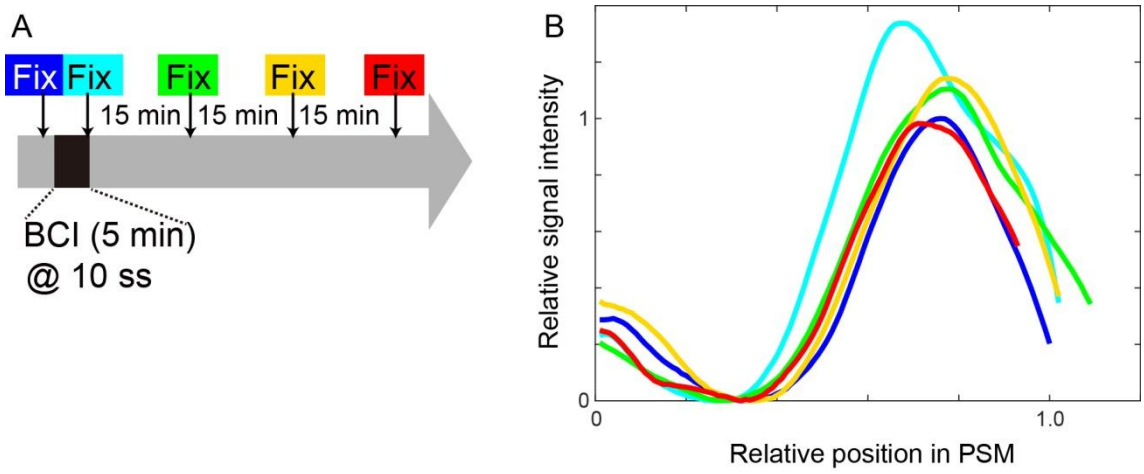


Fig. S1 Time course analysis of dpErk intensity before and after transient BCI treatment. (**A**) Schematic illustration of the experiment. At each time point, control and treated embryos (seven embryos each) were fixed and analyzed. (**B**) Intensity curves were calculated by averaging intensity curves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

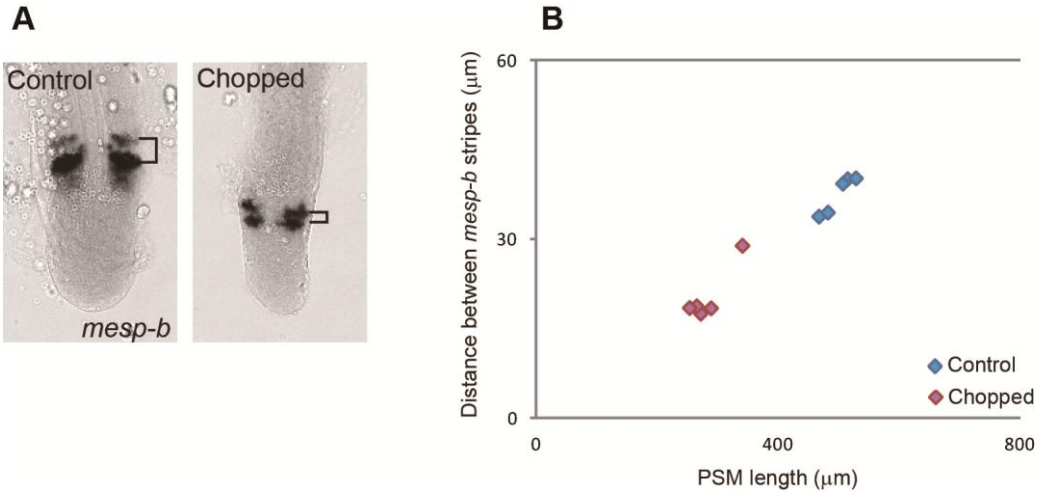


Fig. S2 Scaling of *mesp-b* stripe. (**A**) *in situ* hybridization samples of *mesp-b*. (**B**) PSM length vs distance between *mesp-b* stripes.

