over time for $N = 100$ simulations. We also no longer simulate for 1000 minutes as in previous simulations, rather terminating the simulation when the length of the tissue $L(t)$ is less than x_d (i.e., when $x_a(t) + x_d \geq X_c + R + r_T(t)$). We impose this

constraint so as to prevent asynchrony induced by random addition of new cells biasing our results. It is important to note that we therefore only simulate for a part of the total duration of somitogenesis in zebrafish, and that there may be dynamics in the terminal stages of somitogenesis not captured by this model.

The anterior synchrony and mean anterior frequency over time for $N = 100$ simulations of compaction-extension are shown in figures 5C and 5D, respectively. They are compared with simulations that do not shrink the PSM and maintain constant L, r_T, d_c, and ρ. We find that towards the end of these simulations the clock experiences more noise and fluctuations in synchrony and frequency (figures 5C-D) than the non-shrinking equivalent simulations, however for the majority of the simulation the dynamics of the compacting tissue are broadly comparable with those of the non-compacting tissue, but with synchrony minorly decreasing and frequency becoming noisier towards the end of the simulation (figures 5C-D). As model parameters that are fixed throughout the simulation, such as advection velocity or intrinsic frequency, are thought to change towards the end of somitogenesis (***Schröter et al.*** (***2008***); ***Uriu et al.*** (***2021***)), it is possible that for the latter stages of somitogenesis the results of our simulations are inaccurate. Data for the terminal stages of somitogenesis are extremely limited and we cannot make estimates for model parameters at this stage. Therefore we conclude that for much of somitogenesis in zebrafish the clock is not majorly impacted by the compaction-extension of the tissue.

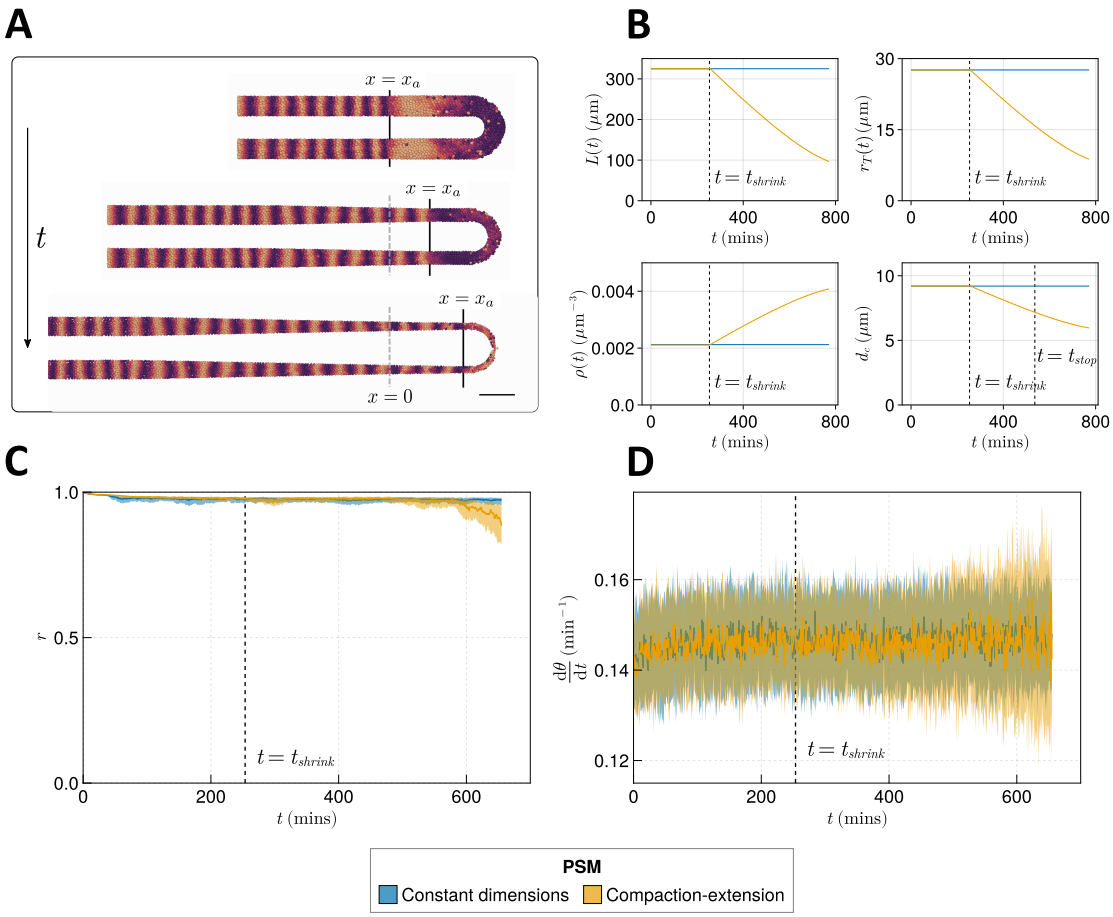


Figure 5. Effect of compaction-extension on clock synchrony and frequency. A Snapshots of an exemplar simulation showing how the PSM shrinks in length and diameter as time progresses. A 100 µm scalebar is shown in the lower right hand side of the figure. **B** Functions for PSM length L, radius r_T, density ρ, and cell diameter d_c, derived from (***Thomson et al.*** (***2021***)) (yellow), and the constant functions (blue) with which the effect of these functions is compared. **C** Anterior synchrony and **D** mean anterior frequency over time, for $N = 100$ simulations. The solid line indicates the median and the inter-quartile range is given by a shaded band either side of this line. Blue shows simulations where the tissue does not undergo compaction-extension after t_{shrink}, and yellow shows simulations where after t_{shrink} the tissue undergoes compaction-extension according to the functions shown in **B**. Results are plotted until the time at which at least one of the replicate simulations encounters a gap in the tissue at the tissue anterior (see methods).

Beyond dynamic changes in density, we also attempted to encode a spatial gradient of decreasing cell density towards the posterior, as reported from chicken (***Bénazéraf et al.*** (***2010***); ***Michaut et al.*** (***2022***)). However, we could not generate

a gradient that quantitatively recaptures that reported by ***Bénazéraf et al.*** (***2010***) (data not shown). Based on our results for changing average cell density ρ_0 (figure 4B-C) it is highly probable that this gradient would affect clock dynamics. We suggest that another model, such as one encoding cell-cell adhesion, may be a more powerful tool for exploring the effect of this gradient, and dynamics at the terminal stages of somitogenesis.

Clock arrest during cell division creates asynchrony

During M phase of the cell cycle, segmentation clock gene expression in zebrafish is known to arrest (***Horikawa et al.*** (***2006***)), causing a lag in nascent daughter cells relative to their neighbours (***Delaune et al.*** (***2012***)). If cell divisions are asynchronous, as appears to be the case in zebrafish (***Kanki and Ho*** (***1997***); ***Steventon et al.*** (***2016***)), then such lags can create asynchrony of oscillations in the tissue (***Murray et al.*** (***2013***)).

As previous models of mitosis and the segmentation clock have been two-dimensional and therefore may not quantitatively recapture the cell mixing and geometry present in zebrafish (***Murray et al.*** (***2013***, 2019)), we sought to investigate the effect of mitosis in the presence of the more accurate three-dimensional geometry and cell mixing. To do so, we simulated cell division by assigning each cell a cell cycle phase τ that increases at a constant rate of $1\ \text{min}^{-1}$, and spawning a new cell adjacent to that cell once that its $\tau \geq T_G + T_M$, where T_G denotes the total time in minutes to complete G1, S, and G2 phases of the cell cycle, and T_M denotes the time spent in M phase, in minutes. Once a daughter cell is generated, the τ for both cells is reset to zero. As daughter cells share the same clock phase after division *in vivo* (***Delaune et al.*** (***2012***)), the daughter cell shares the same clock phase θ as its sibling. To simulate the arrest of the clock during M phase,

if $\tau \in [T_G, T_G + T_M)$ oscillations cease, i.e. $\mathrm{d}\theta_i/\mathrm{d}t = 0$. For simplicity, we allow neighbouring cells to couple their phase to that of a neighbour with arrested phase, as transcriptional arrest is presumably equivalent to the 'off' phase of the segmentation clock. For further details, we refer the reader to the methods section below.

To avoid measuring trivial changes in frequency and synchrony due to arrest of the clock, we measure synchrony and frequency of the clock only in cells where $\tau < T_G$. Therefore our results represent a lower bound on the effect of mitosis on the clock. Using the estimated values of T_G and T_M for zebrafish (see methods for calculation), we find that the presence of cell division creates asynchrony as reported previously (***Murray et al.*** (***2013***)), and has no identifiable effect on frequency (figure 6B-C). Fixing $T_G + T_M = 187.5$ mins and varying T_M shows a T_M-dependent trend where synchrony decreases with increasing T_M (figure 6E). After $T_M = 15$ mins the synchrony recovers slightly but remains noisy (figure 6E). No noticeable trend is observed for the frequency (figure 6 figure supplement 1A). We observe similar results when neighbouring cells are forbidden from coupling their phase to neighbours in M phase (figure 6 supplement 2).

In previous simulations (figure 6), cell division occurs up to the wavefront where segments are pre-patterned ($x = x_a$). However if division is confined to the posterior (***Bouldin et al.*** (***2014***)) cells may have time to re-synchronise before reaching the wavefront. To test whether this is possible within the timescales of zebrafish cell movement, division, and clock dynamics, we define a point in space $x = x_a + x_{div}$ anterior to which (i.e. for $x < x_a + x_{div}$) we halt cell cycle progression by fixing $\tau = 0$ mins in all cells. Performing simulations with $x_{div} = 25$μm, 175μm, 325μm, we see that while for $x_{div} = 25$μm we observe defects in synchrony, for

$x_{div} = 175$μm, $x_{div} = 325$μm, ssynchrony is rescued (figure 6D). This suggests that there exists some point along the AP axis of the PSM, posterior to which cell division can occur without affecting the anterior dynamics of the clock. To try and resolve the value of this point and how it depends on the length of M-phase T_M, we systematically enumerated values of x_{div} from the tissue anterior to the tissue posterior, and T_M across the values tested in figure 6E. We find that for $T_M < 15$ mins division can occur up to the tissue anterior without a major defect in synchrony (figure 6 figure supplement 3A). For $T_M > 15$ mins simulations with division show high synchrony at the anterior although not so high as those with $T_M < 15$ mins, where the value of x_{div} such that clock synchrony is preserved is displaced to the posterior (figure 6 figure supplement 3A). However, only for $T_M = 15$ mins is this value majorly displaced to the posterior, with $x_{div} \approx 150$μm being the anterior-most value of x_{div} such that synchrony is preserved ($r \approx 1$) (figure figure supplement 3A). This result is perhaps intuitive, as the parameters in the model set the clock period to $2\pi/\omega_0 = 30$ mins, and a clock arrest lasting $T_M = 15$ mins would be sufficient to move a cell into antiphase relative to its neighbours. Indeed we see that mitosis is most deleterious to synchrony when M-phase lasts approximately half the length of a clock cycle (figure 6 figure supplement 3B), however for long clock cycles this relationship becomes nonlinear and the most deleterious values of T_M lie between the length of one or two clock cycles (figure 6 figure supplement 3B).

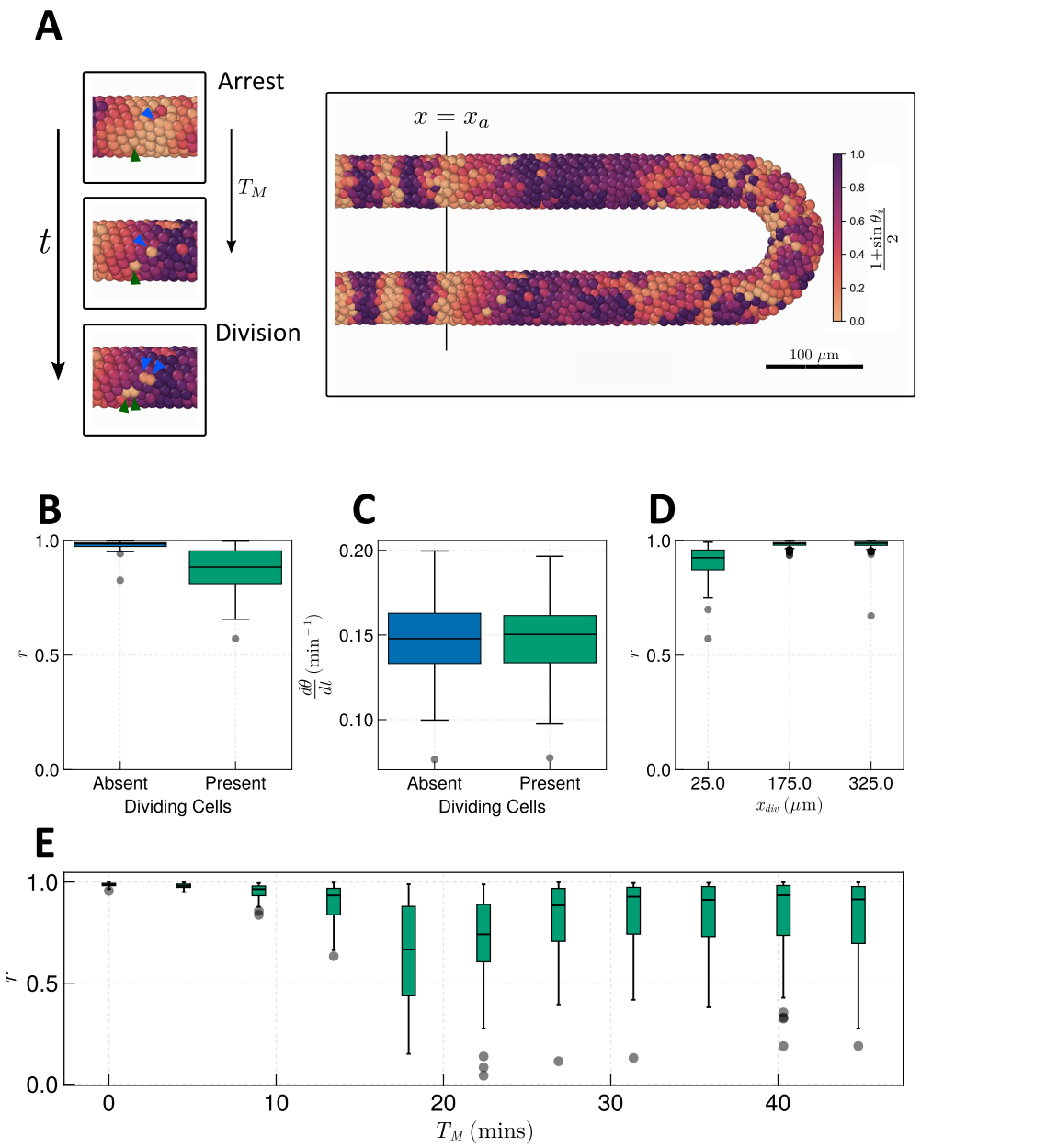


Figure 6. Effect of cell division on clock frequency and synchrony. A Diagram showing how during cell division, clock phase θ arrests, causing a cell to fall out of phase with its neighbour. On the right hand side a still from a simulation for T_M = 15 mins is shown, illustrating how this creates asynchrony of oscillations. The effect of cell division on anterior synchrony and frequency is shown in figures **B** and **C**, respectively, for $T_M = 15$ mins after 1000 mins. $N = 100$. **D** The effect on anterior synchrony when division is restricted to only occur posterior to $x = x_{div}$, for $T_M = 15$ mins. $N = 100$. **E** The effect on anterior synchrony after 1000 mins when T_M is varied. In each case, the total length of the cell cycle is maintained at a constant length, i.e. $T_M + T_G = 187.5$ mins. $N = 100$. To rule out trivial changes in synchrony and frequency, in all analysis here we restrict measurement to non-dividing cells, i.e. cells such that $\tau \in [0, T_G)$.

Zebrafish clock coupling and cell mixing confer robustness to changes in cell ingression

Across many of the morphogenetic scenarios tested here, we have observed a high degree of robustness in clock dynamics with respect to changing cell movements and processes. As shown above, some of this robustness can be attributed to the choice of tissue length L in simulations, and choices of the parameters v_a and v_p governing the advection velocity of cells. However we also expect that the rate of cell mixing (***Uriu et al.*** (***2010***, 2013, 2017)), and the strength of clock coupling, confer robustness too. Within the present model, the global rate of cell mixing is controlled by the parameter v_s and the clock coupling strength is denoted by κ. Notably, the values taken for these parameters have been experimentally measured in zebrafish (***Riedel-Kruse et al.*** (***2007***); ***Herrgen et al.*** (***2010***); ***Uriu et al.*** (***2017***, 2021)). Therefore we were interested in where these parameters lie within the space of parameter pairs that generate clock dynamics robust to changes in morphogenesis - do these results reflect something unique about the zebrafish parameters or can many (κ, v_s) pairs confer robustness?

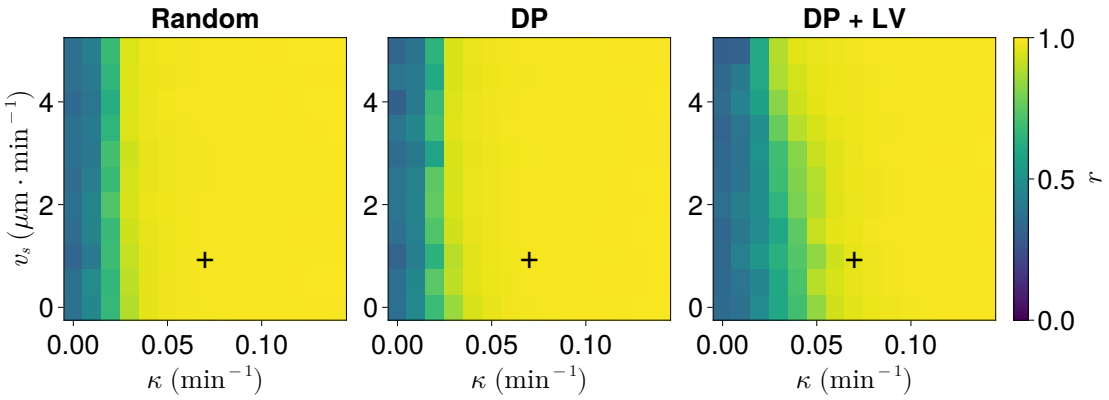


Figure 7. Synchrony for varying coupling strength κ and magnitude of intrinsic cell motion v_s. Anterior synchrony after 1000 mins, for varying maximum magnitude of intrinsic cell motion v_s and clock phase coupling strength κ, for three different scenarios of cell ingression. Each pixel corresponds to the median value of anterior synchrony across $N = 100$ simulations. A black + marks the experimental values for zebrafish, $\kappa = 0.07$ min^{-1}, $v_s = 1$ μm · min^{-1}, derived in ***Riedel-Kruse et al.*** (***2007***) and ***Uriu et al.*** (***2017***) respectively, that are used elsewhere in this paper. All other parameters are held constant at their normal values (see table 1).

To determine this, we take changing cell ingression as an example of varying morphogenesis and for each ingression scenario, systematically enumerate the parameters v_s and κ, studying the anterior synchrony after 1000 mins with N=100 replicates per parameter pair. This reveals a threshold curve of (κ, v_s) pairs below which the clock begins to exhibit asynchrony (figure 7). Comparing across the three scenarios of cell ingression we see that the position of this curve changes with the cell ingression scenario (figure 7), indicating that the space of (κ, v_s) pairs that achieve clock synchrony is constrained by the cell ingression mode present within the tissue. Notably the experimentally-derived parameter pair ($\kappa = 0.07$ min^{-1}, $v_s = 1$ μm · min^{-1}) lies very close to the threshold for the DP+LV case of cell ingression in the κ direction, suggesting that the robustness we have observed in response to changes in cell ingression is sensitive to the choice of κ. However in each case our experimental choices for κ and v_s lie

above the threshold, implying that at least part of the robustness we observed with respect to cell ingression is due to the experimental values for zebrafish being sufficient to confer this robustness.

Finally, we test if this robustness holds in the presence of a phase coupling delay between cells. We add to our simulations the time taken for information about the transcriptional state of a cell's neighbour to be transduced by cell-cell signalling and impact the cell's transcriptional profile. Such delays are important to consider as they can be significant for clock synchrony and frequency (***Lewis*** (***2003***); ***Morelli et al.*** (***2009***)), however they significantly increase the computational complexity of such a model. The value of the cell-cell phase coupling delay has been estimated to be 21 mins in zebrafish (***Herrgen et al.*** (***2010***)), and so to test the effect of delays we repeat the analysis above (figure 7) in the presence of a cell-cell coupling delay (τ) of 21 mins.

We observe that, for $\tau = 21$ mins, oscillations are asynchronous in all three scenarios of cell ingression with the experimentally measured values of κ and v_s, however for larger values of κ (and, to a lesser extent, v_s), we observe synchronous oscillations (figure 8B). Notably, we cannot achieve synchronous oscillations for any of the chosen parameter pairs for the DP+LV scenario of cell ingression (figure 8B), suggesting that ingression of progenitor cells is a far greater source of noise for the clock in the presence of coupling delay than in the instantaneous coupling case. To test this, we varied the length of the PSM L and either restricted the addition of cells to occur within a region of length $100 + R + r_T$ at the tissue posterior or allowed cell addition to occur at any point in the tissue posterior to $x_d = 100\ \mu\text{m}$. In either case the rate of cell addition is assumed to be constant and thus in the latter case we expect the longer tissue domain to

'dilute' the noise effect of cell ingression. We observe no difference between either scenario and find no value of L such that synchronous oscillations are recovered for $\tau = 21$ mins (figure 8 supplements 2 and 3.)

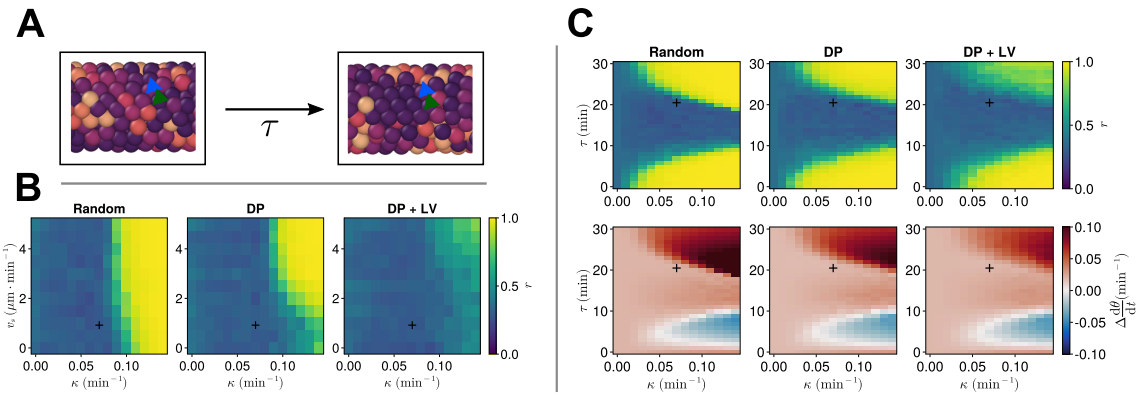


Figure 8. Clock dynamics in the presence of a coupling delay between cells. A Diagram illustrating phase coupling with delay. Cells couple their phase to the phase of their neighbours τ minutes ago. **B** Anterior synchrony in the presence of coupling delay ($\tau = 21$ mins) after 1000 mins, for varying maximum magnitude of intrinsic cell motion v_s and clock phase coupling strength κ, for three different scenarios of cell ingression. Each pixel corresponds to the median value of anterior synchrony across $N = 100$ simulations. A black + marks the experimental values for zebrafish, $\kappa = 0.07$ min^{-1}, $v_s = 1$ μm · min^{-1}, derived in ***Riedel-Kruse et al.*** (***2007***) and ***Uriu et al.*** (***2017***) respectively, that are used elsewhere in this paper. All other parameters are held constant at their normal values (see table 1). **C** Anterior synchrony (**top**) r after 1000 mins, for varying phase coupling delay τ and clock phase coupling strength κ, for three different scenarios of cell ingression. Each pixel corresponds to the median value of anterior synchrony across $N = 100$ simulations. **Bottom**: Difference from expected mean frequency ($\Delta \mathrm{d}\theta/\mathrm{d}t$) at the PSM anterior after 1000 mins, for three different scenarios of cell ingression. Each pixel corresponds to the mean value across $N = 100$ simulations. A black + marks the experimental values for zebrafish, $\kappa = 0.07$ min^{-1}, $\tau = 21$ mins, derived in ***Riedel-Kruse et al.*** (***2007***) and ***Herrgen et al.*** (***2010***) respectively. All other parameters are held constant at their normal values (see table 1).

To explore how the co-evolution of the clock coupling strength κ and the coupling delay τ might constrain the position of cell ingression, we systematically enumerated values of κ as before while co-varying τ, and analysed clock synchrony across the three different scenarios of cell ingression (figure 8C). In agreement with previous modelling work (***Morelli et al.*** (***2009***); ***Herrgen et al.*** (***2010***)) we see that synchrony depends on the value of τ in a repeating manner, with

values of τ closest to half of a clock cycle (~ 30 min) being more asynchronous (figure 8C). These results also predicted that frequency decreases with increasing τ within regions of stability (***Morelli et al.*** (***2009***); ***Herrgen et al.*** (***2010***)). We see evidence of this when we examine the collective frequency at the PSM anterior, where within each region of synchrony the frequency decreases as τ increases (figure 8C). We note that for small values of τ the DP+LV scenario of cell ingression is synchronous, however for larger values of τ oscillations are largely asynchronous, unlike the DP scenario of ingression (figure 8C).

We note that, despite our model failing to achieve synchronous oscillations for $\tau = 21$ mins, only relatively minor increases of τ are required to achieve synchrony in the Random and DP+LV scenarios of cell ingression (figure 8C). This closely resembles the analytical results of ***Herrgen et al.*** (***2010***) who found that the measured delay $\tau = 21$ mins lies towards the extreme end of values of τ for which stable oscillations can occur (***Herrgen et al.*** (***2010***)). Together these results suggest that the zebrafish segmentation clock might exist in a state of criticality, and close to a region of parameter space where oscillations are asynchronous. Overall our results suggest that coupling strength and coupling delay strongly constrain the evolvability of cell ingression, and that robustness to morphogenesis is conferred only for specific combinations of these two parameters.

Discussion

Here we investigated whether morphogenesis of the PSM and the segmentation clock exhibit developmental modularity, as this could explain the high degree of evolvability observed in vertebrate segment number (***Raff*** (***1996***)). We tested a broad range of PSM cellular behaviours and mechanisms that are associated with the elongation of the PSM across vertebrate species, to see if they had an

effect on the dynamics of the segmentation clock. We predict that the dynamics of the zebrafish segmentation clock, specifically synchrony and frequency, are generally robust to changes in these mechanisms such as cell ingression, motility, and (under certain conditions) division. As PSM morphogenesis is independent of clock function (***Schröter and Oates*** (***2010***); ***Forero et al.*** (***2018***)), our results suggest that the clock and morphogenesis of the PSM exhibit developmental modularity.

Our results suggest that a major determinant of this robustness is the clock coupling κ, which in zebrafish appears to be sufficiently strong to confer robustness to changes in cell ingression (figure 7). While synchrony of the clock shows a dependence on cell mixing rate (v_s) and tissue length (L), the dependence of synchrony on these parameters is weak compared to its dependence on coupling strength (figure 7, figure 4 figure supplement **??**), suggesting that this robustness may be more dependent on properties of the segmentation clock than those of the tissue. Due to a lack of quantitative understanding of PSM morphogenesis, it is difficult to predict whether the requirements for robustness imposed on cell mixing or tissue length and density constrain the space of possible morphogenetic processes that can be altered without perturbing the clock. However, our results highlight that these requirements are themselves constrained by the strength of clock phase-coupling, so it is equally possible that vertebrate segmentation clocks have strong phase-coupling generally, and these requirements on morphogenesis are sufficiently relaxed to not constrain the evolution of PSM elongation. Further work, investigating the evolution of phase-coupling strength in vertebrates, and developing accurate quantitative models of PSM morphogenesis, will be necessary to determine this.

It is also worth noting that while in the model formulation κ represents the strength of delta-notch signalling between cells, the experimental value for this parameter is measured at a tissue level, incorporating coupling effects from other pathways e.g. cell mixing (***Riedel-Kruse et al.*** (***2007***)). It is therefore possible that the value of κ measured by ***Riedel-Kruse et al.*** (***2007***) represents an over-estimate of the strength of delta-notch signalling and therefore our results present an over-estimate of robustness to morphogenesis. However our results appear less sensitive to cell mixing (v_s) than κ (figure 7), so we suggest that the contribution of cell mixing to the estimate of κ may be minor. Furthermore, as the precise value of r necessary for correct segment patterning is unknown, and as downstream processes in segmentation may act to correct any defects in clock patterning (***Naganathan et al.*** (***2022***)), somitogenesis may be more robust than we assume here. Therefore, we suggest that our findings are robust to any overestimates of κ by incorporation of cell mixing.

Our results when considering a phase coupling delay between cells (figure 8) suggest that evolution of cell ingression is constrained by both the clock coupling delay between cells, and vice versa. This suggests that the robustness of the segmentation clock to changes in morphogenesis is mediated by the combination of the cell-cell coupling delay and the coupling strength of the clock. While little is known about what determines the coupling strength of the clock, coupling delay is thought to be at least in part controlled by the number of intermediary components involved in signal transduction (***Lewis*** (***2003***); ***Yoshioka-Kobayashi et al.*** (***2020***)). Therefore the morphogenesis of the PSM may constrain the ability of the segmentation clock to undergo developmental system drift, at least within the components of the delta-notch pathway.

We note that our model fails to recover synchrony for the experimentally measured values of κ and τ, and thus may not accurately model the *in vivo* condition in some way. It is possible that the way in which we simulate cell ingression introduces too much noise into the clock, for instance the rate of progenitor cell ingression may be much higher than that *in vivo*, and indeed we note that elongation is driven *in vivo* by a combination of processes, rather than solely ingression, as we implicitly assume here. If this is true, this implies that clock imposes a constraint on the degree to which ingression is capable of driving PSM elongation. This possibility is worthy of further study, and is perhaps better suited to models where PSM elongation is not dependent on cell ingression, as we assume here.

Furthermore, as discussed above, we note that for relatively small increases in the value of τ, e.g. for $\tau = 25$ min, it is possible to achieve synchronous oscillations in both the random and DP scenarios of cell ingression, and partially synchronous oscillations in the DP+LV scenario of cell ingression (figure 8C). This could suggest that the biological parameters controlling delay and cell ingression exist in a state of criticality, i.e. lying close to a region of parameter space where dynamical behaviour abruptly changes or becomes chaotic. Such states can have important implications for evolvability of a biological system (***Kauffman*** (***1993***); ***Verd et al.*** (***2019***)), however it is not clear if such a state is functional in this case. Further study, perhaps with models unlike the mean-field approximation we assume here, may reveal what implication, if any, this has for the evolvability of PSM morphogenesis.

Our analysis is restricted to the modes of PSM morphogenesis that are known to occur in vertebrates, and we cannot rule out the possibility that the robust-

ness we observe merely reflects evolution of mechanisms of PSM elongation under selective pressure to minimise their effect on the clock, and that the clock may not be robust to all possible modes of PSM elongation. Nevertheless, we argue that our results demonstrate there exists a qualitatively distinct set of processes (cell ingression, division, motility, and compaction-extension) that underpin PSM elongation and do not affect clock dynamics. As vertebrate species appear to employ a combination of these processes to varying degrees (***Bénazéraf et al.*** (***2010***); ***Steventon et al.*** (***2016***); ***Mongera et al.*** (***2018***); ***Xiong et al.*** (***2020***); ***Thomson et al.*** (***2021***)), it is possible that evolution alters each of these processes independently to generate diversity in the elongation dynamics of the PSM (***Gomez et al.*** (***2008***); ***Steventon et al.*** (***2016***)). Therefore this relatively small set of processes may be sufficient to explain the high degree of evolvability in vertebrate segment number.

As clock dynamics appear to be robust to changes in morphogenesis, we suggest that the clock and morphogenesis of the PSM comprise two developmental modules within zebrafish, and that this is true more generally whenever the phase coupling of the clock is sufficiently strong, cell mixing sufficiently rapid, or the PSM sufficiently dense and long. We suggest that this may have contributed to the evolution of broad diversity in vertebrate segment number (***Richardson et al.*** (***1998***)), and be a present source of evolvability in vertebrates. We note that the evolution of modularity is an active area of research (***Wagner and Altenberg*** (***1996***); ***Wagner et al.*** (***2007***)), and suggest that studying the evolution of delta-notch signalling in vertebrates, and the evolutionary origin of the PSM and somitogenesis, may be an illuminating model paradigm for this field. We predict that the modularity of the clock and morphogenesis of the PSM permits a synergy between these two processes that heightens the evolvability of segment

number in vertebrates, and that this may be responsible for the high diversity we observe in this trait (***Richardson et al.*** (***1998***)).

Our results suggest that the zebrafish segmentation clock is robust to most changes in PSM morphogenesis, and that this robustness is determined by the strength of clock phase coupling and the phase coupling delay, as well as morphogenetic parameters such as the rate of cell mixing, and the length and density of the PSM. This suggests that, at least in the regions of parameter space where these conditions for robustness hold, PSM morphogenesis is not constrained by the clock and free to evolve, and the clock and morphogenesis of the PSM exhibit developmental modularity. The precise topology of this space remains unclear, and poses an intriguing question for further study. We suggest that this modularity may be responsible for the broad diversity of segment number that we observe in the vertebrates (***Richardson et al.*** (***1998***)), and indeed is a present source of evolvability in the clade. We note that the evolution of modularity is an active area of research (***Wagner and Altenberg*** (***1996***); ***Wagner et al.*** (***2007***)), and while several well-characterised examples of modularity exist (***Monteiro et al.*** (***2003***); ***Mallarino et al.*** (***2011***); ***Verd et al.*** (***2019***)), there are few examples to our knowledge where the existence of such modularity relies on a handful of experimentally tractable parameters. We therefore suggest that studying the evolution of delta-notch signalling in vertebrates, and the evolutionary origin of the PSM and somitogenesis, may be an illuminating model paradigm for this field.

Methods

Computational model of the segmentation clock and PSM

To simulate the movements of cells within the PSM and the dynamics of the segmentation clock, we used the model of ***Uriu et al.*** (***2021***), who sought to explain defective somite patterning after loss of Notch-Delta signalling. Here an abstract phase oscillator Kuramoto model is used to describe intracellular clock oscillations and cell-cell phase coupling. Cells are confined within a three-dimensional tissue domain whose geometry approximates that of the PSM.

This model is suitable for our purposes because it provides a three-dimensional model of cell movements within the PSM, allowing us to alter morphogenesis in order to test hypotheses. Additionally, as the segmentation clock here is described as an abstract phase oscillator, the model has broad applicability to different vertebrate species which are known to differ in the structure and regulation of their segmentation clock gene regulatory networks (***Krol et al.*** (***2011***)). Finally, the validity of the model has been shown in several ways via comparison with data for zebrafish (***Webb et al.*** (***2016***); ***Liao et al.*** (***2016***); ***Uriu et al.*** (***2021***)), and uses parameters experimentally inferred from zebrafish embryos (***Riedel-Kruse et al.*** (***2007***); ***Soroldoni et al.*** (***2014***); ***Uriu et al.*** (***2021***)). This gives an accurate read-out of the effect of changing morphogenesis on the segmentation clock of a vertebrate.

The model is described elsewhere (***Uriu et al.*** (***2021***)), but for completeness we include an account of it here. For clarity, vector variables are shown in **bold**, and scalar variables in normal font.

Tissue geometry and frame of reference

The model assumes that the posterior tip of the PSM is held in an inertial frame of reference, and models the posterior-ward elongation of the PSM by movement of cells towards the tissue anterior.

Cells are confined within a horseshoe-shaped domain resembling the PSM. This domain is comprised of two cylindrical subdomains and one half-toroid subdomain, each representing the lateral sides of the PSM and tailbud, respectively. A schematic can be seen in figure 9. Cells are free to move between each subdomain and leave the tissue at the anterior at $x = x_a$.

The centre of the toroid subdomain is denoted by the coordinate vector $(X_c, Y_c, Z_c)^T$, and the anterior of the PSM is demarcated by the x-coordinate $x = x_a$ (see figure 9). The major radius of the torus is denoted by R and the minor radius (and the radius of the two cylindrical subdomains is denoted by r_T (see figure 9). The resulting length of the PSM in the x direction, L, is therefore given by $L = r_T + R + X_c - x_a$.

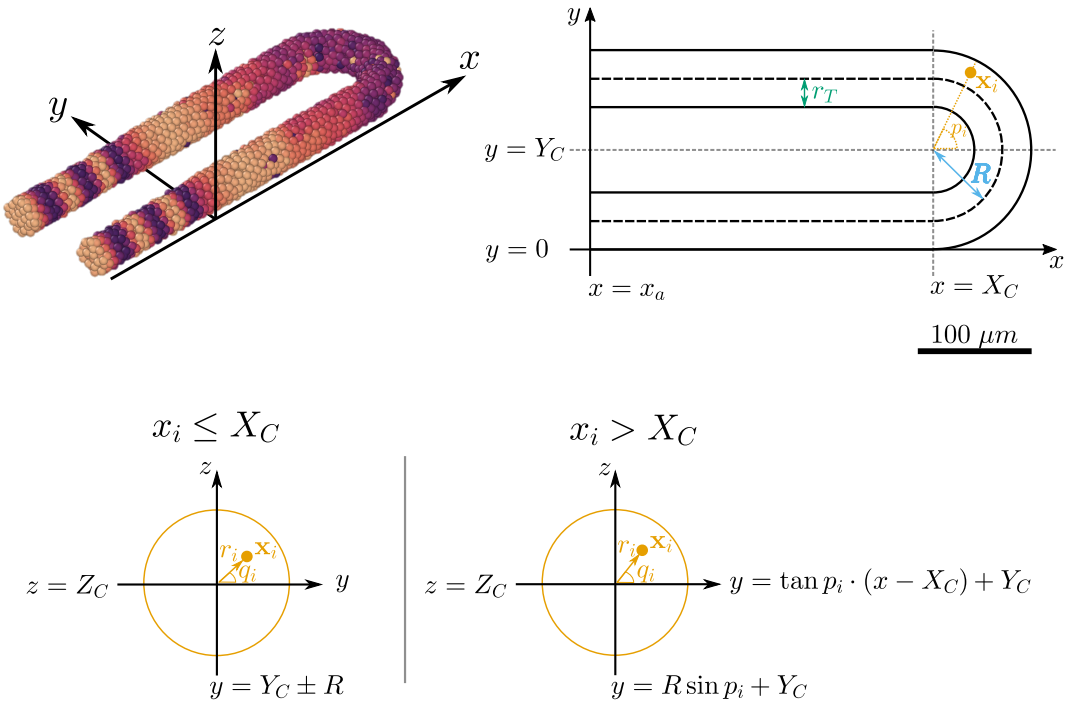


Figure 9. Geometry of the PSM assumed in the present model. Top left: major axes (x, y, z) used in the model. x corresponds to the anterior-posterior axis of the embryo and increases towards the tissue posterior, y corresponds to the left-right axis and increases to the right hand side of the tissue, and z corresponds to the dorsal-ventral axis of the embryo. **Top right:** schematic of the PSM in the xy plane. The PSM is comprised of two cylinders, centred at $y = r_T$ and $y = 2R + r_T$, respectively, with radius r_T. The 'tailbud' is represented as a half-torus subdomain centred at $\mathbf{x} = (X_c, Y_c, Z_c)^T$, with minor radius r_T and major radius R. **Bottom:** cross-sections of the tissue showing how a point $\mathbf{x}_i$ is assigned the polar coordinates r_i and q_i, in both the PSM cylinders ($x_i \leq X_c$) and the half-toroid tailbud ($x_i > X_c$). Adapted from ***Uriu et al.*** (***2021***).

Cell movement

To approximate cell mixing and the advection of cells out of the PSM towards the anterior (in the chosen tailbud-inertial frame of reference), cell movement in the model is described by an equation of motion that incorporates cell advection, cell volume exclusion, random cell motion, and a boundary force confining cells within the tissue domain (equation (1)).

Specifically, each cell is assigned a position vector $\mathbf{x}_i \in \mathbb{R}^3$, where $\mathbf{x}_i = (x_i, y_i, z_i)^T$. This position $\mathbf{x}_i$ corresponds to the cell centre. Cells are assumed to be spheres with diameter d_c. The equation of motion for cell i is

$$\frac{\mathrm{d}\mathbf{x}_i}{\mathrm{d}t} = \mathbf{v}_d(x_i) + v_0(x_i)\mathbf{n}_i(t) + \sum_{j=1, j\neq i}^{N} \mathbf{F}(\mathbf{x}_i, \mathbf{x}_j) + \mathbf{F}_b(\mathbf{x}_i), \tag{1}$$

where $\mathbf{v}_d(x_i)$ represents the advection of the cell towards the PSM anterior, $v_0(x_i)\mathbf{n}_i(t)$ the intrinsic random motion of the cell, $\mathbf{F}(\mathbf{x}_i, \mathbf{x}_j)$ the repulsion force between cell i and cell j, and $\mathbf{F}_b(\mathbf{x}_i)$ the boundary force confining the cell within the tissue.

The advection velocity $\mathbf{v}_d(x_i)$ is given by

$$\mathbf{v}_d(x_i) = (-v_d(x_i), 0, 0)^T, \tag{2}$$

where the function $v_d(x_i)$ describes the local strain rate along the AP axis:

$$v_d(x_i) = \begin{cases} -\frac{v_a - v_p(1-x_q)}{x_q} \cdot \frac{x_i - x_a}{L} + v_a & \frac{x_i - x_a}{L} \leq x_q, \\ -\frac{v_p(x_i - x_a)}{L} + v_p & \frac{x_i - x_a}{L} > x_q, \end{cases} \tag{3}$$

The model assumes that the random cell motility observed within the PSM is intrinsic, i.e. cells would exhibit random motility in isolation from the tissue. This is modelled using the direction vector $\mathbf{n}_i(t)$, which evolves according a random walk on the surface of the unit sphere specified by the following differential equation:

$$\frac{\mathrm{d}\mathbf{n}_i}{\mathrm{d}t} = \sqrt{2D}\xi_{\phi_i}(t)\mathbf{m}_x + \sqrt{2D}\xi_{\varphi_i}(t)\mathbf{m}_y - 2D\mathbf{n}_i(t), \tag{4}$$

where the orthogonal vectors (for $\mathbf{e}_z = (0,0,1)^T$)

$$\mathbf{m}_x = \frac{\mathbf{n}_i(t) \times \mathbf{e}_z}{|\mathbf{n}_i(t) \times \mathbf{e}_z|},\ \mathbf{m}_y = \frac{\mathbf{m}_x \times \mathbf{n}_i(t)}{|\mathbf{m}_x \times \mathbf{n}_i(t)|}, \tag{5}$$

together define a plane tangent to the unit sphere, and $\xi_{\phi_i}(t)$, $\xi_{\varphi_i}(t)$ are white Gaussian noise terms that satisfy, for any t, t', $\langle \xi_{\phi_i}(t) \rangle = \langle \xi_{\varphi_i}(t) \rangle = \langle \xi_{\phi_i}(t) \xi_{\varphi_j}(t) \rangle = 0$, and $\langle \xi_{\phi_i}(t) \xi_{\phi_j}(t) \rangle = \langle \xi_{\varphi_i}(t) \xi_{\varphi_j}(t) \rangle = \delta_{ij} \delta(t - t')$. A derivation for equation (4) is given by ***Uriu et al.*** (***2021***).

To model the increase in random cell motion towards the PSM posterior, the magnitude of intrinsic random cell movement, $v_0(x)$, increases towards the posterior according to equation (6):

$$v_0(x_i) = \frac{v_s}{1 + \left(\frac{1 - \frac{x_i - x_a}{L}}{X_v}\right)^h}, \tag{6}$$

The parameters v_s, X_v, and h determine the maximum magnitude of intrinsic cell motion, profile inflexion point, and steepness of the profile, respectively. By varying these parameters one can create a variety of motility profiles with different shapes and amplitudes (see figure 3B). Here the x-position of cell i is normalised according to the length of the PSM L, and so the profile of cell motility scales with the length of the PSM. This is a valid assumption as the length of the PSM is thought to be partially specified by FGF signalling (***Simsek and Özbudak*** (***2018***)), and the random motility of PSM cells is also under the control of FGF (***Bénazéraf et al.*** (***2010***)).

Cell volume exclusion is encoded by repulsion of two cells if their centres are within a cell diameter d_c of one another, according to the cell-cell repulsion force $\mathbf{F}(\mathbf{x}_j - \mathbf{x}_i)$:

$$\mathbf{F}(\mathbf{x}_j - \mathbf{x}_i) = F(\mathbf{x}_j - \mathbf{x}_i)\frac{\mathbf{x}_j - \mathbf{x}_i}{|\mathbf{x}_j - \mathbf{x}_i|}, \tag{7}$$

where the magnitude of the cell-cell repulsion force $F(\mathbf{x}_j - \mathbf{x}_i)$ is given by

$$F(\mathbf{x}_j - \mathbf{x}_i) = \begin{cases} \mu\left(\frac{|\mathbf{x}_j - \mathbf{x}_i|}{d_c} - 1\right) & |\mathbf{x}_j - \mathbf{x}_i| \leq d_c, \\ 0, & |\mathbf{x}_j - \mathbf{x}_i| > d_c, \end{cases} \tag{8}$$

and the parameter μ controls the strength of cell-cell repulsion throughout the tissue. For computational simplicity, cell-cell adhesion is not encoded within this model.

The form of the boundary force $\mathbf{F}_b(\mathbf{x}_i)$ varies depending on whether the cell i is in one of the lateral PSM cylindrical subdomains or in the half-toroid tailbud (figure 9). If cell i lies within the cylinders, one can use the cylindrical coordinates for cell i, r_i and q_i (see figure 9), to naturally define the boundary force:

$$\mathbf{F}_b(\mathbf{x}_i) = \begin{pmatrix} 0 \\ -\mu_b e^{-\frac{\delta y}{r_b}} \cos q_i \\ -\mu_b e^{-\frac{\delta z}{r_b}} \sin q_i \end{pmatrix}, \tag{9}$$

where

$$\delta y = |(r_T - r_i)\cos q_i|, \tag{10}$$

$$\delta z = |(r_T - r_i)\sin q_i|. \tag{11}$$

The magnitude of the boundary force is denoted by μ_b and the parameter r_b determines the length-scale of the boundary force. In the half-toroid tailbud the position $\mathbf{x}_i$ of each cell i can be described according to the coordinates r_i, p_i,

and q_i:

$$\mathbf{x}_i = \begin{pmatrix} X_c + R\cos p_i + r_i \cos p_i \cos q_i \\ Y_c + R\sin p_i + r_i \sin p_i \cos q_i \\ Z_c + r_i \sin q_i \end{pmatrix}. \tag{12}$$

One can then naturally define

$$\delta x = |(r_T - r_i)\cos p_i \cos q_i|, \tag{13}$$

$$\delta y = |(r_T - r_i)\sin p_i \cos q_i|, \tag{14}$$

$$\delta z = |(r_T - r_i)\sin q_i|, \tag{15}$$

describing the distance of cell i from the tissue surface in the x, y, and z directions, respectively. The boundary force then becomes

$$\mathbf{F}_b(\mathbf{x}_i) = \begin{pmatrix} -\mu_b e^{-\frac{\delta x}{r_b}} \cos p_i \cos q_i \\ -\mu_b e^{-\frac{\delta y}{r_b}} \sin p_i \cos q_i \\ -\mu_b e^{-\frac{\delta z}{r_b}} \sin q_i \end{pmatrix}. \tag{16}$$

Note that in both definitions of $\mathbf{F}_b(\mathbf{x}_i)$, the tissue subdomains are open-ended, i.e. cells are free to move between subdomains, and out of the PSM at the anterior.

For further details we direct the reader to consult ***Uriu et al.*** (***2021***).