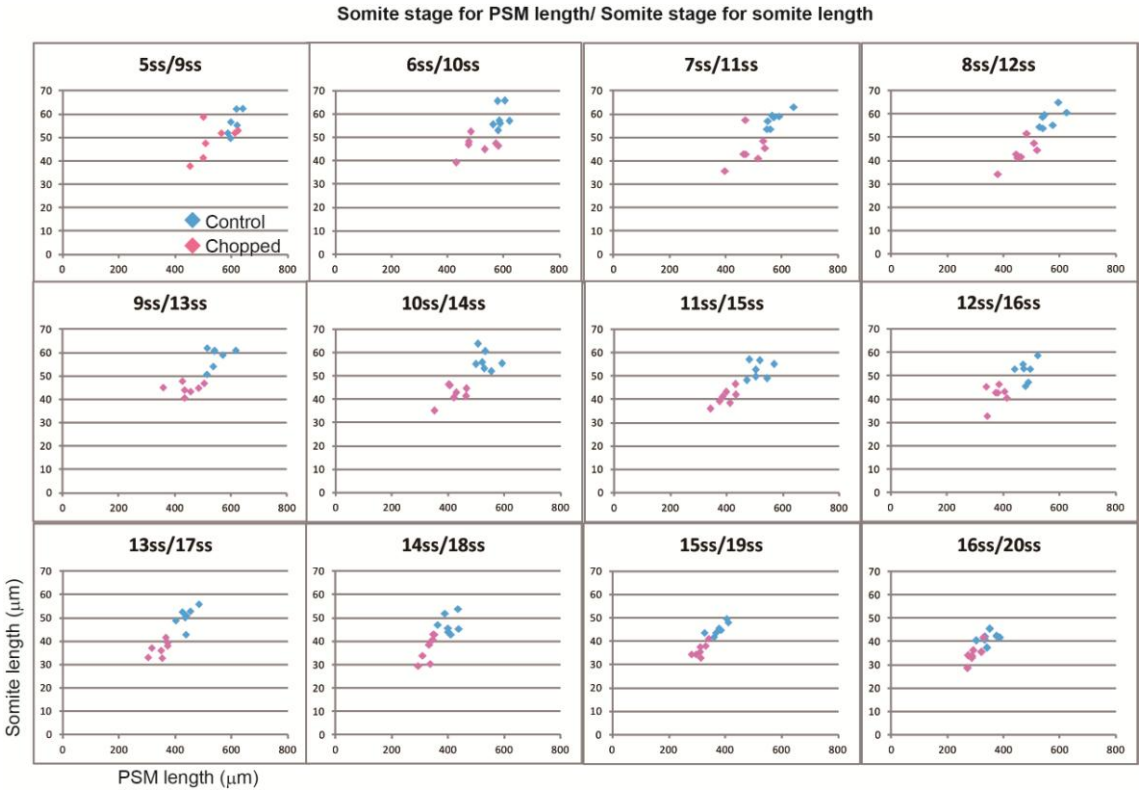


Fig. S3 Somite size vs PSM size between control and chopped embryos from different somite stages.