Cell addition

As cells are removed from the tissue by advection, new cells must be added to replenish the PSM. Loss of cells is monitored by comparing the total density of cells in each of the subdomains with a target density, ρ_0. If the density in any of

these subdomains falls below ρ_0, a new cell is added to that subdomain. If the density of more than one of the subdomains falls below ρ_0 then a cell is only added to the subdomain with the lowest density. The position and phase of the nascent cell both vary across different simulations discussed here. Below is a description of how these quantities vary.

In simulations where cell addition is 'Random', cells are added with a random position within each subdomain, with a random phase, i.e. if a cell is added within one of the cylindrical subdomains, it is given a random $x_i \in [x_a + x_d, X_c]$, a random $r_i \in [0, r_T]$, and $q_i \in [0, 2\pi)$. Alternatively, if the cell is added within the half-toroid subdomain, it is given a random $p_i \in [-\pi/2, \pi/2]$, and random r_i and q_i, defined as above. The cell is also assigned a random phase $\theta_i \in [0, 2\pi)$. This method of cell addition is unbiased in terms of its effect on the model but is a biologically implausible scenario of cell ingression, so we use it here as a negative control.

In simulations where we seek to more accurately model cell ingression, cells are added on the cell surface, i.e. $r_i = r$. Additionally, the surface on which cells can be added is restricted. For the half-toroid subdomain, we restrict cell addition to the dorsal quarter of the subdomain, i.e. $q_i \in [\pi/4, 3\pi/4]$, $p_i \in [-\pi/2, \pi/2]$. For the cylindrical subdomains we restrict addition to the bottom quarter of the surface, i.e. $q_i \in [5\pi/4, 7\pi/4]$. To more accurately model conditions of cell ingression combinations of these three modes of cell addition are employed (figure 2).

In newly added cells the direction vector **n** is initialised with a random vector of unit length. For simulations where cell division is present, cells are added with a random cell cycle phase $\tau \in [0, T_G + T_M)$.

Coupled oscillator model of the segmentation clock

Here only the phase of each cell's segmentation clock is considered, and oscillation amplitudes are neglected. This simplification is appropriate when oscillator coupling is weak (***Kuramoto*** (***1984***)). Here oscillators have a spatially-dependent intrinsic frequency $\omega(x_i)$, and couple to adjacent cells (defined as those within d_c, a cell's diameter) with a magnitude controlled by the scalar value κ:

$$\frac{\mathrm{d}\theta_i}{\mathrm{d}t} = \omega(x_i) + \frac{\kappa}{N_i(t)} \sum_{|\mathbf{x}_j - \mathbf{x}_i| \leq d_c} \sin(\theta_j - \theta_i) + \sqrt{2D_\theta}\xi_{\theta_i}(t). \tag{17}$$

Oscillations are noisy, with phase noise intensity D_θ. The term $\xi_{\theta_i}(t)$ describes gaussian noise as before. $N_i(t)$ is the number of cells within a radius of length d_c of cell i at time t, i.e. $N_i(t) = |\{j \mid |\mathbf{x}_j(t) - \mathbf{x}_i(t)| \leq d_c\}|$.

The intrinsic frequency $\omega(x_i)$ decreases toward the PSM anterior, creating a frequency gradient and travelling phase waves that move toward the anterior (***Soroldoni et al.*** (***2014***)). The intrinsic frequency is defined as:

$$\omega(x_i) = \omega_0\left(\sigma + (1-\sigma)\cdot\frac{1 - e^{-\frac{k(x-x_a)}{L}}}{1 - e^{-k}}\right), \tag{18}$$

where $\omega_0 > 0$ corresponds to the frequency of cells at the posterior of the tissue, $\sigma > 0$ the fold-change in frequency at the anterior compared to the posterior, and $k > 0$ controls the shape of the gradient. For simplicity we assume that $\omega(x_i)$ is constant over time, as cells at the PSM posterior are thought to maintain a constant frequency over time (***Giudicelli et al.*** (***2007***); ***Webb et al.*** (***2016***)).

To simulate the effect of the wavefront, where the clock phase is interpreted by cells to pre-pattern somite boundaries (***Cooke and Zeeman*** (***1976***)), beyond $x =$

x_a oscillations arrest, forming stable patterns of phase (figure 2A), i.e. $\mathrm{d}\theta_i/\mathrm{d}t = 0$ if $x_i < x_a$.

Modelling the segmentation clock with a phase coupling delay

For simulations where we seek to model the effects of a coupling delay between cells, the phase of cells evolves according to

$$\frac{\mathrm{d}\theta_i}{\mathrm{d}t} = \omega(x_i) + \frac{\kappa}{N_i(t)} \sum_{|\mathbf{x}_j - \mathbf{x}_i| \leq d_c} \sin(\theta_j(t-\tau) - \theta_i(t)) + \sqrt{2D_\theta}\xi_{\theta_i}(t), \tag{19}$$

where τ is a constant representing the temporal delay in exchange of phase information between cells, after ***Morelli et al.*** (***2009***). In such simulations we initialise the phase of each newly added cell with the initial conditions $\theta_i(t \leq t_{add}) = 0$, where t_{add} is the time at which cell i was added to the tissue.

Cell division and segmentation clock arrest

During M phase of the cell cycle, the segmentation clock arrests (***Horikawa et al.*** (***2006***); ***Delaune et al.*** (***2012***)). To simulate this, we assign each cell i a cell cycle phase $\tau \in [0, T_G + T_M)$, where T_G and T_M are the durations of G1-G2 and M phases in minutes, respectively, and evolve τ according to

$$\tau = (t - t_{addition}) \mod (T_G + T_M), \tag{20}$$

where $t_{addition}$ denotes the time at which the cell was added.

To simulate the arrest of segmentation clock gene expression during M phase of the clock, we evolve θ_i according to

$$\frac{\mathrm{d}\theta_i}{\mathrm{d}t} = H(T_G - \tau_i) \cdot \left(\omega(x_i) + \frac{\kappa}{N_i(t)} \sum_{|\mathbf{x}_j - \mathbf{x}_i| \leq d_c} \sin(\theta_j - \theta_i) + \sqrt{2D_\theta}\xi_{\theta_i}(t) \right), \tag{21}$$

where $H(x)$ is the Heaviside step function:

$$H(x) = \begin{cases} 1 & x > 0, \\ 0 & x \leq 0. \end{cases} \tag{22}$$

For simulations where we forbid cells from coupling their phase to neighbours undergoing mitosis, the evolution of θ_i is given by

$$\frac{\mathrm{d}\theta_i}{\mathrm{d}t} = H(T_G - \tau_i) \cdot \left(\omega(x_i) + \frac{\kappa}{N_i(t)} \sum_{|\mathbf{x}_j - \mathbf{x}_i| \leq d_c,\ \tau_j < T_G} \sin(\theta_j - \theta_i) + \sqrt{2D_\theta}\xi_{\theta_i}(t) \right). \tag{23}$$

If $\tau_i \geq T_G + T_M$ then we consider the cell cycle complete and a new cell is added at a random position of length d_{new} away from cell i, i.e.

$$\mathbf{x}_{N_{total}+1} = \mathbf{x}_i + d_{new} \begin{pmatrix} \cos\varphi \sin\phi \\ \sin\varphi \sin\phi \\ \cos\phi \end{pmatrix},$$

for uniformly distributed random $\phi \in [0, \pi]$, $\varphi \in [0, 2\pi)$ and where N_{total} is the total number of cells within the tissue. Other variables are inherited from cell i, i.e. $n_{N_{total}+1} = n_i$ and $\theta_{N_{total}+1} = \theta_i$. The choice of d_{new} is somewhat arbitrary, however as this distance decreases the magnitude of the cell-cell repulsion force (equation (8)) grows very large, and one must choose a d_{new} sufficiently small to be biologically accurate while large enough to avoid implausible cell velocities due to a very large cell-cell repulsion force. We find that $d_{new} = d_c/10$ is a suitable compromise between these two scenarios.

We are not aware of any direct measurements of T_G within the zebrafish PSM, and so have calculated a value for this parameter based on some assumptions. M phase within the PSM lasts approximately 15 mins, and approximately 8% of cells in the PSM at a single time point are in M phase (***Horikawa et al.*** (***2006***); ***Kanki and Ho*** (***1997***)). Cells in M phase appear to be distributed homogenously in space (***Kanki and Ho*** (***1997***)), and therefore cell cycles in the PSM may be independent, in which case the total duration $T_G + T_M$ is approximately equal to $15/0.08 = 187.5$ mins, and $T_G = 172.5$ mins. Unless otherwise stated, we fix $T_G + T_M = 187.5$ minutes.

Modelling compaction-extension

To model the compaction-extension of the PSM (***Thomson et al.*** (***2021***)), we shrink the PSM radius r_T, and length L, over time, concurrent with an increase in density ρ_0 and a step-wise decrease in cell diameter d_c (see figure 5). This is derived from the observations of ***Thomson et al.*** (***2021***) who observed changes in the PSM height, width, length, density, and cell diameter, during the latter stages of zebrafish somitogenesis. We sought to model this process using the rates of change in these parameters described by ***Thomson et al.*** (***2021***), however as their published rates are in units of change per somite stage (ss), and the rate of somitogenesis is nonlinear in zebrafish (***Schröter et al.*** (***2008***)), we needed an interpolation of somite number over time in order to describe the according change in rate over time.

To this, we extracted the data for somite number over time from ***Schröter et al.*** (***2008***) using the web app WebPlotDigitizer (***Rohatgi*** (***2022***)), and fitted a function

of the form

$$s(t) = 6 + \frac{t}{24.7} - ae^{b(t-300)}, \tag{24}$$

for unknown parameters a, b, using the Julia package LsqFit.jl. We obtained values of $a = 0.5001$ and $b = 0.0049$ using this process. Using this interpolation of somite number over time, we then could change the tissue radius, length, cell diameter, and density, using the values published in ***Thomson et al.*** (***2021***). As this only occurs in the latter stages of somitogenesis (measured only from 16ss onwards), we hold these variables constant before the time at which 16 somites are formed in zebrafish.

The equations describing change in tissue length (L), wavefront position (x_a), and anterior limit of cell addition therefore become:

$$L(t) = \begin{cases} L_0 & s(t) < 16, \\ L_0 - m_a(s(t) - 16) & s(t) \geq 16, \end{cases} \tag{25}$$

$$x_a(t) = \begin{cases} x_{a0} & s(t) < 16, \\ x_{a0} + m_a(s(t) - 16) & s(t) \geq 16, \end{cases} \tag{26}$$

$$x_d(t) = x_a(t) + 100, \tag{27}$$

where $m_a > 0$ denotes the rate at which the PSM shrinks in length, and $L_0 > 0$ and $x_{a0} > 0$ the initial values for L and x_a before shrinking, respectively. The equation describing the change in radius r_T is

$$r_T(t) = \begin{cases} r_0 & s(t) < 16, \\ r_0 - m_r(s(t) - 16) & s(t) \geq 16, \end{cases} \tag{28}$$

where $m_r > 0$ denotes the rate at which the PSM shrinks in radius, and $r_0 > 0$ the initial value for r_T before shrinking. In their work ***Thomson et al.*** (***2021***) reported that the PSM shrinks in height more rapidly at the PSM posterior than at the anterior, however the rate of shrinkage of the PSM in width (along the medio-lateral axis) was the same across the anterior and posterior halves of the PSM (***Thomson et al.*** (***2021***)). As our model assumes the tissue is as wide as it is tall, we take only the rate of shrinkage in the medio-lateral axis in our value for m_r, and calculate m_r by taking the mean of the two (similar) rates published by ***Thomson et al.*** (***2021***) for the anterior and posterior halves of the PSM.

The functions describing change in tissue density ρ and cell diameter d_c are similarly

$$\rho(t) = \begin{cases} \rho_{t0} & s(t) < 16, \\ \rho_{t0} + m_\rho(s(t) - 16) & s(t) \geq 16, \end{cases} \tag{29}$$

and

$$d_c(t) = \begin{cases} d_{c0} & s(t) < 16, \\ d_{c0} - m_{d_c}(s(t) - 16) & s(t) \geq 16, \end{cases} \tag{30}$$

where $m_\rho > 0$ and $m_{d_c} > 0$ represent the rates of change in density and cell diameter, respectively, and $\rho_{t0} > 0$ and $d_{c0} > 0$ the initial values for density and cell diameter, respectively, before shrinking. In their work ***Thomson et al.*** (***2021***) only report the cell diameter for 16ss and 26ss zebrafish embryos, but we choose to

shrink the diameter of cells at the rate they describe, measured between these two timepoints. Our results are insensitive to choice of a function for $d_c(t)$ where cells continue shrinking in diameter at the same rate past 26ss, or not (see figure 5 figure supplement 1).

Initial conditions

To initialise the simulation, N cells are generated with random positions

$$\mathbf{x}_i = \begin{pmatrix} x_i \\ r_T + r_i \cos q_i \\ Z_c + r_i \sin q_i \end{pmatrix},$$

and a further N with positions

$$\mathbf{x}_i = \begin{pmatrix} x_i \\ r_T + 2R + r_i \cos q_i \\ Z_c + r_i \sin q_i \end{pmatrix},$$

and then a further N_{TB} with positions

$$\mathbf{x}_i = \begin{pmatrix} X_c + R\cos p_i + r_i \cos p_i \cos q_i \\ Y_c + R\sin p_i + r_i \sin p_i \cos q_i \\ Z_c + r_i \sin q_i \end{pmatrix},$$

for $x_i \in [x_a, X_c]$, $p_i \in [-\pi/2, \pi/2]$, $q_i \in [0, 2\pi]$, and $r_i \in [0, r_T]$, where $N = \lfloor \rho_0 \pi r_T^2 X_c \rfloor$ and $N_{TB} = \lfloor \rho_0 \pi^2 r_T^2 R \rfloor$. As this generates more cells near the tissue mid-line than at the periphery, in order to homogenise local density the tissue is 'relaxed' for 10 mins, whereby $\mathbf{x}_i$ evolves according to

$$\frac{\mathrm{d}\mathbf{x}_i}{\mathrm{d}t} = v_0(x_i)\mathbf{n}_i(t) + \sum_{j=1, j\neq i}^{N} \mathbf{F}(\mathbf{x}_i, \mathbf{x}_j) + \mathbf{F}_b(\mathbf{x}_i), \tag{31}$$

and $\mathbf{n}_i(t)$ evolves according to equation (4). This evenly distributes cells throughout the tissue.

When cells are initialised prior to relaxation, $n_i(t)$ is assigned for each cell at random, i.e. $\mathbf{n}_i(0) = (\cos\varphi_i \sin\phi_i, \sin\varphi_i \sin\phi_i, \cos\phi_i)^T$ for random $\varphi_i \in [0, 2\pi)$, $\phi_i \in [0, \pi]$. The segmentation clock phase θ_i is also assigned as a constant value $\theta_i(0) = 3\pi/2$. For simulations with a coupling delay, we initialise clock phase with $\theta_i(t \leq 0) = 3\pi/2$.

Parameter values

Parameter values, and their source, are shown in Table 1.

Symbol	Value	Units	Source(s)	Symbol	Value	Units	Source(s)
X_c	300	μm	*Uriu et al.* (*2021*)	L	$X_c - x_a + R + r_T$	μm	*Uriu et al.* (*2021*)
Y_c	85	μm	*Uriu et al.* (*2021*)	x_a	0	μm	*Uriu et al.* (*2021*)
Z_c	25	μm	*Uriu et al.* (*2021*)	x_d	100	μm	*Uriu et al.* (*2021*)
R	60	μm	*Uriu et al.* (*2021*)	x_q	0.7	-	*Uriu et al.* (*2021*)
r_T	25	μm	*Uriu et al.* (*2021*)	d_c	11	μm	*Uriu et al.* (*2021*)
v_a	1.67	μm·min^{-1}	*Uriu et al.* (*2021*)	D	0.026	min^{-1}	*Uriu et al.* (*2021*)
v_p	3	μm·min^{-1}	*Uriu et al.* (*2021*)	μ	8.71	μm·min^{-1}	*Uriu et al.* (*2021*)
X_v	0.4	-	*Uriu et al.* (*2021*)	μ_b	20	μm·min^{-1}	*Uriu et al.* (*2021*)
h	3	-	*Uriu et al.* (*2021*)	r_b	1	μm	*Uriu et al.* (*2021*)
v_s	1	μm·min^{-1}	*Uriu et al.* (*2021*)	ρ_0	0.0015	μm^{-3}	*Uriu et al.* (*2021*)
L_0	325	μm	*Thomson et al.* (*2021*)	T_G	187.5 - 15	min	*Kanki and Ho* (*1997*), *Horikawa et al.* (*2006*)
m_a	14.1	μm ·somite^{-1}	*Thomson et al.* (*2021*)	T_M	15	min	*Horikawa et al.* (*2006*)
r_0	27.6	μm	*Thomson et al.* (*2021*)	κ	0.07	min^{-1}	*Riedel-Kruse et al.* (*2007*)
m_r	1.1625	μm ·somite^{-1}	*Thomson et al.* (*2021*)	ω_0	0.2094	min^{-1}	*Soroldoni et al.* (*2014*), *Uriu et al.* (*2021*)
d_{c0}	9.2	μm	*Thomson et al.* (*2021*)	σ	0.66	-	*Uriu et al.* (*2021*)
$m_{d_{c0}}$	0.2	μm ·somite^{-1}	*Thomson et al.* (*2021*)	k	3.07	-	*Uriu et al.* (*2021*)
ρ_{r0}	0.002123	μm^{-3}	*Thomson et al.* (*2021*)	$D\theta$	0.0013	min^{-1}	*Riedel-Kruse et al.* (*2007*)
m_ρ	0.000121	μm ·somite^{-1}	*Thomson et al.* (*2021*)	τ	21	min	*Herrgen et al.* (*2010*)

Table 1. Parameter values used for simulations, unless otherwise stated.

Implementation

Equations for cell position (Equation 1) and phase (Equations **??**) are solved from the initial conditions outlined above across a 1000 min time-span using an forward Euler scheme in Julia v1.8.2, using a time step dt = 0.01 mins. Code for simulation and analysis can be found at https://github.com/jewh/ModularityPSMClock.

Analysis and metrics

Here we measure clock dynamics using two metrics. The first is the synchrony of the segmentation clock across cells, denoted r, calculated by

$$re^{i\psi} = \frac{1}{N}\sum_{j=1}^{N} e^{i\theta_j}, \tag{32}$$

which is sometimes referred to as the Kuramoto phase order parameter. When $r \approx 0$, oscillations are asynchronous, and when $r \approx 1$, oscillations are synchronous.

The second metric is the average (mean) frequency of cells in a given domain, where the frequency $\mathrm{d}\theta/\mathrm{d}t$ is simply the value of equation 17. In cases where we are interested in changes in frequency upon varying model parameters, we find it convenient to define

$$\Delta\frac{\mathrm{d}\theta}{\mathrm{d}t} = \frac{\mathrm{d}\theta}{\mathrm{d}t} - \frac{1}{d_c}\int_{x_a}^{x_a+d_c} \omega(x)\mathrm{d}x, \tag{33}$$

in order to highlight changes in frequency.

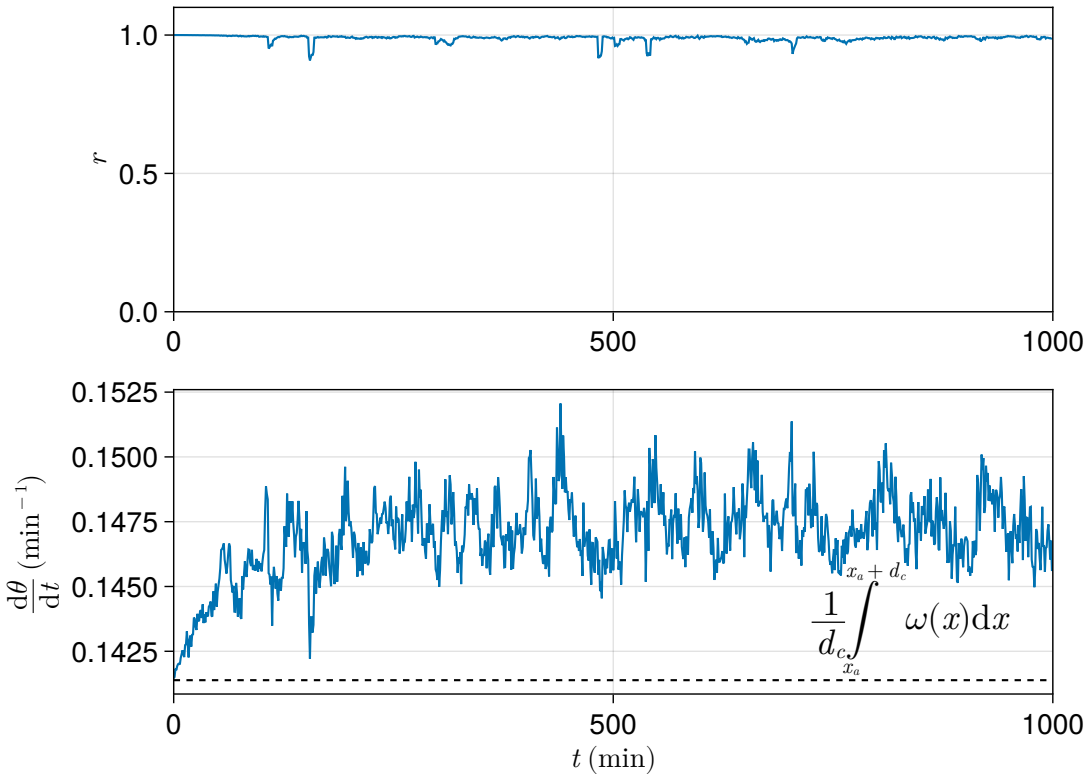


Figure 10. Exemplar dynamics of synchrony and mean anterior frequency. Top Trace of synchrony r for a d_c-wide domain of cells at the left-hand anterior edge of the PSM over 1000 mins. Data drawn from a single simulation with random cell addition, using parameters as per ***Uriu et al.*** (***2021***). **Bottom** Trace of mean frequency $\mathrm{d}\theta/\mathrm{d}t$ for a d_c-wide domain of cells at the left-hand anterior edge of the PSM over 1000 mins. Data drawn from the same simulation as above. Plotted with a black dashed line is the average intrinsic frequency $\omega(x)$ across the domain, calculated using the formula shown.

In order to analyse the output of a simulation in terms of one value, for many analyses we calculate r, $\mathrm{d}\theta/\mathrm{d}t$, or $\Delta\mathrm{d}\theta/\mathrm{d}t$, for the cells within one cell diameter of the anterior (i.e. $x_a \leq x_i \leq x_a + d_c$). We do this because at the wavefront oscillations arrest and pre-pattern somites, and therefore the frequency and synchrony of oscillations immediately adjacent to the PSM anterior will determine the somite length and precision of pre-patterning. We thus regard the values of r, $\mathrm{d}\theta/\mathrm{d}t$, and $\Delta\mathrm{d}\theta/\mathrm{d}t$ as proxies for phenotype exhibited by the simula-

tion. As discussed above, unless stated otherwise we measure clock dynamics after 1000 mins, as we find this sufficient to ensure that the dynamics of the clock have reached a steady state that we deem would be reached in vivo (see figure 10).

Acknowledgments

We thank Koichiro Uriu for sharing his code. We thank Usha Kadiyala for comments regarding estimates for T_G. We thank Ben Steventon for feedback on the project. The authors would like to acknowledge the use of the University of Oxford Advanced Research Computing (ARC) facility (http://dx.doi.org/10.5281/zenodo.22558) in carrying out this work. JH was supported by a Natural Environment Research Council DTP studentship (grant NE/S007474/1). This work was supported by a grant from the Simons Foundation (MP-SIP-00001828, REB).

References

Banavar SP, Carn EK, Rowghanian P, Stooke-Vaughan G, Kim S, Campàs O. Mechanical control of tissue shape and morphogenetic flows during vertebrate body axis elongation. Scientific Reports. 2021; 11(1):1–14. doi: 10.1038/s41598-021-87672-3.

Bouldin CM, Snelson CD, Farr GH, Kimelman D. Restricted expression of cdc25a in the tailbud is essential for formation of the zebrafish posterior body. Genes & Development. 2014; 28:384–395. doi: 10.1101/gad.233577.113.

Bénazéraf B, Francois P, Baker RE, Denans N, Little CD, Pourquié O. A random cell motility gradient downstream of FGF controls elongation of an amniote embryo. Nature. 2010; 466:248–52. doi: 10.1038/nature09151.

Cooke J, Zeeman EC. A clock and wavefront model for control of the number of repeated structures during animal morphogenesis. Journal of Theoretical Biology. 1976; 58(2):455–476. doi: 10.1016/s0022-5193(76)80131-2.

Delaune EA, François P, Shih NP, Amacher SL. Single-cell-resolution imaging of the impact of notch signaling and mitosis on segmentation clock dynamics. Developmental Cell. 2012; 23(5):995–1005. doi: 10.1016/j.devcel.2012.09.009.

Forero LL, Narayanan R, Huitem LF, VanBergen M, Apschner A, Peterson-Maduro J, Logister I, Valentin G, Morelli LG, Oates AC, Schulte-Merker S. Segmentation of the zebrafish axial skeleton relies on notochord sheath cells and not on the segmentation clock. eLife. 2018; 7. doi: 10.7554/eLife.33843.

Giudicelli F, Özbudak EM, Wright GJ, Lewis J. Setting the tempo in development: an investigation of the zebrafish somite clock mechanism. PLoS biology. 2007; 5(6):e150. doi: 10.1371/journal.pbio.0050150.

Gomez C, Özbudak EM, Wunderlich J, Baumann D, Lewis J, Pourquié O. Control of segment number in vertebrate embryos. Nature. 2008; 454(7202):335–339. doi: 10.1038/nature07020.

Gomez C, Pourquié O. Developmental control of segment numbers in vertebrates. Journal of Experimental Zoology Part B: Molecular and Developmental Evolution. 2009; 312(6):533–544. doi: 10.1002/jez.b.21305.

Harima Y, Takashima Y, Ueda Y, Ohtsuka T, Kageyama R. Accelerating the tempo of the segmentation clock by reducing the number of introns in the Hes7 gene. Cell Reports. 2013; 3(1):1–7. doi: 10.1016/j.celrep.2012.11.012.

Herrgen L, Ares S, Morelli LG, Schroter C, Julicher F, Oates AC. Intercellular coupling regulates the period of the segmentation clock. Current Biology. 2010; 20. doi: 10.1016/j.cub.2010.06.034.

Horikawa K, Ishimatsu K, Yoshimoto E, Kondo S, Takeda H. Noise-resistant and synchronized oscillation of the segmentation clock. Nature. 2006; 441:719–723. doi: 10.1038/nature04861.

Kanki JP, Ho RK. The development of the posterior body in zebrafish. Development. 1997 02; 124(4):881–893. doi: 10.1242/dev.124.4.881.

Kauffman SA. The origins of order: Self-organization and selection in evolution. Oxford University Press, USA; 1993.

Krol AJ, Roellig D, Dequéant ML, Tassy O, Glynn E, Hattem G, Mushegian A, Oates AC, Pourquié O. Evolutionary plasticity of segmentation clock networks. Development. 2011; 138(13):2783–2792. doi: 10.1242/dev.063834.

Kuramoto Y. Chemical turbulence. In: *Chemical oscillations, waves, and turbulence* Springer; 1984.p. 111–140.

Lawton AK, Nandi A, Stulberg MJ, Dray N, Sneddon MW, Pontius W, Emonet T, Holley SA. Regulated tissue fluidity steers zebrafish body elongation. Development. 2013; 140(3):573–582. doi: 10.1242/dev.090381.

Lewis J. Autoinhibition with transcriptional delay: A simple mechanism for the zebrafish somitogenesis oscillator. Current Biology. 2003; 13(16):1398–1408. doi: 10.1016/s0960-9822(03)00534-7.

Liao BK, Jörg DJ, Oates AC. Faster embryonic segmentation through elevated Delta-Notch signalling. Nature Communications. 2016; 7(1):1–12. doi: 10.1038/ncomms11861.

Mallarino R, Grant PR, Grant BR, Herrel A, Kuo WP, Abzhanov A. Two developmental modules establish 3D beak-shape variation in Darwin's finches. Proceedings of the National Academy of Sciences. 2011; 108(10):4057–4062. doi: 10.1073/pnas.1011480108.

Mara A, Schroeder J, Chalouni C, Holley SA. Priming, initiation and synchronization of the segmentation clock by deltaD and deltaC. Nature Cell Biology. 2007; 9(5):523–530. doi: 10.1038/ncb1578.

Michaut A, Mongera A, Gupta A, Serra M, Rigoni P, Lee JG, Duarte F, Hall AR, Mahadevan L, Guevorkian K, Pourquié O. Activity-driven extracellular volume expansion drives vertebrate axis elongation. bioRxiv. 2022; doi: 10.1101/2022.06.27.497799.

Mongera A, Rowghanian P, Gustafson HJ, Shelton E, Kealhofer DA, Carn EK, Serwane F, Lucio AA, Giammona J, Campàs O. A fluid-to-solid jamming transition underlies vertebrate body axis elongation. Nature. 2018; 561(7723):401–405. doi: 10.1038/s41586-018-0479-2.

Monteiro A, Prijs J, Bax M, Hakkaart T, Brakefield PM. Mutants highlight the modular control of butterfly eyespot patterns. Evolution & development. 2003; 5(2):180–187. doi: 10.1046/j.1525-142X.2003.03029.x.

Morelli LG, Ares S, Herrgen L, Schröter C, Jülicher F, Oates AC. Delayed coupling theory of vertebrate segmentation. HFSP journal. 2009; 3(1):55–66. doi: 10.2976/1.3027088.

Morin-Kensicki EM, Melancon E, Eisen JS. Segmental relationship between somites and vertebral column in zebrafish. Development. 2002; 129. doi: 10.1242/dev.129.16.3851.

Murray PJ, Maini PK, Baker RE. Modelling oscillator synchronisation during vertebrate axis segmentation. In: *Pattern Formation in Morphogenesis* Springer; 2013.p. 95–105. doi: 10.1007/978-3-642-20164-6_9.

Murray PJ, Carrieri FA, Dale JK. Cell cycle regulation of oscillations yields coupling of growth and form in a computational model of the presomitic mesoderm. Journal of Theoretical Biology. 2019; 481.

Naganathan SR, Popović M, Oates AC. Left–right symmetry of zebrafish embryos requires somite surface tension. Nature. 2022; 605(7910):516–521. doi: 10.1038/s41586-022-04646-9.

Oates AC, Morelli LG, Ares S. Patterning embryos with oscillations: structure, function and dynamics of the vertebrate segmentation clock. Development. 2012; 139(4):625–639. doi: 10.1242/dev.063735.

Palmeirim I, Henrique D, Ish-Horowicz D, Pourquié O. Avian hairy gene expression identifies a molecular clock linked to vertebrate segmentation and somitogenesis. Cell. 1997; 91:639–648. doi: 10.1016/s0092-8674(00)80451-1.

Raff RA. The shape of life – genes, development and evolution of animal forms. University of Chicago Press; 1996.

Richardson MK, Allen SP, Wright GM, Raynaud A, Hanken J. Somite number and vertebrate evolution. Development. 1998; 125(2):151–160. doi: 10.1242/dev.125.2.151.

Riedel-Kruse IH, Müller C, Oates AC. Synchrony dynamics during initiation, failure, and rescue of the segmentation clock. Science. 2007; 317(5846):1911–1915. doi: 10.1126/science.1142538.

Rohatgi A, Webplotdigitizer: Version 4.6; 2022. https://automeris.io/WebPlotDigitizer.

Schröter C, Herrgen L, Cardona A, Brouhard GJ, Feldman B, Oates AC. Dynamics of zebrafish somitogenesis. Developmental dynamics. 2008; 237(3):545–553. doi: 10.1002/dvdy.21458.

Schröter C, Ares S, Morelli LG, Isakova A, Hens K, Soroldoni D, Gajewski M, Jülicher F, Maerkl SJ, Deplancke B, Oates AC. Topology and dynamics of the zebrafish segmentation clock core circuit. PLOS Biology. 2012 07; 10(7):1–20. doi: 10.1371/journal.pbio.1001364.

Schröter C, Oates AC. Segment number and axial identity in a segmentation clock period mutant. Current Biology. 2010; 20(14):1254–1258. doi: 10.1016/j.cub.2010.05.071.

Simsek MF, Chandel AS, Saparov D, Zinani OQ, Clason N, Özbudak EM. Periodic inhibition of Erk activity drives sequential somite segmentation. Nature. 2023; 613(7942):153–159. doi: 10.1038/s41586-022-05527-x.

Simsek MF, Özbudak EM. Spatial fold change of FGF signaling encodes positional information for segmental determination in zebrafish. Cell Reports. 2018; 24(1):66–78. doi: 10.1016/j.celrep.2018.06.023.

Soroldoni D, Jörg DJ, Morelli LG, Richmond DL, Schindelin J, Jülicher F, Oates AC. A Doppler effect in embryonic pattern formation. Science. 2014; 345(6193):222–225. doi: 10.1126/science.1253089.

Steventon B, Duarte F, Lagadec R, Mazan S, Nicolas JF, Hirsinger E. Species-specific contribution of volumetric growth and tissue convergence to posterior body elongation in vertebrates. Development. 2016; 143(10):1732–1741. doi: 10.1242/dev.126375.

Thomson L, Muresan L, Steventon B. The zebrafish presomitic mesoderm elongates through compaction-extension. Cells & Development. 2021; 168:203748. doi: 10.1016/j.cdev.2021.203748.

Uriu K, Ares S, Oates AC, Morelli LG. Dynamics of mobile coupled phase oscillators. Physical Review E. 2013; 87(3):032911. doi: 10.1103/PhysRevE.87.032911.

Uriu K, Bhavna R, Oates AC, Morelli LG. A framework for quantification and physical modeling of cell mixing applied to oscillator synchronization in vertebrate somitogenesis. Biology Open. 2017; 6(8):1235–1244. doi: 10.1242/bio.025148.

Uriu K, Liao BK, Oates AC, Morelli LG. From local resynchronization to global pattern recovery in the zebrafish segmentation clock. eLife. 2021; 10:e61358. doi: 10.7554/eLife.61358.

Uriu K, Morelli LG. Collective cell movement promotes synchronization of coupled genetic oscillators. Biophysical Journal. 2014; 107(2):514–526. doi: 10.1016/j.bpj.2014.06.011.

Uriu K, Morishita Y, Iwasa Y. Random cell movement promotes synchronization of the segmentation clock. Proceedings of the National Academy of Sciences. 2010; 107(11):4979–4984. doi: 10.1073/pnas.0907122107.

Venzin OF, Oates AC. What are you synching about? Emerging complexity of Notch signaling in the segmentation clock. Developmental Biology. 2020; 460(1):40–54. doi: 10.1016/j.ydbio.2019.06.024.

Verd B, Monk NA, Jaeger J. Modularity, criticality, and evolvability of a developmental gene regulatory network. Elife. 2019; 8:e42832. doi: 10.7554/eLife.42832.

Wagner G, Pavlicev M, Cheverud J. The road to modularity. Nature Reviews Genetics. 2007; 8. doi: 10.1038/nrg2267.

Wagner GP, Altenberg L. Perspective: Complex adaptations and the evolution of evolvability. Evolution. 1996; 50. doi: 10.1111/j.1558-5646.1996.tb02339.x.

Webb AB, Lengyel IM, Jörg DJ, Valentin G, Jülicher F, Morelli LG, Oates AC. Persistence, period and precision of autonomous cellular oscillators from the zebrafish segmentation clock. eLife. 2016; 5:e08438. doi: 10.7554/eLife.08438.

Xiong F, Ma W, Bénazéraf B, Mahadevan L, Pourquié O. Mechanical coupling coordinates the co-elongation of axial and paraxial tissues in avian embryos. Developmental Cell. 2020; 55(3):354–366. doi: 10.1016/j.devcel.2020.08.007.

Yoshioka-Kobayashi K, Matsumiya M, Niino Y, Isomura A, Kori H, Miyawaki A, Kageyama R. Coupling delay controls synchronized oscillation in the segmentation clock. Nature. 2020; 580(7801):119–123. doi: 10.1038/s41586-019-1882-z.

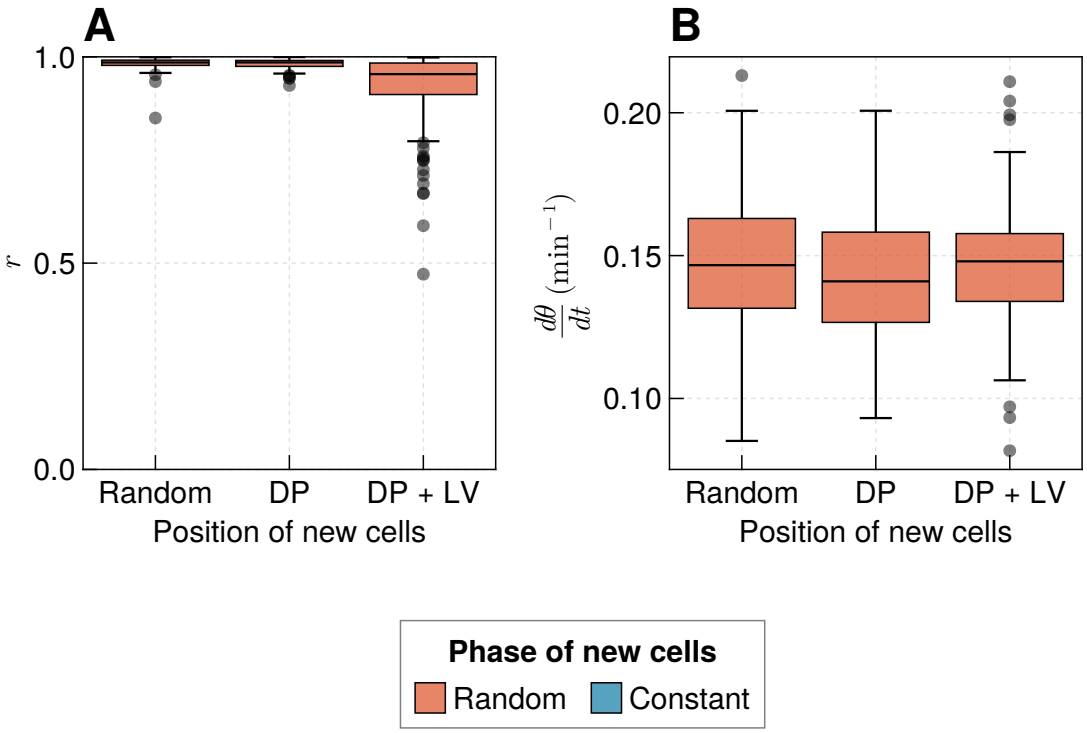


Figure 2—figure supplement 1. Effect of varying cell ingression position with random initial phases. **A** Oscillation synchrony (r) at the PSM anterior after 1000 mins, for the three tested scenarios of cell ingression, with ingressing cells having random initial phase $\theta_i \in [0, 2\pi)$. $N = 100$ simulations. **B** Mean frequency of oscillations for cells at the PSM anterior after 1000 mins of simulation, for the three tested scenarios of cell ingression, with ingressing cells having random initial phase $\theta_i \in [0, 2\pi)$. $N = 100$ simulations.

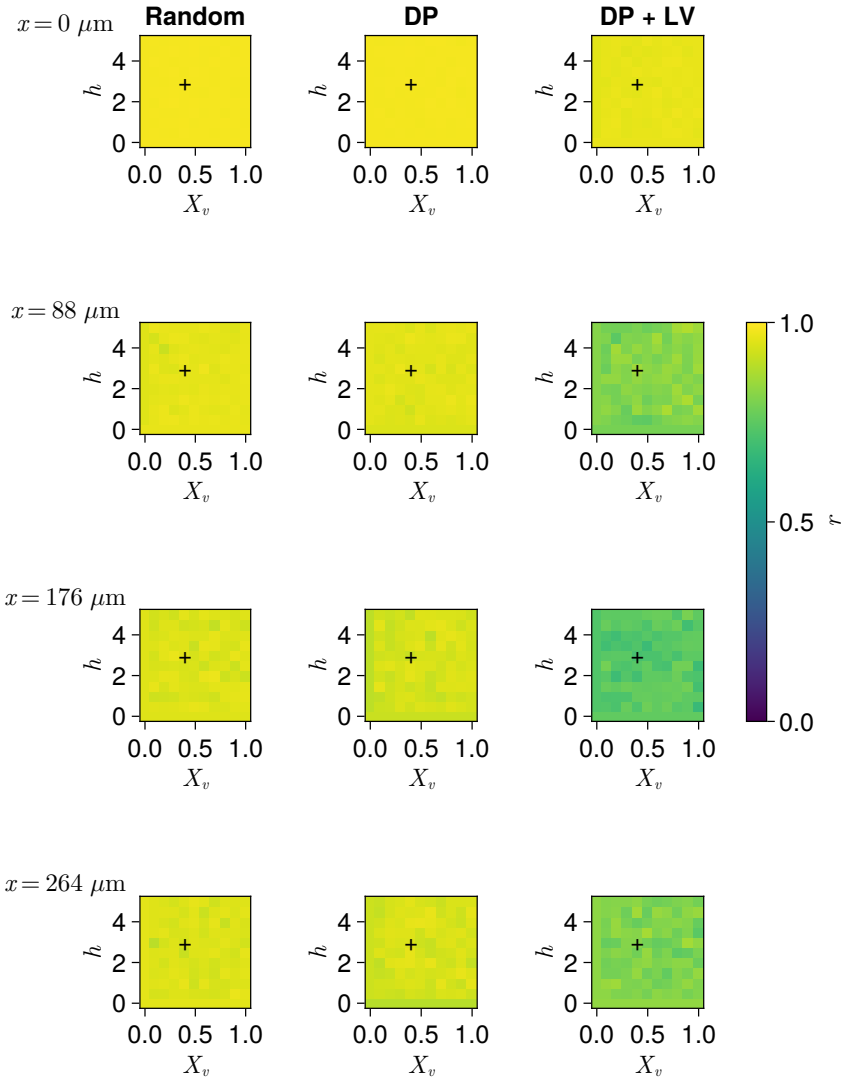


Figure 3—figure supplement 1. Clock synchrony along the x-axis for varying profiles of cell motility. Clock synchrony r for different motility profiles $v_0(\mathbf{x})$ specified by combinations of X_v and h, across three different scenarios of cell ingression. Each row corresponds to synchrony in different cell diameter (d_c) -wide subsection of the PSM, with the x-coordinate of this subsection shown on the left-hand side of the figure. Pixels display the synchrony at the PSM anterior after 1000 mins of simulation, averaged across $N = 100$ simulations. A black + corresponds to the parameter pair used elsewhere in this paper, unless otherwise stated ($X_v = 0.4, h = 3$). Parameter ranges used are $X_v \in \{0, 0.1, 0.2, \dots, 1\}$ and $h \in \{0, 0.5, 1, \dots, 5\}$.

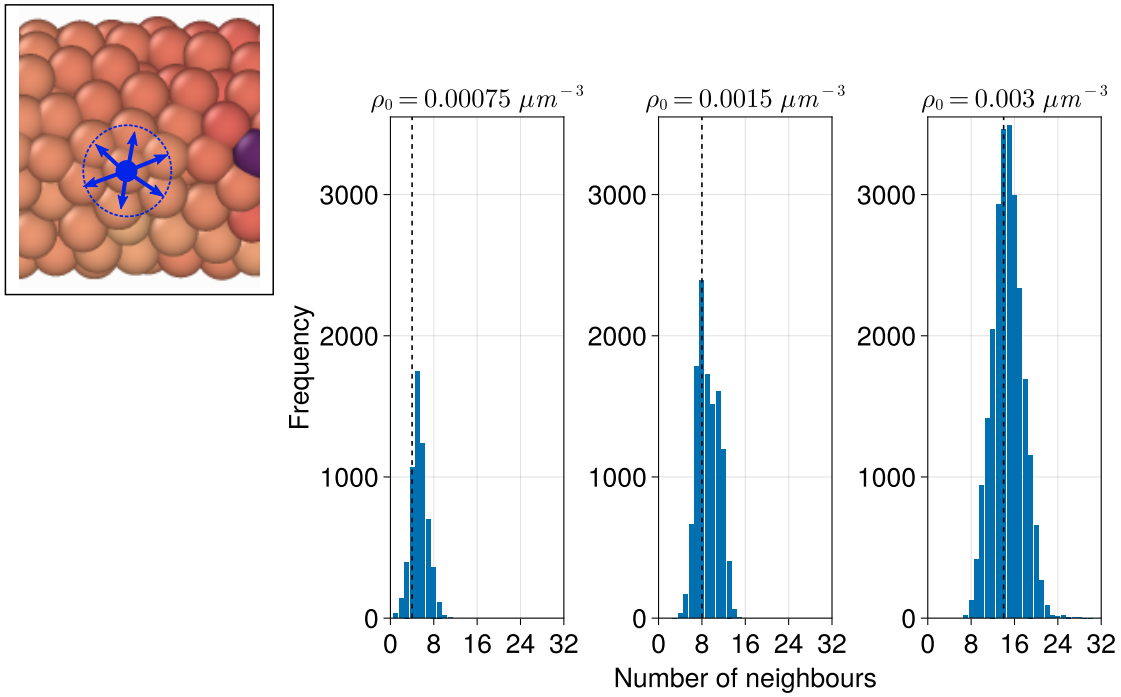


Figure 4—figure supplement 1. Distribution of the number of neighbours for each cell when varying the density (ρ_0) of the tissue. Distributions are calculated from final timepoints of individual simulations with random cell ingression with $\rho_0 = 0.00075, 0.0015, 0.003\ \mu m^{-3}$ by calculating the number of non-self cells within a cell's radius (d_c) of each cell (see inset, left). The median of each distribution is shown with a black dashed line.

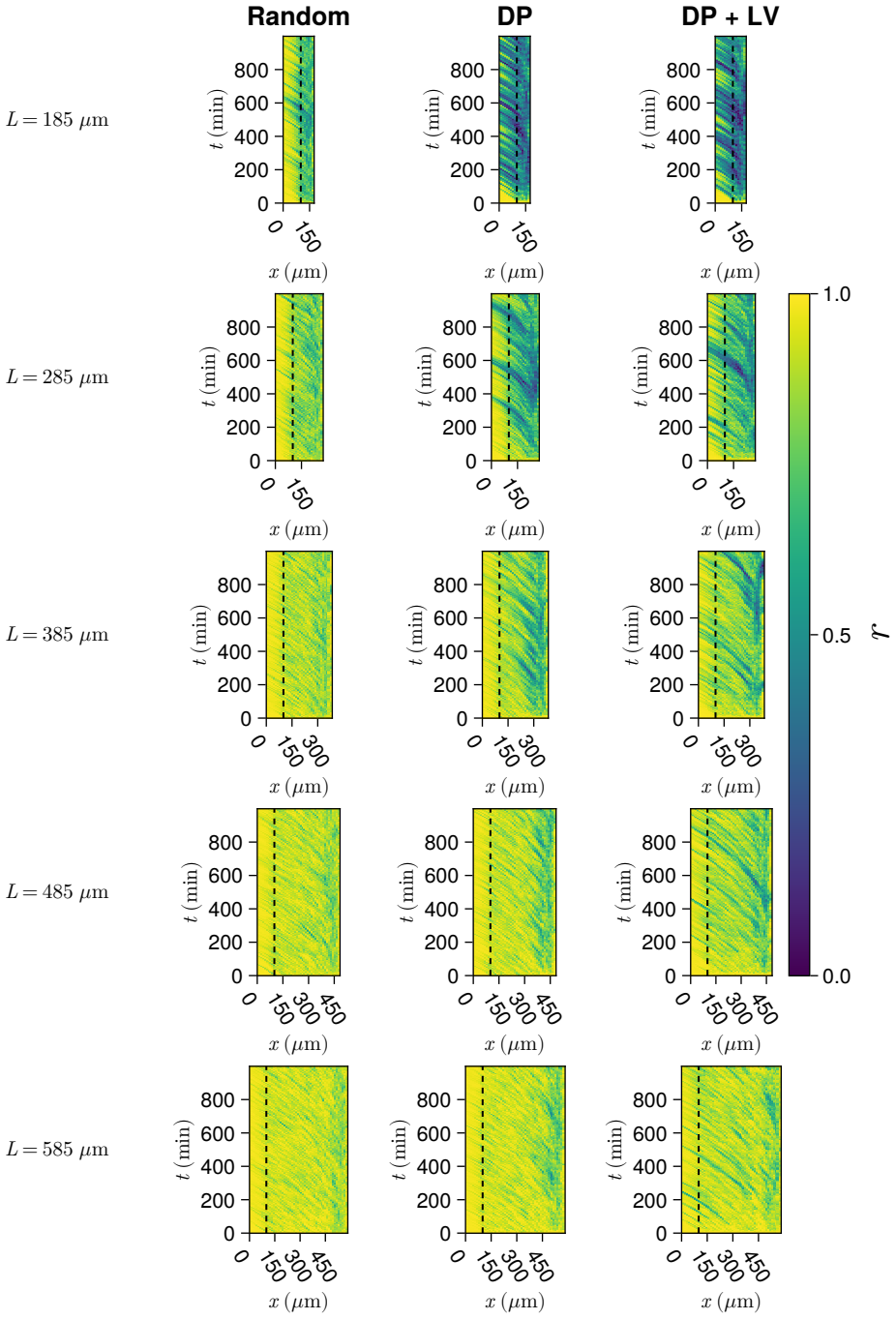


Figure 4—figure supplement 2. Synchrony dynamics for varying tissue length L. Kymographs of synchrony along the x-axis on the left-hand side of the PSM for single simulations from each of the three scenarios of cell ingression tested. Tissue length L for each simulation is shown on the left hand side. The anterior limit of cell addition x_d is fixed at $x_d = 100$, and is annotated with a black dashed line.

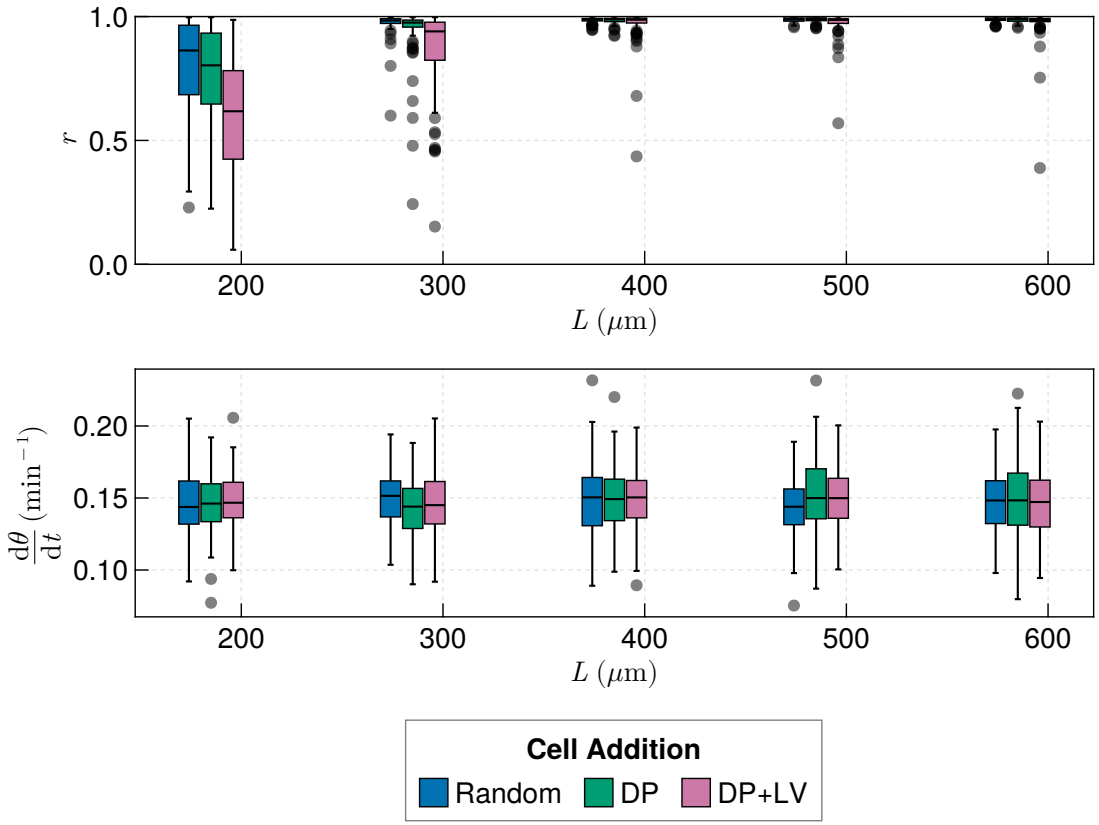


Figure 4—figure supplement 3. Anterior synchrony for varying tissue length L, with varying anterior limit of cell addition x_d. Top Anterior synchrony after 1000 mins for changing tissue length L and varying position of cell ingression, where x_d is fixed at $x_d = L - R - r_T - 100$. $N = 100$ simulations. **Bottom** Anterior mean frequency after 1000 mins for changing tissue length L and varying position of cell ingression where x_d is fixed at $x_d = L - R - r_T - 100$. $N = 100$ simulations.

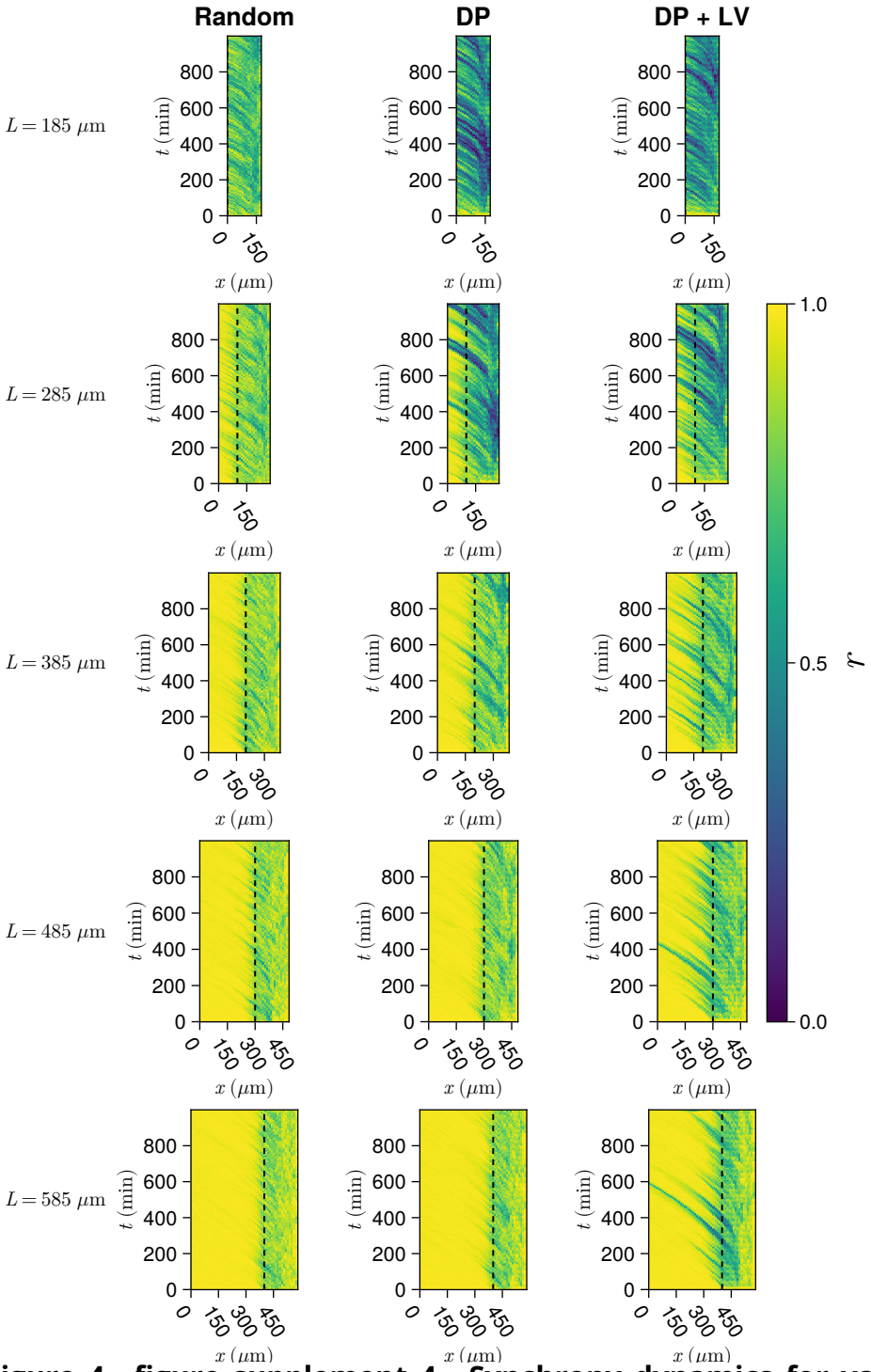


Figure 4—figure supplement 4. Synchrony dynamics for varying tissue length L, with varying anterior limit of cell addition x_d. Kymographs of synchrony along the x-axis on the left-hand side of the PSM for single simulations from each of the three scenarios of cell ingression tested. Tissue length L for each simulation is shown on the left hand side. The anterior limit of cell addition x_d is fixed at $x_d = L - R - r_T - 100$, and is annotated with a black dashed line.

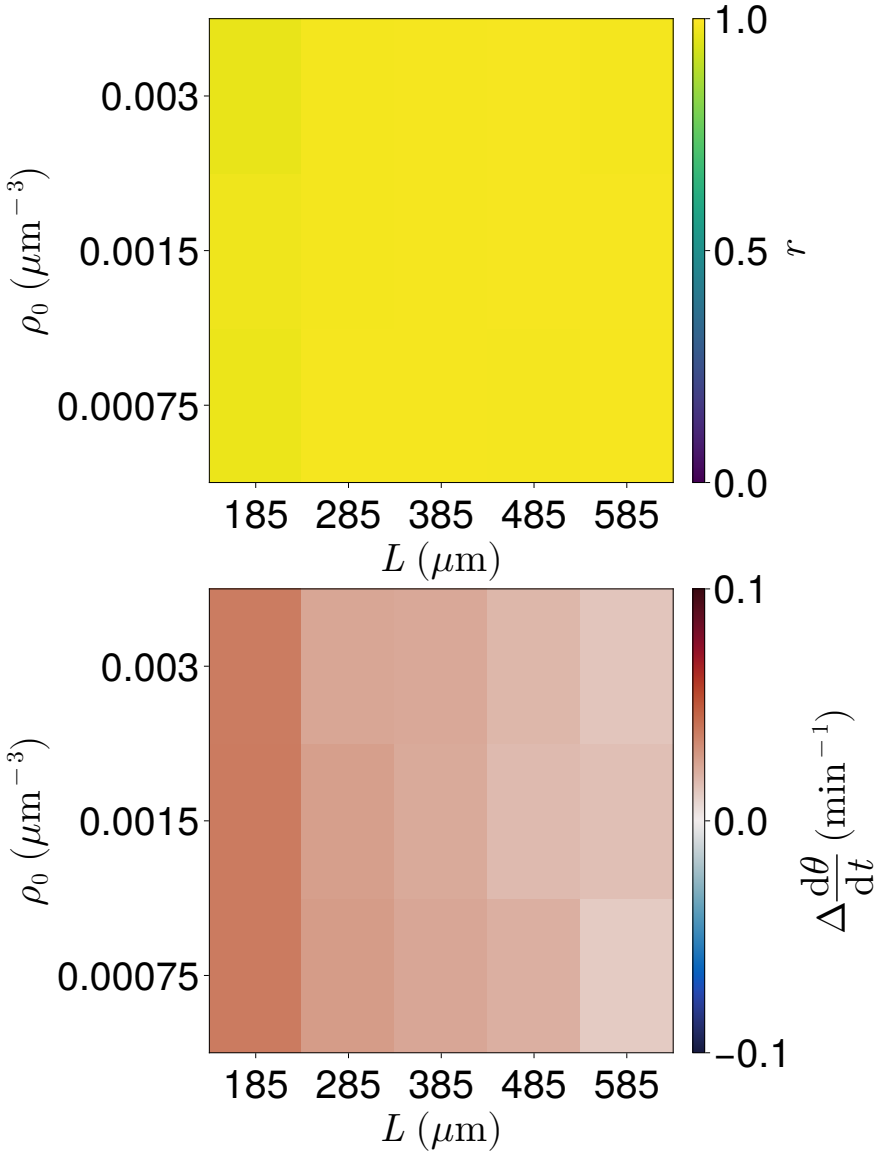


Figure 4—figure supplement 5. Clock dynamics for co-varying density ρ_0 and length L. Median synchrony r (**top**) and difference from expected mean frequency $\Delta\mathrm{d}\theta/\mathrm{d}t$ (**bottom**) at the anterior of the tissue after $t = 999$ mins, for co-varying values of tissue density ρ_0 and tissue length L. Cell ingression is at random positions and with random phase $\theta \in [0, 2\pi)$. For smaller values of ρ_0 gaps in the tissue can occur and no cells occur in the d_c-wide region at the tissue anterior at the final timestep. Such simulations where this occur are excluded from this analysis. As a result for each parameter pair the sample size N is $91 \leq N \leq 100$.

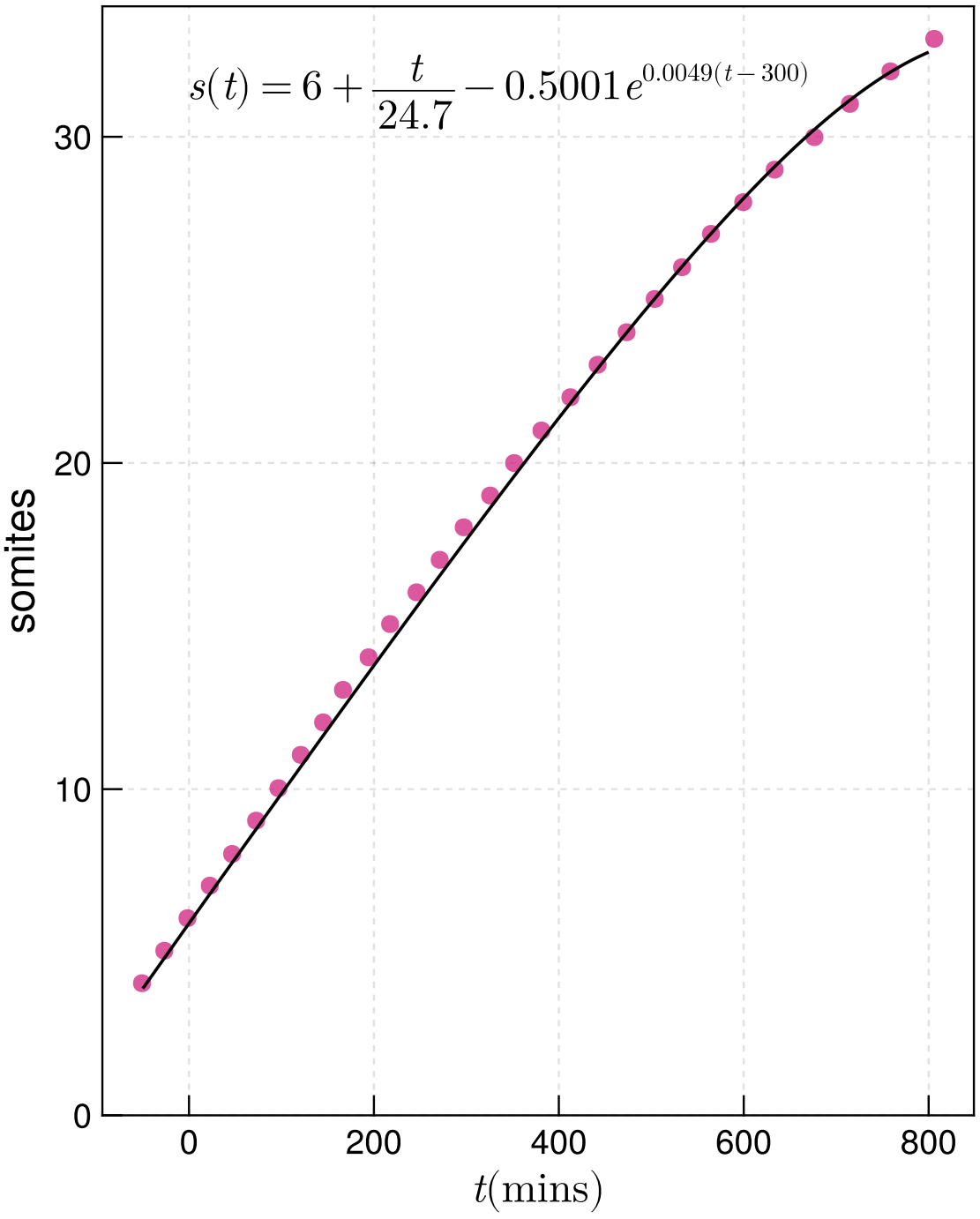


Figure 5—figure supplement 1. Zebrafish somitogenesis, after Schröter et al. 2008. Fitted function interpolating the data for zebrafish somitogenesis from ***Schröter et al.*** (***2008***) (black), against the raw data (magenta).

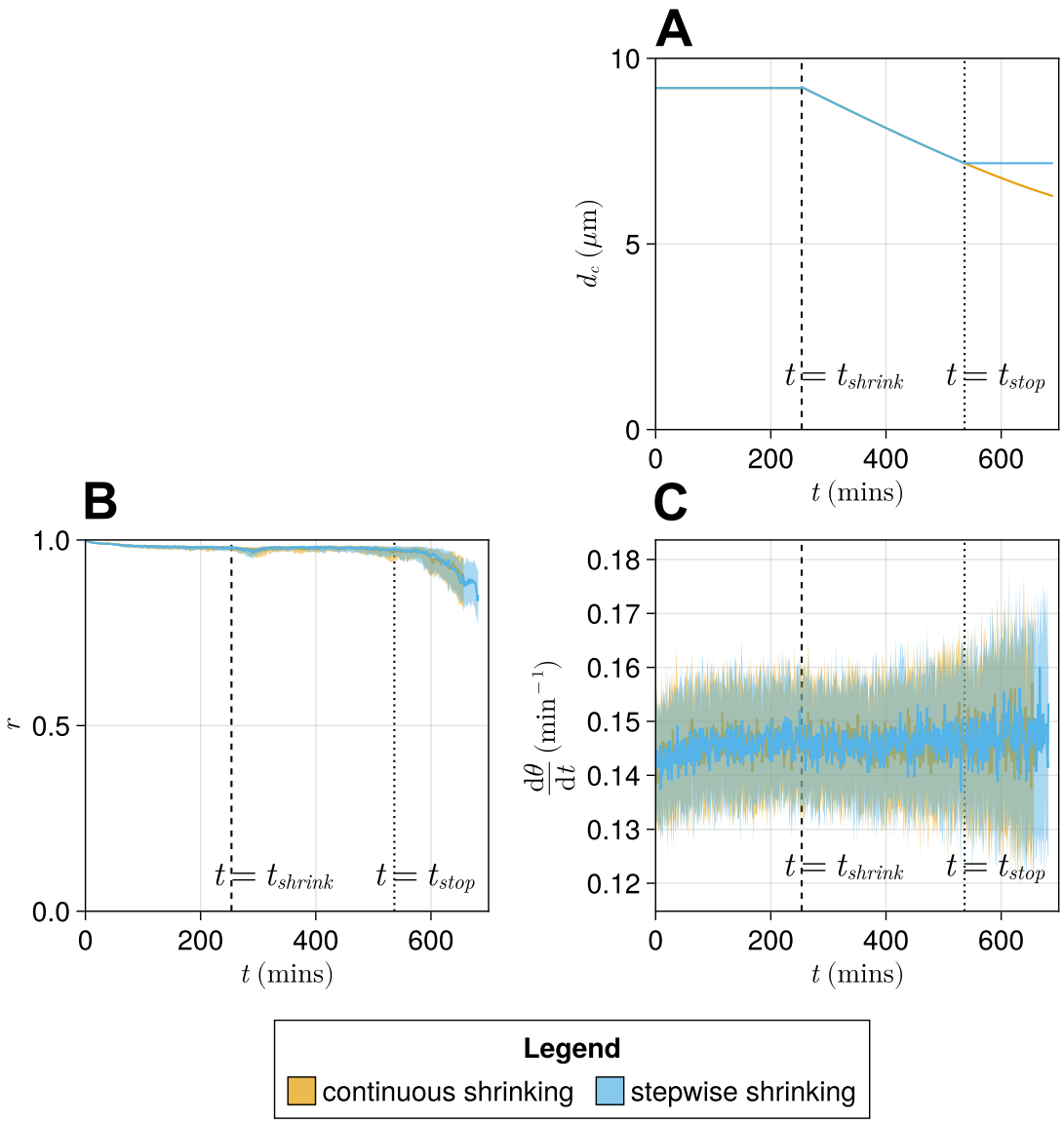


Figure 5—figure supplement 2. Comparison of a 'stepwise' function for decreasing cell diameter d_c against a continuously decreasing function. **A** The stepwise (blue) and continuous (yellow) functions compared here. The intercept and gradient for these functions are derived from data from ***Thomson et al.*** (***2021***). **B** Anterior synchrony (r) over time for stepwise shrinking cells (blue) and continuously shrinking cells (yellow). Dark line shows the median of $N = 100$ simulations, and the shaded area either side of this line shows the inter-quartile range (IQR). Results are plotted until the time at which at least one of the replicate simulations encounters a gap in the tissue at the tissue anterior (see methods). **C** Mean anterior frequency ($\mathrm{d}\theta/\mathrm{d}t$) over time for stepwise shrinking cells (blue) and continuously shrinking cells (yellow). Dark line shows the median of $N = 100$ simulations, and the shaded area either side of this line shows the IQR. Results are plotted until the time at which at least one of the replicate simulations encounters a gap in the tissue at the tissue anterior.

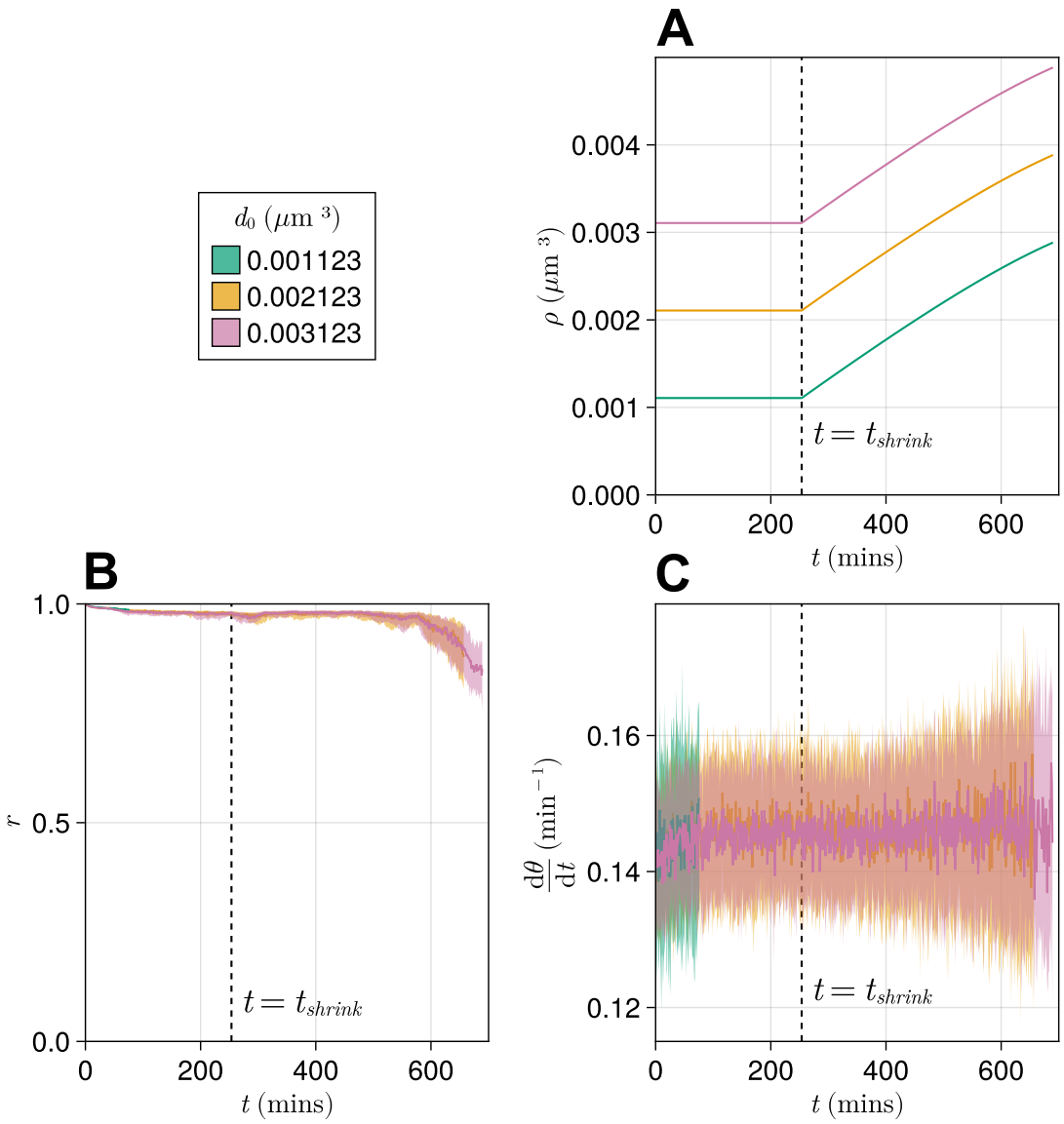


Figure 5—figure supplement 3. Effect of starting tissue density d_0 on anterior synchrony and frequency over time, for a compacting tissue. **A** Functions showing tissue density $\rho(t)$ for different intercept values d_0. **B** Anterior synchrony (r) over time. Dark line shows the median of $N = 100$ simulations, and the shaded area either side of this line shows the inter-quartile range (IQR). Results are plotted until the time at which at least one of the replicate simulations encounters a gap in the tissue at the tissue anterior (see methods). **C** Mean anterior frequency ($d\theta/dt$) over time. Dark line shows the median of $N = 100$ simulations, and the shaded area either side of this line shows the IQR. Results are plotted until the time at which at least one of the replicate simulations encounters a gap in the tissue at the tissue anterior.

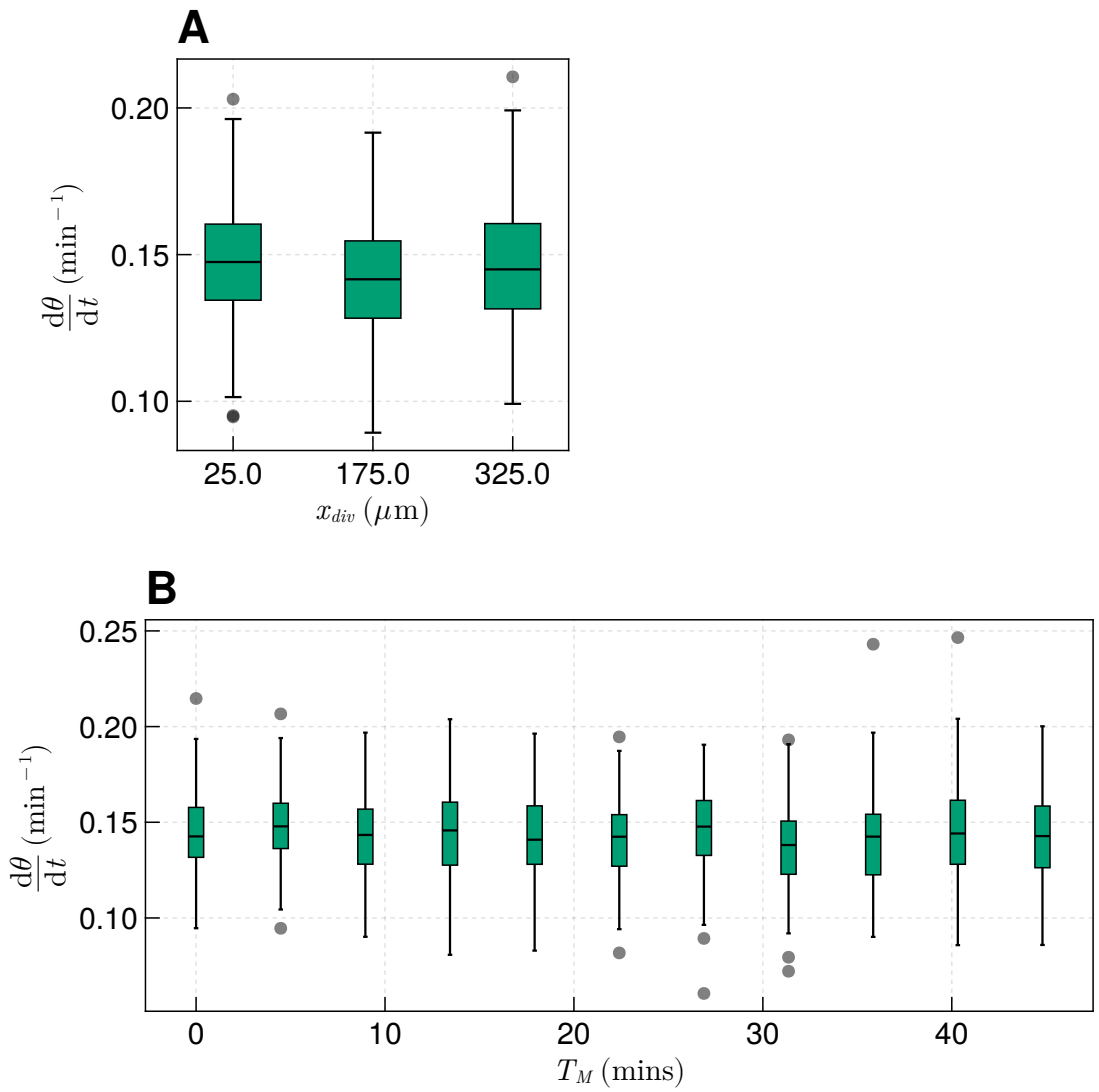


Figure 6—figure supplement 1. The effect of cell division on frequency. A Effect on median frequency when varying x_{div} after 1000 mins for $N = 100$ simulations. **B** Effect on median frequency when varying T_M after 1000 mins for $N = 100$ simulations.

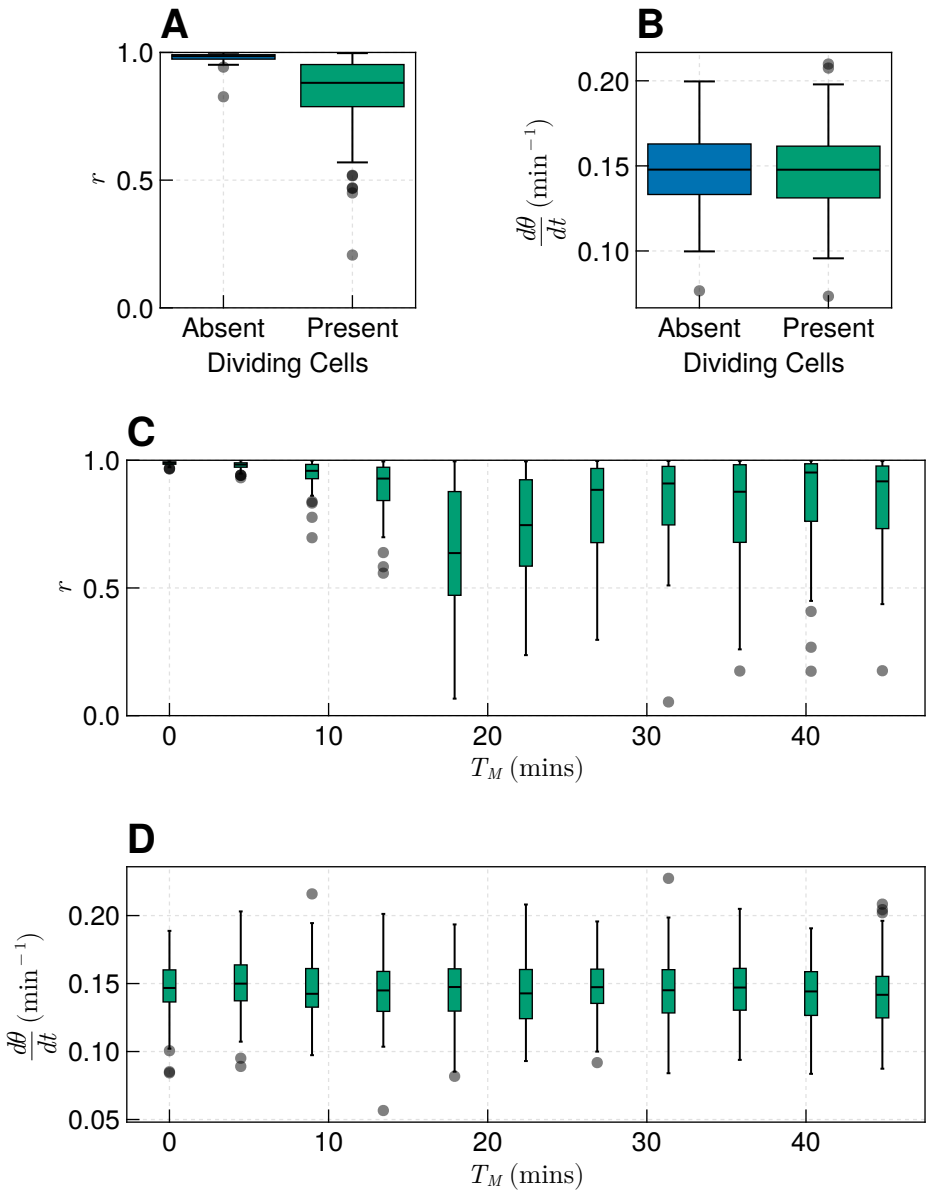


Figure 6—figure supplement 2. Effect of cell division where clock coupling is forbidden during M-phase. A Synchrony at the PSM anterior after 1000 mins in simulations with or without dividing cells, for $T_M = 15$ mins. Neighbouring cells were forbidden from coupling to neighbours in M phase. $N = 100$. **B** Frequency at the PSM anterior after 1000 mins in simulations with or without dividing cells, for $T_M = 15$ mins. Neighbouring cells were forbidden from coupling to neighbours in M phase. $N = 100$. **C** Synchrony at the PSM anterior after 1000 mins in simulations with or without dividing cells, for varying T_M. Neighbouring cells were forbidden from coupling to neighbours in M phase. $N = 100$ for each value of T_M. **D** Frequency at the PSM anterior after 1000 mins in simulations with or without dividing cells, for varying T_M. Neighbouring cells were forbidden from coupling to neighbours in M phase. $N = 100$ for each value of T_M.

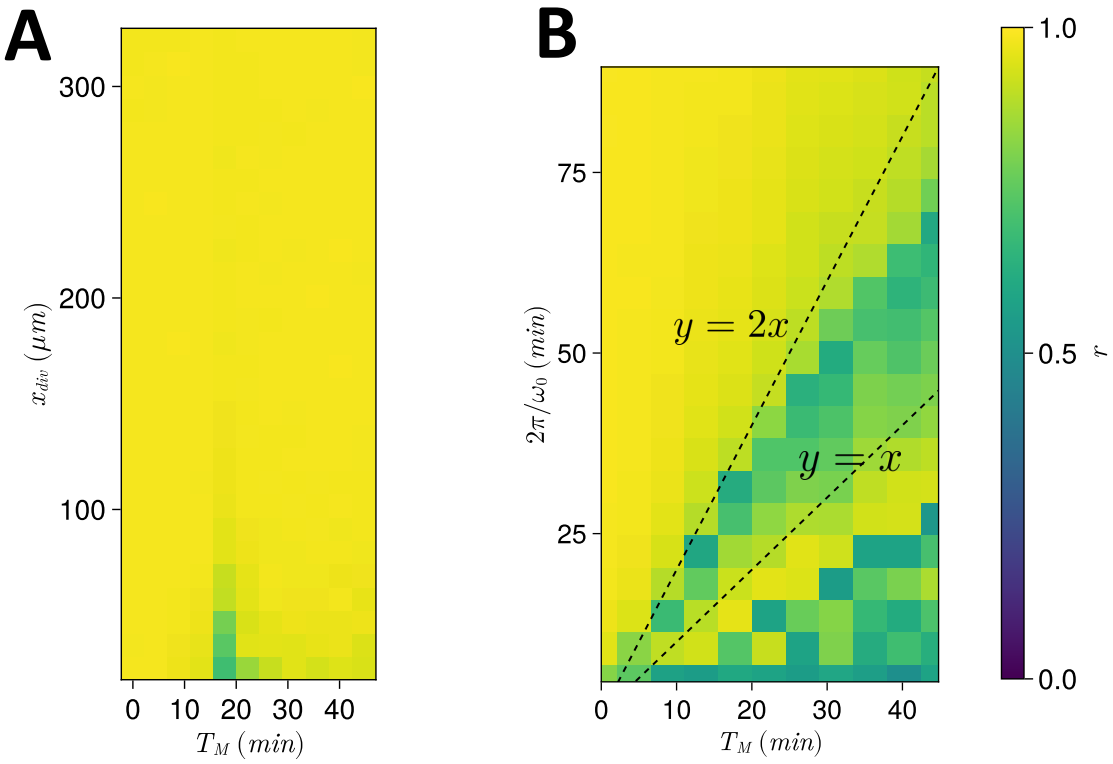


Figure 6—figure supplement 3. Spatiotemporal constraints on T_M. **A** Dependence of anterior synchrony on x_{div} and T_M. The median anterior synchrony after 1000 mins for $N = 100$ simulations is plotted. **B** Dependence of anterior synchrony on T_M and intrinsic clock frequency ω_0. The median anterior synchrony after 1000 mins for $N = 100$ simulations is plotted.

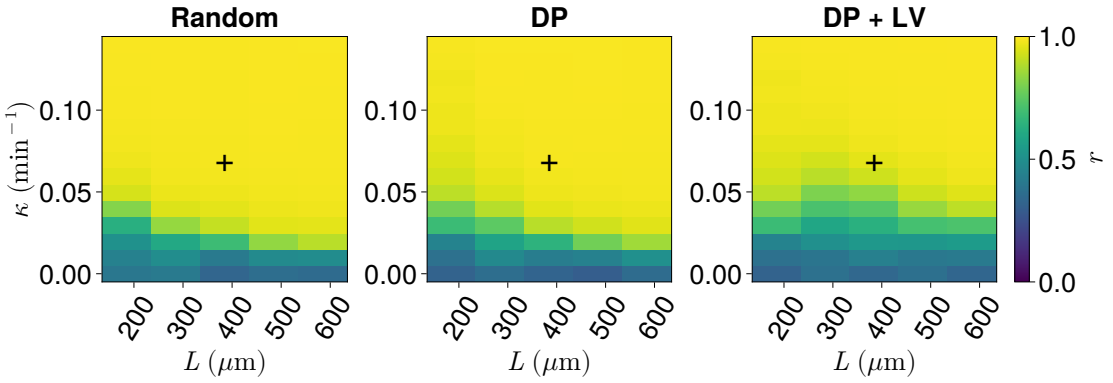


Figure 7—figure supplement 1. Synchrony for varying coupling strength κ and tissue length L. Effect of co-varying coupling strength κ, tissue length L, and position of cell ingression on the median anterior synchrony of $N = 100$ simulations. A black + corresponds to the parameter pair $L = 385$ μm, $\kappa = 0.07$ min^{-1}.

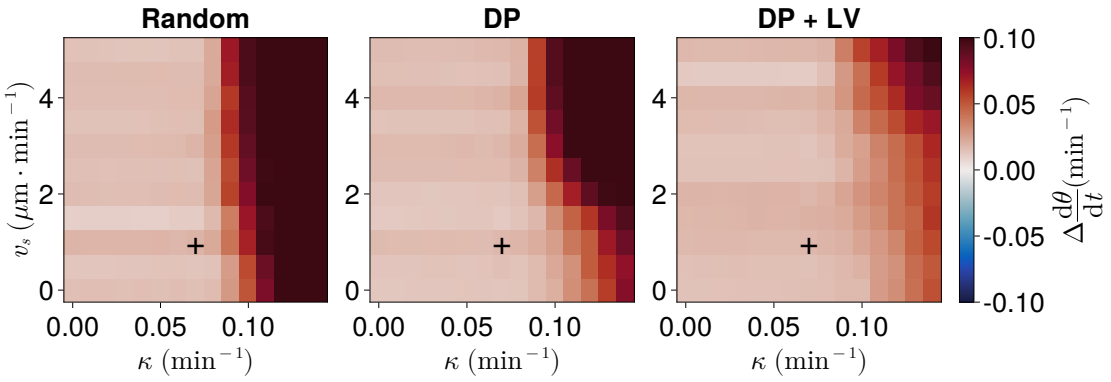


Figure 8—figure supplement 1. Frequency in the presence of coupling delay, for varying coupling strength κ and magnitude of intrinsic cell motion v_s. Difference from expected mean frequency ($\Delta \text{d}\theta/\text{d}t$) at the PSM anterior after 1000 mins, for varying maximum magnitude of intrinsic cell motion v_s and clock phase coupling strength κ, for three different scenarios of cell ingression. Each pixel corresponds to the median value of anterior synchrony across $N = 100$ simulations. A black + marks the experimental values for zebrafish, $\kappa = 0.07$ min^{-1}, $v_s = 1$ μm · min^{-1}, derived in ***Riedel-Kruse et al.*** (***2007***) and ***Uriu et al.*** (***2017***) respectively, that are used elsewhere in this paper. All other parameters are held constant at their normal values (see table 1).

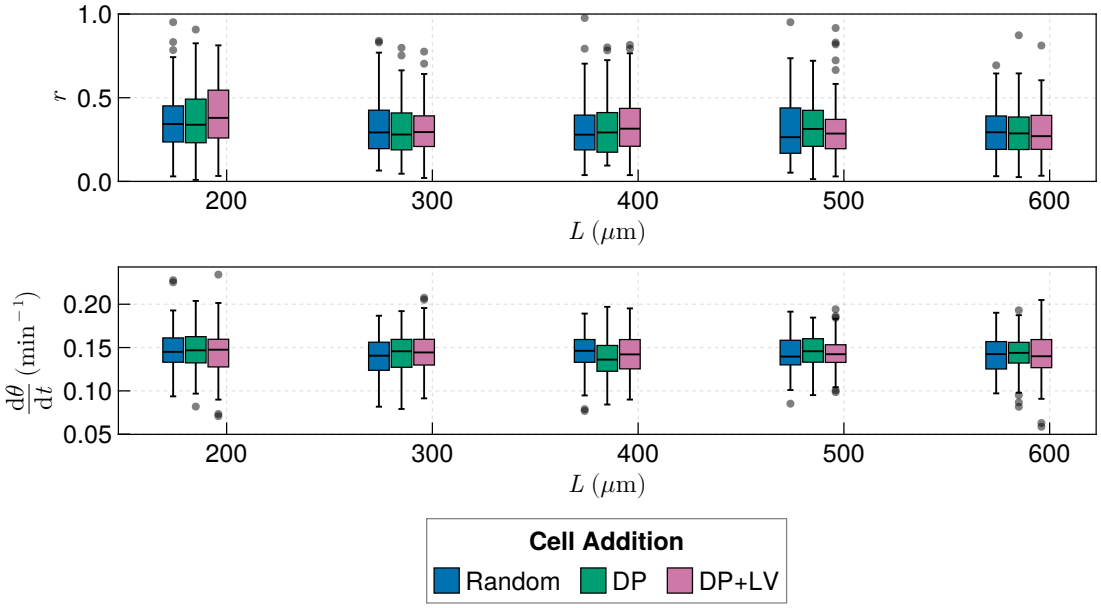


Figure 8—figure supplement 2. Clock dynamics in the presence of a coupling delay, for varying tissue length L. Top: Anterior synchrony after 1000 mins for changing tissue length L and varying position of cell ingression. x_d is fixed at x_d = 100 μm. $N = 100$ simulations. **Bottom:** Anterior mean frequency after 1000 mins for changing tissue length L and varying position of cell ingression. $N = 100$ simulations.

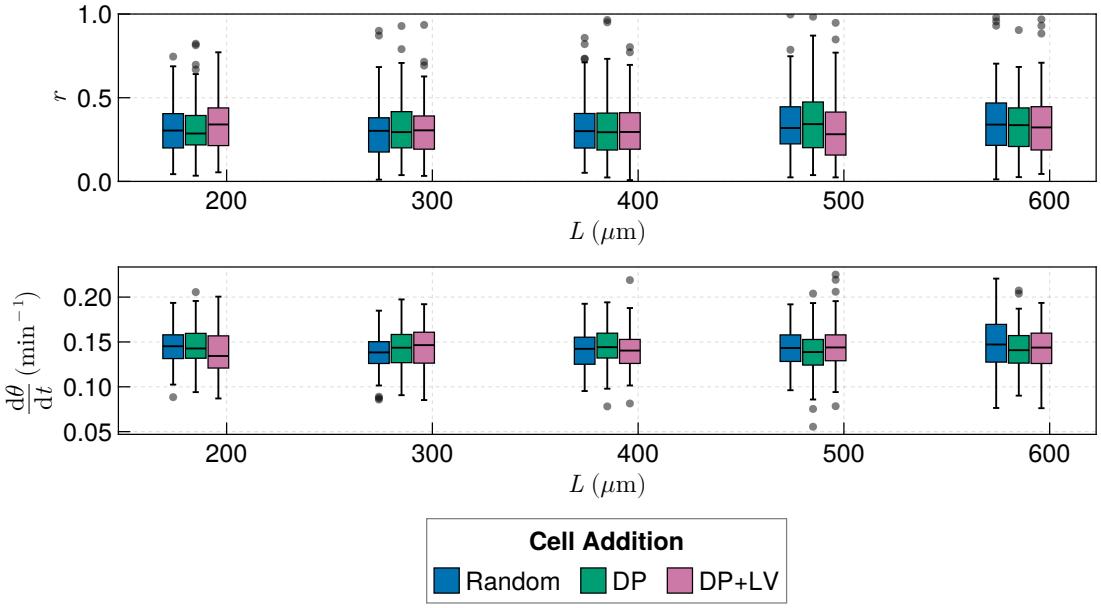


Figure 8—figure supplement 3. Clock dynamics in the presence of a coupling delay, for varying tissue length L and anterior limit of cell addition x_d. Top: Anterior synchrony after 1000 mins for changing tissue length L and varying position of cell ingression. x_d is set to $x_d = L - R - r_T - 100$. $N = 100$ for all simulations, except the DP+LV dataset for $L = 585$ μm , where $N = 99$. **Bottom:** Anterior mean frequency after 1000 mins for changing tissue length L and varying position of cell ingression. x_d is set to $x_d = L - R - r_T - 100$. $N = 100$ for all simulations, except the DP+LV dataset for $L = 585$ μm, where $N = 99$.

My work in this chapter has demonstrated that morphogenesis of the PSM and the segmentation clock show developmental modularity, for most of the known mechanisms of PSM elongation. However several questions remain. For instance, is unclear if this modularity is preserved when we challenge the clock with combinations of these morphogenetic processes, as occurs *in vivo*. Additionally, I have shown that the evolvability of PSM morphogenesis is increased by the clock coupling strength but constrained by the clock coupling delay. It is therefore unclear how the clock and morphogenesis of the PSM co-evolve, as it is possible that the constraining effect of delayed coupling biases evolution of clock coupling strength and PSM morphogenesis along particular trajectories in morphospace. Finally it is unclear if this modularity has actually facilitated the rapid evolution of vertebral number in any vertebrate lineage, or if this is something that is peculiar to zebrafish. These questions are addressed in the following two chapters.

3

Navigability of a fitness landscape in somitogenesis

Abstract

In the previous chapter, I demonstrated that the synchrony of the Zebrafish segmentation clock is robust to variation in the morphogenetic processes associated with elongation of the pre-somitic mesoderm (PSM), namely the ingression of progenitor cells, anterior-posterior profiles of cell motility, cell division, and compaction-extension of the tissue. However PSM elongation is achieved by a combination of these processes, and it is not clear from the previous analysis to which combinations of morphogenetic processes clock synchrony is robust. As synchronous clock oscillations are essential for vertebral development, the synchrony of the segmentation clock and the morphogenesis of the PSM can be thought to comprise a fitness landscape. Here 'valleys' of low synchrony correspond to regimes of morphogenesis inaccessible to evolution, and thus evolution is constrained to proceed along 'plateaus' or 'ridges' in this hypothetical multi-dimensional landscape. The navigability of this landscape, or the capacity for evolution to traverse across different modes of morphogenesis without disrupting clock synchrony, therefore determines the evolvability of PSM morphogenesis. However, the structure of this landscape is unclear. Additionally I have shown that this landscape is likely only to be navigable for specific combinations of clock coupling strength and delay, but it is unclear if the mutual constraints of the clock and morphogensis bias evolution of these two processes along specific evolutionary trajectories.