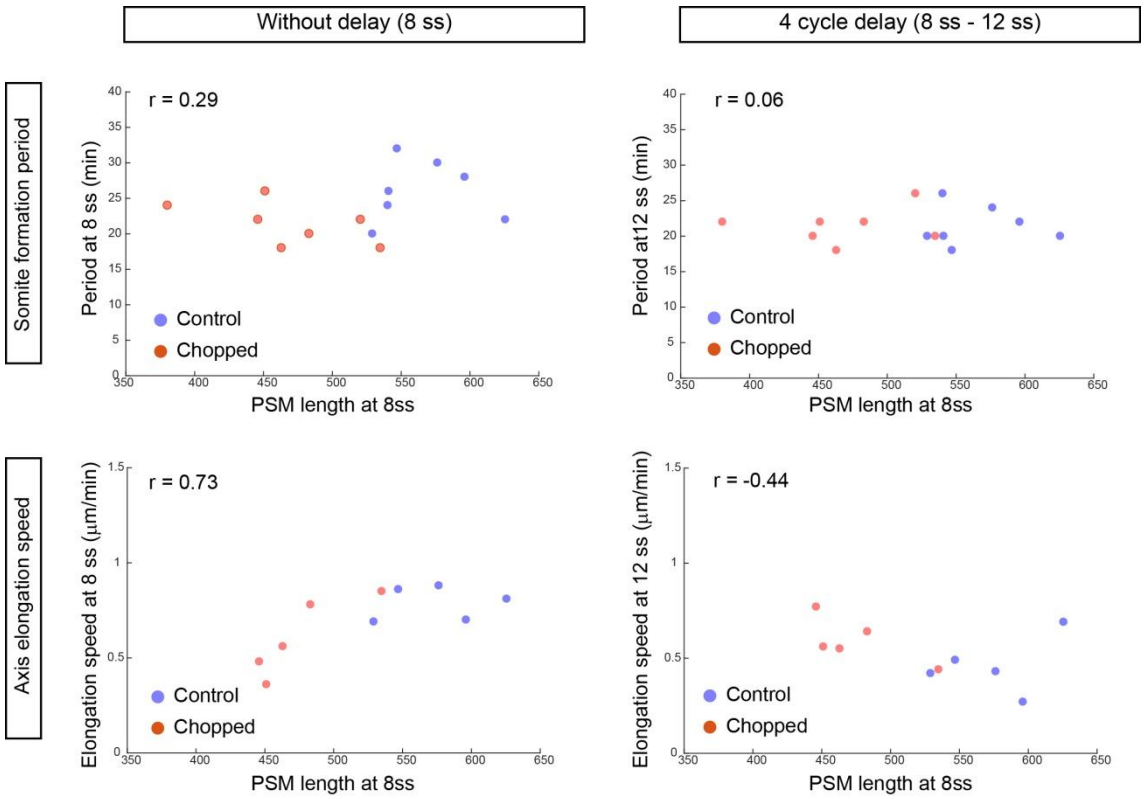


Fig. S4 Clock period and axis elongation speed with and without 4 cycle delay. Neither clock period nor axis elongation speed scale with PSM length even when the time delay is taken into consideration.

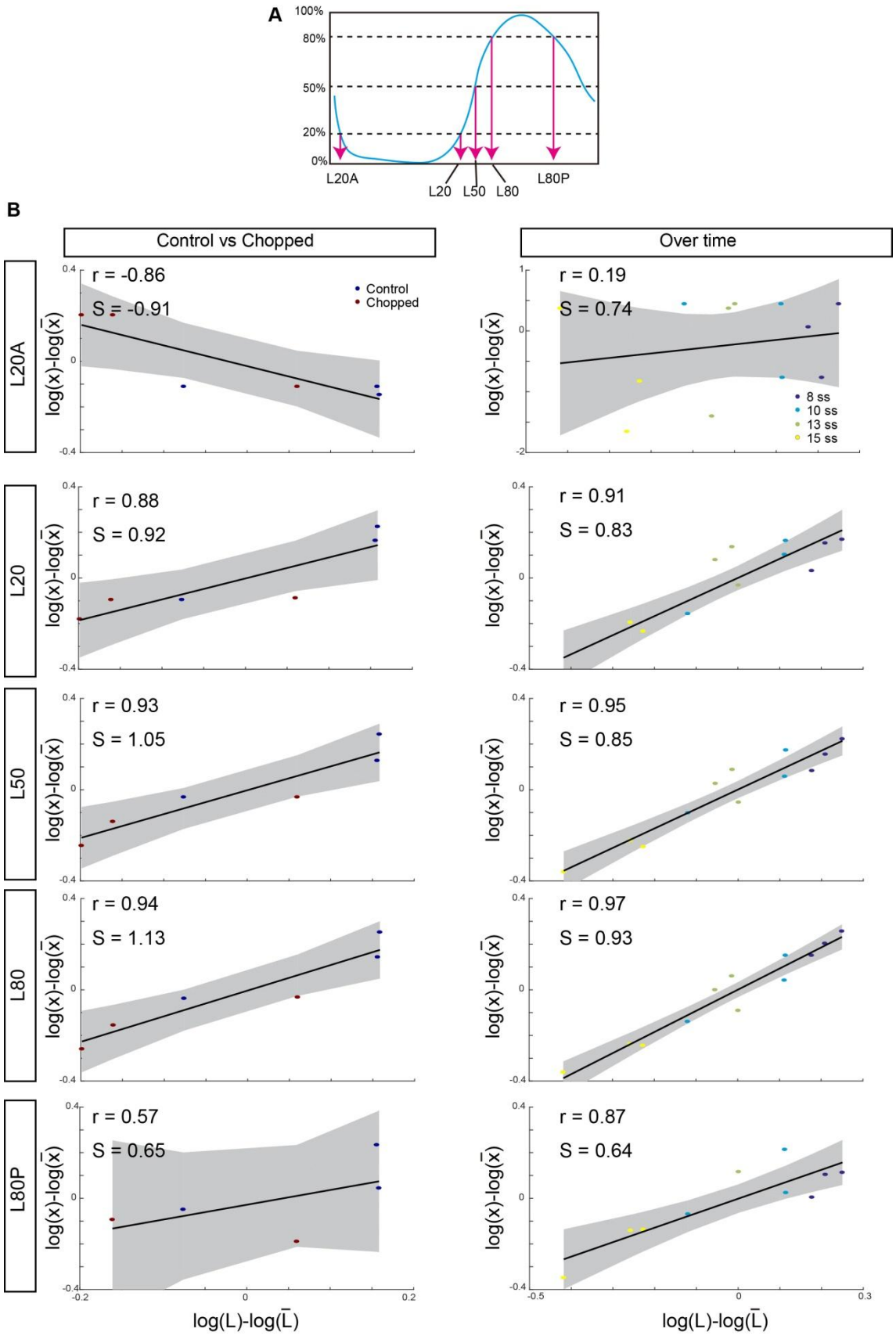


Fig. S5 dpErk scaling for different thresholds. (**A**) Schematic illustration of position names. (**B**) For each PSM, log-deviation in the position $\log(x / \overline{x}) = \log(x) - \log(\overline{x})$ is plotted against log-deviation in the PSM length $\log(\mathrm{L}/\overline{\mathrm{L}}) = \log(\mathrm{L}) - \log(\overline{\mathrm{L}})$, for different dpErk threshold intensities. The scaling coefficient S is obtained by linear regression (95% confidence interval on the slope is shown in gray). Correlation r is shown. In both cases (control vs chopped, and over time), L20, L50 and L80 scales with PSM length more than other positions. For detail, see (Hamaratoglu et al., 2011).