Here I adapt the model of clock dynamics and PSM cell movements in the previous chapter to simulate general mechanisms of PSM morphogenesis, combining cell ingression, division, motility, and tissue compaction-extension to varying degrees. I attempt to resolve the topology of the fitness landscape for clock synchrony across the space of possible PSM morphogeneses spanned by these four processes. I find that clock synchrony is largely invariant across this space, and that the space of PSM morphogeneses with synchronous clock oscillations is largely contiguous, suggesting that the evolution of PSM morphogenesis is unconstrained within this landscape. Navigability of this landscape is heightened by the strength of clock coupling and constrained by the presence of delays in cell-cell communication. However these constraints are escapable under co-evolution of the segmentation clock and morphogenesis of the PSM, suggesting that the evolution of these two processes is highly interdependent.

Author's note: *This chapter is not solely my work. The code and algorithm for the exploration of the model's parameter space by a random walking process were developed by my colleague James Malone, however I performed and designed the analysis. All work presented here was conducted myself.*

3.1 Introduction

The evolvability of vertebral number is likely to be underpinned by the evolvability of PSM elongation, which may act to create the diversity of duration in somitogenesis we observe across the vertebrate clade [44, 46, 113]. In the previous chapter I explored the effect of different morphogenetic processes that are known to be associated with elongation of the PSM on the synchrony and frequency of the segmentation clock, and found that, when each process was tested in isolation, the Zebrafish segmentation clock was broadly robust to variation in morphogenesis. However, *in vivo* these processes do not occur independently of one another, as I assumed in the previous chapter, and they co-evolve over evolutionary time. In the previous chapter I was unable to estimate the effect on the dynamics of the clock when different morphogenetic processes co-occur in the PSM, or what these dynamics would be across the space of possible PSM morphogeneses. This is significant as the way in which dynamics of the clock vary across this space is likely to determine the evolvability of PSM elongation.

To illustrate this, one could imagine that the synchrony of clock oscillations in the presence of a given set of morphogenetic processes (together making up a 'morphogenetic regime') at least partially determines the evolutionary fitness of that regime and the genotypes that correspond to it. Such an assumption is realistic as defects in clock synchrony are associated with disruption of somite patterning and polarity [49, 54, 82], and clock-less Zebrafish mutants exhibit vertebrae with incorrect and variable polarity [91], suggesting an essential role for clock synchrony in vertebral development. The value of clock synchrony across the space of possible morphogenetic regimes therefore forms a multi-dimensional 'fitness landscape' [29, 203]. The structure of such landscapes can be instrumental in determining the evolvability of a trait, as evolution has to proceed along high-dimensional 'peaks' and 'plateaus' of fitness to reach new phenotypes, and can be prevented from accessing new phenotypes by 'valleys' of low fitness [30, 33]. The 'navigability' of this landscape is the capacity for evolution to proceed along fitness peaks or plateaus to discover new, fit phenotypes [33], and can be thought of as being equivalent

to the evolvability of PSM morphogenesis. It is unclear what such a landscape would look like for the synchrony of clock dynamics across the space of possible PSM morphogeneses, but its structure is likely to be highly influential in determining the evolvability of PSM morphogenesis and vertebral number.

Here I have attempted to resolve the topology of this fitness landscape across the space of possible morphogenetic regimes for the PSM, in order to better understand how evolvable PSM morphogenesis is with respect to synchrony of the clock. To do so, I first note, as I did in the first chapter, that across the vertebrates the PSM is only known to be elongated by a handful of mechanisms, namely the addition of progenitor cells via cell ingression [113, 114, 116, 204], anterior-posterior profiles of cell motility [114, 119, 120, 122], growth via cell division [46, 113], and compaction-extension of the tissue [124]. Additionally I note that these processes co-occur to differing degrees within the PSM of different vertebrate species, for instance the degree of growth from cell division varies markedly across the vertebrates [46, 113], and all four processes contribute to elongation of the Zebrafish PSM [113, 114, 120, 124, 204]. Therefore, it appears that evolution traverses a space where the PSM is elongated by a combination of these four processes, and does so by varying the contribution of each to elongation (figure 3.1). Therefore, for my work in this chapter I have modified the model from the previous chapter [54] to combine all of these four morphogenetic processes in the PSM, allowing me to explore general combinations of the four processes and study their corresponding effect on clock dynamics. Here I aim to resolve the structure of the fitness landscape of different morphogenetic regimes using a combination of two approaches. In the first, I will use my modified model to elucidate the macroscopic structure of this landscape by systematically sampling the parameter space of different morphogenetic regimes at a coarse granularity and evaluating the synchrony across this space. In the second, I will attempt to demonstrate the navigability of the landscape by attempting to perform 'walks' across the synchronous regions of the model's parameter space, initiating these from four initial conditions that each correspond to a morphogenetic regime where elongation is majorly driven by only one of the four processes I have considered (figure 3.1).

This second analysis will provide a connection with my previous chapter, where I showed that clock synchrony is robust to variation in each of these processes, by testing if it is possible to navigate across the landscape from one of these simplistic scenarios for PSM elongation to another (figure 3.1).

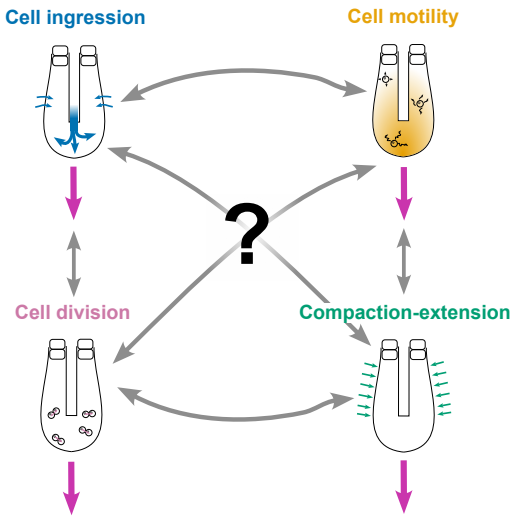


Figure 3.1: Can evolution traverse parameter space between different strategies for elongating the PSM? The PSM is elongated (magenta arrow) by at least four qualitatively distinct processes — ingression of progenitor cells (blue), spatial profiles of intrinsic cell motility (orange), cell division (purple), and compaction-extension (green). Is it possible for evolution to traverse parameter space between each of these processes in isolation without perturbing the clock? How is this affected by co-evolution of the clock?

3.2 Results

3.2.1 Modelling a general mechanism for PSM elongation

Here I have adapted the methodology in the previous chapter to simulate and co-vary morphogenetic processes within the PSM, allowing me to explore how the mechanisms underlying elongation of the PSM might evolve. The model from the previous chapter was modified so that each of these processes could be continuously varied (see methods), by introducing two new parameters x_g and α (figure 3.2). The first of these parameters, x_g, allows one to continuously vary the mode of progenitor cell ingression from where ingression of progenitor cells with constant phase only occurs at the dorso-posterior surface of the PSM (for $x_g = X_c$),

to where cells can also ingress along the ventral surface of the PSM (for $x_g < X_c$). This allows exploration of the possible scenarios for cell ingression between the two extreme values that I considered in the previous chapter, where cells with constant phase were added either only on the dorso-posterior surface of the PSM, or also ventrally, along the length of the two lateral halves of the PSM. Varying x_g allows one to vary this length, as x_g corresponds to the anterior-most position at which new cells appear on the ventral surface of the PSM, where for $x_g \geq X_C$ new cells are added only on the dorsal surface of the posterior PSM. Here, as in the previous model, the appearance of new cells in the PSM is assumed to be density-dependent. Anterior to x_g and up until the anterior limit of cell addition, x_d, new cells are added at random positions with random phase, to maintain cellular density of the PSM.

The second parameter α is introduced to allow control of how much compaction-extension occurs within the PSM. Here I have modified the model so that the rates at which the PSM shrinks in length and width, the rate that the cell diameter decreases, and the rate at which the tissue density increases (figure 3.2), are all scaled by a common factor α. Therefore one can increase the rate of compaction-extension that occurs within the PSM (relative to the rates given by [124]) by setting $\alpha > 1$, decrease the rates by setting $\alpha < 1$, and remove compaction-extension from the morphogenetic regime entirely by setting $\alpha = 0$. Such scaling has been chosen for convenience, removing the need to alter four different parameters. Additionally these parameters are linked biologically, as e.g. the cell density of the PSM increases with a decrease in PSM volume, so such an assumption is biologically realistic [124].

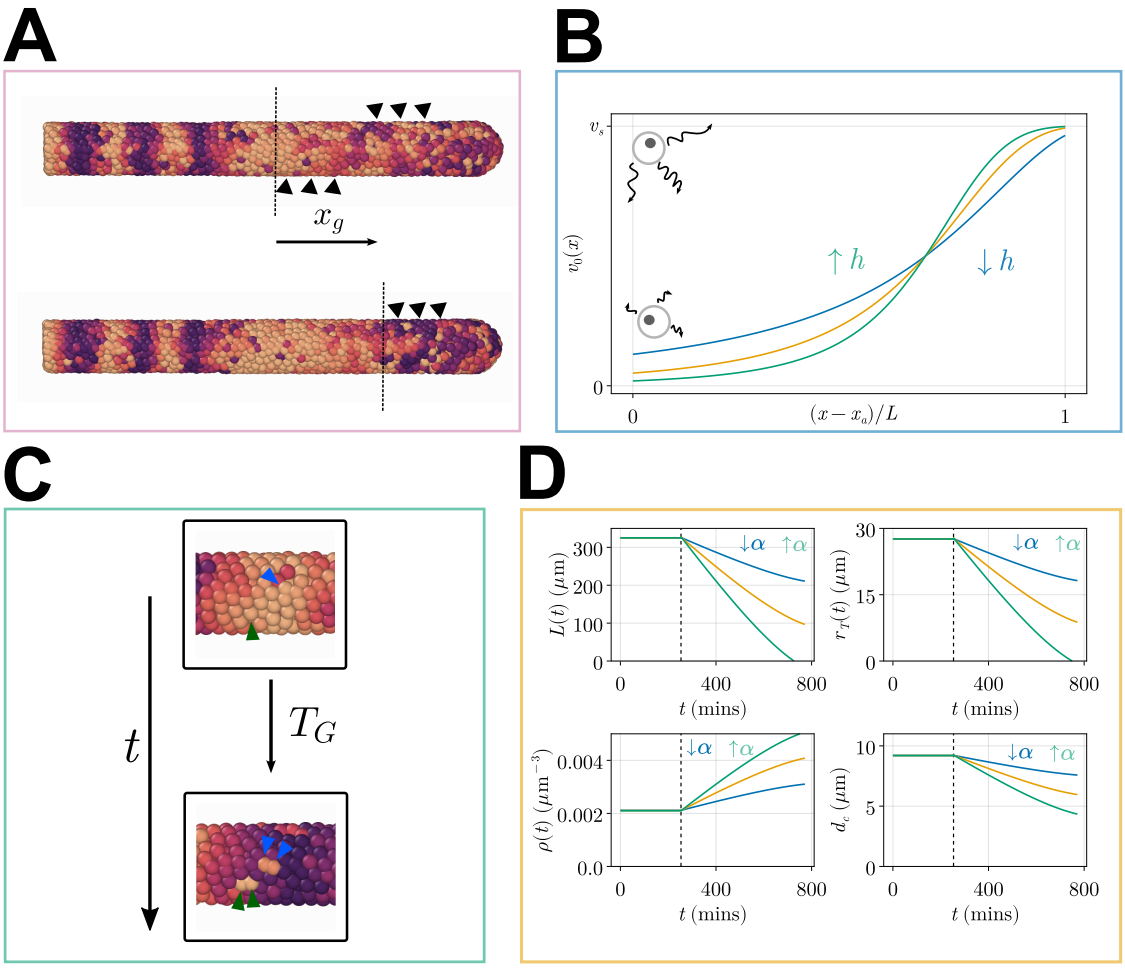


Figure 3.2: Modelling a general mechanism for PSM elongation. A The parameter x_g can be varied continuously to transition the cell ingression regime from lateral-ventral and dorso-posterior ingression of progenitor cells (black arrowheads) to only dorso-posterior ingression (black arrowheads). **B** The parameter h can be varied to create profiles of cell motility ($v_s(x)$) with varying steepness. **C** The amount of cell division within the PSM can be adjusted by varying the total length of the G1, S, and G2 phases of the cell cycle T_G. **D** The rate at which tissue length L, radius r_T, and cell diameter d_c decrease and tissue density ρ increases over time can be co-varied by changes to the parameter α.

Varying the degree of cell division in the model, or the steepness of the cell motility profile, does not require any further modification of the previous model. The steepness of the cell motility profile is already given by the parameter h, and by varying this one can transition between graded profiles of cell motility such as those present in the Chicken embryo [119, 122] to steeper, abrupt anterior-posterior changes in cell motility such as that seen in Zebrafish [120] (figure 3.2). The degree of cell division in the PSM can similarly be varied by altering the combined duration of the G1, S, and G2 phases of the cell cycle, T_G (figure 3.2), and thus for longer values of T_G

fewer cell divisions occur. Co-variation of these four parameters can therefore span the space of possible PSM elongation strategies known from the vertebrate clade (figure 3.1).

3.2.2 Structure of a fitness landscape for general mechanisms of PSM elongation

To determine the topology of the fitness landscape across the space of possible morphogenetic regimes for PSM elongation, I systematically varied the parameters x_g, h, T_G, and α and evaluated the median synchrony of $N = 50$ simulations (figure 3.3). Due to the computational complexity of this model, the parameter space for these four parameters must be sampled with a relatively large step size Δ between parameters. For x_g, h, T_G, and α these were Δx_g = 40 μm, Δh = 0.4, ΔT_G = 40 mins, and $\Delta\alpha$ = 0.4 respectively. Values of these four parameters were considered for ranges that appear plausible based on the parameters for Zebrafish, specifically here I consider the synchrony across $x_g \in [100, 400]$, $h \in [0, 5]$, $T_G \in [187.5, 387.5]$, and $\alpha \in [0, 2]$. This analysis reveals that the synchrony landscape for the space of these four parameters is largely homogeneous, with synchrony being broadly equal for many different parameter sets, only increasing marginally with more rapid rates of compaction-extension α (figure 3.3). Varying the coupling strength κ reveals a similar trend across different coupling strengths $\kappa \in [0, 0.14]$, suggesting that the topology of this synchrony landscape does not change majorly depending on the value of κ (supplementary figures 3.1, 3.2).

Overall synchrony is lower than one would expect from the results in the previous chapter (figure 3.3), however this is likely due to the combination of various processes that were known to have deleterious effects on synchrony, such as cell division or lateral-ventral ingression of progenitor cells. Surprisingly we observe that higher rates of compaction-extension α have greater synchrony at the end of simulation (figure 3.3). This is likely due to a gradual decay of synchrony over time in these simulations due to the co-occurence of these noise-generating processes, and because a simulation is terminated whenever the PSM becomes too short (see methods), simulations with larger α terminate sooner before synchrony has decayed

(supplementary figure 3.3). Additionally we see that synchrony is marginally higher when cells divide more rapidly ($T_G = 172.5$ min) than when they divide slowly (e.g. $T_G = 372.5$ min). This is surprising when one considers my result in the previous chapter that cell division can create asynchrony in clock oscillations, as a more rapid cell cycle implies that cells are more likely to divide and create noise before their clock phase is read-out by the wavefront, thus reducing synchrony of the established somite pattern. However in the previous chapter I did not consider the effect on synchrony effected by cell division in the presence of cell ingression, as I assume here. While cell division has a greater effect on clock synchrony than ingression of non-oscillating progenitor cells (c.f. figure 2B, chapter 1, and figure 6B, chapter 1) when compared in isolation, as I assume ingression to be density-dependent in this model it is likely that cell division has a buffering effect against the noise created by cell ingression. Specifically, more rapid cell divisions increase the density of the PSM and prevent the addition of new cells via ingression.

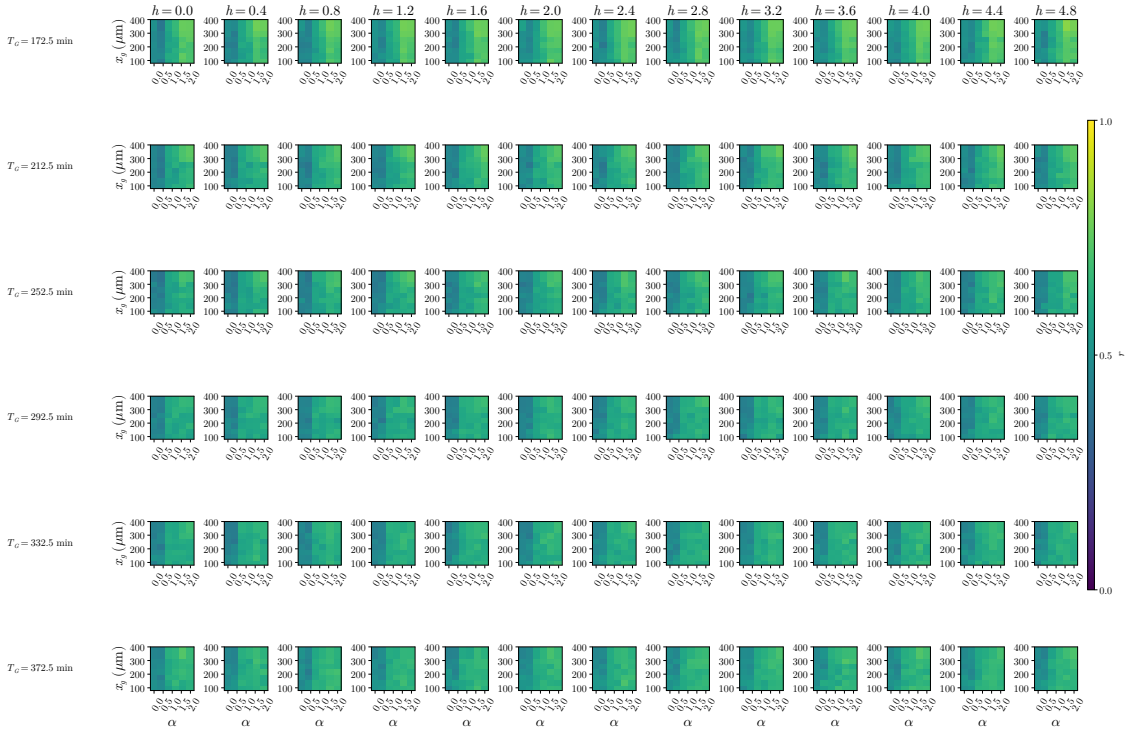


Figure 3.3: Clock synchrony across the space of possible PMS elongation strategies. Synchrony of the clock r at the PSM anterior at the end of a simulation, for varying anteriormost position of ventral cell ingression x_g, steepness of cell motility profile h, duration of G1, S, and G2 phases of the cell cycle T_G, and rate of tissue compaction α. Each pixel represents the median value of $N = 50$ simulations.

It is unknown what the threshold value of synchrony r_c is for the formation of a correctly patterned somite [49], and therefore it is difficult to say from this analysis how much of this landscape is accessible to evolution. However given the noisy nature of clock oscillations *in vivo* [101], and the existence of downstream corrective mechanisms that canalise somitogenesis [205], it is likely that $r_c < 1$. To understand how navigability of this landscape depends on this threshold value, and what in general the navigability of this landscape is, I treat this landscape as a network of parameter sets $\mathcal{T} = (x_g, h, T_G, \alpha)$. Specifically, I consider the landscape to be a network where the vertices $\mathcal{T}$ correspond to a median synchrony $r(\mathcal{T}) \geq r_c$ across $N = 50$ simulations. Any two vertices $\mathcal{T}$ and $\mathcal{T}'$ are connected if they lie next to one another in parameter space, that is for $\mathcal{T} = (x_g, h, T_G, \alpha)$ and $\mathcal{T}' = (x_g', h', T_G', \alpha')$, the vertices $\mathcal{T}$ and $\mathcal{T}'$ are connected if $\mathcal{T}' = (x_g \pm \Delta x_g, h, T_G, \alpha)$ or $\mathcal{T}' = (x_g, h \pm \Delta h, T_G, \alpha)$ or $\mathcal{T}' = (x_g, h, T_G \pm \Delta T_G, \alpha)$ or $\mathcal{T}' = (x_g, h, T_G, \alpha \pm \Delta\alpha)$. Describing the fitness landscape

in this way allows one to assess the navigability directly, as one can determine if it is possible for evolution to traverse the set of parameters corresponding to synchronous dynamics by studying the paths in the resulting network. Additionally, one can identify the existence of disconnected or isolated regions of parameter space with synchronous dynamics that may be inaccessible for evolution (these correspond to 'peaks' or 'plateaus' in the fitness landscape). An example network showing disconnected regions of parameter space is shown in figure 3.4A. The number of these regions is significant for the navigability of the landscape, as the existence of many such regions implies the landscape is poorly navigable, as many possible modes of PSM morphogenesis are inaccessible to evolution.

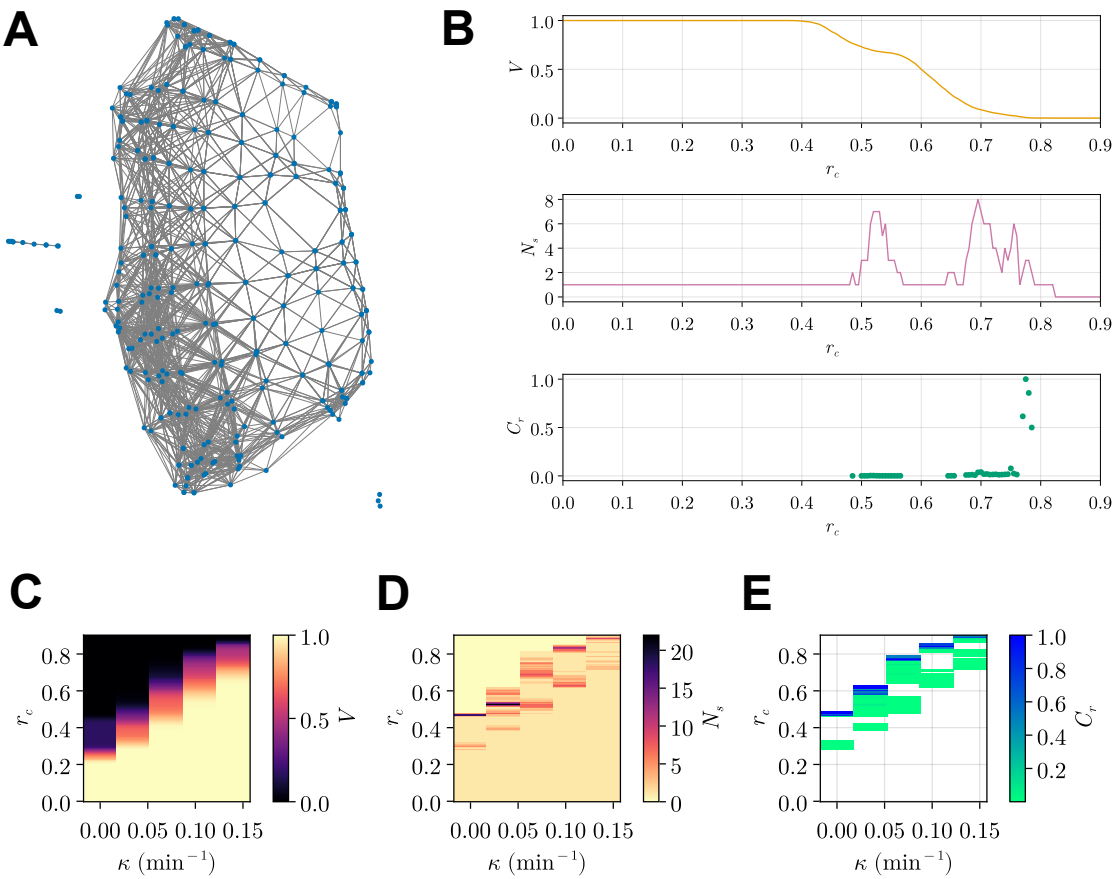


Figure 3.4: Navigability of the fitness landscape for PSM elongation for different synchrony thresholds for somite patterning. **A** SG-t-SNE embedding of the network describing the fitness landscape for $r_c = 0.7$. Blue vertices correspond to parameters with synchronous clock dynamics, and are connected by grey lines if they lie adjacent to one another in parameter space. Embedding was generated using the SG-t-SNE-π software in Julia v1.8.2 [206]. **B** Top: The proportion of parameter space with synchrony $r \geq r_c$, V, for varying r_c. Middle: The number of disconnected clusters N_s for varying threshold r_c. Bottom: The ratio of the number of elements in the second largest cluster to the number of elements in the largest cluster (C_r), for varying threshold r_c. **C** The proportion of parameter space with synchrony $r \geq r_c$, V, for varying r_c and coupling strength κ. **D** The number of disconnected clusters N_s for varying threshold r_c and coupling strength κ. **E** The ratio of the number of elements in the second largest cluster to the number of elements in the largest cluster (C_r), for varying threshold r_c and coupling strength κ.

To assess the navigability of this landscape for varying r_c I define the metric V, which corresponds to the number of different parameter sets tested in this simulation that exhibit $r \geq r_c$, divided by the total number of parameter sets tested (3744). We see that, as expected, as r_c increases, V decreases, and the size of parameter space accessible to evolution is constrained by the value of r_c (figure 3.4B). When we consider the number of disconnected synchronous regions of parameter space N_s, we see that for $r_c < 0.48$ the landscape is one large connected

space and thus fully navigable (figure 3.4B). However for $r_c \geq 0.48$ we see that the synchronous regions of parameter space become isolated with the appearance of up to eight isolated regions for $r_c \approx 0.7$ (figure 3.4B). To evaluate the size of these isolated regions relative to the overall number of synchronous parameter sets, I define C_r, the ratio of the number of elements in the second largest isolated region to the number of elements in the largest. Using this metric we can see that for most r_c the size of the smaller isolated regions is miniscule relative to the size of the largest cluster, except for $r_c \approx 0.77$ where the space of synchronous parameter sets breaks into two to three large disconnected regions (figure 3.4B). Beyond this point the landscape becomes fully navigable again but very small ($V \approx 0$), eventually containing zero elements at $r_c \approx 0.82$ (figure 3.4B). This suggests that, except for these largest values of r_c, the landscape remains largely navigable, as most synchronous parameter sets are connected as part of one large cluster. This trend is present across different values of coupling strength κ, with the value of r_c at which V begins to decline or where the landscape is fractured into multiple disconnected regions increasing with increasing κ (figure 3.4C-E), suggesting that increasing κ increases the contiguity (and thus navigability) of parameter space for this model.

However, I have not considered how the navigability of this landscape is affected by the presence of a coupling delay between cells. In the previous chapter I showed that the introduction of such a delay constrains the ability for morphogenesis to change without perturbing the synchrony of the clock. Therefore when one considers a more realistic model of the segmentation clock incorporating such a delay, the fitness landscape for synchrony may not be so navigable. However, performing such systematic simulations across parameter space is computationally demanding, and large aggregates of simulations such as those shown in figure 3.3 can take days to weeks to complete. Due to this high computational burden, I was not able to replicate such a survey of parameter space incorporating a coupling delay τ into the model, which further increases the computational complexity of the model.

3.2.3 Computationally navigating a fitness landscape for general mechanisms of PSM elongation

Additionally, due to the constraints imposed by the high computational complexity of these models, I have only been able to conduct a macroscopic study of the structure of the fitness landscape for synchrony. However it is unclear if there are finer-scale regions of parameter space that correspond to asynchronous dynamics and so decrease the navigability of this landscape. Therefore as a second test of the navigability of this landscape I performed an exploration of parameter space at higher resolution. To reduce the computational complexity of the model, I reduced the PSM to a cylinder and so reducing the number of cells that are simulated (see methods for details). Additionally I adopted a new approach for exploring parameter space that avoids such large-scale exhaustive explorations of parameter space, this being a more computationally tractable method that simulates multiple independent 'walks' across parameter space. Here, I test given sets of parameters and see whether that parameter set corresponds to clock oscillations with synchrony exceeding r_c. Additionally I set $r_c = 0.7$, which captures the noise that one sees in wild-type expression of segmentation clock genes [101], and corresponds to the r_c with the largest number of disconnected regions of parameter space for our default coupling strength $\kappa = 0.07\ \text{min}^{-1}$ (figure 3.4B). While larger values of r_c have proportionately larger disconnected regions of parameter space, the total size of the parameter space is much smaller (figure 3.4B). Therefore $r_c = 0.7$ appears to be the value that maximises complexity of this landscape's structure and is the most likely to challenge my previous finding that this landscape is navigable.

This walking method proceeds as follows. For a given parameter set $\mathcal{T}$, a new $\mathcal{T}'$ is selected from the neighbourhood of $\mathcal{T}$ in parameter space (this neighbourhood is defined more rigorously in the methods section below). If $r(\mathcal{T}') \geq r_c$, $\mathcal{T}'$ is selected as the next 'step' in a 'walk' and a new $\mathcal{T}''$ is selected from the neighbourhood of $\mathcal{T}'$. The synchrony of the clock for the parameter set $\mathcal{T}''$ is then tested as above, repeating the process. In this manner it is possible to traverse or 'walk' across parameter space. However if $r(\mathcal{T}') < r_c$ up to four other $\mathcal{T}'$ may be proposed

from the neighbourhood of $\mathcal{T}$. If after testing five $\mathcal{T}'$ in total no $\mathcal{T}'$ such that $r(\mathcal{T}') \geq r_c$ has been found then the walk is terminated. This constraint on the number of repetitions is to reduce the computational burden of the method, and similarly I limit walks to be no more than 150 steps in total. A pseudocode description of this algorithm is given in the methods section.

As I am interested in the value of synchrony across the space of possible combinations of x_g, h, T_G, and α, I initialise such walks from one of four initial conditions that each correspond to an extreme case of PSM morphogenesis. Specifically these conditions are where elongation is driven only by cell ingression (this initial condition is termed 'ingression'), where it is driven only by an anterior-posterior profile of cell motility ('motility'), where it is only driven by cell division ('division'), and where it is only driven by compaction-extension ('compaction-extension'). Initialising walks from these four extreme scenarios for PSM elongation can demonstrate whether the landscape for synchrony is traversable, or if there exist regions of asynchrony in between. Additionally such initial conditions provide a connection with the results of the previous chapter, which showed that synchrony is robust to variation in the one process in each of these scenarios, but not whether continuous paths across parameter space exist that can connect them without perturbing synchrony, and thus whether the clock and morphogenesis of the PSM are truly modular. Details of the intial conditions are given in the methods section below.

Performing $N = 100$ such walks from each of the four initial conditions we see that, for the model with instantaneous coupling, much of parameter space is explored (figure 3.5A,B), and that it is possible to traverse parameter space between the 'division-only' and 'motility-only' initial conditions (figure 3.5A). However it is unclear whether it is possible to reach the 'ingression-only' or 'compaction-only' parameter sets from other initial conditions. Accordingly, the parameter space for parameters h and T_G is broadly sampled, but parameter space is poorly sampled for intermediate values of x_g and larger values of α. These distributions of x_g and α could be explained by a poor coverage of parameter space by these walks, as there appears to be a large region of parameter space unvisited by walks initiating from these initial conditions

(supplementary figure 3.4), and it may be possible that by increasing the length of walks this region may be traversable. However it is not possible to rule out the existence of an asynchronous region of parameter space here, and further analysis is required. Nevertheless there are multiple possible paths between the 'division' and 'motility' initial conditions with the majority having a length of approximately 12 iterations (figure 3.5C). Overall these results appear similar to figure 3.3, suggesting that parameter space is broadly navigable to evolution.

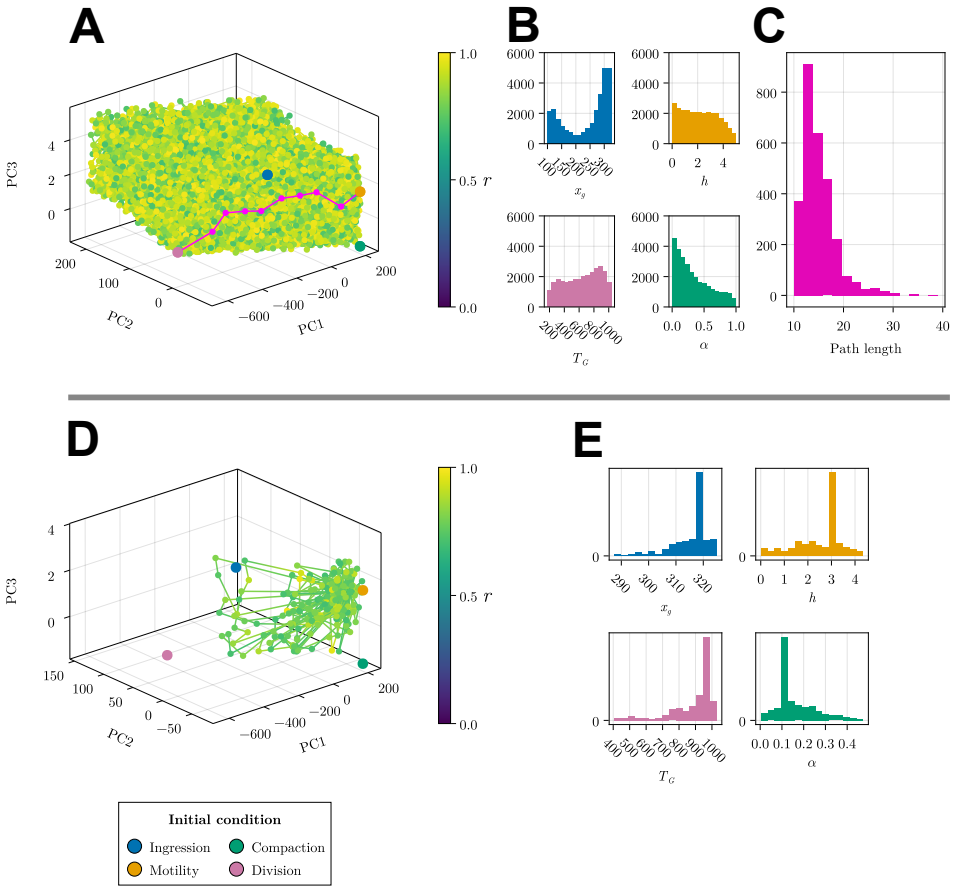


Figure 3.5: Random walks exploring parameter space. A PCA projection of the results of $N = 400$ independent random walks ($N = 100$ from each initial condition), for instantaneous coupling ($\tau = 0$ mins) coloured according to the synchrony r at the end of the simulation. Individual parameter sets visited on a walk are shown as circles and coloured according to their clock synchrony. Lines connecting these circles show the connection of any two parameter sets as part of a walk. Locations of the four initial conditions are shown in blue, orange, green, and mauve dots (see legend, bottom). The shortest path across parameter space between the 'division' and 'motility' initial conditions is plotted in magenta. **B** Distributions of parameters visited as part of the $N = 400$ independent walks. **C** Distribution of path lengths between the 'division' and 'motlity' initial conditions. **D** PCA projection of the results of $N = 400$ independent walks ($N = 100$ from each initial condition), for delayed coupling ($\tau = 30$ mins) coloured according to the synchrony r at the end of the simulation. Individual parameter sets visited on a walk are shown as circles and coloured according to their clock synchrony. Lines connecting these circles show the connection of any two parameter sets as part of a walk. Locations of the four initial conditions are shown in blue, orange, green, and mauve dots (see legend, bottom). **E** Distributions of parameters visited as part of the $N = 400$ independent walks.

To test how the introduction of a coupling delay τ affects the navigability of the fitness landscape, I introduce a coupling delay of $\tau = 30$ mins into the model. It should be noted that the clock coupling delay in the Zebrafish segmentation clock is estimated to be $\tau = 21$ mins [139],

however in the previous chapter I showed that this value is marginally too low to achieve synchrony of clock oscillations with this model. In this previous analysis $\tau = 30$ mins was predicted to allow synchronous oscillations to occur even in the presence of lateral-ventral ingression of PSM progenitor cells, and therefore I have chosen this value to explore the effects of delayed coupling on the navigability of this landscape. Repeating the exploration of parameter space with $\tau = 30$ mins, we see that navigability is significantly curtailed (figure 3.5D). Here we see that for the 'ingression', 'division', and 'compaction-extension' initial conditions synchrony cannot be achieved, reflected by large peaks in the distribution for each parameter at the values corresponding to the various initial conditions (figure 3.5E). For the 'motility' initial condition there exists a neighbourhood of synchronous parameter space, however the exploration of the rest of parameter space appears to be constrained compared to the instantaneous coupling case (figure 3.5A,D). This suggests that, similar to the results in the previous chapter, the coupling delay acts as a significant constraint to the evolution of PSM morphogenesis.

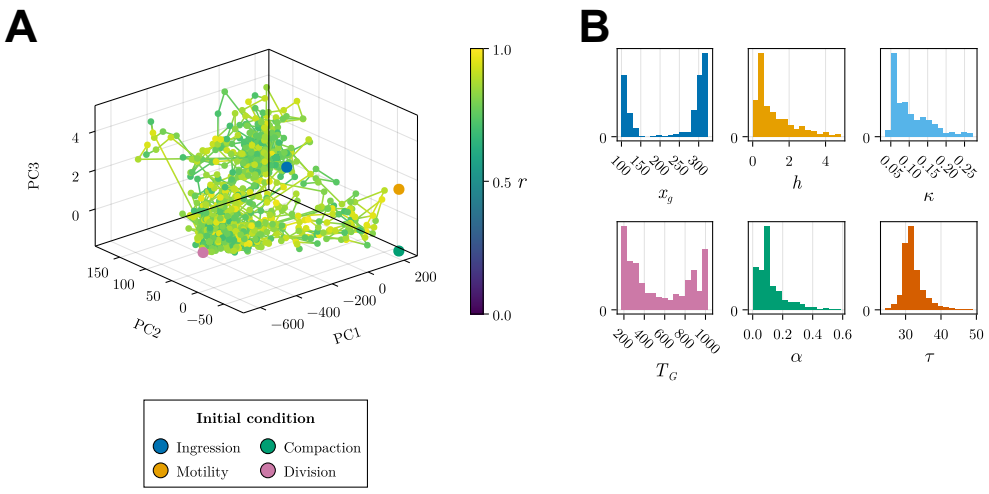


Figure 3.6: Random walks across a space of PSM morphogeneses and clock coupling. A PCA projection of the results of $N = 400$ independent heuristic searches ($N = 100$ from each initial condition), for varying coupling strength κ and coupling delay τ, coloured according to the synchrony r at the end of the simulation. Locations of the four initial conditions are shown in blue, orange, green, and mauve dots (see legend, bottom). The shortest path across parameter space between the 'division' and 'motility' initial conditions is plotted in magenta. **B** Distributions of parameters visited as part of the $N = 400$ independent walks.

However it is unrealistic to assume that the coupling strength and coupling delay remain

constant across evolutionary time, as these processes will co-evolve with the morphogenesis of the PSM. Allowing these parameters to co-evolve with morphogenesis of the PSM may make the landscape more navigable. To test this I repeated the above analysis but allowing coupling strength κ and coupling delay τ to vary along with x_g, h, T_G, and α. Simulations are initiated from the same initial conditions as the previous analyses, with $\kappa = 0.07$ min^{-1} and $\tau = 30$ min. Here we see that the parameter space for x_g, h, T_G, and α is explored more thoroughly than when κ and τ are fixed (figure 3.6). This suggests that via evolution of the segmentation clock it is possible to escape the constraints imposed on the evolution of PSM morphogenesis by the clock, and evolve new regimes of morphogenesis.

3.3 Discussion

Here I have shown that when cell-cell phase coupling of the segmentation clock is assumed to be instantaneous, the synchrony of the segmentation clock is largely invariant across a large space of possible modes of PSM elongation (figure 3.3), and that the parameter space corresponding to these synchronous dynamics is largely contiguous (figure 3.4), suggesting that it can easily be explored by evolution. Additionally it is possible to traverse this parameter space to reach different scenarios of PSM elongation (figure 3.5A-C), further verifying my prediction of the modularity of the segmentation clock and morphogenesis of the PSM.

My results highlight the importance of the canalisation of somite patterning in determining the evolvability of the morphogenesis of the PSM. When the threshold clock synchrony for correctly formed somites is decreased, the total number of parameter sets that can correctly pattern somites increases, as expected (figure 3.4B). Because my results show that for almost any value of r_c these parameter sets form a contiguous space, this threshold r_c is therefore a major determinant of the evolvability of PSM morphogenesis. As I mentioned previously, the true value of r_c is unknown, and quantification of this value is needed to fully understand the evolvability of PSM morphogenesis and by extension of vertebral number. It is also unclear how canalisation would occur, though oscillations of segmentation clock genes are noisy, even when

cells are coupled to their neighbours [101, 200], so presumably noise-buffering mechanisms exist. Naganathan *et al.* recently showed that surface tension of the somite buffers against imprecision in the anterior-posterior length of somites, ensuring symmetrically sized somites along the body axis [205]. As the variability in somite length was shown to be increased upon disruption of notch signalling [205], it is likely these imprecisions are at least partially due to the noise of the clock. Such corrective mechanisms in somitogenesis have been poorly studied to date, but further investigation could reveal much regarding the evolvability of PSM morphogenesis.

It is unclear in general whether developmental modularity evolves via neutral processes or via selection [207]. As I show here that canalisation of somite patterning promotes modularity of the segmentation clock and PSM morphogenesis, this could argue for evolution of modularity as a secondary effect of selection. Specifically, it seems plausible that canalising processes can arise via selection for robust somite patterning, given the fundamental importance of somitogenesis for organisation of the vertebrate body plan. To my knowledge the evolution of these processes has not been investigated, and further study is likely to be illuminating for our understanding of the evolution of modularity.

Here I have assumed that the fitness of morphogenesis of the PSM can be approximated by the synchrony of the segmentation clock, however fitness is likely to also be determined by other properties. For instance, the morphogenetic processes elongating the PSM must not only elongate the PSM but couple this elongation to that of surrounding tissues [121]. We cannot predict whether this is the case for the general mechanisms of PSM elongation I consider here. Additionally I have not been able to consider whether the different morphogenetic regimes simulated in this model affect the rate of PSM elongation. This rate could be significant for the fitness of the resulting phenotype as this would be predicted to alter the length of the formed somites [86]. It would also be desirable to use a model that can recapitulate the clock synchrony that we observe *in vivo*, and it is unclear how this model should be altered such that this is achieved. Additionally different morphogenetic processes in the PSM may appear or disappear over the course of development, for instance in Zebrafish the elongation of the PSM

is majorly driven early in somitogenesis by cell ingression [113, 120, 204], however by the end of somitogenesis this ingression has ceased and compaction-extension of the tissue is the major driver of PSM elongation [124]. Due to computational constraints I have not been able to explore such variations in the deployment of specific morphogenetic processes, but it is possible that for the parameters describing the Zebrafish segmentation clock, as I assume here, it may be unrealistic to assume clock synchrony can be achieved in the presence of both cell ingression and compaction-extension of the PSM.

What is novel about my results here is that I have shown that the navigability of the fitness landscape for PSM morphogenesis is enhanced by the clock coupling strength and constrained by the magnitude of the coupling delay between cells. In my previous chapter I was able to predict this based on comparisons of how the synchrony landscape for varying κ and τ changed across three different scenarios for cell ingression, however it is only with these results that I have been able to show that this holds for morphogenesis of the PSM in general. It is unclear how these two parameters can co-evolve with the morphogenesis of the PSM and if there exist specific manifolds within parameter space for this model along which synchrony of the clock is preserved. These are likely to be highly significant for the evolvability of somitogenesis as evolution is constrained to these manifolds and as such their structure could create biases in evolution.

Finally, I note that there is insufficient data to determine to what degree the space of possible PSM morphogeneses has been explored by evolution, so it is difficult to evaluate whether my predictions about this landscape are supported by experimental studies. Hopefully as imaging techniques improve, the diversity of PSM morphogenesis across the vertebrate clade will become clearer, as will whether the morphogenesis of the PSM is as freely evolvable as I predict here.

3.3.1 Future work

There are several open questions here that I would like to address in future work. The most pertinent is resolving the structure of the parameter space of PSM morphogenesis, clock

coupling strength, and coupling delay, that preserves synchronous clock dynamics. It will be interesting to compare these computational results with experimental data to determine whether clock dynamics and PSM morphogenesis in extant species conform to this parameter space, to determine whether this has indeed influenced the evolution of somitogenesis.

Another interesting avenue of research is to explore the evolutionary potential of somitogenesis when either the clock is more evolvable than PSM morphogenesis, or vice versa. Although it is unclear if this is the case in nature, it seems plausible that one of these two processes is more evolvable than the other, and if this were the case it could have interesting implications. For instance, given my finding that the evolution of clock parameters allows evolution to explore the parameter space for PSM morphogenesis (figure 3.6A, B), a weakly evolvable clock could force young evolutionary lineages to diversify their vertebral number using a constrained set of morphogenetic processes over short timescales, with gradual evolution of the clock permitting greater exploration of morphogenesis at longer timescales. It may be possible to determine the minimal values of κ and τ such that these constraints are broken and this landscape is fully navigable.

Indeed, another question would be how this navigability is affected by the way in which model parameters are mutated. Here, as it is unknown whether any one of cell ingression, cell motility, cell division, or compaction-extension is more or less evolvable than the others, in my random walk simulations I have adjusted the magnitude of change in each parameter under mutation based on the magnitude of each parameter in Zebrafish. However as above, it is likely that each of these parameters shows differing evolvability in nature. An interesting future avenue of study would be to adjust the magnitude of mutation in each of these parameters and see how this affects the navigability of the fitness landscape for synchrony. It may be that certain parameters have a lesser effect on navigability than others, and so may be less relevant to evolvability.

3.4 Methods

3.4.1 Modelling PSM morphogenesis

Here I adapt the model of Uriu *et al.* used in chapter 1 to simulate up to four morphogenetic processes co-occuring in the PSM.

Cell ingression

In my previous analysis, I studied two different scenarios of cell ingression; one where cells with constant phase ingressed into the PSM from the dorsal surface of the posterior of the PSM, and one where this co-occurred with ingression from the ventral surface of the lateral halves of the PSM (scenarios 'DP' and 'DP+LV', chapter 1 figure 2). To allow incremental evolution between these two scenarios, I introduce the parameter $x_g \in [x_d, x_a + L - R - r_T]$ which divides the PSM into a region $x < x_g$ in which cells are added at random positions and with random phase $\theta \in [0, 2\pi)$, and $x_a + L - R - r_T \geq x \geq x_g$ where new cells are added with constant phase $\theta = 0$ and on the ventral surface of the PSM. For $x \geq x_a + L - R - r_T$ cells are added onto the dorsal surface.

For simulations where the PSM is reduced to a single tubular domain, for $x \geq x_g$ cells are added on the dorsal surface of the PSM with constant phase, and for $x_d \leq x \leq x_g$ cells are added at random positions within the PSM with random phase.

Tissue compaction

To allow variation of the overall rate of tissue compaction, and to reduce the dimensionality of parameter space, I introduce the parameter α which scales the rate that tissue length L, anterior of the PSM x_a, PSM radius r_T, tissue density ρ, and cell diameter d_c by a constant value. Explicitly, the equations describing the evolution of these variables are:

$$L(t) = \begin{cases} L_0 & s(t) < 16, \\ L_0 - \alpha \cdot m_a \cdot (s(t) - 16) & s(t) \geq 16, \end{cases} \tag{3.1}$$

$$x_a(t) = \begin{cases} x_{a0} & s(t) < 16, \\ x_{a0} + \alpha \cdot m_a \cdot (s(t) - 16) & s(t) \geq 16, \end{cases} \tag{3.2}$$

$$r_T(t) = \begin{cases} r_0 & s(t) < 16, \\ r_0 - \alpha \cdot m_r \cdot (s(t) - 16) & s(t) \geq 16, \end{cases} \tag{3.3}$$

$$\rho(t) = \begin{cases} \rho_{t0} & s(t) < 16, \\ \rho_{t0} + \alpha \cdot m_\rho \cdot (s(t) - 16) & s(t) \geq 16, \end{cases} \tag{3.4}$$

and

$$d_c(t) = \begin{cases} d_{c0} & s(t) < 16, \\ d_{c0} - \alpha \cdot m_{d_c} \cdot (s(t) - 16) & s(t) \geq 16; \end{cases} \tag{3.5}$$

where $s(t)$ is the number of somites at time t, and this function along with all other parameters are as defined in chapter 1. As previously, if the length of the PSM $L(t) \leq x_d + R + r_T$ the simulation is terminated, as the PSM is now too short to permit cell addition.

Tissue geometry

Here, two different geometries for the PSM are defined. One, a horseshoe-shaped domain similar to that used previously in chapter 1, was achieved by using the same definitions of a tissue boundary force as outlined in that chapter, which are omitted here for brevity.

In the second geometry, the PSM is comprised of an $X_C - x_a$ long cylindrical subdomain with

radius r_T, bounded at the posterior end by a half-spherical subdomain with radius r_T. In these simulations the boundary force confining a cell i within the cylindrical subdomain of the PSM is given by

$$\mathbf{F}_b(\mathbf{x}_i) = \begin{pmatrix} 0 \\ -\mu_b e^{-\frac{\delta y}{r_b}} \cos q_i \\ -\mu_b e^{-\frac{\delta z}{r_b}} \sin q_i \end{pmatrix}, \tag{3.6}$$

where as before

$$\delta y = |(r_T - r_i) \cos q_i|, \tag{3.7}$$

$$\delta z = |(r_T - r_i) \sin q_i|, \tag{3.8}$$

for the cylindrical coordinates describing the position of cell i, r_i and q_i. As previously the magnitude of the boundary force is denoted by μ_b and the parameter r_b determines the length-scale of the boundary force. If cell i is within the half-spheroidal domain at the posterior, the position of i, $\mathbf{x}_i$, is given by

$$\mathbf{x}_i = \begin{pmatrix} X_c + r_i \cos p_i \cos q_i \\ Y_c + r_i \sin p_i \cos q_i \\ Z_c + r_i \sin q_i \end{pmatrix}. \tag{3.9}$$

for the spherical coordinates $r_i \in [0, r_T]$, $p_i \in [-\pi/2, \pi/2]$, and $q_i \in [0, \pi/2]$. One can then define, as before,

$$\delta x = |(r_T - r_i) \cos p_i \cos q_i|, \tag{3.10}$$

$$\delta y = |(r_T - r_i) \sin p_i \cos q_i|, \tag{3.11}$$

$$\delta z = |(r_T - r_i)\sin q_i|, \tag{3.12}$$

describing the distance of cell i from the tissue surface in the x, y, and z directions, respectively. The boundary force in the half-spheroid domain is therefore

$$\mathbf{F}_b(\mathbf{x}_i) = \begin{pmatrix} -\mu_b e^{-\frac{\delta x}{r_b}} \cos p_i \cos q_i \\ -\mu_b e^{-\frac{\delta y}{r_b}} \sin p_i \cos q_i \\ -\mu_b e^{-\frac{\delta z}{r_b}} \sin q_i \end{pmatrix}. \tag{3.13}$$

3.4.2 Random walk exploration of parameter space

To allow exploration of parameter space, $N = 100$ parallel walkers were initiated from one of a set of initial conditions (detailed below), and were allowed to 'walk' for a maximum of 150 iterations. At each iteration, a new parameter set $\mathcal{T}' = (x_g', h', T_G', \alpha')$ is proposed to the walker where $x_g' = x_g + \epsilon_{x_g}$, $h' = h + \epsilon_h$, $T_G' = T_G + \epsilon_{T_G}$, and $\alpha' = \alpha + \epsilon_\alpha$, for some randomly chosen $\epsilon_{x_g} \in [-10, 10]$, $\epsilon_h \in [-1, 1]$, $\epsilon_{T_G} \in [-100, 100]$, and $\epsilon_\alpha \in [-0.1, 0.1]$, and where $\mathcal{T} = (x_g, h, T_G, \alpha)$ is the parameter set for the previous iteration or the initial condition. The walker performs a simulation and then 'accepts' $\mathcal{T}'$ if the simulation yields oscillations with synchrony $r \geq 0.7$ at the end of the simulation. If $\mathcal{T}'$ is 'rejected', then a new $\mathcal{T}'$ is proposed and tested in the same manner. This is repeated up to 5 times until a proposed $\mathcal{T}'$ is accepted. If after 5 repetitions no acceptable $\mathcal{T}'$ has been found the walk is terminated. If $\mathcal{T}'$ is accepted, the walk proceeds to the next iteration and a new parameter set $\mathcal{T}''$ is proposed to the walker.

It is worth noting that here I assume a uniform distribution for the various ϵ terms. While different distributions could be used, a uniform distribution has been chosen because it is unknown what the underlying distribution would look like for 'mutation' in these four parameters.

Pseudocode implementation

Here I have provided a pseudocode description of the algorithm used to carry out random walk explorations of parameter space:

$\mathcal{T} \leftarrow \mathcal{T}_0$

$n \leftarrow 1$

while $n \leq 150$ **do**

$m \leftarrow 1$

while $m \leq 5$ **do**

if $r(\mathcal{T}) \geq r_c$ **then**

$\mathcal{T} \leftarrow \mathcal{T}'$

$m \leftarrow 6$

end if

$\mathcal{T} \leftarrow \mathcal{T}'$

$m \leftarrow m + 1$

end while

$n \leftarrow n + 1$

end while

here $\mathcal{T}, \mathcal{T}', r_c$ are defined as above. The function $r(\mathcal{T})$ measures the synchrony at the anterior of the PSM at the end of the simulation, using the same methodology as the previous chapter. $\mathcal{T}_0$ represents the initial condition for the walk, and the possible values for this are given below.

Initial conditions

Walkers were initiated from a set of initial parameters. To ensure that heuristic searches spanned the parameter space of the model, these were chosen to simulate any one of the four extremes possible:

1. **'Ingression':** Progenitor cell ingression is occuring along the ventral surface of the PSM, however there is no anterior-posterior change in cell motility, cell divisions are rare, and the PSM is not compacting.
2. **'Motility':** Progenitor cell ingression is occuring only at the posterior tip of the PSM, there is a steep anterior-posterior change in cell motility, cell divisions are rare, and the PSM is not compacting.
3. **'Division':** Progenitor cell ingression is occuring only at the posterior tip of the PSM, there is no anterior-posterior change in cell motility, cell divisions are common, and the PSM is not compacting.
4. **'Compaction-extension':** Progenitor cell ingression is occuring only at the posterior tip of the PSM, there is no anterior-posterior change in cell motility, cell divisions are common, and the PSM is compacting.

These parameter sets can be thought of as corresponding to the scenarios where PSM elongation is driven only by ingression of cellular material, anterior-posterior profiles of cell motility, cell division, or compaction-extension respectively. The parameter sets to which these four initial conditions correspond can be seen below in table 3.1.

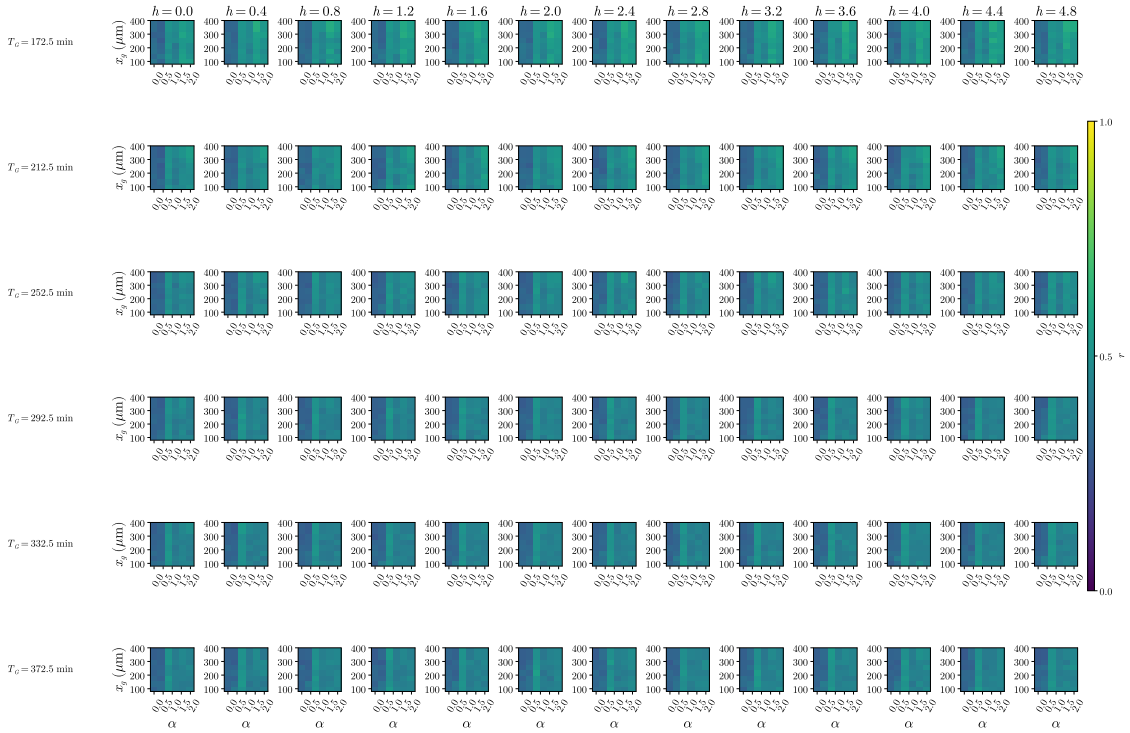
	Ingression	Motility	Division	Compaction-extension
x_g (μm)	100	X_c	X_c	X_c
h	0	3	0	0
T_G (mins)	1035	1035	172.5	1035
α	0	0	0	1

Table 3.1: Initial conditions for heuristic searches. Parameter sets corresponding to the different initial conditions used in figures 3.5 and 3.6. All other parameters, unless otherwise stated in the text, where as described in the previous chapter.

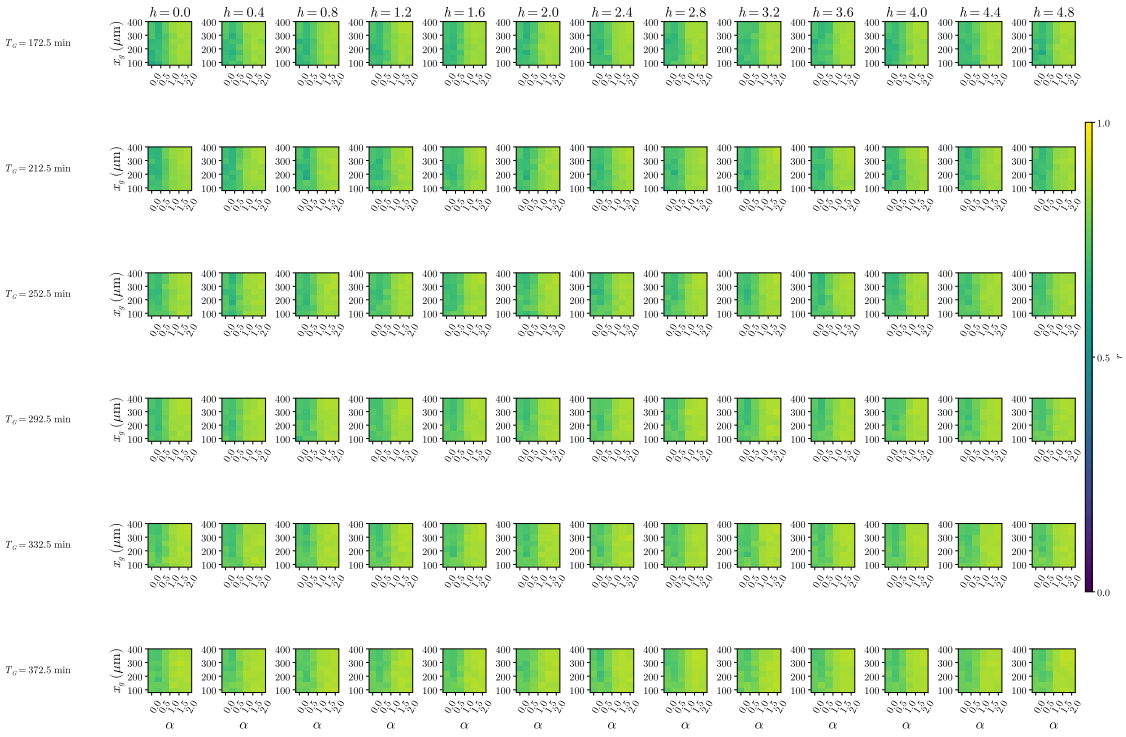
3.4.3 Implementation

The equations detailed in the previous chapter describing the evolution of clock phase and cell position were solved using an Euler-Maruyama method with timestep $dt = 0.01$ mins. Simulations were allowed to run for a maximum of 690 mins. As detailed in the previous chapter these simulations were performed using custom code written in Julia v1.8.2. The code for this model is open access and can be viewed at https://github.com/jewh/ModularityPSMClock.

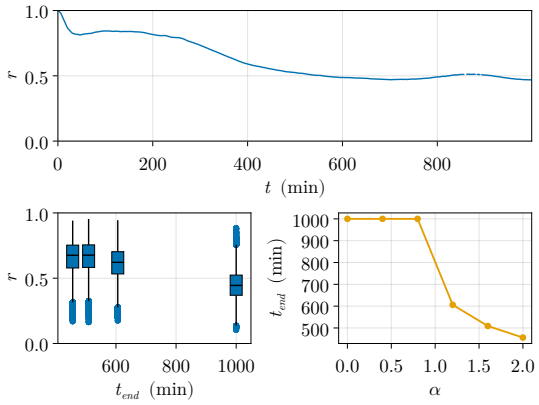
3.5 Supplementary material



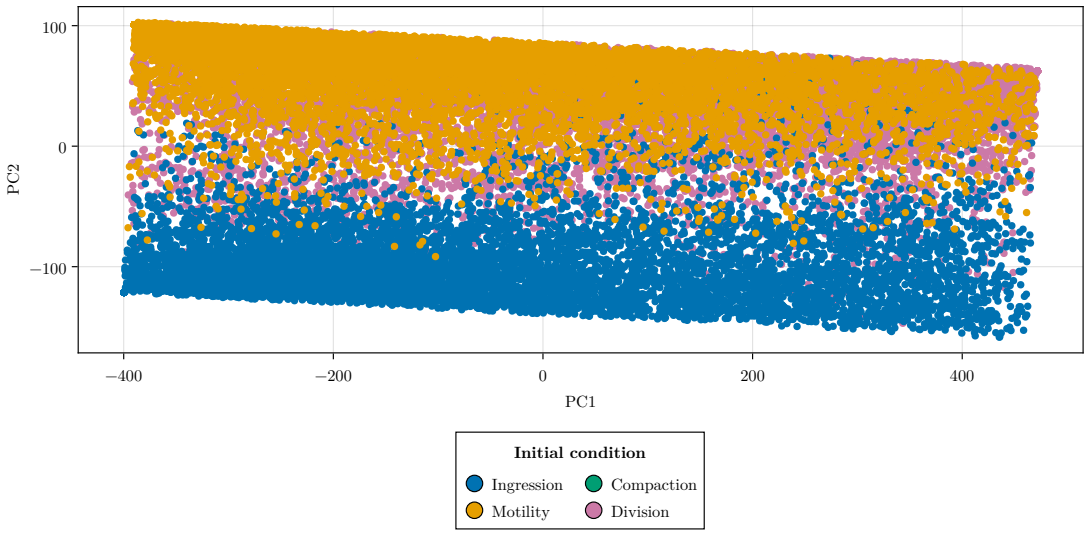
Supplementary Figure 3.1: Clock synchrony across the space of possible PMS elongation strategies, $\kappa = 0.035$ min^{-1}. Synchrony of the clock r at the PSM anterior at the end of a simulation, for varying anteriormost position of ventral cell ingression x_g, steepness of cell motility profile h, duration of G1, S, and G2 phases of the cell cycle T_G, and rate of tissue compaction α. Each pixel represents the median value of $N = 50$ simulations.



Supplementary Figure 3.2: Clock synchrony across the space of possible PMS elongation strategies, $\kappa = 0.14$ min^{-1}. Synchrony of the clock r at the PSM anterior at the end of a simulation, for varying anteriormost position of ventral cell ingression x_g, steepness of cell motility profile h, duration of G1, S, and G2 phases of the cell cycle T_G, and rate of tissue compaction α. Each pixel represents the median value of $N = 50$ simulations.



Supplementary Figure 3.3: Effect of compaction-extension rate α on synchrony. Top: Median synchrony over time for $\alpha = 0$, $N = 31200$ simulations. **Bottom left:** Synchrony r at the time at which a simulation is terminated, t_{end}. **Bottom right:** Time at which a simulation is terminated, t_{end}, for varying rate of compaction-extension α.



Supplementary Figure 3.4: Exploration of parameter space by random walks. PCA projection of the results of $N = 400$ independent random walks ($N = 100$ from each initial condition), for instantaneous coupling ($\tau = 0$ mins). Individual parameter sets visited on a walk are shown as circles and coloured according to the initial condition for that walk.

4

Quantitative biology of the segmentation clock in the Lake Malawi Cichlidae *Rhamphochromis* sp. 'chilingali' and *Astatotilapia calliptera*

Abstract

The Lake Malawi Cichlids *Rhamphochromis* sp. 'chilingali' and *Astatotilapia calliptera* have rapidly evolved different numbers of vertebrae (30 in *A. calliptera*, 36 in *R.* sp. 'chilingali') over the past 460,000 years with minimal changes to the genome. They therefore present an excellent opportunity to study the evolvability of somitogenesis, by suggesting which components of somitogenesis change most rapidly in evolution to generate diversity in vertebral number, and which properties of somitogenesis enable this. In this chapter, I characterise the dynamics of somitogenesis in these species, and establish a range of experimental procedures to allow the quantification of the noise, coupling strength, and coupling delay of the segmentation clock in these species using chemical inhibition of notch signalling. I find that the difference in vertebral number between the two species is driven by changes in the duration of somitogenesis and not in its frequency, and by treating explanted pre-somitic mesoderm tissue from both species with the notch inhibitor DAPT I find that both species exhibit similar values for coupling strength, coupling delay, and noise. The noise of the clock is similar to that reported from Zebrafish, and when the difference in developmental rate between cichlids and Zebrafish is accounted for, so is the delay. The estimated coupling strength for cichlids is a quarter of that reported for Zebrafish, however this estimate is imprecise and modelling suggests that the true value is likely nearer to that of Zebrafish, suggesting that cichlids may exhibit developmental modularity of the clock and PSM morphogenesis.

Author's note: *Not all the work presented in this chapter is my own. The protocol for culture of cichlid explanted PSM tissue is based on a protocol developed by my colleague Dr Georgina Stooke-Vaughan. All other work presented here is my own.*

4.1 Introduction

While mathematical models of development are a powerful tool for understanding evolvability [1, 7], it is also possible to understand evolvability via comparative studies of development focusing on closely related species, to suggest which components of the developmental process are most flexible under evolutionary pressure. While understanding the influence of development over evolvability is a major research goal within evo-devo, most evo-devo research focuses on organisms that are significantly diverged, and so it is difficult to identify which differences in

development are causal or neutral in the evolution of a trait, or the order in which changes in development occur over evolutionary time. This is particularly true of somitogenesis, which has largely focused on comparative studies between teleost fish, reptiles, birds, and mammals [45, 46, 100], all of which diverged in the middle Silurian approximately 400 million years ago [208]. Additionally there has only been a minor exploration of which processes change in evolution to create a diversity of vertebral number (discussed more thoroughly in the literature review at the beginning of this thesis) [46].

With this in mind, the Lake Malawi cichlids *Rhamphochromis* sp. 'chilingali' and *Astatotilapia calliptera* (see figure 4.1B) present an excellent model system in which to study the evolvability of vertebral number in vertebrates. The two species have evolved relatively rapidly, diverging ~460,000 years ago, and share approximately ≥99.75% of their genome [209]. Despite this high degree of conservation, the two species show a large difference in the number of vertebrae in the vertebral column, with *A. calliptera* possessing on average 30 vertebrae, and *R.* sp. 'chilingali' 36 (figure 4.1B) (CV Bucklow, *pers. comm.*). Comparative study of somitogenesis in these two species has the potential to reveal which components of somitogenesis, such as gene expression or morphogenetic processes driving elongation of the PSM, are capable of rapid change in evolution to generate diversity in vertebral number, and thus reveal sources of evolvability.

In the previous two chapters I have shown that the two central components of somitogenesis, its duration and frequency, exhibit developmental modularity, and I hypothesise that this may have facilitated the evolution of diverse vertebral numbers in vertebrates. However, this prediction was based on data from Zebrafish [49], and it is unclear if this property holds in general for vertebrates, or if it has facilitated the evolution of divergent vertebral numbers in a vertebrate clade. Due to their close relatedness and divergent vertebral number, *A. calliptera* and *R.* sp. 'chilingali' present an excellent model system to test this hypothesis. In this chapter, I examine the developmental origins of the difference in vertebral number between these two species, and parameterise a model of the segmentation clock in each species to see if this clade exhibits modularity of the frequency and duration of somitogenesis, and if this has facilitated

the evolution of divergent vertebral number.

4.1.1 The natural history of Lake Malawi Cichlids

As neither species is often used in embryology or evo-devo it is likely that their biology is perhaps unfamiliar to the reader, so I have included here a short introduction to cichlid embryology and their biology. The eponymous Lake Malawi (also known as Lake Nyasa) is a large rift valley lake in East Africa, home to at least 700 species of haplochromine cichlids [210]. This 'species flock' comprises an adaptive radiation, with all extant species being descended from an *Astatotilapia*-like ancestor that lived approximately 460,000 years ago [209]. Similarly dramatic and diverse radiations of haplochromines have also evolved in the other great east african lakes, Lakes Tanganyika and Victoria, and the east african cichlids are a canonical example of adaptive radiations and rapid evolution [211]. The cause of cichlids' unusually rapid diversification is still an open question and is an area of active research [211–213].

Astatotilapia calliptera is a small (∼10 cm long) haplochromine cichlid from the Lake Malawi catchment. It is a widespread generalist and found in shallow weedy riverine or lacustrine habitats throughout the catchment [214, 215]. Like many haplochromine cichlids, this species exhibits sexual dimorphism, with males exhibiting a bright colourful breeding dress, and females a more subdued sandy hue (figure 4.1B, C). Presumably due to its generalist nature and affinity for small shallow bodies of water, the species is very amenable to laboratory culture and breeds readily in captivity.

Rhamphochromis sp. 'chilingali' is an undescribed species of *Rhamphochromis* known from Lake Chilingali, a satellite lake of Lake Malawi [216]. Many species of this genus exhibit similarly high numbers of vertebrae, however they generally occur in the pelagic zone of the lake and can grow to large sizes (e.g. *R. woodi*, 40 cm [217]), making them troublesome to maintain in captivity. Unlike most of its cogeners, *R.* sp. 'chilingali' is small (average length 10.6 cm in the wild [216], but can grow to twice this size in captivity), and naturally occurs in shallow water habitats, therefore making it more amenable to laboratory culture. Throughout its history

Lake Malawi has undergone cycles of rising and falling water levels, creating 'satellite lakes' with isolated populations of cichlids which can undergo allopatric speciation [216, 218]. Lake Chilingali, the satellite lake to which *Rhamphochromis* sp. 'chilingali' is endemic, is a shallow lake on the western shore of Lake Malawi [214]. In recent years the lake has suffered from encroachment by surrounding vegetation and sedimentation due to the construction of a dam, and consequently *R*. sp. 'chilingali' is now thought to be extinct in the wild and exists only in research facilities and aquaria [214]. Other haplochromine taxa recorded from Lake Chilingali include an endemic species of *Lethrinops*, *Pseudocrenilabrus philander*, *Serranochromis robustus*, and *Astatotilapia calliptera*; as well as a variety of other lacustrine fish species such as catfish, cyprinidae, and mormyridae [214]. *R.* sp. 'chilingali' breeds readily in captivity, however not so reliably as *A. calliptera* and can be sensitive to changes in culture conditions such as diet.

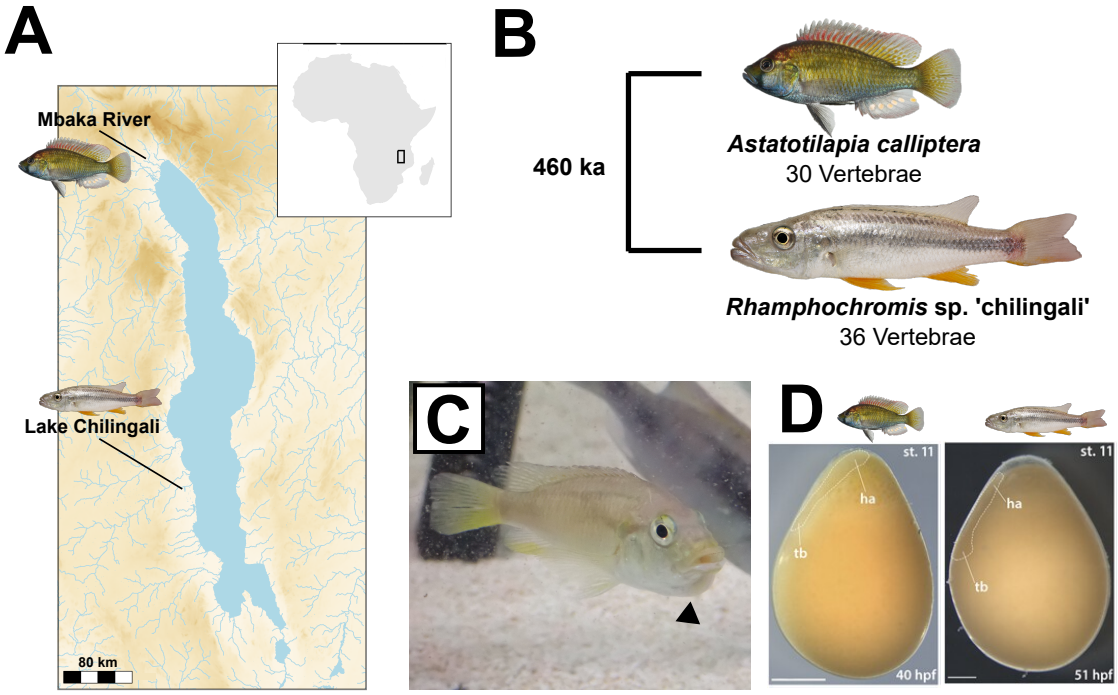


Figure 4.1: A cichlid model system for the rapid evolution of vertebral number, *Astatotilapia calliptera* and *Rhamphochromis* sp. 'chilingali'. A Map of the Lake Malawi watershed, with an inset showing its position within the continent of Africa. The origins of the fish used in this study are annotated, with the stock of *Astatotilapia calliptera* being descended from fish from the Mbaka river in Tanzania at the northern end of the lake, and *Rhamphochromis* sp. 'chilingali' being endemic to the small Lake Chilingali in Malawi on the eastern coast of the lake. Photo credits: *A. calliptera*, Howard Jelks, U.S. Geological Survey. *R.* sp. 'chilingali', Aleksandra Marconi. **B** Males of each species studied here, *A. calliptera* and *R.* sp. 'chilingali', with the number of vertebrae annotated. Photo credits: *A. calliptera*, Howard Jelks, U.S. Geological Survey. *R.* sp. 'chilingali', Aleksandra Marconi. **C** Mouthbrooding female *A. calliptera*. The diagnostic distension of the lower jaw indicating mouthbrooding is highlighted with a black arrowhead. **D** Somitogenesis-stage embryos of *A. calliptera* (left) and *R.* sp. 'chilingali' (right). The outline of the embryo is shown by a white dashed line. The head anlage and tailbud are annoated with 'ha' and 'tb' respectively. White scale bar indicates 1 mm. These images of cichlid embryos were taken from [146] and reproduced under a Creative Commons CC BY license; © 2023 The Authors. Evolution & Development published by Wiley Periodicals LLC.

Cichlid Embryology

Both species are mouthbrooders, with females protecting embryos in their mouths (figure 4.1C) from spawning onwards until the fry are fully developed. Cichlid eggs are large and yolky (figure 4.1D), reflecting the sizeable energy contribution of a cichlid mother to her offspring, with the yolk fully occupying the volume of the chorion [146, 219, 220], unlike most other teleost species [74, 144, 157, 221]. The clutch size of mouthbrooding haplochromines is also

typically lower than that of free-spawning species, associated with a reduction in size of the left ovary in this group [222]. Nevertheless, the clutch size of both species treated here is large enough to be of practical use for embryology; a typical clutch of *A. calliptera* is ∼30 embryos, or ∼50 for *R.* sp. 'chilingali'. The availability of genetic tools and laboratory techniques for cichlid embryology is currently fairly limited, but is rapidly growing [211], and a staging series for the development of *A. calliptera* and *R.* sp. 'chilingali' has recently been published [146], as well as a protocol for CRISPR/Cas9 gene editing in *A. calliptera* [223]. There is also a body of literature addressing embryology in Nile Tilapia (*Oreochromis niloticus*), a cichlid, due to its importance in aquaculture [219, 224, 225].

Stage	Zebrafish (hpf)	Medaka (hpf)	Cichlids (hpf)
Blastula	3	8	12
Gastrula	6	15	24
Segmentation	16	41	48
Pharyngula	24	96	72

Table 4.1: The timing of teleost development. Approximate time of appearance of various stages of early development in Lake Malawi Cichlids (based on data from *Astatotilapia calliptera*, *Rhamphochromis* sp. 'chilingali', and *Tropheops* sp. 'mauve' [146]), compared to the model teleosts Zebrafish (data from [74]) and Medaka (data from [157])

Like most teleost species the pace of development of Lake Malawi cichlids is slower than that of Zebrafish (Table 4.1), the standard model teleost. This is also true of somitogenesis. Somite formation in both *A. calliptera* and *R.* sp. 'chilingali' has been studied by Marconi *et al.*, who found that the rate of somitogenesis differed slightly between the two species: 0.814 somites·hr^{-1} and 0.886 somites·hr^{-1} respectively [146], compared to 2.4 somites·hr^{-1} for Zebrafish [76]. They also found that the minimum number of somites formed at any point in somitogenesis in either species was four, suggesting that these four somites form simultaneously at the beginning of somitogenesis [146]. Marconi *et al.* also reported that the overall duration of somitogenesis, defined here as the time from the appearance of the neural keel and the end of gastrulation to the cessation of somitogenesis, was the same in both species, approximately 35 hrs [146]. As the measured rates and this predicted duration cannot account for the difference in somite numbers between the two species, Marconi *et al.* suggested that somitogenesis in *R.* sp. 'chilingali' is

much more rapid than *A. calliptera* in its early stages, and that this accounts for the difference in somite number between the two species [146]. However, the data for these early stages is limited in their dataset, and statistical support for this inference is weak, and it remains unclear how the somitogenesis of the two species differs to create the difference in vertebral (and somite) number.

4.1.2 Quantification of the segmentation clock using chemical inhibitors

In the preceding two chapters, the dynamics of the segmentation clock have been approximated by a Kuramoto-type model of a network of coupled oscillators [54], in order to explore the relationship of the segmentation clock and morphogenesis of the PSM. One of the attractive features of such models for evo-devo is that they assume little about the underlying gene regulatory architecture of the segmentation clock, which can vary in evolution [45]. Additionally, the parameters for such models are often relatively simple to measure from data such as *in situ* hybridisations (ISHs) and live-imaging of developing embryos [49, 139, 142], allowing measurement of model parameters in non-model species.

Using morpholino injections and treatments with the γ-secretase inhibitor DAPT to disrupt notch signalling, Riedel-Kruse *et al.* were able to parameterise a kuramoto model to describe the Zebrafish segmentation clock [49]. This was achieved by performing *in situ* hybridisations (ISHs) on treated Zebrafish embryos targeting genes marking segment boundaries, and scoring each embryo according to their anterior limit of defect (ALD), the first somite which exhibits disrupted gene expression due to inhibitor or morpholino treatment of the embryo [49, 90]. As each cell is assumed to be a noisy autonomous oscillator, the synchrony of the segmentation clock is expected to decay exponentially in the absence of notch-delta signalling, and posterior to the ALD all somites are expected to be defectively patterned. By varying the quantity of chemical inhibitor or morpholino delivered to the embryo and recording the ALD, it is possible to fit a relationship between the ALD and inhibitor quantity that estimates the values of κ, the coupling strength (or equivalently a proxy for the strength of notch-delta signalling), and

$D\theta$, the noise of the clock [49]. Additionally, Herrgen *et al.* were able to estimate the coupling delay τ, the time taken for information about the phase of the segmentation clock in a cell to be passed to the clock in its neighbour, by measuring the decrease in clock frequency upon loss of notch signalling due to DAPT treatment [139].

4.1.3 Aims and objectives

The first aim for this chapter is to resolve the origin of the difference in somite number between *A. calliptera* and *R.* sp. 'chilingali'. Specifically, to determine if this is driven by changes in duration, changes in somitogenesis frequency, or a combination of the two. Additionally, I sought to sample early time points of somitogenesis in both species and test the prediction of Marconi *et al.* that *R.* sp. 'chilingali' forms somites more rapidly early in somitogenesis [146]. Understanding of this will inform future work on this species pair to understand the developmental origin of evolvability in vertebral number.

The second aim for this chapter is to measure the clock coupling strength κ, coupling delay τ, and clock noise $D\theta$ parameters for a kuramoto model of the segmentation clock in both species of cichlid, to see if the modularity of somitogenesis duration and frequency is a feature of species other than Zebrafish, and, depending on the results of the first aim, if this modularity has facilitated the evolution of divergent vertebral number. To do this I used small molecule inhibitors to disrupt notch signalling in the cichlid PSM, favouring this approach over morpholinos due to better quantitative precision of inhibitor delivery compared to microinjection of morpholinos, and the difficulty in obtaining 1–8 cell stage cichlid embryos appropriate for microinjection [226, 227].

4.2 Results

4.2.1 Dynamics of somitogenesis in *Rhamphochromis* sp. 'chilingali' and *Astatotilapia calliptera*

In order to determine the rate and duration of somitogenesis in *Rhamphochromis* sp. 'chilingali' and *Astatotilapia calliptera*, I fixed embryos of each species along a time course and studied the number of somites. The opacity and volume of the yolk, and the position of the embryo nested within the yolk with little to no vertical profile above the yolk surface (figure 4.1D), makes live-imaging of somitogenesis with standard brightfield microscopy (c.f. [76]) impossible. The embryo can be stained with vital dyes and visualised within the chorion [146], however imaging the full process of somitogenesis is challenging as the growing embryo curves around the yolk meaning the embryo must be repositioned multiple times during imaging (if one is using standard imaging techniques). Therefore I decided to take an approach relying on fixing embryos and counting somites along a timecourse. To allow comparison of timing between the two species, I collected clutches at the one-cell stage. After extraction, embryos were cultured in a heated room at 28°C and fixed at hourly intervals within the desired time window to capture somitogenesis. Once fixed, embryos were dissected and the nuclei were stained with DAPI, then somites were counted under a fluorescence microscope.

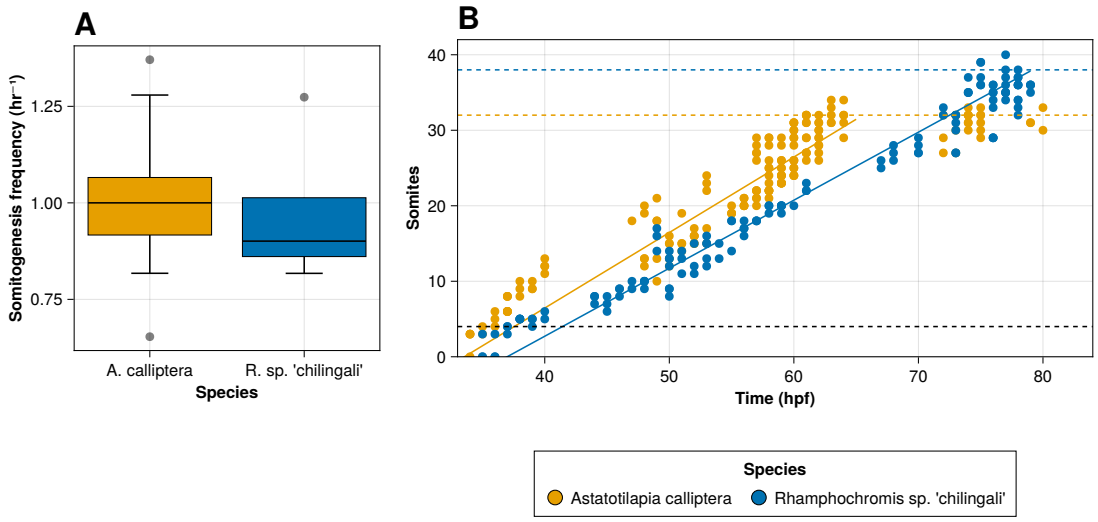


Figure 4.2: Dynamics of somitogenesis in *Rhamphochromis* sp. 'chilingali' and *Astatotilapia calliptera*. **A** Linear rate of somite formation for individual clutches of cichlid embryos. N = 9 clutches for *A. calliptera*, N = 4 clutches for *R.* sp. 'chilingali'. **B** Somite formation over time (hpf). Trendline for *A. calliptera* is $S = t - 33.547$, trendline for *R.* sp. 'chilingali' is $S = 0.901t - 33.316$. The minimum number of somites reported by Marconi *et al.* (4) is annotated with a black dashed line. The typical maximum number of somites formed in each species, 32 and 38 for *A. calliptera* and *R.* sp. 'chilingali' respectively (S. Taylor, *pers. comm.*), are annotated with dashed lines coloured according to the species. N = 178 individual embryos for *A. calliptera*, N = 135 individual embryos for *R.* sp. 'chilingali'.

The resulting data are displayed in figure 4.2B. I struggled to achieve a linear regression to the data in both species that agreed with the distribution of rates within clutches (figure 4.2A), likely due to noise created by uncertainty regarding the true timing of fertilisation (see methods). Thus, I estimated the species-specific rate of somitogenesis based on the median of rates of somitogenesis within clutches (figure 4.2A), specifically 1.000 somites·hr^{-1} (95% CI: [0.686 somites·hr^{-1}, 1.352 somites·hr^{-1}]) in *A. calliptera* and 0.901 somites·hr^{-1} (95% CI: [0.822 somites·hr^{-1}, 1.248 somites·hr^{-1}]) in *R.* sp. 'chilingali'. These median values do not show any statistically significant difference (Mann-Whitney U-test, $p = 0.5035$, Mann-Whitney U-statistic $= 23.0$). Using these estimated rates I fitted linear trendlines for each species, $S = t - 33.547$ for *A. calliptera* and $S = 0.901t - 33.316$ for *R.* sp. 'chilingali', where S is the number of somites and t is the time in hpf. Additionally, unlike Marconi *et al.*, I observe embryos with three somites in each species, but no less (figure 4.2B), suggesting that three, not four, somites form rapidly at the beginning of somitogenesis. Based on these estimates we can

predict that somitogenesis lasts ∼29 hrs in *A. calliptera* and 10 hrs more (∼39 hrs) in *R.* sp. 'chilingali'. We can also predict that somitogenesis in *R.* sp. 'chilingali' initiates at $\sim$ 40 hpf whereas *A. calliptera* initiates somitogenesis earlier, at approximately 36.5 hpf. This latter result suggests that gastrulation lasts longer in *R.* sp. 'chilingali' than in *A. calliptera*.

4.2.2 Disrupting clock synchrony using the γ-secretase inhibitor DAPT

After Riedel-Kruse *et al.*, I sought to use the γ-secretase inhibitor DAPT to disrupt notch signalling in the PSM and induce asynchrony of the clock [49]. Titrating embryos with increasingly saturating concentrations of DAPT can be used to estimate the clock coupling strength (referred to here as κ, in line with previous chapters), the delay in coupling (τ), and the phase noise of oscillators in the PSM ($D\theta$) [49, 139].

Analogous to the method of Riedel-Kruse *et al.*, I incubated cichlid embryos with their chorions intact in the presence of DAPT (see methods). Embryos of *A. calliptera* were collected at the blastula or gastrula stage (12–24 hpf) and incubated under DAPT until pharyngula stage (4 dpf) in order to observe defects in segment patterning (assessed via xirp2a expression, in line with previous studies [49, 54]). Embryos treated with DAPT developed outwardly normally and the migration and formation of melanophores from the mid-somitogenesis stage (48 hpf) onwards appeared initially normal. However at 3 dpf treated embryos showed a reduction in the quantity and size of melanophores across the yolk and trunk body. By 4 dpf this difference was stark, and DAPT-treated embryos were obvious to the naked eye. I observed this trend across a variety of concentrations (15 µM, 20 µM, 25 µM, 30 µM, 50 µM, 100 µM), and across these concentrations the severity of the phenotype correlated with the concentration of DAPT. A knock-down in melanophores is an expected effect of DAPT based on treatments in Zebrafish, where DAPT treatment has been shown to kill melanophores [228, 229], suggesting that DAPT is able to successfuly inhibit γ-secretase and disrupt notch signalling within the cichlid embryo.

To quantify the ALD upon tratment with DAPT I studied the expression of the segment boundary gene xirp2a using Hybridisation Chain Reaction v3.0 (HCR) [230]. Examination

of xirp2a expression in 4 dpf embryos revealed no defects in segment patterning along the trunk body, although segment boundaries appeared to become disordered in the most posterior two somites (figure 4.3, white arrowhead). Additionally examination of deltaC expression, a cyclically expressed segmentation clock gene that exhibits a 'salt and pepper' expression pattern when notch signalling is disrupted [90], showed no obvious defects in the PSM (figure 4.3). Some possible desynchronisation of deltaC oscillations was observed in the posterior of the PSM (figure 4.3, white arrowhead), however deltaC expression in the somite was unchanged in the DAPT treatment group versus the control, with deltaC expression being restricted to the posterior of the somite (figure 4.3). This suggests that somite polarity is unperturbed and therefore it does not appear that treatment with DAPT in this manner disrupts notch signalling in the PSM, despite the knock-down in melanophores observed from 3 dpf onwards.

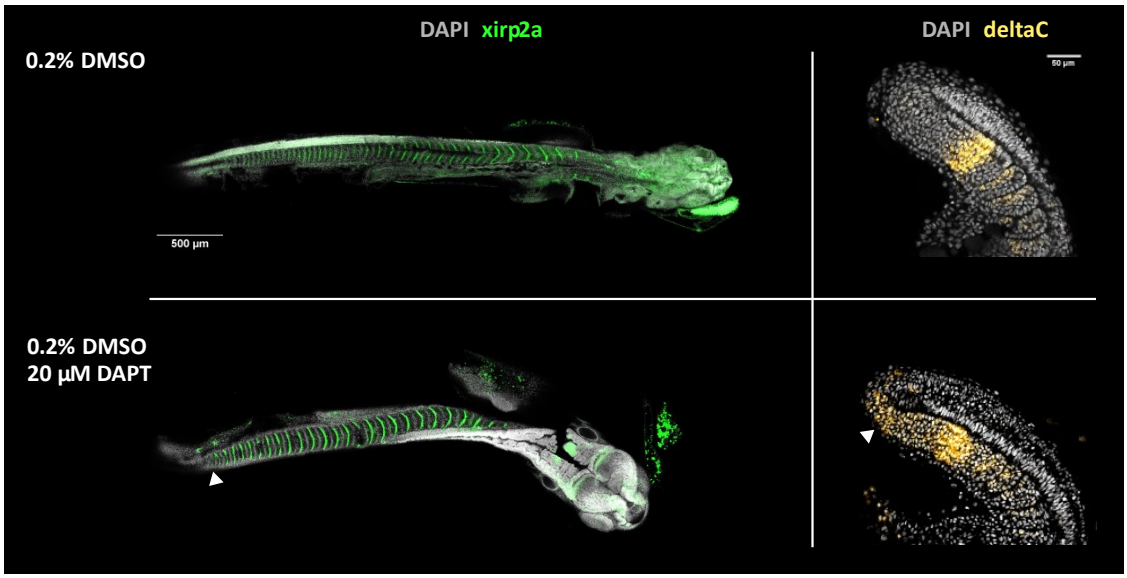


Figure 4.3: The effect of DAPT treatment on cichlid embryos. Left: HCR for xirp2a (green) in 4 dpf *Astatotilapia calliptera* embryos, either treated with just DMSO (top), or 20 µM DAPT (bottom), with the chorion intact. **Right:** HCR of deltaC (yellow) in 2 dpf *Astatotilapia calliptera* embryos, either treated with just DMSO (top), or 20 µM DAPT (bottom), with the chorion intact.

While this could imply that notch signalling is dispensible for cichlid somitogenesis (as is the case for some arthropod species [231]), or at least that the clock is robust to the perturbation of notch signalling by DAPT, this may not necessarily be the case. The chorion and the membranes that surround the embryo thin over the course of development [224]. This is significant because

it could imply that the apparent lack of effect of DAPT on somitogenesis (which ends at approximately 3 dpf, figure 4.2), and the knock-down in melanophores observed after this point, are due to increasing penetrance of DAPT through a thinning chorion, rather than notch-independence of somitogenesis. Therefore, I sought to test this hypothesis by developing methods to dechorionate the embryos of Lake Malawi cichlids.

Chemical dechorionation of cichlid embryos

Mechanical dechorionation of the cichlid embryo is extremely challenging, as the chorion is thick and the yolk is pressed flush against it (figure 4.1D), and therefore any mechanical force sufficient to tear or penetrate the chorion is very likely to damage the yolk. This is almost always fatal to the embryo, as the high internal pressure of the yolk rapidly ejects material upon injury. The embryo then perishes due to lack of nutrition or infection from the fouling of the culture media. Therefore in order to avoid a need for mechanical dechorionation I attempted to establish a chemical dechorionation protocol for *A. calliptera* and *R.* 'chilingali', based on existing protocols for Nile Tilapia *Oreochromis niloticus* [224].

The first method I attempted was the use of hatching enzyme to induce dechorionation, as this leads to the least mortality out of the available chemical dechorionation protocols for Nile Tilapia [224]. To obtain hatching enzyme, one cultures pre-hatching embryos (approximately 5 dpf in cichlids [146]) until hatching, and retains the media. Early stage embryos can then be dechorionated by incubation under this media, which contains the enzyme(s) produced by embryos to break the chorion and induce hatching [224]. In general this method is unreliable as it depends on the quantity of embryos cultured and the volume of culture media, however for some species like Medaka it is possible to extract a concentrate containing the enzyme from pre-hatching embryos [232]. Unfortunately the number of embryos required ($\sim$5000 embryos for 15 mL of enzyme [233]) is prohibitive for practical use in cichlids. Additionally due to the large quantity of yolk in cichlid eggs, extraction of such a concentrate does not to appear to be practical, as the mass of yolky material prevents effective filtering of the concentrate [224].

Therefore for my experiments I obtained hatching enzyme by culturing a large number of embryos in a small volume around the moment of hatching.

Using the media from hatchling embryos I was able to successfully remove the chorion of gastrula-stage embryos after overnight incubations under hatching media at 28°C (after their overnight culture the embryos were thus at segmentation stage [146]). Hatching enzyme appeared to be cross-reactive between the two species, with dechorionation being achieved in cultures of *A. calliptera* treated with *R.* sp. 'chilingali' media, and vice versa. Additionally the reactivity of the hatching enzyme was not adversely affected by freezing at -20°C, similar to hatching enzymes obtained from other teleost species [224], with fresh hatching media frozen overnight then thawed the next day for use being capable of dechorionating embryos. However, this method was inconsistent and I was unable to replicate dechorionation on further attempts, or ensure a consistent availability of fresh hatching media.