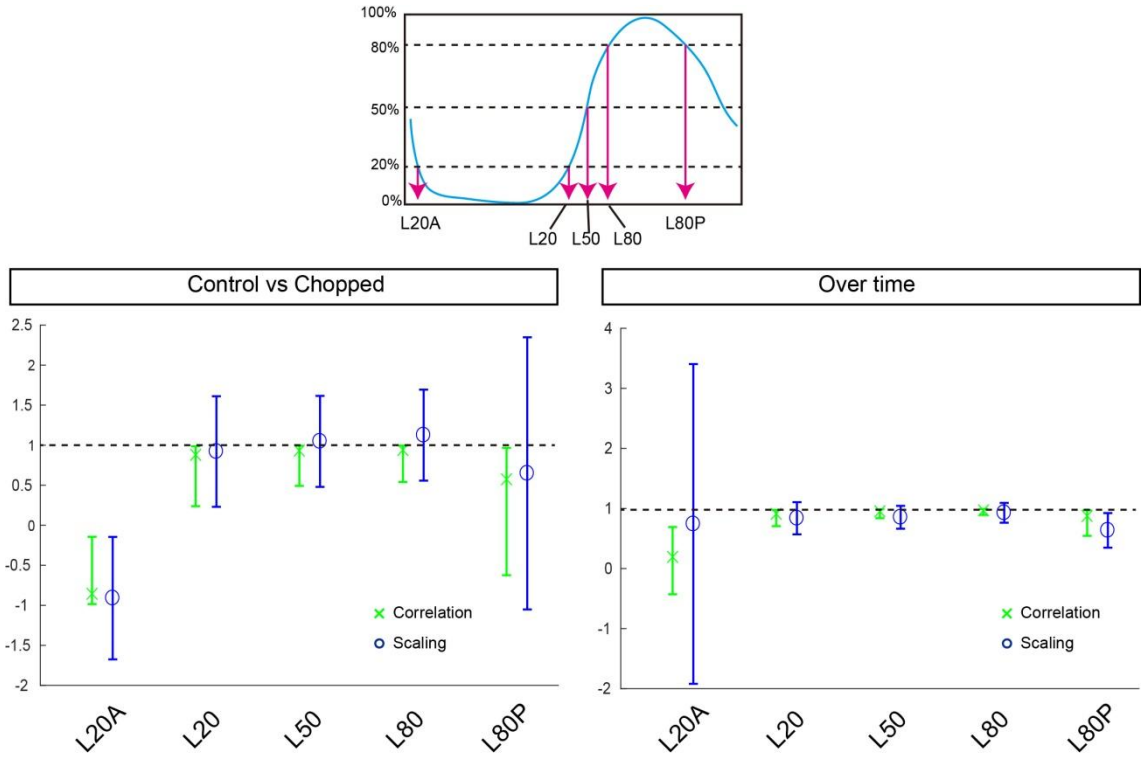


Fig. S6 Correlation coefficient and scaling coefficient for dpErk gradient. dpErk correlation (x) and scaling (o) for several gradient positions. Error bars represent the 95% confidence intervals. At L20, L50 and L80, the correlation and scaling coefficients are closer to 1, compared to other positions, which is consistent with Fig. S5. For detail, see (Hamaratoglu et al., 2011).