I therefore turned to a pronase-based protocol for chemical dechorionation of cichlid embryos. Pronase is the term used to describe a preparation of proteinases with broad reactivity isolated from *Streptomyces griseus*, used in dechorionation protocols for a variety of species across the animal kingdom, including Nile Tilapia [224]. Using pronase I was able to dechorionate *A. calliptera* gastrulae (see supplementary methods for a full protocol) and culture these embryos overnight to segmentation stage (10 ss). Unfortunately most embryos did not survive, and I was only able to culture one untreated control embryo and one embryo treated with 50 μM DAPT, with all DMSO-treated embryos perishing. The cause of this mortality is unclear, and could perhaps reflect over-digestion of the chorion with pronase, or merely the fragility of such early embryos. Morrison *et al.* reported similar mortality in pronase-treated *O. niloticus* embryos, and found that embryos dechorionated at the gastrula stage often lost most of their yolks during culture [224]. The large and fragile cichlid yolk is easily damaged during handling and without mechanical support of the chorion it is likely that small cuts or abrasions during handling cause the yolk to collapse and disintegrate.

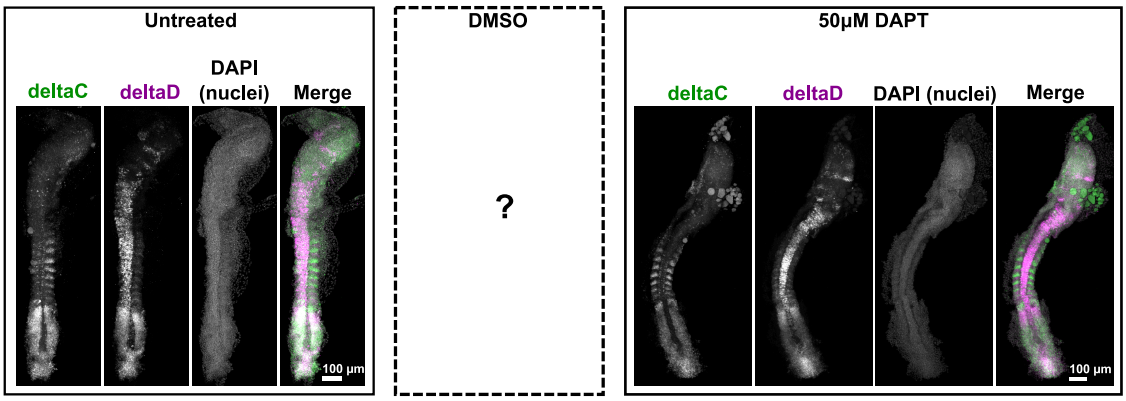


Figure 4.4: Effect of DAPT treatment on delta expression in dechorionated *Astatotilapia calliptera* embryos. HCR targeting deltaC (green) and deltaD (magenta), co-stained with DAPI (grey) to show embryo morphology. Left-hand panel shows an untreated embryo, and the right-hand an embryo that was cultured under 50 µM of DAPT. DMSO-treated embryos perished during culture and as such no embryo is shown in this panel.

Compared to the untreated control, the DAPT-treated embryo exhibited no differences in expression of deltaC, or its paralog deltaD (figure 4.4), and appeared to show normal expression of both genes at the posterior and anterior of the somite respectively, and no evidence of 'salt and pepper' expression in the PSM (figure 4.4). The expression pattern of deltaD is non-oscillatory in Zebrafish [77], however it is an essential component of notch signalling in somitogenesis and this pattern is disrupted upon perturbation of notch signalling [77, 90, 189, 200]. This result suggests that either the cichlid segmentation clock is robust to disruption by DAPT, or that it is not the chorion, but rather the vitelline envelope surrounding the embryo which creates a barrier to inhibitor entry.

Using detergents to increase small molecule penetrance to the embryo

Treatment of embryos with small molecule inhibitors can also be a challenge for other teleost species, such as Medaka, which possesses a thick chorion and layers of inner membranes that prevent the uptake of small molecules, even Li^{+} ions [234]. These inner membranes can be disrupted with weak solutions (0.01–0.03% v/v) of the detergent Triton X-100 (Tx), allowing uptake of small molecules into the embryo [234]. To test whether this can increase the penetration of DAPT into the cichlid embryo, I cultured *R.* sp. 'chilingali' blastulae according

to the standard protocol (see methods) overnight in the presence of either 0.01% or 0.03% Tx, as well as either 0.2% v/v DMSO, 0.2% v/v DMSO and 20 μM DAPT, or no additional treatment.

I found that after overnight culture 3/4 of untreated control embryos were correctly developed, whereas no embryos treated with either DMSO or DMSO and DAPT had survived. No embryo treated with 0.03% Tx had survived whereas 1/2 of embryos treated only with 0.01% Tx had correctly developed. This suggested that Tx was toxic to the embryos, and possibly that increased penetrance of DMSO into the embryo was also toxic. It is also possible that the toxicity of Tx is not only due to its disruption of biological membranes but the penetrance of tank media into the embryo, which is not isotonic to the intracellular environment of the embryo. To troubleshoot this, I performed a subsequent experiment varying the culture media used for overnight culture, using either tank media or the isotonic culture medium L-15, with or without 0.01% v/v Tx. I also varied the developmental stage at which treatment began, either at the 2–4 cell stage, or segmentation stage (two different clutches of *A. calliptera*), to test whether my earlier results perhaps reflected fragility of early-stage embryos to chemical treatment.

Treatment	Stage	Survival
Tank media + 0.01% Triton X-100	2–4 cells	0/3
Tank media	2–4 cells	1/3
L-15 + 0.01% Triton X-100	2–4 cells	0/3
L-15	2–4 cells	3/3
Tank media + 0.01% Triton X-100	Segmentation	0/3
Tank media	Segmentation	3/3
L-15 + 0.01% Triton X-100	Segmentation	0/3
L-15	Segmentation	3/3

Table 4.2: Effect of detergent Triton X-100 and culture medium on embryo survival. Survival of *Astatotilapia calliptera* embryos after overnight incubation at 28°C with a given treatment solution, sorted by developmental stage at which treatment began.

The results of this experiment are shown in table 4.2. We see that, after overnight culture, all Tx-treated embryos were dead, at either stage. Indeed I observed that the addition of Tx is rapidly fatal to cichlid embryos, as 3/3 of Tank media + 0.01% Tx embryos appeared to be dead after only 90 mins of culture. Therefore I abandoned this protocol in favour of alternative

means of disrupting synchrony of the segmentation clock.

Disrupting clock synchrony using the γ-secretase inhibitor LY411575

As an alternative to DAPT I trialled the use of an alternative γ-secretase inhibitor, LY411575, to disrupt synchrony of the segmentation clock. LY411575 has two attractive properties for this purpose, namely it is a more potent inhibitor of γ-secretase than DAPT [235], and it is more soluble in aqueous solution (personal observation), meaning that inhibitor is not lost from the solution to precipitation, and proportionately less inhibitor needs to penetrate the chorion and inner membranes to create an effect on the embryo.

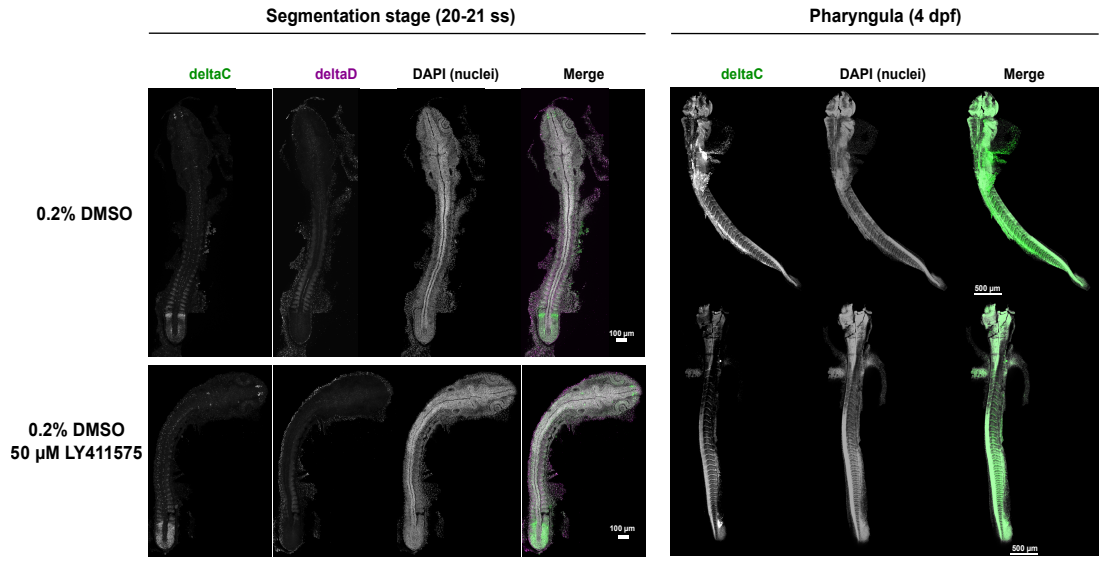


Figure 4.5: Effect of LY411575 on segment patterning in *Astatotilapia calliptera*. Left HCR of segmentation stage *Astatotilapia calliptera* embryos staining for the genes deltaC (green) and deltaD (magenta). Nuclei (DAPI staining) are shown in grey. Bottom row shows an embryo treated with LY411575, and the top row a control embryo treated with only DMSO. Maximum intensity z-projections of z-stacked images. **Right** HCR of late pharyngula (4 dpf) *Astatotilapia calliptera* embryos staining for the gene xirp2a (green). Nuclei (DAPI staining) are shown in grey. Bottom row shows an embryo treated with LY411575, and the top row a control embryo treated with only DMSO. Maximum intensity z-projections of z-stacked images.

Treating *A. calliptera* blastulae with 50 µM of LY411575 and culturing as previously (see methods), we observe a knock-down in melanophores at the pharyngula stage (4 dpf), but not at segmentation stage (2 dpf), as observed for DAPT. However, I observe no disruption

to xirp2a expression at 4 dpf or deltaC expression in segmentation-stage embryos (figure 4.5). Therefore similar to DAPT it seems likely that LY411575 is not penetrating the chorion and inner membranes surrounding the embryo and therefore is not a better alternative inhibitor in this case.

Quantifying clock noise and coupling strength in explanted PSMs

In order to circumvent the issues posed by the inner membranes and the chorion surrounding the embryo, I attempted to disrupt clock synchrony in explanted PSMs, to conclusively test if DAPT can perturb cichlid somitogenesis. To my knowledge this has not been attempted in cichlids before, and constitutes a technological advance in the field. To ensure that the PSM was isolated intact, the PSM was dissected off with three somites attached (see methods). This was easier to achieve at earlier stages of somitogenesis (8–20 ss) where the somites are large and easily counted. Explanted PSMs were then cultured for up to 12 hrs in an L-15 -based culture media at 28°C (see methods). It is possible to culture whole embryos isolated from the yolk using this method for at least 12 hrs, however the embryo has a tendency to curl up and adhere to itself, making imaging challenging.

To first test if DAPT disrupts notch signalling in cichlid PSMs, I applied increasing concentrations of DAPT (30 μM, 50 μM, 100 μM) to explanted PSMs of *A. calliptera* and incubated the explants for 4 hrs to minimise abnormalities associated with their dissection from the rest of the embryo. Imaging the expression of deltaC and deltaD using HCR v3.0 we observe that at least for 100 μM DAPT, the polarity of deltaC and deltaD expression in the somites is disrupted, with a loss of the typically neat expression shown in the control where deltaC is expressed in the posterior of the somite and deltaD in the anterior (figure 4.6). This suggests that DAPT effectively disrupts the synchrony of the segmentation clock in cichlids, and verifies that the apparent lack of effect from DAPT on segmentation when applied to intact embryos is due to the chorion and vitelline envelope acting as barriers to DAPT entry.

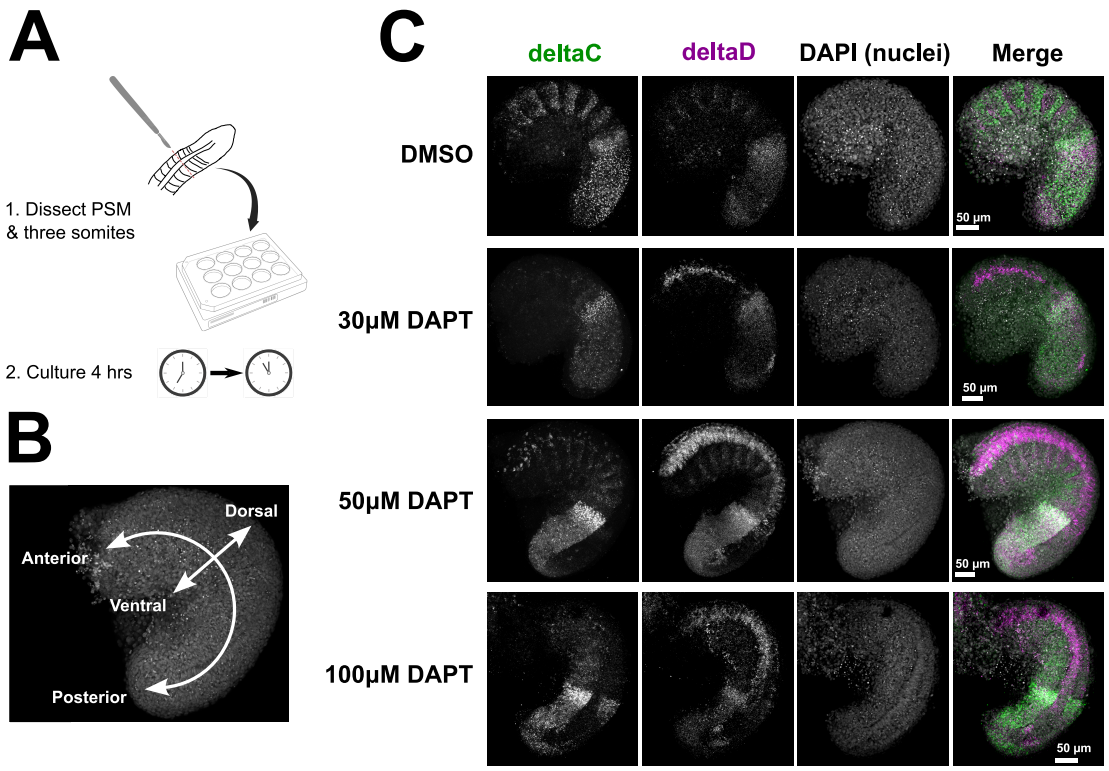


Figure 4.6: Effect of DAPT on deltaC and deltaD expression in explanted PSMs of *Astatotilapia calliptera*. A Cartoon showing experimental protocol. **B** Max-intensity z-projection of DAPI stain (nuclei) of an explanted PSM, showing the typical 'comma' morphology of explants, with the anterior-posterior and dorsal-ventral axes annotated. **C** HCR for deltaC (green) and deltaD (magenta) in explanted PSMs of *Astatotilapia calliptera* after 4 hrs of culture. Maximum intensity z-projections of z-stacked images.

Now that I have established a method for successfully disrupting notch signalling in the cichlid PSM, I can now begin to measure the ALD for various concentrations of DAPT and measure κ, τ, and $D\theta$. As I was conscious that DAPT treatment could perhaps knock-down expression of deltaC in the formed somites and preclude the accurate measurement of ALD, I decided to use myoD as a marker of somite boundaries, as this exhibits similar phenotypes to xirp2a when notch signalling is disrupted [236, 237]. Additionally, I increased the culture time from 4 hrs to 8 hrs, to incease the range of DAPT concentrations for which we can observe a response in ALD [49]. Culturing explants for longer (12 hrs) would allow for a greater range of response, but culturing for this long leads to abnormal morphology and the additional length of the explant often causes it to self-adhere and curl up, precluding effective imaging, and so 8 hrs was

chosen as a compromise. Culturing explanted PSMs of *A. calliptera* and *R.* sp. 'chilingali' for 8 hrs according to the standard protocol and varying the concentration of DAPT, I was able to generate a curve of ALD against DAPT concentration (figure 4.7B). Similar to the results of Riedel-Kruse *et al.*, we see that ALD decreases with increasing DAPT concentration, reaching an asymptotic value of $ALD_\infty \approx 5$ (figure 4.7B).

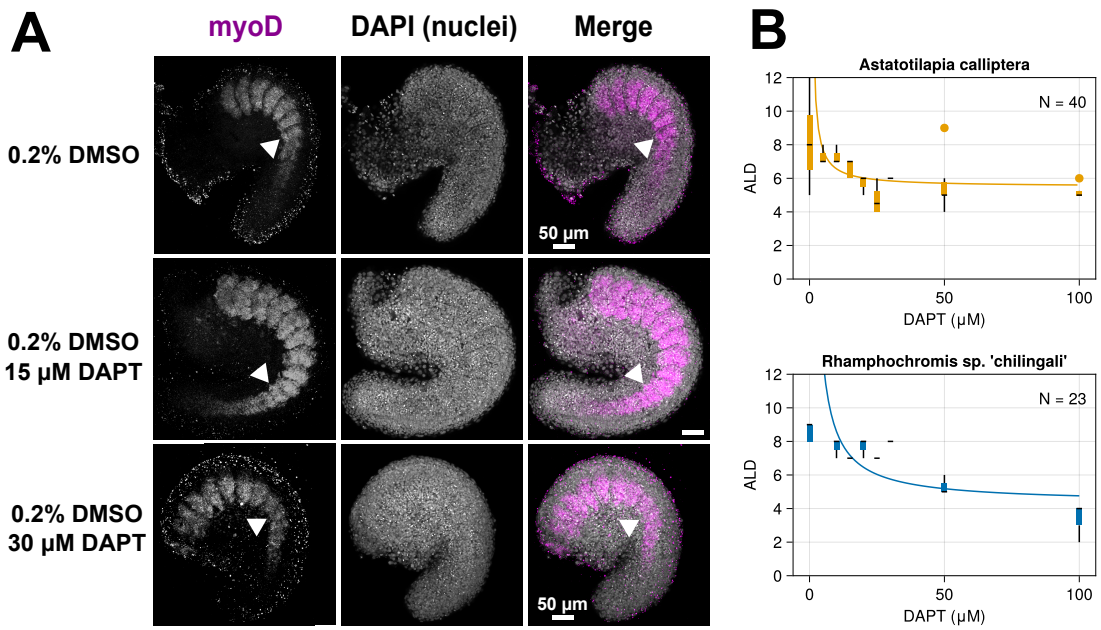


Figure 4.7: Quantifying the effect of notch inhibition on segment patterning in cichlids. A HCR for myoD (magenta) in explanted PSMs of *Astatotilapia calliptera* after 8 hrs of culture, for varying concentrations of notch inhibitor DAPT Nuclei (DAPI staining) shown in grey. Maximum intensity z-projections of z-stacked images. Posterior-most correctly patterned segment ($ALD - 1$) for each explant is annotated with a white arrowhead. **B** Anterior limit of defect (ALD) plotted against DAPT concentration for data collected from explanted PSMs of *Astatotilapia calliptera* (top) and *Rhamphochromis* sp. 'chilingali' (bottom). Solid line in each plot shows equation 4.3 fitted to data for each species.

After Riedel-Kruse *et al.* [49], we can use these results to estimate the noise of the clock $D\theta$ using the equation

$$\sqrt{2D\theta} = \frac{2\alpha \ln(r_0/r_c)}{ALD_\infty}, \tag{4.1}$$

where α is the rate of somite formation, r_0 and r_c represent the initial synchrony of the

segmentation clock before perturbation and the critical value of synchrony at which a defective segment is formed respectively, and ALD_∞ the asymptotic value of ALD achieved for saturating doses of DAPT.

The asymptotic value ALD_∞ was estimated by taking the mean value of ALD for explants treated with 50 μM and 100 μM of DAPT. For *A. calliptera* I estimate $ALD_\infty = 5.5$ and for *R.* 'chilingali' I estimate $ALD_\infty = 4.3$. Based on my previous estimates of α in each species (figure 4.2), I thus estimate the value of $\sqrt{2D\theta}$ to be $\sqrt{2D\theta} = 0.909$ hr^{-1} in *A. calliptera* and $\sqrt{2D\theta} = 1.040$ hr^{-1} in *R.* 'chilingali', or equivalently that $D\theta = 0.0069$ min^{-1} for *A. calliptera* and $D\theta = 0.0090$ min^{-1} for *R.* sp. 'chilingali'. To ensure comparability of my results with that of Riedel-Kruse *et al.*, and because the true values of r_0 and r_c are unknown, we follow the method of Riedel-Kruse *et al.* and assume that $\ln(r_0/r_c) \approx 2.5$ [49]. This assumption is based on the observation that for either a 3- or 100-fold difference between r_0 and r_c, $\ln(r_0/r_c) \approx 2.5 \pm 1$ [49]. My results for the noise of the clock are very similar to those of Riedel-Kruse *et al.*, who report $\sqrt{2D\theta} \approx 0.8$ hr^{-1} in Zebrafish [49].

The coupling strength κ can be estimated using the equations

$$n_c = n_0 \cdot \left(\frac{\kappa}{\sqrt{2D\theta}} - 1\right) \tag{4.2}$$

and

$$ALD = \frac{ALD_\infty \cdot (n + n_0)}{n - n_c}, \tag{4.3}$$

where n is the concentration of DAPT, n_0 the concentration of DAPT that halves notch signalling, and n_c the critical concentration of DAPT at which we would expect to observe defects in segmentation, or equivalently the smallest concentration of DAPT for which the ALD is finite. By fitting equation 4.3 against a series of concentrations and corresponding ALD it is possible to estimate values of n_0 and n_c (figure 4.7). By performing this fitting (see

methods) we estimate $n_0 = 1.542$ μM (95% CI: [0.329 μM, 1.828 μM]) for *A. calliptera* and $n_0 = 9.820$ μM (95% CI: [9.461 μM, 10.579 μM]) for *R.* sp. 'chilingali', $n_c = 0.321$ μM (95% CI: [0.005 μM, 1.050 μM]) for *A. calliptera* and $n_c = 0$ μM (95% CI: [0 μM, 0 μM]) for *R.* sp. 'chilingali'. Substituting into equation 4.2 yields $\kappa = 0.018$ min^{-1} (95% CI: [0.017 min^{-1}, 0.073 min^{-1}]) for *A. calliptera* and $\kappa = 0.017$ min^{-1} for *R.* sp. 'chilingali'. Note that as $n_c = 0$ for *R.* sp. 'chilingali' I was unable to construct a distribution for κ using boostrapped values of n_c and n_0 and as such could not calculate the confidence intervals for this value.

The coupling delay τ can be estimated according to solutions of the equation

$$\frac{2\pi}{T_{WT}} = \frac{2\pi}{T_A} - \kappa \cdot \sin(\frac{2\pi \cdot \tau}{T_{WT}}), \tag{4.4}$$

where T_{WT} represents the collective period of the segmentation clock, T_A the cell-autonomous period of the segmentation clock, and κ the coupling strength as defined above [139]. The value of T_A can be estimated from the period of somitogenesis when notch signalling is lost, using the relationship

$$T = \frac{n \cdot T_A + Q \cdot T_{WT}}{n + Q}, \tag{4.5}$$

where as before n represents the concentration of DAPT, T_{WT} the collective period of somitogenesis in the unperturbed (wild-type) embryo, and Q is given by

$$Q = n_0 \cdot (1 + \kappa \cdot T_A), \tag{4.6}$$

for n_0 defined as above [139].

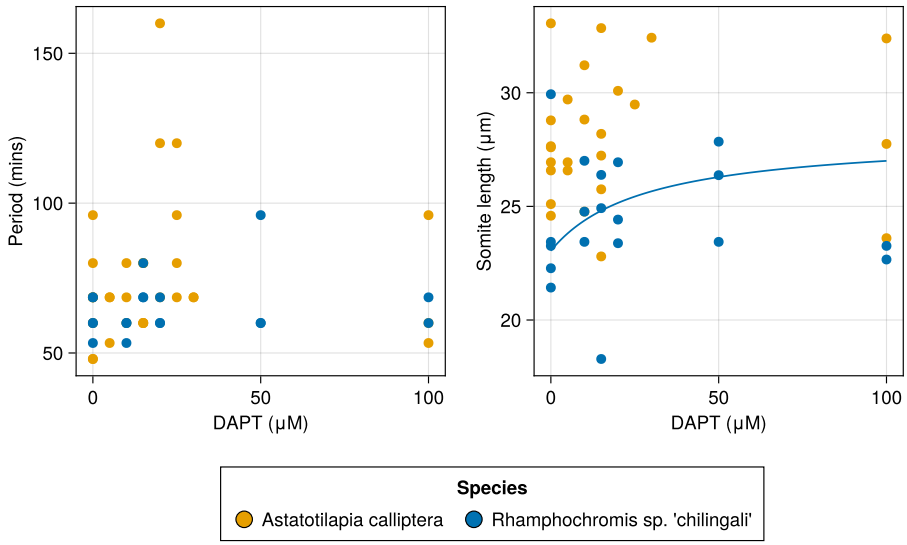


Figure 4.8: Effect of notch inhibition on somitogenesis period and somite length. Left: Period of somitogenesis in PSM explants, calculated as the total number of somites formed after dissection divided by the incubation time (8 hrs), for varying concentrations of the notch signalling inhibitor DAPT. **Right:** Anterior-posterior length of the most posterior somite for varying concentrations of DAPT. Solid line represents fit of equation 4.5 to the data (see methods).

To infer the frequency of somitogenesis in cultured explants I counted the number of somites visible in the explant and subtracted three from the total (the number of somites attached to the PSM at dissection), and then divided this number by the time taken for culture (8 hrs). Converting this number to a period in minutes and attempting to fit equation 4.5 to the period across varying concentrations of DAPT (figure 4.8), I was not able to obtain satisfactory fits to the data in either species. However, if we assume that somite length is directly proportional to the clock period, we can expect somite length to change in the same way upon disruption of notch signalling [139]. If we substitute the terms T_A, T_{WT}, and T in equation 4.5 with the terms S_A, S_{WT}, and S that represent the somite length in the absence of a coupling delay, somite length in the unperturbed embryo, and somite length in response to a given level of notch inhibitor n respectively, and fit equation 4.5 to the anterior-posterior length of the posterior-most somite (referred to as 'somite length' hereafter for brevity) (figure 4.8), I estimate S_A = 28.114 μm and Q = 28.456 μm in *R.* sp. 'chilingali'. No satisfactory fit for the data from *A. calliptera* could be found.

From these data the median somite length for the unperturbed embryo (S_{WT}) is 22.77 μm. Assuming that somite length scales proportionately with somitogenesis period, then $T_A = T_{WT} \cdot (S_A/S_{WT})$. As $S_A/S_{WT} \approx 1.23$ therefore we expect $T_A \approx 82$ mins in *R.* sp. 'chilingali' and assuming this value scales by the same proportion in *A. calliptera*, $T_A \approx 74$ mins in that species. It is worth noting that this estimated value of $S_A/S_{WT} \approx 1.23$ is in line with the previous estimate of Herrgen *et al.* who estimated $T_A/T_{WT} \approx 1.18$ in Zebrafish [139]. However is it not possible to estimate τ based on this value as solutions to equation 4.4 are only defined for

$$|2\pi/T_{WT} - 2\pi/T_A| \leq \kappa, \tag{4.7}$$

which $T_A \approx 1.23T_{WT}$ fails to satisfy for my estimated values of κ and T_{WT}. Therefore to estimate the delay I assume $T_A \approx 1.207T_{WT}$, the largest possible value of T_A such that equation 4.7 is satisfied for my estimates of κ and T_{WT}. Multiple T_{WT}-periodic solutions for equation 4.4 exist, however I restrict my solutions to only consider $0 < \tau \leq T_{WT}$ for computational simplicity. Substituting these values into equation 4.4 I estimate $\tau \approx 46$ mins in *A. calliptera*, and $\tau \approx 54$ mins for *R.* sp. 'chilingali'.

4.2.3 Modelling cichlid somitogenesis

At a first glance, the estimated values of κ may appear surprisingly low, as previous results would suggest that such a value is too low to achieve clock synchrony, with my previous results suggesting that synchrony is only achievable for $\kappa > 0.03$ (figure 2.8). However, one must account for the fact that the segmentation clock in cichlids is slower than Zebrafish, as too is the overall rate of development and thus presumably the advection velocity of cells towards the wavefront. Thus, this may grant the clock enough time in a given clock cycle to achieve synchrony before the clock phase is read-out by the wavefront.

Therefore, to test the dynamics for these predicted values we adjust the advection velocity $v_d(x)$,

frequency of the clock ω_0, and length of the PSM L, all of which differ between Zebrafish and cichlids and can impact the synchrony of the segmentation clock (see chapter 2). As we lack detailed data of the morphogenesis of the cichlid PSM, we cannot explicitly measure the advection velocity within the PSM, nor if the local strain rate varies along the PSM, as the model of Uriu *et al.* assumes for Zebrafish [54] (see chapter 2). Therefore, we assume a constant strain rate across the PSM and assume that the advection velocity $v_d(x)$ is given by

$$v_d(x) = v_a \cdot (1 - \frac{x - x_a}{L}), \tag{4.8}$$

for all $x_a \leq x \leq x_a + L$.

To estimate v_a we note that somite length $S \approx v_d(x_a) \cdot T$ where T is the period of the segmentation clock [138]. As 15 ss, the nascent somite is approximately 45 μm long in *A. calliptera* and 41 μm in *R.* sp. 'chilingali' (S Taylor, *pers. comm.*). Therefore, using my estimates of the period of somitogenesis in each species (60 mins and 67 mins respectively), I estimate the advection velocity at the PSM anterior (v_a) to be 0.76 μm·min^{-1} and 0.61 μm·min^{-1} for each species, respectively. Additionally the PSM at 15 ss is approximately 400 μm long in *A. calliptera* and 500 μm in *R.* sp. 'chilingali' (S Taylor, *pers. comm.*). As we lack quantitative data on the morphogenesis of the cichlid PSM the remaining parameters in the model were assumed to be the same as in Zebrafish (see table 2.1). Setting the intrinic frequency $\omega_0 = 2\pi/T_{WT}$ to $\omega_0 = 0.105$ min^{-1} and $\omega_0 = 0.094$ min^{-1} for each species respectively, we see that in neither species is the estimated coupling strength $\kappa \approx 0.018$ min^{-1} capable of producing synchronous oscillations in the presence of the measured level of noise and a coupling delay of either $\tau = 46$ mins or $\tau = 54$ mins for *A. calliptera* and *R.* sp. 'chilingali' respectively (figure 4.9A). However, a $2.5\times$ increase in κ (κ = 0.05 min^{-1}), and a corresponding increase in τ according to equation 4.4 ($\tau \approx 56$ mins) is sufficient to achieve synchronous oscillations in these conditions (figure 4.9B). Due to the noisy nature of the data and wide range of estimates for κ (at least in *A. calliptera*), it therefore seems likely that the true value of κ for both species

is greater than $0.018\ \text{min}^{-1}$.

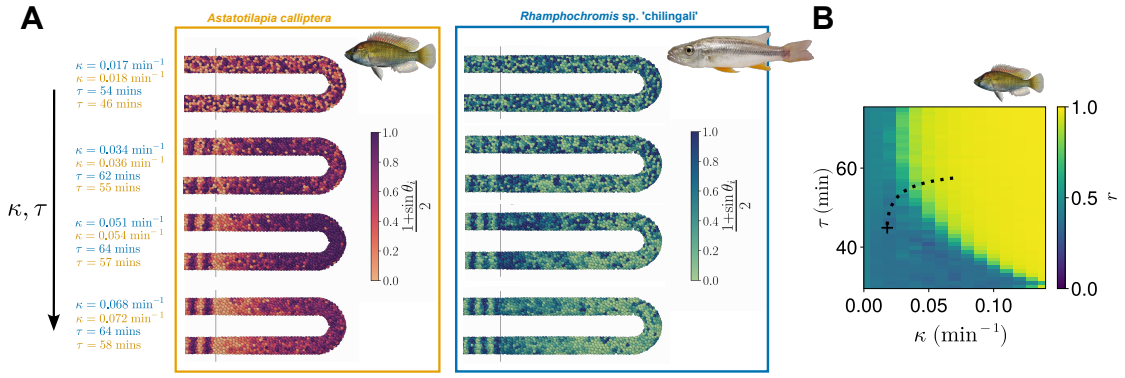


Figure 4.9: Model of the cichlid segmentation clock. A Stills of simulations of the PSM and clock dynamics for the cichlids *Astatotilapia calliptera* and *Rhamphochromis* sp. 'chilingali', after 840 mins. Coupling strength κ and delay τ for *A. calliptera* and *R.* sp. 'chilingali' are shown in orange and blue, respectively. Parameter sets used in this figure: *A. calliptera* = $\{\omega_0 = 0.105\ \text{min}^{-1}, D\theta = 0.0069\ \text{min}^{-1}, v_a = 0.76\ \mu\text{m}\cdot\text{min}^{-1}, x_q = 1, L = 400\ \mu\text{m}\}$. *R.* sp. 'chilingali' = $\{\omega_0 = 0.094\ \text{min}^{-1}, D\theta = 0.009\ \text{min}^{-1}, v_a = 0.59\ \mu\text{m}\cdot\text{min}^{-1}, x_q = 1, L = 500\ \mu\text{m}\}$. **B** Synchrony at the anterior of the PSM after 1000 mins for varying coupling strength κ and coupling delay τ, using the parameter set for *A. calliptera*. Each pixel is coloured according to the median synchrony of $N = 100$ simulations for that pair of values. The estimated coupling strength and coupling delay for *A. calliptera* ($\kappa = 0.018\ \text{min}^{-1}$, $\tau = 46$ mins) is annotated with a black +. A black dashed line marks the solutions to equation 4.4 for $\kappa \in [0.017\ \text{min}^{-1}, 0.073\ \text{min}^{-1}]$.

Species	κ (min^{-1})	$D\theta$ (min^{-1})	τ (min)	ω_0 (min^{-1})	v_a μm·min^{-1}
Danio rerio (values taken from refs [49, 54, 139])	0.07	0.0053*	21	0.2094	1.67
Astatotilapia calliptera	0.018	0.0069	46	0.105	0.76
Rhamphochromis sp. 'chilingali'	0.017	0.0090	54	0.094	0.61

Table 4.3: Summary of the estimated values for parameters of a model of the segmentation clock. Parameters for Zebrafish (*Danio rerio*) derived from previous work and used in the preceding two chapters, and the values inferred for the Lake Malawi cichlids *Astatotilapia calliptera* and *Rhamphochromis* sp. 'chilingali' in this chapter. *This is the value estimated from experimental data by Riedel-Kruse *et al.* [49]. Note that as this measurement incorporates noise due to morphogenesis, Uriu *et al.* estimate the true noise of the clock to be 0.0013 min^{-1} [54].

4.3 Discussion

Here I show, contrary to previous work [146], that the difference in somite number between *Astatotilapia calliptera* and *Rhamphochromis* sp. 'chilingali' is underpinned by differences in the duration of somitogenesis, and find no evidence to support that this is driven by changes in frequency. As in teleosts the correspondence between somites and vertebrate is known to be

one-to-one [72], I therefore conclude that the difference in vertebral number between these two species is underpinned by a difference in the duration of somitogenesis.

It is interesting to note the similarity of my results to those of Kingsley *et al.*, who found that differences in the duration of somitogenesis, likely mediated by changes in the regulation of HoxD13, underpin the difference in caudal vertebral number in different populations of the Deer mouse *Peromyscus maniculatus* [15]. Due to the short evolutionary timescale involved in this case (~10,000 years [15]), it is tempting to speculate that my results and these suggest that changes in duration are more the more evolvable route for changing vertebral number. Indeed, extreme prolongation of somitogenesis appears to be possible, with *Nemichthys* Snipe eels and multiple species of Salamander apparently adding new vertebrae throughout their lifetime [11, 238], although it is unclear if this reflects a true prolongation of somitogenesis or merely a gradual maturation of precursor vertebral elements [238]. The large numbers of vertebrae formed in the eels *Anguilla japonica* and *Muraenesox cinereus* also appear to be underpinned by a prolongation of somitogenesis [145], and indeed the Anguilliformes have repeatedly increased their vertebral number across the clade by addition of new caudal vertebrae [57], possibly indicative of repeated changes in duration. In part because the developmental control of the duration of somitogenesis is not well understood, it is unclear why duration should be more readily evolvable than frequency.

It may be notable in this regard that the egg size of these species differs, with *R.* sp. 'chilingali' having eggs that are up to twice the volume of *A. calliptera* [146]. It has been observed that greater nutritional availability to the embryo, such as in internally developing species like mice or species with large yolks such as birds or dogfish, correlates with greater volumetric growth of the PSM [113]. Additionally conjoined trout embryos form fewer somites, suggesting that division of nutrition between two embryos may limit the elongation of the embryo [194]. It has been suggested that this greater volumetric growth of the PSM allows for greater elongation of the posterior body and the formation of more somites [113, 115]. Future work on the development of *A. calliptera* and *R.* sp. 'chilingali' should examine the volumetric growth of the PSM to

see whether the difference in duration of somitogenesis between these two species can be explained by differences in volumetric growth. If this were the case, it is possible that the adoption of a piscivorous diet by *Rhamphochromis* (in contrast with its generalist *Astatotilapia* ancestor [209]) facilitated the evolution of divergent vertebral number by allowing greater nutritional contribution to the embryo.

It is worth noting that the evolvability of vertebral number in response to changes in duration is contingent on the value of the frequency, i.e. changes in duration need to be of equivalent magnitude to the period of the clock to have an impact on vertebral number. Accordingly, the species with faster clocks relative to their overall pace of development are more likely to exhibit an evolvable vertebral number in response to small changes in duration. Interestingly we see that cichlids are in this respect 'unremarkable', i.e. when one scales the frequency of somitogenesis by the overall pace of development (see Introduction), one sees that the frequency in cichlids does not appear to be unusually fast relative to the overall pace of development (figure 4.10). Due to their rapid diversification relative to other vertebrates, it is tempting to regard cichlids as somehow special, however with regards to their somitogenesis, it is unclear if they are unique compared to other vertebrates.

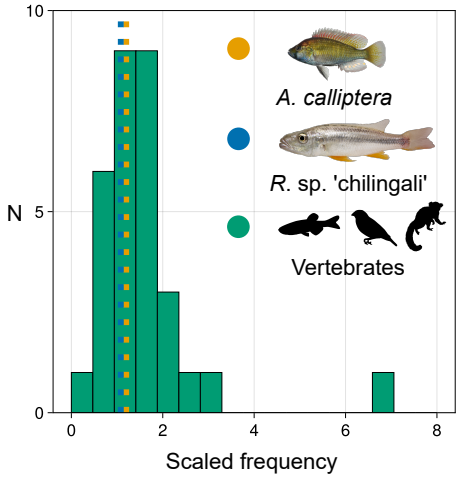


Figure 4.10: The development-adjusted somitogenesis frequency of Lake Malawi cichlids in the context of the vertebrates. Distribution of the frequency of somitogenesis in vertebrates, scaled by the overall pace of development (see Introduction). For raw data and references see table 1.1. Frequencies for the Lake Malawi cichlids *Astatotilapia calliptera* and *Rhamphochromis* sp. 'chilingali' are plotted with dashed lines in orange and blue, respectively. Photo credits: *A. calliptera*, Howard Jelks, U.S. Geological Survey. *R.* sp. 'chilingali', Aleksandra Marconi.

Indeed the segmentation clock of cichlids appears to share several quantitative features with Zebrafish. For instance, I have shown here that the clock in cichlids is approximately as noisy as that in Zebrafish (figure 4.7). Jenkins *et al.* predict that the noise in Hes/Her oscillations is driven by stochastic dissociation of Hes/Her dimer proteins from DNA, and that while the clock is tolerant of up to a tenfold increase in the level of this noise, this is only when the period of the clock is significantly lengthened to compensate [239]. While it is likely that my results reflect the conservation of DNA-binding affinity within the Teleostei, they may therefore also reflect limitations imposed on the clock, such as selection on the clock period to maintain a sufficiently large number of vertebrae for locomotion. Furthermore, as here I have measured the clock noise at a tissue-level, my estimated values of $D\theta$ for cichlids also incorporates noise from morphogenetic sources such as cell division [49, 200]. This conservation of noise may thus also reflect conserved features of PSM morphogenesis within the Teleostei.

The length of the clock coupling delay relative to the length of the clock cycle also appears to be conserved between cichlids and Zebrafish. If we consider $\tau \approx 56$ mins for *A. calliptera*, as appears to be the minimum such value such that we observe synchronous oscillations in our model (figure 4.9B), then $\tau/T_{WT} \approx 0.93$. This is similar to the results of Herrgen *et al.*, who estimated that $\tau/T_{WT} \approx 0.9$ in Zebrafish [139]. As the difference in T_{WT} between cichlids and Zebrafish is likely largely driven by the differences in the overall pace of development (see table 4.1) [47], the differences in τ between cichlids and Zebrafish can thus be largely accounted for by the difference in developmental rate. It is unclear from the present data whether the clock coupling strength κ is conserved between cichlids and Zebrafish. However given the broad confidence intervals for my estimates of κ (at least in *A. calliptera*), my computational results suggesting that for synchronous oscillations κ must be such that $\kappa \geq 0.05$ min^{-1} in *A. calliptera* (figure 4.9B), and the conservation of noise and delay, it seems plausible that the clock coupling strength is conserved with that of Zebrafish ($\kappa = 0.07$ min^{-1}). However, further quantification of the cichlid segmentation clock will be necessary to determine whether this is the case.

While by using explanted PSMs I have been able to successfully infer parameters of a model describing the segmentation clock, my experiments have revealed problems with this approach. The total number of somites formed in each explant is extremely variable (figure 4.8), and reflects the varying period for which an explanted PSM will continue to form somites. While this could perhaps be reflective of premature death of explants, it could also reflect varying capacity for elongating the PSM and continuing somitogenesis. Indeed I notice that dissection skill is a major determinant of explants' viability, and explants where the posterior tip of the PSM is damaged fail to elongate or do so asymmetrically. This noise could therefore be introduced by dissection. Additionally removal of surrounding tissues from the explant appears to reduce the viability of explants, possibly indicative of elongation being driven by cellular material drawn from surrounding tissue. The cichlid PSM also reduces in size over the course of somitogenesis (S Taylor, *pers. comm.*), which could introduce noise as due to cichlids' small clutch sizes, I was unable to perform dissections at only a specific somite stage, instead dissecting PSMs from cichlid embryos ranging from 8–20 ss. Therefore some of this variation in the total number of somites may reflect the varying quantity of cellular material the PSM posesses at dissection, and thus its capacity to elongate. While this experimental system is useful for performing drug perturbations of cichlid development, future work using such explants should focus on identifying consistent protocols for breeding and dissection that maximise the reproducibility of explant development.

Is modularity of the segmentation clock and PSM morphogenesis conserved between cichlids and Zebrafish? Without explicitly testing this using the computational framework developed in the previous chapters, we cannot be sure, however the apparent conservation of parameters of the segmentation clock might suggest that this is the case. A direct comparison is complicated by the fact that the rate of cell advection v_a in cichlids is slower than in Zebrafish, and the rate of clock oscillations ω_0 is also slower, and thus I cannot directly compare the results in this chapter with the simulations in the previous two chapters. However, if we assume, as in the introductory chapter of this thesis, that $T_{WT} \propto T_P$ and $v_a \propto 1/T_P$ where T_P is the time

taken for a species to reach the phylotypic stage of development (defined as the appearance of the pharyngeal arches), then we can scale these parameters to the pace of development, specifically T_{WT}/T_P and $v_a \cdot T_P$. Scaling these parameters in this manner we see that the values for Zebrafish and cichlids are largely similar (table 4.4). As the coupling delay τ scales with T_{WT}, and clock noise $D\theta$ is conserved between cichlids and Zebrafish, and κ is plausibly conserved between cichlids and Zebrafish as well, it seems likely therefore that my computational results in the previous chapters hold for cichlids and the segmentation clock and morphogenesis of the PSM exhibit developmental modularity in this clade.

Species	T_P (hpf)	T_{WT}/T_P	$v_a \cdot T_P$
Danio rerio (values taken from refs [49, 54, 74, 139])	24	1.02	40.08
Astatotilapia calliptera	72	0.83	54.72
Rhamphochromis sp. 'chilingali'	72	0.93	43.92

Table 4.4: Comparing segmentation clocks and elongation in cichlid and Zebrafish. Clock period T_{WT} and advection velocity v_a, normalised by the rate of development (measured by the time taken to reach phylotypic stage T_P). Parameters for Zebrafish (*Danio rerio*) derived from previous work and used in the preceding two chapters, and the values inferred for the Lake Malawi cichlids *Astatotilapia calliptera* and *Rhamphochromis* sp. 'chilingali' in this chapter.

As there does not appear to be any difference in somitogenesis frequency between *A. calliptera* and *R.* sp. 'chilingali' (figure 4.2A), the difference in vertebral number between the two species appears to be solely driven by differences in the duration of somitogenesis. Therefore it does not appear that modularity of the segmentation clock and PSM morphogenesis has facilitated the rapid evolution of divergent vertebral number in this species pair, as we would predict such modularity enables co-evolution of both the clock and PSM morphogenesis to generate diversity in vertebral number. Nevertheless it is possible that robustness of the segmentation clock to varying PSM morphogenesis has enabled the evolution of divergent duration of somitogenesis in these two species. Further study of the mechanism of PSM elongation and control of duration in these two species will resolve whether this is the case. As noted above by comparison with the work of Kingsley *et al.*, it may be that duration is more readily evolvable than the frequency of somitogenesis. If this is the case, then modularity of the segmentation clock and duration of somitogenesis may not be relevant for the evolvability of vertebral number over

short timescales, as I have considered here for this pair of cichlid species. Precisely what this means for the evolvability of vertebral number is unclear and exactly how much this modularity has contributed to diversity in vertebral number will only become clearer with further study of somitogenesis across different vertebrate species.

It is hoped that the results outlined within this chapter both stimulate further work quantifying the segmentation clock in this species pair, which present a very interesting model system to understand the evolution of somitogenesis. Additionally, this quantitative approach is likely applicable to many other non-model organisms, and could allow a broader understanding of segmentation clock dynamics across the tree of life. Hopefully, a combination of the two approaches will allow a deeper understanding of the causes of evolvability in vertebral number and its distribution across the vertebrates.

4.4 Methods

All animal work was conducted following approval by the Departmental Animal Welfare Ethical Review Body (AWERB) for the Department of Biology at the University of Oxford.

4.4.1 Animal husbandry and collection of embryos

Breeding stock populations of mixed sex (approximately 1♂: 1♀) *Astatotilapia calliptera* of the 'Mbaka' race, and of *Rhamphochromis* sp. 'chilingali' were maintained in 227 L (100 cm × 70 cm × 40 cm) or 193 L (120 cm × 50 cm × 40 cm) or 178 L (116 cm × 48 cm × 40 cm) glass fish tanks in a room heated to 24°C with a 12:12 hr light:dark cycle. *A. calliptera* were fed a diet of JBL novo Malawi flakes for algal-feeding cichlids, supplemented with weekly enrichment feeds of thawed frozen Brine shrimp. *R.* sp. 'chilingali' were fed a diet of JBL novo Tanganyika flakes for predatory cichlids, supplemented with weekly enrichment feeds of thawed frozen bait fish. Additionally, we found the regular feeding of *R.* sp. 'chilingali' females with surplus embryos from other tanks and unfertilised eggs to be essential to ensure reliable breeding in this species. Using this husbandry setup we typically obtained 3–5 clutches of *A. calliptera* a week

from a stock population of ∼100 fish, and 1 clutch of *R.* sp. 'chilingali' embryos a week from a stock population of ∼30 fish. By removing subdominant non-breeding males from a tank it is possible to reduce husbandry burden and increase food availability for females, however when the density of fish in a tank is lowered males can often become extremely aggressive and fight to the death, so we found keeping subdominant males with breeding stocks to be useful for preventing mortality.

Embryos were collected by catching mouthbrooding females (see figure 4.1) in a small net and then isolating them in a bucket containing approximately 1 L of tank water. The fish was then caught with the hands and held at a steep angle such that its head is pointing towards the bucket. Then, using a 1 mL plastic pasteur pipette, water was squirted into the fish's mouth, rinsing the embryos into the bucket of water below. To minimise stress to the fish, the fish is released back into the bucket after every three rinses and the bucket is covered to create a dark environment. Once all embryos are removed the female is returned to her tank and embryos can be collected for culture or experiments.

4.4.2 Embryo culture

Clutches of embryos were cultured at 28°C in either 50 mL conical falcon tubes with gentle rocking or 90 mm petri dishes containing ∼35 mL of tank media with 0.00001% w/v Methylene Blue to prevent fungal growth. To minimise effects from overcrowding I cultured no more than 30 embryos per container, and to prevent fouling of the water dead or dying embryos were removed daily. The media was replaced at least daily, or more depending on the degree of fouling of the media.

4.4.3 Quantifying somitogenesis in *Rhamphochromis* sp. 'chilingali' and *Astatotilapia calliptera*

The timing of somitogenesis onset and completion was determined by culturing embryos from the one-cell stage. One-cell stage embryos were obtained by either checking fish tanks at hourly

intervals for new mouthbrooding females, by extracting after spawning has completed, or by chance when performing regular embryo extractions. Embryos were staged using a BelOMO 10x Triplet Jewelers Loupe and regularly checked to ensure correct development and accurate staging.

Embryos were cultured in 50 mL conical falcon tubes under circa 35 mL of fish tank media (tank water + 0.00001% w/v Methylene Blue) on a rocker to achieve aeration. The media was changed at least once every 24 hours, or more depending on the fouling of the media. To ensure accurate results, embryo culture and staging took place in the same temperature-controlled tank room where the adult fish were housed. Embryos were fixed at hourly intervals using 4% Formaldehyde in PBS, and then kept in this solution overnight at 4°C. After overnight fixation embryos were transferred to PBSTw and dissected, then again stained overnight with DAPI. Embryos were then mounted on glass slides and somites were counted under a fluorescence scope.

4.4.4 Drug treatment of cichlid embryos

Aliquots of 10 mM and 50 mM DAPT in DMSO were prepared from solid DAPT powder and stored at -20°C until use. To prevent precipitation of DAPT when added to solution, the desired volume of DAPT solution was first added to media (to a final total volume of 1000 µL) in a 1.5 mL Eppendorf tube and vortexed vigorously for 60 seconds until no precipitate was visible. This solution was then quickly added to the remaining volume of media for culture in the container, and mixed. Aliquots of 10 mM and 50 mM LY411575 in DMSO were prepared and added to culture media as above, although I observed that the vortexing was often unnecessary and LY411575 formed little to no precipitate, being readily soluble in aqeuous solution. In all treatments, DMSO was added as necessary to ensure that the final concentration of DMSO in the solution was 0.2% v/v.

4.4.5 Chemical dechorionation of cichlid embryos

Cichlid embryos were chemically dechorionated using pronase, according to a protocol which can be seen in the supplementary methods. Pronase (P5147, Lot #s 0000300367, 0000357104) was used and stored at -20°C in powdered form until use.

4.4.6 *in situ* visualisation of gene expression using Hybridisation Chain Reaction

Hybridisation Chain Reaction (HCR) v3.0 [230] was used for multiplexed fluorescent imaging of mRNA transcripts. Embryos or explanted tissues were first fixed overnight at 4°C in 4% v/v Formaldehyde in PBS. Long-term storage of fixed cichlid embryos for HCR is possible in 100% Methanol at -20°C, and is sometimes recommended for HCRs to prevent buoyancy of samples. However I avoided this as long-term methanol storage makes embryos very brittle and prone to damage during dissection. After overnight fixation, embryos were rinsed of formaldehyde solution and stained using HCR according to a standard protocol (see supplementary methods).

Fluoresence images of HCR-stained embryos or explants were obtained using a Zeiss 880 LSM Confocal Microscope with a $10\times$ or $20\times$ objective, using the Zen Black software.

4.4.7 Culture of PSM explants

Embryos were mechanically dechorionated and dissected off of the yolk using #5 Dumont forceps, and the PSM and three somites removed from the rest of the embryo by cutting the posterior embryo using a 0.4 mm $\times$ 12 mm needle (see supplementary methods — protocols). Dissected PSMs were transferred to an agarose-coated dish using a glass pasteur pipette and incubated under dissection media according to a standard protocol (see supplementary methods — protocols).

4.4.8 Models and parameter fitting

Fitting of equations 4.3 and 4.5 to data was achieved using the LsqFit package in Julia v1.8.2. The details of the mode of the segmentation clock and morphogenesis of the PSM used in figure 4.9 were exactly as those outlined in Chapter 2, and so have been omitted here for brevity.

4.5 Supplementary methods — protocols

The following are a set of protocols that I have used for the experimental work outlined in this chapter.

4.5.1 Cichlid embryo culture

Cichlid tank water (makes 100 L):

- 50 g SeaChem Cichlid Lake Salt
- 15 g SeaChem Tanganyika Buffer
- to 100 L deionised H_2O

4.5.2 Hybridisation Chain Reaction

This method is adapted from the protocol published by Bruce *et al.* [240]. As fixed cichlid embryos are very sticky and adhere readily to plastic, we found it useful to carry out the staining in a 30 mm wide × 6 mm deep glass staining dish. This can be sealed to prevent evaporation by placing drops of 5× SSCTw at each corner of the dish then placing a 40 mm glass square firmly down over the dish to seal the dish. The dish can then be sealed with parafilm to hold it in place.

Probes

Probe mixtures were designed by, and purchased from, Molecular Instruments. As only the genome of *A. calliptera* is annotated and publicly available, probes were designed based on

annotated gene sequences available on Ensembl. Gene transcripts and their Ensembl IDs are shown in table 4.5 below.

Gene	Ensembl Transcript ID
xirp2a	ENSACLT00000039276
myoD	ENSACLT00000002710
deltaC	ENSACLT00000018560
deltaD	ENSACLT00000013473

Table 4.5: Gene transcripts used to design HCR probes.

Solutions

Below are recipes for buffers and solutions used in this protocol. Probe hybridisation buffer (Lot #s BPH01726 & BPH03126), probe wash buffer (Lot #s BPW01626 & BPW03626), and amplification buffer (Lots #s BAM01726 & BAM04326), and hairpin solutions were all purchased from Molecular Instruments.

PBSTw:

- PBS (pH 7.4)
- 0.1% Tween-20

5× SSCTw (makes 100 mL):

- 25 mL 20× SSCT
- 1 mL Tween-20
- to 100 mL distilled H_2O

Detergent solution (makes 500 mL):

- 50 mL 10% SDS
- 12.5 mL 20% Tween
- 25 mL 1 M Tris-HCl (pH 7.5)

- 1 mL 0.5 M EDTA (pH 8.0)
- 15 mL 5 M NaCl
- to 500 mL distilled H_2O

Protocol

Unless otherwise stated, each step is carried out at room temperature (18°C).

1. 1 × 10 mins wash in 500 µL PBSTw.
2. 2 × 5 mins wash in 500 µL PBSTw.
3. Permeabilise sample with 500 µL of detergent solution for 30 mins. Pre-warm hybridisation buffer to 37°C.
4. Pre-hybridise samples in probe hybridisation buffer (200 µL for explants and young embryos, 500 µL for 4 dpf embryos) at 37°C for 30 mins.
5. Prepare probe solution by combining 2 pmol (2 µL of 1 µM probe mixture solution) of each desired probe mixture with of hybridisation buffer, to a final volume of 100 µL. Scale this solution up accordingly if more liquid is required. Incubate solution at 37°C until step 4 is complete.
6. Cover samples in 100 µL of probe solution. More volume may be needed for 4 dpf embryos, if the embryos are covered enough solution has been added. Leave for 72 hrs at 37°C, taking care to seal the staining vessel to prevent evaporation (see above).
7. Pre-heat probe wash buffer to 37°C, and amplificiation buffer to room temperature.
8. 4 × 15 mins wash with probe wash buffer (200 µL for explants and young embryos, 500 µL for 4 dpf embryos) at 37°C.
9. 3 × 5 mins wash in 500 µL 5× SSCTw.
10. Pre-amplify samples in amplificiation buffer (200 µL for explants and young embryos,

500 µL for 4 dpf embryos) at room temperature for 30 mins.

11. Prepare hairpin solution by heat-shocking 6 pmol (2 µL of 3 µM stock) of each hairpin at 95°C for 90 s, then cooling at 21°C for 30 mins in the dark. Combine with amplification buffer to a final volume of 100 µL. Scale this solution up accordingly if more liquid is required.

12. Cover samples in 100 µL of hairpin solution. More volume may be needed for 4 dpf embryos, if the embryos are covered enough solution has been added. Leave overnight (12–16 hrs) at room temperature in the dark, taking care to seal the staining vessel to prevent evaporation (see above).

13. 3 × 5 mins wash in 500 µL 5× SSCTw.

14. Cover samples in 500 µL of 5× SSCTw + 1 µg/mL DAPI (0.5 µL of 1 mg/mL stock solution). Leave overnight at 4°C in the dark.

15. 3 × 5 mins wash in 500 µL PBSTw

16. Transfer to 90% glycerol in PBS in 25% concentration increments. Mount and image.

Explants were mounted in 10 µL of 90% glycerol under a #1 coverslip supported by two #1 coverslip struts.

4.5.3 Pronase chemical dechorionation of cichlid embryos

This protocol is adapted from Morrison *et al.* [224].

Solutions

20 mg/mL Pronase solution (makes 10 mL):

- 200 mg Pronase (P5147 Sigma-Aldrich, Lot #s 0000300367 & 0000357104)
- 10 mL tank water (pH 8.8)

Protocol

1. Cover 5–8 embryos in 5 mL of 20 mg/mL pronase solution.

2. Incubate embryos under pronase at 28°C for 2–3 hrs, checking regularly.

   Pronase reactivity can be batch-specific and the time taken for dechorionation can vary accordingly. This time will likely need optimisation depending on the batch of pronase used.

3. After approximately 2 hrs, the embryos should appear to be glossy as the lamellae surrounding the chorion have been digested away. From this point test if the chorion can be easily removed with forceps — if the chorion is sufficiently digested then it should slide off the embryo with minimal mechanical force. Alternatively, this can be tested by moving embryos up and down in a fire-polished glass pipette — if the chorion is digested, this action should free the embryo from the chorion.

4. Once one or two embryos appear to have sufficiently digested chorions, or several embryos begin to show yolk protruding through the chorion, rinse the embryos 3× in fresh tank water. Save the pronase solution as it can be re-used for future dechorionations. Store at 4°C and discard after a week.

5. Mechanically dechorionate embryos either using forceps or by pipetting up and down with a fire-polished glass pipettee as described above.

6. Using a fire-polished glass pipette and taking care that dechorionated embryos do not make contact with the air-water interface, transfer dechorionated embryos to an agarose-coated well containing explant culture medium (see below). Incubate at 28°C. Discard any embryos with damage to the yolk as they will slowly release yolk over time and foul the media.

4.5.4 PSM explant dissection and culture

This is a protocol for the culture of explanted cichlid PSMs, however the same protocol can be successfully adapted for the culture of whole embryos that have been dissected off of the yolk.

Solutions

Explant culture medium (makes 30 mL):

- 900 µL Foetal Bovine Serum
- 150 µL Penicillin-Streptomycin solution (10,000 units Penicillin & 10 mg Streptomycin per mL, 0.1 µm filtered.)
- 3 µL Methylene Blue solution (1% w/v)
- to 30 mL L-15 medium

Formaldehyde fixation solution (makes 10 mL):

- 1 mL 40% Formaldehyde v/v solution
- 9 mL Phospho-buffered saline solution (for recipe see protocol for Hybridisation Chain Reaction)

Materials and Equipment

- Two pairs of sharp #5 forceps. It is important that the ends of the forceps can touch one another.
- Needle (0.4 mm × 12 mm)
- Plastic 60 mm petri dish
- Dissecting microscope
- 12-well plate coated in 1% w/v agarose for culture. To coat, add molten agarose to each well, swirl, then quickly drain the molten agarose.

- 270 mm glass pasteur pipette

Protocol

1. Gather embryos, as detailed elsewhere, at the 8–20 somite stage.
2. Transfer embryos to a 60-mm petri dish containing L-15 medium warmed to 18°C.
3. Using forceps, dissect the embryo away from the yolk. Once the yolk has been punctured the pressure within the yolk will cause yolk material to spill out of the puncture rapidly. The posterior part of the embryo is extremely fragile and care must be taken that the ejection of yolk does not break the posterior embryo against the chorion.
4. Using the edge of a 0.4 mm × 12 mm needle like a blade, cut across the boundary between the third- and fourth-nearest somites to the PSM in one smooth motion.
5. Using a glass pasteur pipette, transfer the explanted PSM and three attached somites to an agarose-coated well in a 12-well plate, containing 1 mL of culture medium plus any drug treatments applied to the explant. Take care to transfer the explant in a minimal volume of fluid. With fine control of the pipette once the pipette tip is immersed in the media the explant should 'drop out' of the pipette into the well.
6. Repeat this process until there are at most three explants per well. Take care to avoid explants making contact with one another, as they will adhere to one another and make imaging challenging.
7. Place the plate in an incubator set to 28°C.
8. After the desired incubation time, fix the explants by transferring them using a glass pasteur pipette as described above into 200 µL of 4% v/v formaldehyde in PBS in a 200 µL PCR tube. For fixation leave overnight at 4°C.

5

Discussion

5.1 Aims, approach, and findings

In this thesis I aimed to identify sources of evolvability in the developmental processes that determine vertebral number. Specifically, I have chosen to focus on the relationship between the dynamics of the segmentation clock, which controls the rate of somitogenesis, and the morphogenesis of the PSM, which is associated with changes in the duration of somitogenesis [46, 113]. I chose to focus on this question for several reasons. The first was that the relationship between these two processes had not been studied in detail, despite evidence that morphogenesis can affect the dynamics of the clock [200, 201, 241, 242]. The second was that the two processes appear to co-vary in evolution to generate diversity in vertebral number [46]. Given the apparent dependence of clock dynamics on PSM morphogenesis, it was unclear exactly how these two processes could co-evolve, and whether the clock constrains the evolution

of PSM morphogenesis and may bias the evolution of vertebral number. Another reason is that the genomic architecture of the segmentation clock gene regulatory network is complex and still not fully characterised [45, 109]. While the core regulatory logic of the Hes/Her feedback circuit and its evolution have been well-characterised [48, 79, 96, 98–100], many features of the clock remain unexplained such as the entrainment and function of oscillations in Wnt signalling [191]. While understanding the precise regulatory architecture of the clock will be vital to understanding the evolvability of vertebral number, the extensive experimental work required is far beyond the scope of this thesis. A similar problem exists for understanding the evolvability of the duration of somitogenesis, the causal mechanism for which is much less understood. Therefore, given these constraints of current knowledge, studying the relationship between the morphogenesis of the PSM and the segmentation clock appeared to be the most achievable way of furthering our understanding of the evolvability of vertebral number.

In the first two chapters of this thesis I have adopted a computational approach and have shown that the synchrony and frequency of the segmentation clock are largely robust to changes in the morphogenesis of the PSM, for the correct combination of clock coupling strength and coupling delay. As perturbing the clock appears to have no effect on morphogenesis of the PSM, these two processes appear to exhibit developmental modularity for the correct combination of clock parameters. This could be a source of evolvability by allowing diverse numbers of vertebrae to be formed by combinatorial changes in the duration and frequency of somitogenesis.

For this approach I have used a computational model of cell movements in the PSM and clock dynamics [54]. However, as discussed previously, this analysis is limited in that the three-dimensional model of cell movements I have used does not explicitly create elongation of the PSM. Here I have demonstrated modularity by testing various combinations of patterns of cell ingression, cell motility, cell division, and tissue compaction, and tested whether the synchrony of the clock is robust to variation in these processes. However because the model I have used does not generate PSM elongation, I have no idea whether a given combination

of these processes is biologically feasible. Additionally, one must not only consider whether such processes can elongate the PSM, but also whether they can drive and coordinate the elongation of surrounding tissues [121]. This is significant as my analysis therefore ignores possible constraints and biases in the evolution of the morphogenesis of the PSM, which will alter whether modularity of the clock and morphogenesis is relevant to evolution or not.

Additionally there are some properties of the segmentation clock which I have not considered that could affect its response to varying PSM morphogenesis, for instance that the dynamics of the segmentation clock appear to be responsive to mechanical force [243], or that the segmentation clock appears to be coupled to the cell cycle, with mitosis and cell divisions occuring preferentially in the 'off' phase of the clock cycle [200, 241, 242]. Therefore the relationship between the segmentation clock and morphogenesis of the PSM may be more nuanced than I assume here.

Nevertheless this model is the best available for investigating the effect of PSM morphogenesis on the dynamics of the clock, and I have shown that it can yield novel insights into somitogenesis. While computational models exist that can explain elongation of the PSM, they are usually only two-dimensional [52, 118, 119, 121], and so cannot quantitatively recapture the cell mixing in the PSM which has been shown to be important for maintaining synchrony of the clock [140, 201, 202]. To my knowledge the model used here is the only three-dimensional model of the PSM that has been extensively parameterised based on experimental data (in this case, for Zebrafish) [54]. Given the need to predict the effect of PSM elongation on the surrounding tissues [121], it is likely that testing whether a given combination of morphogenetic processes corresponds to robust PSM elongation is computationally intractable in three dimensions, and a complementary approach may be needed using two-dimensional models to explore this.

For my final chapter in this thesis I took an experimental approach to understanding evolvability, by a comparative study of two closely related species of cichlid, *Astatotilapia calliptera* and *Rhamphochromis* sp. 'chilingali'. These species present an excellent system to study which components of somitogenesis change most rapidly during evolution, as they have differing

numbers of vertebrae (30 and 36) and diverged only 460,000 years ago [209]. These present an excellent model system as the short evolutionary timescale for interspecific divergence aids in determining whether an evolutionary change in development was causal or neutral in effecting change on the number of vertebrae. An experimental approach was chosen to complement the modelling work and test predictions about sources of evolvability. There are few alternatives to the model system I have chosen. One would be the prairie and forest populations of Deer mice which have evolved a difference of four vertebrae in the last 10,000 years [15], however as these are placental mammals it is difficult to observe their somitogenesis or apply chemical perturbations to disrupt development. Another is the latitudinally distinct populations of Medaka from mainland Japan, which have evolved a difference of approximately one vertebra over the past 4–6 million years [60, 158]. This system is attractive as Medaka is a well-established model species with a large body of protocols and is amenable to live imaging of somitogenesis [158]. However, the evolutionary timescale involved is much longer than that of cichlids, and the difference in the number of vertebrae is much less (one). Additionally the large difference in vertebral number allows us to suggest mechanisms for effecting major change in the number of vertebrae. Furthermore, protocols for the laboratory culture and experimental manipulation of these species have been published [146, 223]. Taking these facts into account, this species pair presents the best possible model system to study the evolvability of vertebral number.

Here I have shown that the segmentation clock and morphogenesis of the PSM exhibit developmental modularity, that is that variation in either process has little to no effect on the other. This modularity depends on the parameters of the segmentation clock, most notably the cell-cell coupling strength and the coupling delay. Zebrafish appears to exhibit developmental modularity based on these parameters, which appear to be conserved between Zebrafish and the cichlids *A. calliptera* and *R.* sp. 'chilingali'. However, the difference in vertebral number between these two cichlid species is driven solely by changes in the duration of somitogenesis, rather than a combination of changes to frequency and duration, so it is unclear if this modularity has facilitated the rapid evolution of differences in vertebral number.

5.2 Modularity and the evolvability of vertebral number

How well does this modularity explain the evolvability of vertebral number in the vertebrates? As discussed in the introduction, it appears that many evolutionary changes in vertebral number are associated with changes in both the duration of somitogenesis and its frequency. This suggests that many evolutionary transitions in vertebral number may have been facilitated or accelerated by the modularity of frequency and duration. Here I have shown that this modularity appears to be conserved between Zebrafish and Lake Malawi cichlids, two lineages which diverged at the base of the Clupeocephala approximately 220 million years ago [208]. It therefore seems probable that the values I have measured here for coupling strength, noise, and coupling delay of the clock are ancestral to this clade. Notably the Clupeocephala contains approximately 90% of extant Actinopterygians (24,000 species) [66], or approximately 34% of extant vertebrate species [244]. The diversity of vertebral number in this clade is extremely high, ranging from as few as 16 vertebrae in the Oceanic sunfishes (Molidae) to 240 vertebrae in the Quillfish (*Ptilichthys goodei*) [66]. Based on my results, it is probable that the evolution of this diversity has been facilitated by the modular organisation of somitogenesis.

The origin of the Clupeocephala occurred in the mid-Triassic after the divergence of the Eels (Anguilliformes) from the rest of the Actinopterygii at the Permian-Triassic boundary (figure 5.1A) [208], and therefore my results cannot say anything about the existence of developmental modularity in Eels. As previously mentioned, this group is extremely diverse in its vertebral number, largely driven by changes in the number of caudal vertebrae and thus possibly this vertebral diversity is underpinned by changes in the duration of somitogenesis [57]. If that is indeed the case modularity may not explain the evolvability of vertebral number in this group, however robustness of the segmentation clock to varying morphogenesis of the PSM may have had a permissive role in evolution.

Indeed it is unclear when and how this modularity arose in vertebrate evolution, and an understanding of how deeply conserved this property is necessary for a quantitative accounting of its effect on evolvability of vertebral number across the clade. We can imagine four scenarios

for the evolution of modularity (figure 5.1B). In the first, modularity arose in the lineage leading to the Clupeocephala and modularity is peculiar to the somitogenesis of that clade. In the second, this property arose somewhere between the origin of the Clupeocephala and the divergence of the tetrapods and the Actinopterygii in the Silurian [208], and is a unique feature of Actinopterygian development (figure 5.1B, green). The third scenario is that this feature is ancestral to the osteichthyes and arose between the divergence of the cyclostomes and gnathostomes in the Cambrian [183] and the divergence of the tetrapods and Actinopterygii (figure 5.1B, pink). The fourth and final scenario is that this feature is ancestral to the vertebrates and arose in the lineage leading to the vertebrate ancestor (figure 5.1B, blue). As the quantitative framework I have used to describe the segmentation clock makes minimal assumptions about the structure of the segmentation clock, the approach I have taken to quantify the segmentation clock in Lake Malawi cichlids is likely to be applicable to any vertebrate species. Therefore, I would advocate for a research program quantifying the coupling strength and coupling delay across the vertebrates, to discriminate between these scenarios and understand when and where this modularity arose.

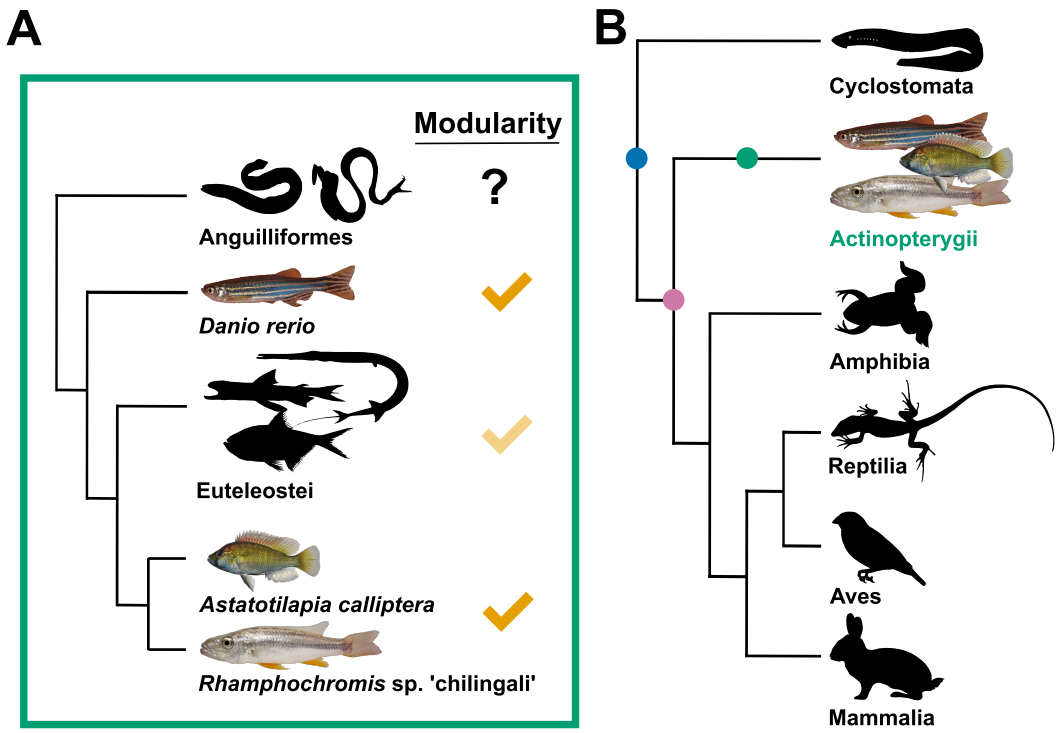


Figure 5.1: Modularity across the vertebrate clade. A: Modularity of the segmentation clock and morphogenesis of the PSM appear to be conserved across the Euteleostei, based on data from Zebrafish (*Danio rerio*) and the Lake Malawi cichlids *Astatotilapia calliptera* and *Rhamphochromis* sp. 'chilingali'. It is unclear if this property is conserved with the Eels (Anguilliformes). Photo credits: *A. calliptera*, Howard Jelks, U.S. Geological Survey. *R.* sp. 'chilingali', Aleksandra Marconi. **B:** Three scenarios for the evolution of modularity of the segmentation clock and morphogenesis, across the vertebrates. Blue: Modularity is an ancestral feature of vertebrates, Pink: modularity is ancestral only to the gnathostomes, and Green: modularity is specific to the Actinopterygii.

5.3 The evolution of modularity

While developmental modularity has been widely recognised as a source of evolvability, how modularity evolves, and indeed how evolvability evolves, remains a major open question within evolutionary biology [3, 5, 6, 207]. Specifically, it is unclear if modularity arises by neutral processes, such as via gene duplication and subsequent differentiation of functions between paralagous genes, or if modularity arises by natural selection, via its interaction with other traits or as a result of historic adaption for changing environments [207]. The modularity of the segmentation clock and morphogenesis of the PSM could be an extremely illuminating study

system for this question, as this modularity is contingent on only a handful of experimentally tractable parameters, allowing for measurement of modularity across the vertebrate phylogeny and thus allowing one to study its evolution (figure 5.1B).

Experimentally increasing the copy number of the deltaD gene in Zebrafish has been shown to increase coupling strength [189], implying a role for gene duplication in the evolution of clock robustness to morphogenesis, and perhaps a neutral mode of evolution for modularity. I note that we have so far only demonstrated modularity in teleosts, a lineage which has undergone three whole-genome duplications (WGDs) since the divergence of the vertebrate and urochordate lineages [245], compared to two WGDs in the rest of the vertebrates [183, 184]. Therefore an interesting question is to ask if one of these three duplication events conferred robustness to the clock, and if there is a general increase in clock robustness to morphogenesis with each duplication. In both mouse and Zebrafish, two delta ligands are required for segmentation, specifically Dll1 and Dll3 in mouse and deltaC and deltaD in Zebrafish [77]. This is in spite of the presence of five delta genes in Zebrafish versus three in mammals [246]. Interestingly in both species each paralogous gene appears to have a distinct role in segmentation and the two genes do not show functional redundancy [77, 127, 247]. Assuming that coupling strength is proportional to the copy number of delta genes, there is no evidence based on the complement of delta genes in Zebrafish somitogenesis that the teleost-specific WGD has increased the coupling strength within this group relative to the mammals. Given the functional redundancy of the two paralogous genes in both species, it is also unclear if the two vertebrate WGDs have had a lasting effect on the coupling strength of the clock in the extant vertebrates. However, the genetic and physical determinants of coupling strength remain unclear, so ultimately the only way in which this could be examined is by experimental quantification of coupling strength across the vertebrates (figure 5.1B).

It is also unclear whether the coupling delay evolves via a neutral or selective mechanism. Given the sensitivity of clock synchrony to changes in coupling delay [92], and the presumably stochastic nature of such delays, it seems likely that there would be a selective advantage in

increasing or decreasing the coupling delay away from bifurcation points where the clock transitions between synchronous and asynchronous states [93, 139]. Additionally because coupling delays are causal in setting the emergent frequency of clock oscillations in the PSM [93, 139], the selective pressures that effect changes in the frequency of somitogenesis are likely to drive evolution of delay. The developmental origins of this delay are poorly understood, but it has been shown that overexpression of the mindbomb ligase shortens the coupling delay in the Zebrafish segmentation clock [139], likely due an increase in the rate of NICD cleavage [248].

It is unclear how the robustness of the segmentation clock to noise has evolved. There has been much research into the role of notch signalling in maintaining synchronous oscillations in the face of noise [49, 77, 82, 84], however it is unclear whether the degree of robustness to noise we observe in Zebrafish and cichlids is an ancestral feature of the vertebrate segmentation clock or something that has evolved within this group. It is worth noting that the evolution of somitogenesis appears to predate the evolution of the PSM; the cephalochordate Amphioxus lacks mesenchymal tissue in the developing tail and forms the posterior somites directly out of a small body of epithelial cells in the tailbud [249]. While it is unclear if somitogenesis in Amphioxus is driven by a segmentation clock mechanism as in vertebrates, Hes/Her homologs are expressed in the presumptive somitic mesoderm and their expression depends on notch signalling [237, 250], suggesting that the segmentation clock gene circuit predates the vertebrates. As the embryos of lampreys, representatives of the most basal vertebrate clade [183], possess a PSM [113, 251, 252], the PSM likely evolved over the intervening ∼70 million years between the divergence of the cephalochordate and vertebrate lineages in the Ediacaran and the divergence of the lampreys with the jawed vertebrates in the Cambrian [183]. It is therefore tempting to suggest that the evolution of the PSM was permitted by the existence of an already robust gene regulatory network for the segmentation clock, and that this modularity may be ancestral to the vertebrates. However, further characterisation of somitogenesis in both lamprey and amphioxus will be necessary to determine whether this is the case.

It is currently unclear what has driven the evolution of modularity of the segmentation clock

and morphogenesis, however from the arguments above it seems likely that it has arisen either via a wholly neutral process or as a secondary artefact of selection for robustness against noise. Nevertheless, my work here establishes the segmentation clock as a novel model system to understand the evolution of evolvability, and it is my hope that this thesis will inspire future work in addressing this question.

References

1. Goodwin, B. C. Development and evolution. *Journal of Theoretical Biology* **97,** 43–55. doi:10.1016/0022-5193(82)90275-2 (1982).

2. Smith, J. M., Burian, R., Kauffman, S., Alberch, P., Campbell, J., Goodwin, B., Lande, R., Raup, D. & Wolpert, L. Developmental constraints and evolution: a perspective from the Mountain Lake conference on development and evolution. *The Quarterly Review of Biology* **60,** 265–287. doi:10.1086/414425 (1985).

3. Wagner, G. P. & Altenberg, L. Perspective: Complex adaptations and the evolution of evolvability. *Evolution* **50,** 967–976. doi:10.1111/j.1558-5646.1996.tb02339.x. (1996).

4. Kirschner, M. & Gerhart, J. Evolvability. *Proceedings of the National Academy of Sciences* **95,** 8420–8427. doi:10.1073/pnas.95.15.8420 (1998).

5. Pigliucci, M. Is evolvability evolvable? *Nature Reviews Genetics* **9,** 75–82. doi:10.1038/nrg2278 (2008).

6. Payne, J. L. & Wagner, A. The causes of evolvability and their evolution. *Nature Reviews Genetics* **20,** 24–38. doi:10.1038/s41576-018-0069-z (2019).

7. Oster, G. & Alberch, P. Evolution and bifurcation of developmental programs. *Evolution* **36,** 444–459. doi:10.1111/j.1558-5646.1982.tb05066.x (1982).

8. Conrad, M. The geometry of evolution. *BioSystems* **24,** 61–81. doi:10.1016/0303-2647(90)90030-5 (1990).

9. Verd, B., Monk, N. A. & Jaeger, J. Modularity, criticality, and evolvability of a developmental gene regulatory network. *eLife* **8,** e42832. doi:10.7554/eLife.42832 (2019).

10. Mosleh, S., Choi, G. P. T., Musser, G. M., James, H. F., Abzhanov, A. & Mahadevan, L. Beak morphometry and morphogenesis across avian radiations. *Proceedings of the Royal Society B: Biological Sciences* **290,** 20230420. doi:10.1098/rspb.2023.0420 (2023).

11. Beebe, W. & Crane, J. Deep sea fishes of the Bermuda oceanographic expeditions. Family Nemichthyidae. *Zoologica: scientific contributions of the New York Zoological Society* **22,** 349–383. doi:10.5962/p.184688 (1937).

12. Neat, F. & Campbell, N. Proliferation of elongate fishes in the deep sea. *Journal of Fish Biology* **83,** 1576–1591. doi:10.1111/jfb.12266 (2013).

13. Galis, F., Carrier, D. R., van Alphen, J., van der Mije, S. D., Dooren, T. J. M. V., Metz, J. A. J. & ten Broek, C. M. A. Fast running restricts evolutionary change of the vertebral column in mammals. *Proceedings of the National Academy of Sciences* **111,** 11401–11406. doi:10.1073/pnas.1401392111 (2014).

14. Lindell, L. The evolution of vertebral number and body size in snakes. *Functional Ecology* **8,** 708–719. doi:10.2307/2390230 (1994).

15. Kingsley, E. P., Hager, E. R., Lassance, J.-M., Turner, K. M., Harringmeyer, O. S., Kirby, C., Neugeboren, B. I. & Hoekstra, H. E. Adaptive tail-length evolution in deer mice is associated with differential Hoxd13 expression in early development. *Nature Ecology & Evolution* **8,** 791–805. doi:10.1038/s41559-024-02346-3 (2024).

16. Dawkins, R. in *Artificial life* 201–220 (Routledge, 1989). ISBN: 9780429032769.

17. Alberch, P. in *Evolution and development* 313–332 (Springer, 1982). doi:10.1007/978-3-642-45532-2_15.

18. Alberch, P. Ontogenesis and Morphological Diversification. *American Zoologist* **20,** 653–667. doi:10.1093/icb/20.4.653 (1980).

19. Salazar-Cuidad, I. Why call it developmental bias when it is just development? *Biology Direct* **16,** 3. doi:10.1186/s13062-020-00289-w (2021).

20. Alberch, P. The logic of monsters: Evidence for internal constraint in development and evolution. *Geobios* **22,** 21–57. doi:https://doi.org/10.1016/S0016-6995(89)80006-3 (1989).

21. Alberch, P. From genes to phenotype: dynamical systems and evolvability. *Genetica* **84,** 5–11. doi:10.1007/BF00123979 (1991).

22. Wagner, A. Robustness and evolvability: a paradox resolved. *Proceedings of the Royal Society B: Biological Sciences* **275,** 91–100. doi:10.1098/rspb.2007.1137 (2008).

23. Pigliucci, M. Evolution of phenotypic plasticity: where are we going now? *Trends in Ecology & Evolution* **20,** 481–486. doi:https://doi.org/10.1016/j.tree.2005.06.001 (2005).

24. West-Eberhard, M. J. Developmental plasticity and the origin of species differences. *Proceedings of the National Academy of Sciences* **102,** 6543–6549. doi:10.1073/pnas.0501844102 (2005).

25. Wessinger, C. A. & Hileman, L. C. Accessibility, constraint, and repetition in adaptive floral evolution. *Developmental biology* **419,** 175–183. doi:10.1016/j.ydbio.2016.05.003 (2016).

26. Bódi, Z., Farkas, Z., Nevozhay, D., Kalapis, D., Lázár, V., Csörgő, B., Nyerges, Á., Szamecz, B., Fekete, G., Papp, B., *et al.* Phenotypic heterogeneity promotes adaptive evolution. *PLoS biology* **15,** e2000644. doi:10.1371/journal.pbio.2000644 (2017).

27. Galhardo, R. S., Hastings, P. J. & Rosenberg, S. M. Mutation as a stress response and the regulation of evolvability. *Critical reviews in biochemistry and molecular biology* **42,** 399–435. doi:10.1080/10409230701648502 (2007).

28. Masel, J. & Trotter, M. V. Robustness and Evolvability. *Trends in Genetics* **26,** 406–414. doi:https://doi.org/10.1016/j.tig.2010.06.002 (2010).

29. Wright, S. The Roles of Mutation, Inbreeding, crossbreeding and Selection in Evolution. *Proceedings of the XI International Congress of Genetics* **8,** 209–222 (1932).

30. Salazar-Ciudad, I. & Marín-Riera, M. Adaptive dynamics under development-based genotype–phenotype maps. *Nature* **497,** 361–364. doi:10.1038/nature12142 (2013).

31. Zheng, J., Guo, N. & Wagner, A. Selection enhances protein evolvability by increasing mutational robustness and foldability. *Science* **370,** eabb5962. doi:https://doi.org/10.1126/science.abb5962 (2020).

32. Sarkisyan, K. S., Bolotin, D. A., Meer, M. V., Usmanova, D. R., Mishin, A. S., Sharonov, G. V., Ivankov, D. N., Bozhanova, N. G., Baranov, M. S., Soylemez, O., *et al.* Local fitness landscape of the green fluorescent protein. *Nature* **533,** 397–401. doi:10.1038/nature17995 (2016).

33. Papkou, A., Garcia-Pastor, L., Escudero, J. A. & Wagner, A. A rugged yet easily navigable fitness landscape. *Science* **382,** eadh3860. doi:10.1126/science.adh3860 (2023).

34. Hayden, E. J., Ferrada, E. & Wagner, A. Cryptic genetic variation promotes rapid evolutionary adaptation in an RNA enzyme. *Nature* **474,** 92–95. doi:10.1038/nature10083 (2011).

35. Ratcliff, W. C., Fankhauser, J. D., Rogers, D. W., Greig, D. & Travisano, M. Origins of multicellular evolvability in snowflake yeast. *Nature communications* **6,** 6102. doi:10.1038/ncomms7102 (2015).

36. Levin-Reisman, I., Ronin, I., Gefen, O., Braniss, I., Shoresh, N. & Balaban, N. Q. Antibiotic tolerance facilitates the evolution of resistance. *Science* **355,** 826–830. doi:10.1126/science.aaj2191 (2017).

37. Santos-Moreno, J., Tasiudi, E., Kusumawardhani, H., Stelling, J. & Schaerli, Y. Robustness and innovation in synthetic genotype networks. *Nature Communications* **14,** 2454. doi:10.1038/s41467-023-38033-3 (2023).

38. Dichtel-Danjoy, M.-L. & Félix, M.-A. Phenotypic neighborhood and micro-evolvability. *Trends in Genetics* **20,** 268–276. doi:10.1016/j.tig.2004.03.010 (2004).

39. Abzhanov, A., Kuo, W. P., Hartmann, C., Grant, B. R., Grant, P. R. & Tabin, C. J. The calmodulin pathway and evolution of elongated beak morphology in Darwin's finches. *Nature* **442,** 563–567. doi:10.1038/nature04843 (2006).

40. Mallarino, R., Grant, P. R., Grant, B. R., Herrel, A., Kuo, W. P. & Abzhanov, A. Two developmental modules establish 3D beak-shape variation in Darwin's finches. *Proceedings of the National Academy of Sciences* **108,** 4057–4062. doi:10.1073/pnas.1011480108 (2011).

41. al-Mosleh, S., Choi, G. P. T., Abzhanov, A. & Mahadevan, L. Geometry and dynamics link form, function, and evolution of finch beaks. *Proceedings of the National Academy of Sciences* **118,** e2105957118. doi:10.1073/pnas.2105957118 (2021).

42. Raff, R. A. *The shape of life – genes, development and evolution of animal forms* ISBN: 0-226-70266-9 (University of Chicago Press, 1996).

43. Wotton, K. R., Jiménez-Guri, E., Crombach, A., Janssens, H., Alcaine-Colet, A., Lemke, S., Schmidt-Ott, U. & Jaeger, J. Quantitative system drift compensates for altered maternal inputs to the gap gene network of the scuttle fly Megaselia abdita. *eLife* **4,** e04785. doi:10.7554/eLife.04785.001 (2015).

44. Gomez, C. & Pourquié, O. Developmental control of segment numbers in vertebrates. *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution* **312,** 533–544. doi:10.1002/jez.b.21305 (2009).

45. Krol, A. J., Roellig, D., Dequéant, M.-L., Tassy, O., Glynn, E., Hattem, G., Mushegian, A., Oates, A. C. & Pourquié, O. Evolutionary plasticity of segmentation clock networks. *Development* **138,** 2783–2792. doi:10.1242/dev.063834 (2011).

46. Gomez, C., Özbudak, E. M., Wunderlich, J., Baumann, D., Lewis, J. & Pourquié, O. Control of segment number in vertebrate embryos. *Nature* **454,** 335–339. doi:10.1038/nature07020 (2008).

47. Lázaro, J., Costanzo, M., Sanaki-Matsumiya, M., Girardot, C., Hayashi, M., Hayashi, K., Diecke, S., Hildebrandt, T. B., Lazzari, G., Wu, J., *et al.* A stem cell zoo uncovers intracellular scaling of developmental tempo across mammals. *Cell Stem Cell* **30,** 938–949 (2023).

48. Lewis, J. Autoinhibition with transcriptional delay: A simple mechanism for the zebrafish somitogenesis oscillator. *Current Biology* **13,** 1398–1408. doi:10.1016/s0960-9822(03)00534-7 (2003).

49. Riedel-Kruse, I. H., Müller, C. & Oates, A. C. Synchrony dynamics during initiation, failure, and rescue of the segmentation clock. *Science* **317,** 1911–1915. doi:10.1126/science.1142538 (2007).

50. Baker, R. E., Schnell, S. & Maini, P. K. Mathematical models for somite formation. *Current topics in developmental biology* **81,** 183–203. doi:10.1016/S0070-2153(07)81006-4 (2008).

51. Terry, A. J., Sturrock, M., Dale, J. K., Maroto, M. & Chaplain, M. A. A spatio-temporal model of Notch signalling in the zebrafish segmentation clock: conditions for synchronised oscillatory dynamics. *PLoS One* **6,** e16980. doi:10.1371/journal.pone.0016980 (2011).

52. Hester, S. D., Belmonte, J. M., Gens, J. S., Clendenon, S. G. & Glazier, J. A. A multi-cell, multi-scale model of vertebrate segmentation and somite formation. *PLoS computational biology* **7,** e1002155. doi:10.1371/journal.pcbi.1002155 (2011).

53. Cotterell, J., Robert-Moreno, A. & Sharpe, J. A local, self-organizing reaction-diffusion model can explain somite patterning in embryos. *Cell systems* **1,** 257–269. doi:10.1016/j.cels.2015.10.002 (2015).

54. Uriu, K., Liao, B.-K., Oates, A. C. & Morelli, L. G. From local resynchronization to global pattern recovery in the zebrafish segmentation clock. *eLife* **10,** e61358. doi:10.7554/eLife.61358 (2021).

55. François, P. & Mochulska, V. Waves, patterns, bifurcations: A tutorial review on the vertebrate segmentation clock. *Physics Reports* **1080,** 1–104. doi:https://doi.org/10.1016/j.physrep.2024.05.002 (2024).

56. Ward, A. B. & Brainerd, E. L. Evolution of axial patterning in elongate fishes. *Biological Journal of the Linnean Society* **90,** 97–116. doi:10.1111/j.1095-8312.2007.00714.x (2007).

57. Mehta, R. S., Ward, A. B., Alfaro, M. E. & Wainwright, P. C. Elongation of the body in eels. *Integrative and Comparative Biology* **50,** 1091–1105. doi:10.1093/icb/icq075 (2010).

58. Oliver, M. K. African cichlid fishes: morphological data and taxonomic insights from a genus-level survey of supraneurals, pterygiophores, and vertebral counts (Ovalentaria, Blenniiformes, Cichlidae, Pseudocrenilabrinae). *Biodiversity Data Journal* **12,** e130707. doi:10.3897/BDJ.12.e130707 (2024).

59. Bucklow, C., Genner, M., Turner, G., Maclaine, J., Benson, R. & Verd, B. A whole-body micro-CT scan library that captures the skeletal diversity of Lake Malawi cichlid fishes. *Scientific Data* **11,** 984. doi:10.1038/s41597-024-03687-1 (2024).

60. Yamahira, K. & Nishida, T. Latitudinal variation in axial patterning of the Medaka (Actinopterygii: Adrianichthyidae): Jordan's rule is substantiated by genetic variation in abdominal vertebral number. *Biological Journal of the Linnean Society* **96,** 856–866. doi:10.1111/j.1095-8312.2008.01150.x (2009).

61. Adams, C. & Maitland, P. Arctic charr in Britain and Ireland–15 species or one? *Ecology of Freshwater Fish* **16,** 20–28. doi:10.1111/j.1600-0633.2006.00179.x (2007).

62. McDowall, R. M. Variation in vertebral number in galaxiid fishes (Teleostei: Galaxiidae): a legacy of life history, latitude and length. *Environmental Biology of Fishes* **66,** 361–381. doi:10.1023/A:1023902922843 (2003).

63. Tibblin, P., Berggren, H., Nordahl, O., Larsson, P. & Forsman, A. Causes and consequences of intra-specific variation in vertebral number. *Scientific Reports* **6,** 26372. doi:10.1038/srep26372 (2016).

64. Campbell, C. S., Adams, C. E., Bean, C. W., Pilakouta, N. & Parsons, K. J. Evolvability under climate change: Bone development and shape plasticity are heritable and correspond with performance in Arctic charr (Salvelinus alpinus). *Evolution & Development* **23,** 333–350. doi:https://doi.org/10.1111/ede.12379 (2021).

65. Betancur-R, R., Wiley, E. O., Arratia, G., Acero, A., Bailly, N., Miya, M., Lecointre, G. & Orti, G. Phylogenetic classification of bony fishes. *BMC evolutionary biology* **17,** 1–40. doi:10.1186/s12862-017-0958-3 (2017).

66. Nelson, J. S., Grande, T. C. & Wilson, M. V. *Fishes of the World* ISBN: 9781119174844. doi:10.1002/9781119174844 (John Wiley & Sons, 2016).

67. R Core Team. *R: A Language and Environment for Statistical Computing* R Foundation for Statistical Computing (Vienna, Austria, 2022).

68. Paradis, E. & Schliep, K. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* **35,** 526–528. doi:10.1093/bioinformatics/bty633 (2019).

69. Revell, L. J. phytools 2.0: an updated R ecosystem for phylogenetic comparative methods (and other things). *PeerJ* **12,** e16505. doi:10.7717/peerj.16505 (2024).

70. Christ, B. & Ordahl, C. P. Early stages of chick somite development. *Anatomy and embryology* **191,** 381–396. doi:10.1007/bf00304424 (1995).

71. Richardson, M. K., Allen, S. P., Wright, G. M., Raynaud, A. & Hanken, J. Somite number and vertebrate evolution. *Development* **125,** 151–160. doi:10.1242/dev.125.2.151 (1998).

72. Morin-Kensicki, E. M., Melancon, E. & Eisen, J. S. Segmental relationship between somites and vertebral column in zebrafish. *Development* **129,** 3851–3860. doi:10.1242/dev.129.16.3851 (2002).

73. Criswell, K. E. & Gillis, J. A. Resegmentation is an ancestral feature of the gnathostome vertebral skeleton. *eLife* **9,** e51696. doi:10.7554/eLife.51696 (2020).

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

75. Gorodilov, Y. N. Description of the early ontogeny of the Atlantic salmon, Salmo salar, with a novel system of interval (state) identification. *Environmental Biology of fishes* **47,** 109–127. doi:10.1007/BF00005034 (1996).

76. Schröter, C., Herrgen, L., Cardona, A., Brouhard, G. J., Feldman, B. & Oates, A. C. Dynamics of zebrafish somitogenesis. *Developmental dynamics* **237,** 545–553. doi:10.1002/dvdy.21458 (2008).

77. Venzin, O. F. & Oates, A. C. What are you synching about? Emerging complexity of Notch signaling in the segmentation clock. *Developmental Biology* **460,** 40–54. doi:10.1016/j.ydbio.2019.06.024 (2020).

78. Bessho, Y., Hirata, H., Masamizu, Y. & Kageyama, R. Periodic repression by the bHLH factor Hes7 is an essential mechanism for the somite segmentation clock. *Genes & development* **17,** 1451–1456. doi:10.1101/gad.1092303 (2003).