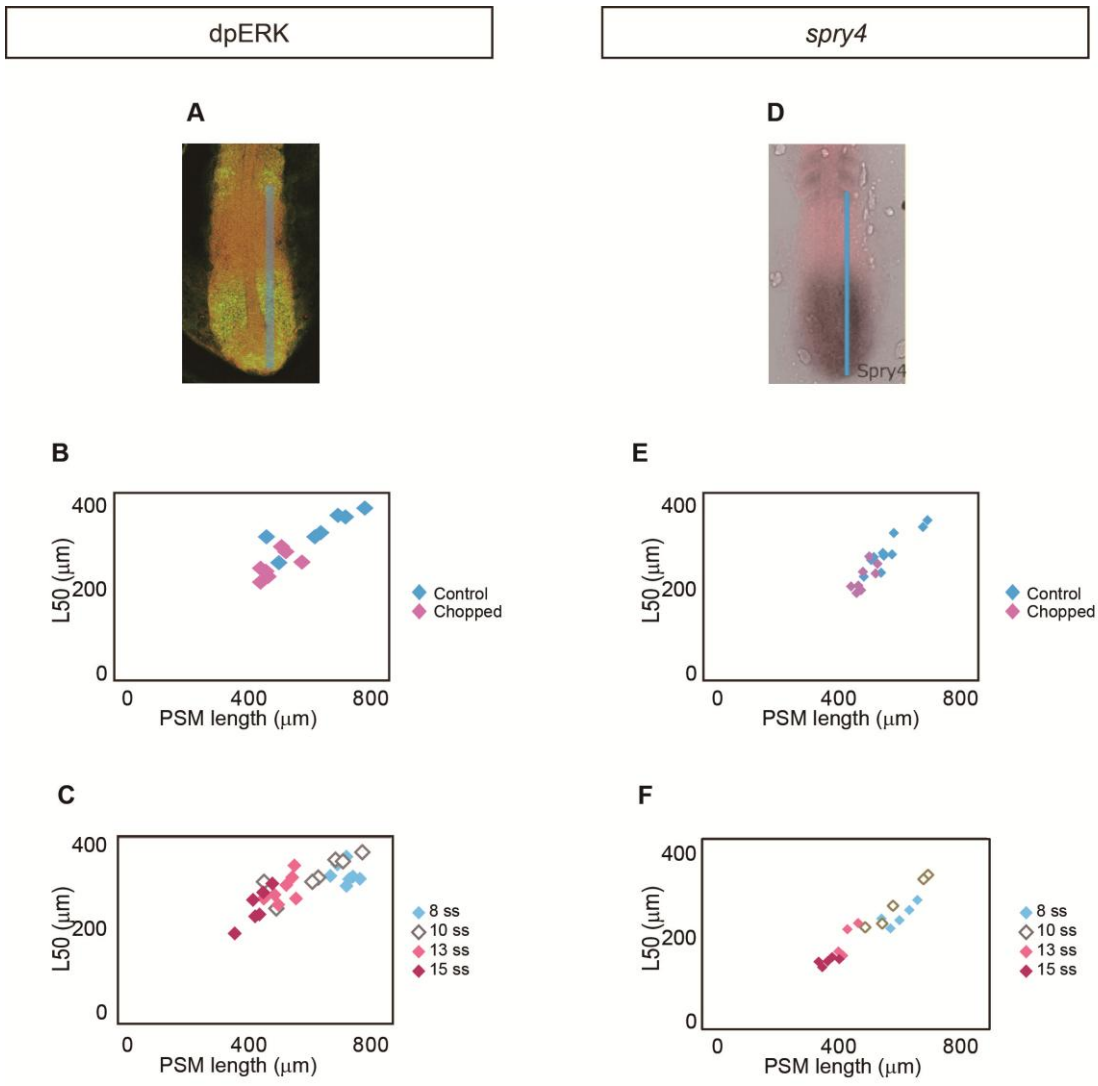


Fig. S7 Fgf activity scaling detected by dpErk and *spry4*. (**A** to **C**) L50 analysis on dpErk. (**D** to **F**) L50 analysis on *spry4* expression. (**A**) dpErk activity was detected by immunostaining. (**D**) *spry4* mRNA was detected by in situ hybridization. (**B** and **C**, **E** and **F**) L50 was calculated using *spry4* in situ hybridization samples similarly as in Fig. 3**I**. For both dpErk and *spry4*, L50 was found to scale with PSM length both between control and chopped embryos (**B** and **E**) and over time (**C** and **F**)

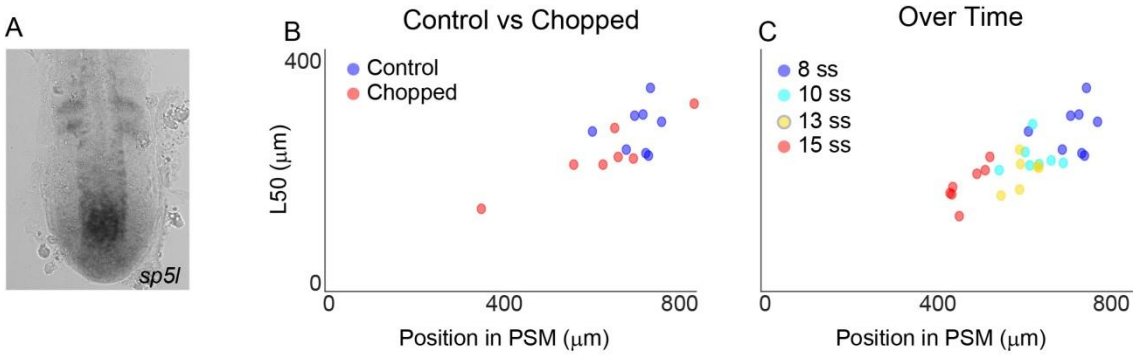


Fig. S8 Wnt signal scales with PSM length. (**A**) *sp5l* mRNA was detected by in situ hybridization. (**B** and **C**) L50 was calculated using *sp5l* in situ hybridization samples similarly as in Fig. 3**P**. L50 was found to scale with PSM length both between control and chopped embryos (**B**) and over time (**C**)

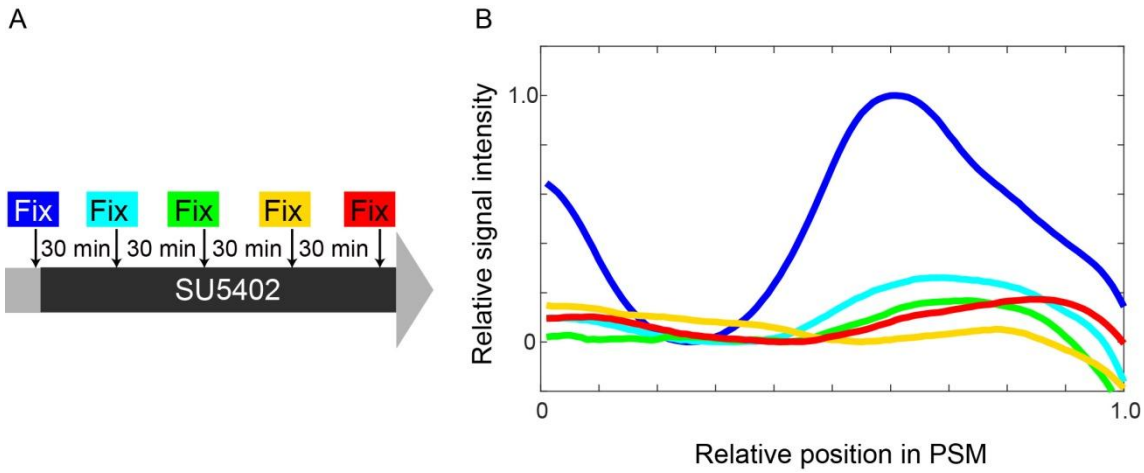


Fig. S9 Fgf activity gradient following SU5402 treatment (**A**) Schematic illustration of the experiment. At each time point, control and treated embryos (seven embryos each) were fixed and analyzed. (**B**) Intensity curves were calculated by averaging intensity curves of treated embryos. Relative signal intensity (y axis) was determined by using scaling factors: (maximum intensity of treated embryos) / (maximum intensity of control embryos at the corresponding time points). Relative position (x axis) was determined by normalizing positions in the PSM in treated embryos by averaged PSM length of control embryos at the corresponding time points. Colors correspond to the colors in (**A**). dpErk intensity drops immediately after onset of SU5402 treatment, and remains almost the same level over the course of experiment. Somites were still formed with relatively flat gradient of dpErk (Fig. 4L and M), possibly suggesting that dpErk is not the direct readout of the wavefront activity (see also discussion in the main text).

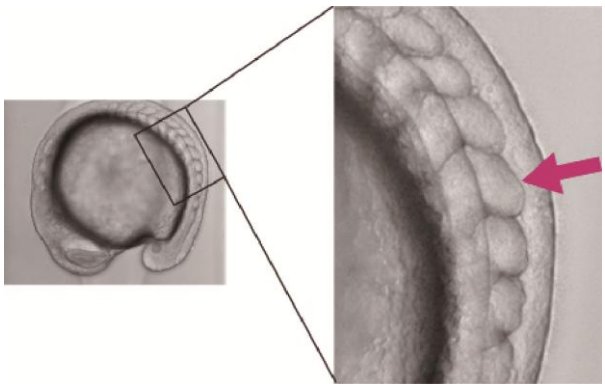


Fig. S10 Long-term SU5402 treatment under constant dark condition The embryos were treated with SU5402 at low concentration (16μM) for 4 hrs with the light completely blocked. One or two larger somites were formed (magenta arrow in the right panel) several cycles after initiation of the treatment (10 out of 11).