79. Harima, Y., Takashima, Y., Ueda, Y., Ohtsuka, T. & Kageyama, R. Accelerating the tempo of the segmentation clock by reducing the number of introns in the Hes7 gene. *Cell Reports* **3,** 1–7. doi:10.1016/j.celrep.2012.11.012 (2013).

80. Bone, R. A., Bailey, C. S. L., Wiedermann, G., Ferjentsik, Z., Appleton, P. L., Murray, P. J., Maroto, M. & Dale, J. K. Spatiotemporal oscillations of Notch1, Dll1 and NICD are coordinated across the mouse PSM. *Development* **141,** 4806–4816. doi:10.1242/dev.115535 (2014).

81. Isomura, A. & Kageyama, R. Progress in understanding the vertebrate segmentation clock. *Nature Reviews Genetics,* 1–15. doi:10.1038/s41576-025-00813-6 (2025).

82. Horikawa, K., Ishimatsu, K., Yoshimoto, E., Kondo, S. & Takeda, H. Noise-resistant and synchronized oscillation of the segmentation clock. *Nature* **441,** 719–723. doi:10.1038/nature04861 (2006).

83. Webb, A. B., Lengyel, I. M., Jörg, D. J., Valentin, G., Jülicher, F., Morelli, L. G. & Oates, A. C. Persistence, period and precision of autonomous cellular oscillators from the zebrafish segmentation clock. *eLife* **5,** e08438. doi:10.7554/eLife.08438 (2016).

84. Keskin, S., Devakanmalai, G. S., Kwon, S. B., Vu, H. T., Hong, Q., Lee, Y. Y., Soltani, M., Singh, A., Ay, A. & Özbudak, E. M. Noise in the vertebrate segmentation clock is boosted by time delays but tamed by notch signaling. *Cell Reports* **23,** 2175–2185.e4. doi:https://doi.org/10.1016/j.celrep.2018.04.069 (2018).

85. Rohde, L. A., Bercowsky-Rama, A., Valentin, G., Naganathan, S. R., Desai, R. A., Strnad, P., Soroldoni, D. & Oates, A. C. Cell-autonomous timing drives the vertebrate segmentation clock's wave pattern. *eLife* **13,** RP93764. doi:10.7554/eLife.93764 (2024).

86. Cooke, J. & Zeeman, E. A clock and wavefront model for control of the number of repeated structures during animal morphogenesis. *Journal of Theoretical Biology* **58,** 455–476. doi:10.1016/s0022-5193(76)80131-2 (1976).

87. Palmeirim, I., Henrique, D., Ish-Horowicz, D. & Pourquié, O. Avian hairy gene expression identifies a molecular clock linked to vertebrate segmentation and somitogenesis. *Cell* **91,** 639–648. doi:10.1016/s0092-8674(00)80451-1 (1997).

88. Pourquié, O. A brief history of the segmentation clock. *Developmental Biology* **485,** 24–36. doi:10.1016/j.ydbio.2022.02.011 (2022).

89. Bessho, Y., Sakata, R., Komatsu, S., Shiota, K., Yamada, S. & Kageyama, R. Dynamic expression and essential functions of Hes7 in somite segmentation. *Genes & development* **15,** 2642–2647. doi:10.1101/gad.930601 (2001).

90. Oates, A. C. & Ho, R. K. Hairy/E (spl)-related (Her) genes are central components of the segmentation oscillator and display redundancy with the Delta/Notch signaling pathway in the formation of anterior segmental boundaries in the zebrafish. *Development* **129,** 2929–2946. doi:10.1242/dev.129.12.2929 (2002).

91. Forero, L. L., Narayanan, R., Huitem, L. F., VanBergen, M., Apschner, A., Peterson-Maduro, J., Logister, I., Valentin, G., Morelli, L. G., Oates, A. C. & Schulte-Merker, S. Segmentation of the zebrafish axial skeleton relies on notochord sheath cells and not on the segmentation clock. *eLife* **7,** e33843. doi:10.7554/eLife.33843 (2018).

92. Yoshioka-Kobayashi, K., Matsumiya, M., Niino, Y., Isomura, A., Kori, H., Miyawaki, A. & Kageyama, R. Coupling delay controls synchronized oscillation in the segmentation clock. *Nature* **580,** 119–123. doi:10.1038/s41586-019-1882-z (2020).

93. Morelli, L. G., Ares, S., Herrgen, L., Schröter, C., Jülicher, F. & Oates, A. C. Delayed coupling theory of vertebrate segmentation. *HFSP journal* **3,** 55–66. doi:10.2976/1.3027088 (2009).

94. Hirata, H., Bessho, Y., Kokubu, H., Masamizu, Y., Yamada, S., Lewis, J. & Kageyama, R. Instability of Hes7 protein is crucial for the somite segmentation clock. *Nature genetics* **36,** 750–754. doi:10.1038/ng1372 (2004).

95. Matsuda, M., Hayashi, H., Garcia-Ojalvo, J., Yoshioka-Kobayashi, K., Kageyama, R., Yamanaka, Y., Ikeya, M., Toguchida, J., Alev, C. & Ebisuya, M. Species-specific segmentation clock periods are due to differential biochemical reaction speeds. *Science* **369,** 1450–1455. doi:10.1126/science.aba7668 (2020).

96. Takashima, Y., Ohtsuka, T., González, A., Miyachi, H. & Kageyama, R. Intronic delay is essential for oscillatory expression in the segmentation clock. *Proceedings of the National Academy of Sciences* **108,** 3300–3305. doi:10.1073/pnas.1014418108 (2011).

97. Shimojo, H., Isomura, A., Ohtsuka, T., Kori, H., Miyachi, H. & Kageyama, R. Oscillatory control of Delta-like1 in cell interactions regulates dynamic gene expression and tissue morphogenesis. *Genes & development* **30,** 102–116. doi:10.1101/gad.270785.115 (2016).

98. Schröter, C., Ares, S., Morelli, L. G., Isakova, A., Hens, K., Soroldoni, D., Gajewski, M., Jülicher, F., Maerkl, S. J., Deplancke, B. & Oates, A. C. Topology and dynamics of the zebrafish segmentation clock core circuit. *PLOS Biology* **10,** 1–20. doi:10.1371/journal.pbio.1001364 (2012).

99. Trofka, A., Schwendinger-Schreck, J., Brend, T., Pontius, W., Emonet, T. & Holley, S. A. The Her7 node modulates the network topology of the zebrafish segmentation clock via sequestration of the Hes6 hub. *Development* **139,** 940–947. doi:10.1242/dev.073544 (2012).

100. Gajewski, M., Elmasri, H., Girschick, M., Sieger, D. & Winkler, C. Comparative analysis of her genes during fish somitogenesis suggests a mouse/chick-like mode of oscillation in medaka. *Development Genes and Evolution* **216,** 315–332. doi:10.1007/s00427-006-0059-6 (2006).

101. Giudicelli, F., Özbudak, E. M., Wright, G. J. & Lewis, J. Setting the tempo in development: an investigation of the zebrafish somite clock mechanism. *PLoS biology* **5,** e150. doi:10.1371/journal.pbio.0050150 (2007).

102. Eckalbar, W. L., Lasku, E., Infante, C. R., Elsey, R. M., Markov, G. J., Allen, A. N., Corneveaux, J. J., Losos, J. B., DeNardo, D. F., Huentelman, M. J., *et al.* Somitogenesis in the Anole lizard and Alligator reveals evolutionary convergence and divergence in the amniote segmentation clock. *Developmental Biology* **363,** 308–319. doi:10.1016/j.ydbio.2011.11.021 (2012).

103. Tsiairis, C. D. & Aulehla, A. Self-organization of embryonic genetic oscillators into spatiotemporal wave patterns. *Cell* **164,** 656–667. doi:10.1016/j.cell.2016.01.028 (2016).

104. Ay, A., Holland, J., Sperlea, A., Devakanmalai, G. S., Knierer, S., Sangervasi, S., Stevenson, A. & Özbudak, E. M. Spatial gradients of protein-level time delays set the pace of the traveling segmentation clock waves. *Development* **141,** 4158–4167. doi:10.1242/dev.111930 (2014).

105. Cinquin, O. Repressor dimerization in the zebrafish somitogenesis clock. *PLoS Computational Biology* **3,** e32. doi:10.1371/journal.pcbi.0030032 (2007).

106. Dubrulle, J., McGrew, M. J. & Pourquié, O. FGF signaling controls somite boundary position and regulates segmentation clock control of spatiotemporal Hox gene activation. *Cell* **106,** 219–232. doi:10.1016/s0092-8674(01)00437-8 (2001).

107. Simsek, M. F. & Özbudak, E. M. Spatial fold change of FGF signaling encodes positional information for segmental determination in zebrafish. *Cell Reports* **24,** 66–78. doi:10.1016/j.celrep.2018.06.023 (2018).

108. Naganathan, S. R. & Oates, A. C. Patterning and mechanics of somite boundaries in zebrafish embryos. *Seminars in cell & developmental biology* **107,** 170–178. doi:10.1016/j.semcdb.2020.04.014 (2020).

109. McDaniel, C., Simsek, M. F., Chandel, A. S. & Özbudak, E. M. Spatiotemporal control of pattern formation during somitogenesis. *Science Advances* **10,** eadk8937. doi:10.1126/sciadv.adk8937 (2024).

110. Lauschke, V. M., Tsiairis, C. D., François, P. & Aulehla, A. Scaling of embryonic patterning based on phase-gradient encoding. *Nature* **493,** 101–105. doi:10.1038/nature11804 (2013).

111. Simsek, M. F., Chandel, A. S., Saparov, D., Zinani, O. Q., Clason, N. & Özbudak, E. M. Periodic inhibition of Erk activity drives sequential somite segmentation. *Nature* **613,** 153–159. doi:10.1038/s41586-022-05527-x (2023).

112. Dubrulle, J. & Pourquié, O. fgf8 mRNA decay establishes a gradient that couples axial elongation to patterning in the vertebrate embryo. *Nature* **427,** 419–422. doi:10.1038/nature02216 (2004).

113. Steventon, B., Duarte, F., Lagadec, R., Mazan, S., Nicolas, J.-F. & Hirsinger, E. Species-specific contribution of volumetric growth and tissue convergence to posterior body elongation in vertebrates. *Development* **143,** 1732–1741. doi:10.1242/dev.126375 (2016).

114. Lawton, A. K., Nandi, A., Stulberg, M. J., Dray, N., Sneddon, M. W., Pontius, W., Emonet, T. & Holley, S. A. Regulated tissue fluidity steers zebrafish body elongation. *Development* **140,** 573–582. doi:10.1242/dev.090381 (2013).

115. Mongera, A., Michaut, A., Guillot, C., Xiong, F. & Pourquié, O. Mechanics of anteroposterior axis formation in vertebrates. *Annual Review of Cell and Developmental Biology* **35,** 259–283. doi:10.1146/annurev-cellbio-100818-125436 (2019).

116. Banavar, S. P., Carn, E. K., Rowghanian, P., Stooke-Vaughan, G., Kim, S. & Campàs, O. Mechanical control of tissue shape and morphogenetic flows during vertebrate body axis elongation. *Scientific Reports* **11,** 1–14. doi:10.1038/s41598-021-87672-3 (2021).

117. Zhang, L., Kendrick, C., Jülich, D. & Holley, S. A. Cell cycle progression is required for zebrafish somite morphogenesis but not segmentation clock function. *Development* **135,** 2065–2070. doi:10.1242/dev.022673 (2008).

118. Romanos, M., Salisbury, T., Stephan, S., Lansford, R., Degond, P., Trescases, A. & Bénazéraf, B. Differential proliferation regulates multi-tissue morphogenesis during embryonic axial

extension: Integrating viscous modeling and experimental approaches. *Development* **151,** dev202836. doi:10.1242/dev.202836 (2024).

119. Bénazéraf, B., Francois, P., Baker, R. E., Denans, N., Little, C. D. & Pourquié, O. A random cell motility gradient downstream of FGF controls elongation of an amniote embryo. *Nature* **466,** 248–52. doi:10.1038/nature09151 (2010).

120. Mongera, A., Rowghanian, P., Gustafson, H. J., Shelton, E., Kealhofer, D. A., Carn, E. K., Serwane, F., Lucio, A. A., Giammona, J. & Campàs, O. A fluid-to-solid jamming transition underlies vertebrate body axis elongation. *Nature* **561,** 401–405. doi:10.1038/s41586-018-0479-2 (2018).

121. Xiong, F., Ma, W., Bénazéraf, B., Mahadevan, L. & Pourquié, O. Mechanical coupling coordinates the co-elongation of axial and paraxial tissues in avian embryos. *Developmental Cell* **55,** 354–366. doi:10.1016/j.devcel.2020.08.007 (2020).

122. Michaut, A., Mongera, A., Gupta, A., Serra, M., Rigoni, P., Lee, J. G., Duarte, F., Hall, A. R., Mahadevan, L., Guevorkian, K. & Pourquié, O. Activity-driven extracellular volume expansion drives vertebrate axis elongation. *bioRxiv.* doi:10.1101/2022.06.27.497799 (2022).

123. Keller, R. & Sutherland, A. Convergent extension in the amphibian, Xenopus laevis. *Current topics in developmental biology* **136,** 271–317. doi:10.1016/S1534-5807(02)00197-1 (2020).

124. Thomson, L., Muresan, L. & Steventon, B. The zebrafish presomitic mesoderm elongates through compaction-extension. *Cells & Development* **168,** 203748. doi:10.1016/j.cdev.2021.203748 (2021).

125. Jouve, C., Iimura, T. & Pourquie, O. Onset of the segmentation clock in the chick embryo: evidence for oscillations in the somite precursors in the primitive streak. *Development* **129,** 1107–1117. doi:10.1242/dev.129.5.1107 (2002).

126. Falk, H. J., Tomita, T., Mönke, G., McDole, K. & Aulehla, A. Imaging the onset of oscillatory signaling dynamics during mouse embryo gastrulation. *Development* **149,** dev200083. doi:10.1242/dev.200083 (2022).

127. Mara, A., Schroeder, J., Chalouni, C. & Holley, S. A. Priming, initiation and synchronization of the segmentation clock by deltaD and deltaC. *Nature Cell Biology* **9,** 523–530. doi:10.1038/ncb1578 (2007).

128. Tenin, G., Wright, D., Ferjentsik, Z., Bone, R., McGrew, M. J. & Maroto, M. The chick somitogenesis oscillator is arrested before all paraxial mesoderm is segmented into somites. *BMC Developmental Biology* **10,** 1–12. doi:10.1186/1471-213X-10-24 (2010).

129. Tam, P. P. & Tan, S.-S. The somitogenetic potential of cells in the primitive streak and the tail bud of the organogenesis-stage mouse embryo. *Development* **115,** 703–715. doi:10.1242/dev.115.3.703 (1992).

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

131. Young, T., Rowland, J. E., van de Ven, C., Bialecka, M., Novoa, A., Carapuco, M., van Nes, J., de Graaff, W., Duluc, I., Freund, J.-N., *et al.* Cdx and Hox genes differentially regulate posterior

axial growth in mammalian embryos. *Developmental cell* **17,** 516–526. doi:10.1016/j.devcel.2009.08.010 (2009).

132. Woltering, J. From lizard to snake; behind the evolution of an extreme body plan. *Current genomics* **13,** 289–299. doi:10.2174/138920212800793302 (2012).

133. Monk, N. A. Oscillatory expression of Hes1, p53, and NF-κB driven by transcriptional time delays. *Current biology* **13,** 1409–1413. doi:10.1016/S0960-9822(03)00494-9 (2003).

134. Goodwin, B. C. Oscillatory behavior in enzymatic control processes. *Advances in enzyme regulation* **3,** 425–437. doi:10.1016/0065-2571(65)90067-1 (1965).

135. Ay, A., Knierer, S., Sperlea, A., Holland, J. & Özbudak, E. M. Short-lived Her proteins drive robust synchronized oscillations in the zebrafish segmentation clock. *Development* **140,** 3244–3253. doi:10.1242/dev.093278 (2013).

136. Schröter, C. & Oates, A. C. Segment number and axial identity in a segmentation clock period mutant. *Current Biology* **20,** 1254–1258. doi:10.1016/j.cub.2010.05.071 (2010).

137. Kuramoto, Y. in *Chemical oscillations, waves, and turbulence* 111–140 (Springer, 1984). ISBN: 978-3-642-69689-3. doi:10.1007/978-3-642-69689-3.

138. Jörg, D. J., Morelli, L. G., Soroldoni, D., Oates, A. C. & Jülicher, F. Continuum theory of gene expression waves during vertebrate segmentation. *New Journal of Physics* **17,** 093042. doi:10.1088/1367-2630/17/9/093042 (2015).

139. Herrgen, L., Ares, S., Morelli, L. G., Schroter, C., Julicher, F. & Oates, A. C. Intercellular coupling regulates the period of the segmentation clock. *Current Biology* **20,** 1244–1253. doi:10.1016/j.cub.2010.06.034 (2010).

140. Uriu, K., Bhavna, R., Oates, A. C. & Morelli, L. G. A framework for quantification and physical modeling of cell mixing applied to oscillator synchronization in vertebrate somitogenesis. *Biology Open* **6,** 1235–1244. doi:10.1242/bio.025148 (2017).

141. Ho, C., Jutras-Dubé, L., Zhao, M. L., Mönke, G., Kiss, I. Z., François, P. & Aulehla, A. Nonreciprocal synchronization in embryonic oscillator ensembles. *Proceedings of the National Academy of Sciences* **121,** e2401604121. doi:10.1073/pnas.2401604121 (2024).

142. Murray, P. J., Maini, P. K. & Baker, R. E. The clock and wavefront model revisited. *Journal of Theoretical Biology* **283,** 227–238. doi:10.1016/j.jtbi.2011.05.004 (2011).

143. Di-Poi, N., Montoya-Burgos, J. I., Miller, H., Pourquié, O., Milinkovitch, M. C. & Duboule, D. Changes in Hox genes' structure and function during the evolution of the squamate body plan. *Nature* **464,** 99–103. doi:10.1038/nature08789 (2010).

144. Salis, P., Lee, S.-h., Roux, N., Lecchini, D. & Laudet, V. The real Nemo movie: Description of embryonic development in Amphiprion ocellaris from first division to hatching. *Developmental Dynamics* **250,** 1651–1667. doi:10.1002/dvdy.354 (2021).

145. Kawakami, T., Yamada, Y., Tanaka, S. & Tsukamoto, K. Prolongation of somitogenesis in two anguilliform species, the Japanese eel Anguilla japonica and pike eel Muraenesox cinereus, with refined descriptions of their early development. *Journal of Fish Biology* **90,** 1533–1547. doi:10.1111/jfb.13249 (2017).

146. Marconi, A., Yang, C. Z., McKay, S. & Santos, M. E. Morphological and temporal variation in early embryogenesis contributes to species divergence in Malawi cichlid fishes. *Evolution & Development* **25,** 170–193. doi:10.1111/ede.12429 (2023).

147. Hinaux, H., Pottin, K., Chalhoub, H., Pere, S., Elipot, Y., Legendre, L. & Rétaux, S. A developmental staging table for Astyanax mexicanus surface fish and Pachón cavefish. *Zebrafish* **8,** 155–165. doi:10.1089/zeb.2011.0713 (2011).

148. Tsai, H.-Y., Chang, M., Liu, S.-C., Abe, G. & Ota, K. G. Embryonic development of goldfish (Carassius auratus): a model for the study of evolutionary change in developmental mechanisms by artificial selection. *Developmental dynamics* **242,** 1262–1283. doi:10.1002/dvdy.24022 (2013).

149. Johnston, I. A., Vieira, V. L. & Abercromby, M. Temperature and myogenesis in embryos of the Atlantic herring Clupea harengus. *Journal of Experimental Biology* **198,** 1389–1403. doi:10.1242/jeb.198.6.1389 (1995).

150. Kawakami, T., Okouchi, H., Aritaki, M., Aoyama, J. & Tsukamoto, K. Embryonic development and morphology of eggs and newly hatched larvae of Pacific herring Clupea pallasii. *Fisheries Science* **77,** 183–190. doi:10.1007/s12562-010-0317-4 (2011).

151. Cucchi, P., Sucré, E., Santos, R., Leclère, J., Charmantier, G. & Castille, R. Embryonic development of the sea bass Dicentrarchus labrax. *Helgoland Marine Research* **66,** 199–209. doi:10.1007/s10152-011-0262-3 (2012).

152. Von Herbing Hunt, I., Boutilier, R., Miyake, T. & Hall, B. Effects of temperature on morphological landmarks critical to growth and survival in larval Atlantic cod (Gadus morhua). *Marine Biology* **124,** 593–606. doi:10.1007/BF00351041 (1996).

153. Hall, T. E., Smith, P. & Johnston, I. A. Stages of embryonic development in the Atlantic cod Gadus morhua. *Journal of morphology* **259,** 255–270. doi:10.1002/jmor.10222 (2004).

154. Arakawa, T., Kanno, Y., Akiyama, N., Kitano, T., Nakatsuji, N. & Nakatsuji, T. Stages of embryonic development of the Ice Goby (Shiro-uo), Leucopsarion petersii. *Zoological science* **16,** 761–773. doi:10.2108/zsj.16.761 (1999).

155. Liang, Z., Yi, B., Yu, Z. & Wang, N. Spawning areas and early development of Long spiky-head carp (Luciobrama macrocephalus) in the Yangtze River and Pearl River, China. *Hydrobiologia* **490,** 169–179. doi:10.1023/A:1023426909635 (2003).

156. Fujimoto, T., Kataoka, T., Sakao, S., Saito, T., Yamaha, E. & Arai, K. Developmental stages and germ cell lineage of the loach (Misgurnus anguillicaudatus). *Zoological Science* **23,** 977–989. doi:10.2108/zsj.23.977 (2006).

157. Iwamatsu, T. Stages of normal development in the Medaka Oryzias latipes. *Mechanisms of Development* **121.** Medaka, 605–618. doi:https://doi.org/10.1016/j.mod.2004.03.012 (2004).

158. Seleit, A., Brettell, I., Fitzgerald, T., Vibe, C., Loosli, F., Wittbrodt, J., Naruse, K., Birney, E. & Aulehla, A. Modular control of vertebrate axis segmentation in time and space. *The EMBO Journal* **43,** 4068–4091. doi:10.1038/s44318-024-00186-2 (2024).

159. Koecke, H.-U. Normalstadien der embryonal-entwicklung bei der hausente (Anas bosch as domestic A). *Embryologia* **4,** 55–78. doi:10.1111/j.1440-169X.1958.tb00147.x (1958).

160. Olea, G. B. & Sandoval, M. T. Embryonic development of Columba livia (Aves: Columbiformes) from an altricial-precocial perspective. *Revista Colombiana de Ciencias Pecuarias* **25,** 3–13. doi:10.17533/udea.rccp.324728 (2012).

161. Christ, B., Huang, R. & Wilting, J. The development of the avian vertebral column. *Anatomy and embryology* **202,** 179–194. doi:10.1007/s004290000114 (2000).

162. Yamasaki, M. & Tonosaki, A. Developmental Stages of the Society Finch, Lonchura striata var. dornestica. *Development, Growth & Differentiation* **30,** 515–542. doi:https://doi.org/10.1111/j.1440-169X.1988.00515.x (1988).

163. Herbert, C. A times series of embryonic development stages of the Adelie penguin (Pygoscelis adeliae) from Signy Island, South Orkney Islands. *British Antarctic Survey Bulletin* **14,** 45–67.

164. Onimaru, K., Motone, F., Kiyatake, I., Nishida, K. & Kuraku, S. A staging table for the embryonic development of the brownbanded bamboo shark (Chiloscyllium punctatum). *Developmental Dynamics* **247,** 712–723. doi:10.1002/dvdy.24623 (2018).

165. Richardson, M. K., Admiraal, J. & Wright, G. M. Developmental anatomy of lampreys. *Biological Reviews* **85,** 1–33. doi:https://doi.org/10.1111/j.1469-185X.2009.00092.x (2010).

166. Haldiman, J. T. Bovine somite development and vertebral anlagen establishment. *Anatomia, Histologia, Embryologia* **10,** 289–309. doi:10.1111/j.1439-0264.1981.tb00695.x (1981).

167. Chambers, P. L. & Hearn, J. P. Embryonic, foetal and placental development in the Common marmoset monkey (Callithrix jacchus). *Journal of Zoology* **207,** 545–561. doi:https://doi.org/10.1111/j.1469-7998.1985.tb04951.x (1985).

168. Casteleyn, C., Bakker, J., Breugelmans, S., Kondova, I., Saunders, J., Langermans, J. A. M., Cornillie, P., den Broeck, W. V., Loo, D. V., Hoorebeke, L. V., Bosseler, L., Chiers, K. & Decostere, A. Anatomical description and morphometry of the skeleton of the Common marmoset (Callithrix jacchus). *Laboratory Animals* **46,** 152–163. doi:10.1258/la.2012.011167 (2012).

169. Chang, C.-H. & Jang, C.-M. *On the Processing and Mounting of a Skeleton of a White Rhinoceros, Ceratotherium sinum* in. **17.** https://api.semanticscholar.org/CorpusID:130776191 (2004), 58–69.

170. Tam, P. P. L. The control of somitogenesis in mouse embryos. *Development* **65,** 103–128. doi:10.1242/dev.65.Supplement.103 (1981).

171. Beaudoin, S., Barbet, P. & Bargy, F. Developmental stages in the Rabbit embryo: Guidelines to choose an appropriate experimental model. *Fetal Diagnosis and Therapy* **18,** 422–427. doi:10.1159/000073136 (2003).

172. Proks, P., Stehlik, L., Nyvltova, I., Necas, A., Vignoli, M. & Jekl, V. Vertebral formula and congenital abnormalities of the vertebral column in rabbits. *The Veterinary Journal* **236,** 80–88. doi:https://doi.org/10.1016/j.tvjl.2018.04.016 (2018).

173. Al-Mukhaini, N., Ba-Omar, T., Mahmoud, I. & Albarwani, S. Embryonic staging of the Green Turtles, Chelonia mydas. *Zoology in the Middle East* **51,** 39–50. doi:10.1080/09397140.2010.10638439 (2010).

174. Yntema, C. A series of stages in the embryonic development of Chelydra serpentina. *Journal of morphology* **125,** 219–251. doi:10.1002/jmor.1051250207 (1968).

175. Matsubara, Y., Sakai, A., Kuroiwa, A. & Suzuki, T. Efficient embryonic culture method for the Japanese striped snake, Laphe quadrivirgata, and its early developmental stages. *Development,*

Growth & Differentiation **56,** 573–582. doi:https://doi.org/10.1111/dgd.12157 (2014).

176. Werneburg, I., Hugi, J., Müller, J. & Sánchez-Villagra, M. R. Embryogenesis and ossification of Emydura subglobosa (Testudines, Pleurodira, Chelidae) and patterns of turtle development. *Developmental Dynamics* **238,** 2770–2786. doi:https://doi.org/10.1002/dvdy.22104 (2009).

177. Ward, A. B. & Mehta, R. S. Axial elongation in fishes: Using morphological approaches to elucidate developmental mechanisms in studying body shape. *Integrative and Comparative Biology* **50,** 1106–1119. doi:10.1093/icb/icq029 (2010).

178. Jimenez, Y. E., Lucas, K. N., Long Jr, J. H. & Tytell, E. D. Flexibility is a hidden axis of biomechanical diversity in fishes. *Journal of Experimental Biology* **226,** jeb245308. doi:10.1242/jeb.245308 (2023).

179. Rekaik, H. & Duboule, D. A CTCF-dependent mechanism underlies the Hox timer: relation to a segmented body plan. *Current Opinion in Genetics & Development* **85,** 102160. doi:0.1016/j.gde.2024.102160 (2024).

180. Thomson, R. C., Spinks, P. Q. & Shaffer, H. B. A global phylogeny of turtles reveals a burst of climate-associated diversification on continental margins. *Proceedings of the National Academy of Sciences* **118,** e2012215118. doi:10.1073/pnas.2012215118 (2021).

181. Müller, J., Scheyer, T. M., Head, J. J., Barrett, P. M., Werneburg, I., Ericson, P. G., Pol, D. & Sánchez-Villagra, M. R. Homeotic effects, somitogenesis and the evolution of vertebral numbers in recent and fossil amniotes. *Proceedings of the National Academy of Sciences* **107,** 2118–2123. doi:10.1073/pnas.0912622107 (2010).

182. Lyson, T. R. & Bever, G. S. Origin and Evolution of the Turtle Body Plan. *Annual Review of Ecology, Evolution, and Systematics* **51,** 143–166. doi:https://doi.org/10.1146/annurev-ecolsys-110218-024746 (2020).

183. Yu, D., Ren, Y., Uesaka, M., Beavan, A. J., Muffato, M., Shen, J., Li, Y., Sato, I., Wan, W., Clark, J. W., *et al.* Hagfish genome elucidates vertebrate whole-genome duplication events and their evolutionary consequences. *Nature Ecology & Evolution* **8,** 519–535. doi:10.1038/s41559-023-02299-z (2024).

184. Marlétaz, F., Timoshevskaya, N., Timoshevskiy, V. A., Parey, E., Simakov, O., Gavriouchkina, D., Suzuki, M., Kubokawa, K., Brenner, S., Smith, J. J., *et al.* The hagfish genome and the evolution of vertebrates. *Nature* **627,** 811–820. doi:10.1038/s41586-024-07070-3 (2024).

185. Dequéant, M.-L., Glynn, E., Gaudenz, K., Wahl, M., Chen, J., Mushegian, A. & Pourquié, O. A complex oscillating network of signaling genes underlies the mouse segmentation clock. *science* **314,** 1595–1598. doi:10.1126/science.1133141 (2006).

186. Kageyama, R., Ohtsuka, T. & Kobayashi, T. The Hes gene family: repressors and oscillators that orchestrate embryogenesis. *Development* **134,** 1243–1251. doi:10.1242/dev.000786 (2007).

187. Gazave, E., Guillou, A. & Balavoine, G. History of a prolific family: the Hes/Hey-related genes of the annelid Platynereis. *EvoDevo* **5,** 1–33. doi:10.1186/2041-9139-5-29 (2014).

188. Kuretani, A., Yamamoto, T., Taira, M. & Michiue, T. Evolution of hes gene family in vertebrates: the hes5 cluster genes have specifically increased in frogs. *BMC Ecology and Evolution* **21,** 1–15. doi:10.1186/s12862-021-01879-6 (2021).

189. Liao, B.-K., Jörg, D. J. & Oates, A. C. Faster embryonic segmentation through elevated Delta-Notch signalling. *Nature Communications* **7,** 1–12. doi:10.1038/ncomms11861 (2016).

190. Aulehla, A., Wehrle, C., Brand-Saberi, B., Kemler, R., Gossler, A., Kanzler, B. & Herrmann, B. G. Wnt3a plays a major role in the segmentation clock controlling somitogenesis. *Developmental cell* **4,** 395–406. doi:10.1016/s1534-5807(03)00055-8 (2003).

191. Sonnen, K. F., Lauschke, V. M., Uraji, J., Falk, H. J., Petersen, Y., Funk, M. C., Beaupeux, M., François, P., Merten, C. A. & Aulehla, A. Modulation of phase shift between Wnt and Notch signaling oscillations controls mesoderm segmentation. *Cell* **172,** 1079–1090. doi:10.1016/j.cell.2018.01.026 (2018).

192. Sun, Y.-H., Chen, S.-P., Wang, Y.-P., Hu, W. & Zhu, Z.-Y. Cytoplasmic impact on cross-genus cloned fish derived from transgenic common carp (Cyprinus carpio) nuclei and goldfish (Carassius auratus) enucleated eggs. *Biology of reproduction* **72,** 510–515. doi:10.1095/biolreprod.104.031302 (2005).

193. Sequence of formation and inheritance of meristic variation in the post-cranial axial skeleton of Atlantic salmon (*Salmo salar*). *Journal of Fish Biology* **106,** 954–968. doi:10.1111/jfb.16004 (2025).

194. Garside, E. & Fry, F. A possible relationship between yolk size and differentiation in trout embryos. *Canadian Journal of Zoology* **37,** 383–386. doi:10.1139/z59-044 (1959).

195. Arnold, L. M., Smith, W. D., Spencer, P. D., Evans, A. N., Heppell, S. A. & Heppell, S. S. The role of maternal age and context-dependent maternal effects in the offspring provisioning of a long-lived marine teleost. *Royal Society Open Science* **5,** 170966. doi:10.1098/rsos.170966 (2018).

196. Fowler, J. A. Control of Vertebral Number in Teleosts - an Embryological Problem. *The Quarterly Review of Biology* **45,** 148–167. doi:10.1086/406492 (1970).

197. McDowall, R. M. Jordan's and other ecogeographical rules, and the vertebral number in fishes. *Journal of Biogeography* **35,** 501–508. doi:10.1111/j.1365-2699.2007.01823.x (2008).

198. Reyes Corral, W. D. & Aguirre, W. E. Effects of temperature and water turbulence on vertebral number and body shape in Astyanax mexicanus (Teleostei: Characidae). *PLoS One* **14,** e0219677. doi:10.1371/journal.pone.0219677 (2019).

199. Zhang, W., Scerbo, P., Delagrange, M., Candat, V., Mayr, V., Vriz, S., Distel, M., Ducos, B. & Bensimon, D. Fgf8 dynamics and critical slowing down may account for the temperature independence of somitogenesis. *Communications Biology* **5,** 113. doi:10.1038/s42003-022-03053-0 (2022).

200. Delaune, E. A., François, P., Shih, N. P. & Amacher, S. L. Single-cell-resolution imaging of the impact of notch signaling and mitosis on segmentation clock dynamics. *Developmental Cell* **23,** 995–1005. doi:10.1016/j.devcel.2012.09.009 (2012).

201. Uriu, K., Morishita, Y. & Iwasa, Y. Random cell movement promotes synchronization of the segmentation clock. *Proceedings of the National Academy of Sciences* **107,** 4979–4984. doi:10.1073/pnas.0907122107 (2010).

202. Uriu, K., Ares, S., Oates, A. C. & Morelli, L. G. Dynamics of mobile coupled phase oscillators. *Physical Review E* **87,** 032911. doi:10.1103/PhysRevE.87.032911 (2013).

203. De Visser, J. A. G. & Krug, J. Empirical fitness landscapes and the predictability of evolution. *Nature Reviews Genetics* **15,** 480–490. doi:10.1038/s41559-022-01867-z (2014).

204. Kanki, J. & Ho, R. The development of the posterior body in zebrafish. *Development* **124,** 881–893. doi:10.1242/dev.124.4.881 (1997).

205. Naganathan, S. R., Popović, M. & Oates, A. C. Left–right symmetry of zebrafish embryos requires somite surface tension. *Nature* **605,** 516–521. doi:10.1038/s41586-022-04646-9 (2022).

206. Pitsianis, N., Iliopoulos, A.-S., Floros, D. & Sun, X. *Spaceland Embedding of Sparse Stochastic Graphs* in *IEEE High Performance Extreme Computing Conference* (2019). doi:10.1109/HPEC.2019.8916505.

207. Wagner, G., Pavlicev, M. & Cheverud, J. The road to modularity. *Nature Reviews Genetics* **8,** 921–931. doi:10.1038/nrg2267 (2007).

208. Hughes, L. C., Ortí, G., Huang, Y., Sun, Y., Baldwin, C. C., Thompson, A. W., Arcila, D., Betancur-R, R., Li, C., Becker, L., *et al.* Comprehensive phylogeny of ray-finned fishes (Actinopterygii) based on transcriptomic and genomic data. *Proceedings of the National Academy of Sciences* **115,** 6249–6254. doi:10.1073/pnas.1719358115 (2018).

209. Malinsky, M., Svardal, H., Tyers, A. M., Miska, E. A., Genner, M. J., Turner, G. F. & Durbin, R. Whole-genome sequences of malawi cichlids reveal multiple radiations interconnected by gene flow. *Nature Ecology & Evolution* **2,** 1940–1955. doi:10.1038/s41559-018-0717-x (2018).

210. Turner, G. F., Seehausen, O., Knight, M. E., Allender, C. J. & Robinson, R. L. How many species of cichlid fishes are there in African lakes? *Molecular Ecology* **10,** 793–806. doi:https://doi.org/10.1046/j.1365-294x.2001.01200.x (2001).

211. Santos, M. E., Lopes, J. F. & Kratochwil, C. F. East African cichlid fishes. *EvoDevo* **14,** 1. doi:10.1186/s13227-022-00205-5 (2023).

212. Ronco, F., Matschiner, M., Böhne, A., Boila, A., Büscher, H. H., El Taher, A., Indermaur, A., Malinsky, M., Ricci, V., Kahmen, A., *et al.* Drivers and dynamics of a massive adaptive radiation in cichlid fishes. *Nature* **589,** 76–81. doi:10.1038/s41586-020-2930-4 (2021).

213. De-Kayne, R., Schley, R., Barth, J. M., Campillo, L. C., Chaparro-Pedraza, C., Joshi, J., Salzburger, W., Van Bocxlaer, B., Cotoras, D. D., Fruciano, C., *et al.* Why do some lineages radiate while others do not? Perspectives for future research on adaptive radiations. *Cold Spring Harbor perspectives in biology* **17,** a041448. doi:10.1101/cshperspect.a041448 (2025).

214. Turner, G., Ngatunga, B. & Genner, M. The Natural History of the Satellite Lakes of Lake Malawi. *EcoEvoRxiv.* doi:10.32942/osf.io/sehdq (2019).

215. Turner, G., Ngatunga, B. P. & Genner, M. J. Astatotilapia species (Teleostei, Cichlidae) from Malawi, Mozambique and Tanzania, excluding the basin of Lake Victoria. *EcoEvoRxiv.* doi:10.32942/osf.io/eu6rx (2021).

216. Genner, M. J., Nichols, P., Carvalho, G. R., Robinson, R. L., Shaw, P. W., Smith, A. & Turner, G. F. Evolution of a cichlid fish in a Lake Malawi satellite lake. *Proceedings of the Royal Society B: Biological Sciences* **274,** 2249–2257. doi:10.1098/rspb.2007.0619 (2007).

217. Turner, G. F., Robinson, R. L., Shaw, P. W., Carvalho, G. R. & Snoeks, J. in *The Cichlid Diversity of Lake Malawi/Nyasa/Niassa: Identification: Distribution and Taxonomy* 198–251 (Cichlid Press, 2004).

218. Ivory, S. J., Blome, M. W., King, J. W., McGlue, M. M., Cole, J. E. & Cohen, A. S. Environmental change explains cichlid adaptive radiation at Lake Malawi over the past 1.2 million years. *Proceedings of the National Academy of Sciences* **113,** 11895–11900. doi:10.1073/pnas.1611028113 (2016).

219. Fujimura, K. & Okada, N. Development of the embryo, larva and early juvenile of Nile tilapia Oreochromis niloticus (Pisces: Cichlidae). Developmental staging system. *Development, growth & differentiation* **49,** 301–324 (2007).

220. Woltering, J. M., Holzem, M., Schneider, R. F., Nanos, V. & Meyer, A. The skeletal ontogeny of Astatotilapia burtoni–a direct-developing model system for the evolution and development of the teleost body plan. *BMC Developmental Biology* **18,** 1–23. doi:10.1186/s12861-018-0166-4 (2018).

221. Pospisilova, A., Brejcha, J., Miller, V., Holcman, R., Šanda, R. & Stundl, J. Embryonic and larval development of the Northern pike: An emerging fish model system for evo-devo research. *Journal of Morphology* **280,** 1118–1140. doi:10.1002/jmor.21005 (2019).

222. Axelrod, H. R. & Burgess, W. E. *African cichlids of lakes Malawi and Tanganyika* ISBN: 9780866228879 (TFH Publications, 1989).

223. Clark, B., Elkin, J., Marconi, A., Turner, G. F., Smith, A. M., Joyce, D., Miska, E. A., Juntti, S. A. & Santos, M. E. Oca2 targeting using CRISPR/Cas9 in the Malawi cichlid Astatotilapia calliptera. *Royal Society Open Science* **9,** 220077. doi:10.1098/rsos.220077 (2022).

224. Morrison, C., Pohajdak, B., Henry, M. & Wright Jr, J. Structure and enzymatic removal of the chorion of embryos of the Nile tilapia. *Journal of fish biology* **63,** 1439–1453. doi:10.1111/j.1095-8649.2003.00257.x (2003).

225. Segev-Hadar, A., Slosman, T., Rozen, A., Sherman, A., Cnaani, A. & Biran, J. Genome editing using the CRISPR-Cas9 system to generate a solid-red germline of Nile tilapia (Oreochromis niloticus). *The CRISPR Journal* **4,** 583–594. doi:10.1089/crispr.2020.0115 (2021).

226. Nasevicius, A. & Ekker, S. C. Effective targeted gene 'knockdown'in zebrafish. *Nature genetics* **26,** 216–220. doi:10.1038/79951 (2000).

227. Bill, B. R., Petzold, A. M., Clark, K. J., Schimmenti, L. A. & Ekker, S. C. A primer for morpholino use in zebrafish. *Zebrafish* **6,** 69–77. doi:10.1089=zeb.2008.0555 (2009).

228. Hamada, H., Watanabe, M., Lau, H. E., Nishida, T., Hasegawa, T., Parichy, D. M. & Kondo, S. Involvement of Delta/Notch signaling in zebrafish adult pigment stripe patterning. *Development* **141,** 318–324. doi:10.1242/dev.099804 (2014).

229. Hsu, C.-H., Liou, G.-G. & Jiang, Y.-J. Nicastrin deficiency induces tyrosinase-dependent depigmentation and skin inflammation. *Journal of Investigative Dermatology* **140,** 404–414. doi:10.1016/j.jid.2019.07.702 (2020).

230. Choi, H. M., Schwarzkopf, M., Fornace, M. E., Acharya, A., Artavanis, G., Stegmaier, J., Cunha, A. & Pierce, N. A. Third-generation in situ hybridization chain reaction: multiplexed, quantitative, sensitive, versatile, robust. *Development* **145,** dev165753. doi:10.1242/dev.165753 (2018).

231. Clark, E., Peel, A. D. & Akam, M. Arthropod segmentation. *Development* **146,** dev170480. doi:10.1242/dev.170480 (2019).

232. Yamagami, K. Isolation of a choriolytic enzyme (hatching enzyme) of the teleost, Oryzias latipes. *Developmental Biology* **29,** 343–348. doi:10.1111/j.1440-169x.1982.00135.x (1972).

233. Wakamatsu, Y. *Preparation of Medaka Hatching Enzyme* tech. rep. 9 (Nagoya University, 1997), 49–50.

234. Jung, G., Hug, M., Halter, C., Friesenhengst, A., Walzer, J. & Czerny, T. Diffusion of small molecules into medaka embryos improved by electroporation. *BMC Biotechnology* **13,** 1–14. doi:10.1186/1472-6750-13-53 (2013).

235. Mizutari, K., Fujioka, M., Hosoya, M., Bramhall, N., Okano, H. J., Okano, H. & Edge, A. S. Notch inhibition induces cochlear hair cell regeneration and recovery of hearing after acoustic trauma. *Neuron* **77,** 58–69. doi:10.1016/j.neuron.2012.10.032 (2013).

236. Oates, A. C., Mueller, C. & Ho, R. K. Cooperative function of deltaC and her7 in anterior segment formation. *Developmental biology* **280,** 133–149. doi:10.1016/j.ydbio.2005.01.010 (2005).

237. Onai, T., Aramaki, T., Inomata, H., Hirai, T. & Kuratani, S. On the origin of vertebrate somites. *Zoological Letters* **1,** 1–10. doi:10.1186/s40851-015-0033-0 (2015).

238. Vaglia, J. L., White, K. & Case, A. Evolving possibilities: postembryonic axial elongation in salamanders with biphasic (Eurcyea cirrigera, Eurycea longicauda, Eurycea quadridigitata) and paedomorphic life cycles (Eurycea nana and Ambystoma mexicanum). *Acta Zoologica* **93,** 2–13. doi:10.1111/j.1463-6395.2010.00475.x (2012).

239. Jenkins, R. P., Hanisch, A., Soza-Ried, C., Sahai, E. & Lewis, J. Stochastic regulation of her1/7 gene expression is the source of noise in the zebrafish somite clock counteracted by Notch signalling. *PLOS Computational Biology* **11,** e1004459. doi:10.1371/journal.pcbi.1004459 (2015).

240. Bruce, H. S., Jerz, G., Kelly, S. R., McCarthy, J., Pomerantz, A., Senevirathne, G., Sherrard, A., Sun, D. A., Wolff, C. & Patel, N. H. Hybridization Chain Reaction (HCR) In Situ Protocol. *protocols.io.* doi:10.17504/protocols.io.bunznvf6 (2021).

241. Carrieri, F. A., Murray, P. J., Ditsova, D., Ferris, M. A., Davies, P. & Dale, J. K. CDK1 and CDK2 regulate NICD1 turnover and the periodicity of the segmentation clock. *EMBO reports* **20,** e46436. doi:https://doi.org/10.15252/embr.201846436 (2019).

242. Van Oostrom, M. J., Li, Y. I., Meijer, W. H. M., Noordzij, T. E. J. C., Fountas, C., Timmers, E., Korving, J., Thomas, W. M., Simons, B. D. & Sonnen, K. F. Coupling of cell proliferation to the segmentation clock ensures robust somite scaling. *bioRxiv.* doi:10.1101/2025.01.10.632257 (2025).

243. Hubaud, A., Regev, I., Mahadevan, L. & Pourquié, O. Excitable dynamics and Yap-dependent mechanical cues drive the segmentation clock. *Cell* **171,** 668–682. doi:10.1016/j.cell.2017.08.043 (2017).

244. Wilkinson, B. H., Ivany, L. C. & Drummond, C. N. Estimating vertebrate biodiversity using the tempo of taxonomy – a view from Hubbert's peak. *Biological Journal of the Linnean Society* **134,** 402–422. doi:10.1093/biolinnean/blab080 (2021).

245. Glasauer, S. M. & Neuhauss, S. C. Whole-genome duplication in teleost fishes and its evolutionary consequences. *Molecular genetics and genomics* **289,** 1045–1060. doi:10.1007/s00438-014-0889-2 (2014).

246. Gazave, E., Lapébie, P., Richards, G. S., Brunet, F., Ereskovsky, A. V., Degnan, B. M., Borchiellini, C., Vervoort, M. & Renard, E. Origin and evolution of the Notch signalling pathway: an overview from eukaryotic genomes. *BMC Evolutionary Biology* **9,** 1–27. doi:10.1186/1471-2148-9-249 (2009).

247. Alpay, E. E., Zinani, O. Q., Hu, X., Ay, A. & Özbudak, E. M. DeltaC and DeltaD ligands play different roles in the segmentation clock dynamics. *Nature Communications* **16,** 2413. doi:10.1038/s41467-025-57645-5 (2025).

248. Oates, A. C. Waiting on the Fringe: cell autonomy and signaling delays in segmentation clocks. *Current Opinion in Genetics & Development* **63,** 61–70. doi:https://doi.org/10.1016/j.gde.2020.04.008 (2020).

249. Schubert, M., Holland, L. Z., Stokes, M. D. & Holland, N. D. Three amphioxus Wnt genes (AmphiWnt3, AmphiWnt5, and AmphiWnt6) associated with the tail bud: the evolution of somitogenesis in chordates. *Developmental Biology* **240,** 262–273. doi:10.1006/dbio.2001.0460 (2001).

250. Minguillón, C., Jiménez-Delgado, S., Panopoulou, G. & Garcia-Fernàndez, J. The amphioxus Hairy family: differential fate after duplication. *Development* **130,** 5903–5914. doi:10.1242/dev.00811 (2003).

251. Lara-Ramirez, R., Pérez-González, C., Anselmi, C., Patthey, C. & Shimeld, S. M. A Notch-regulated proliferative stem cell zone in the developing spinal cord is an ancestral vertebrate trait. *Development* **146,** dev166595. doi:10.1242/dev.166595 (2019).

252. Bedois, A. M., Parker, H. J., Price, A. J., Morrison, J. A., Bronner, M. E. & Krumlauf, R. Sea lamprey enlightens the origin of the coupling of retinoic acid signaling to vertebrate hindbrain segmentation. *Nature Communications* **15,** 1538. doi:10.1038/s41467-024-45911-x (2024).