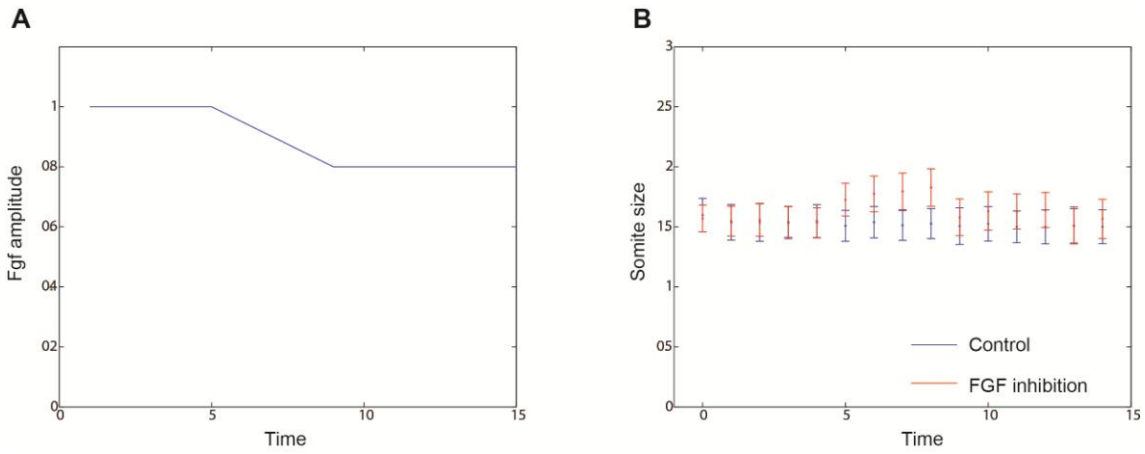


Fig. S11 Increasing Fgf inhibition can cause multiple larger somites. Here using the clock and scaled gradient model, we simulated the situation in which Fgf inhibition occurs increasingly, rather than in a step-wise manner (**A**). (**B**) As a result, multiple larger somites were predicted to be formed,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

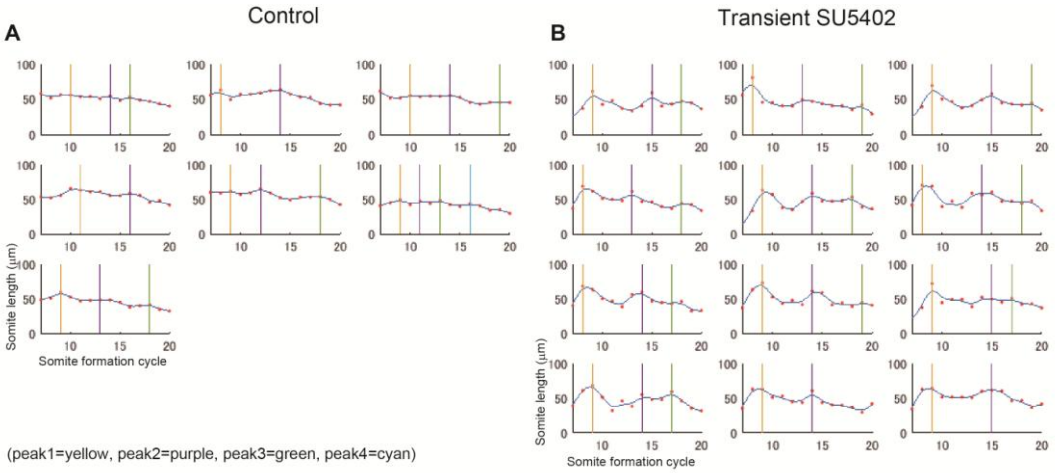


Fig. S12 Somite length change in individual samples. Somite sizes were measured using time-lapse imaging both in control (**A**) and SU5402 treated embryos (**B**). The peaks are detected using matlab function. Note the periodic change in somite length is much more obvious in SU5402 treated embryos, compared to control embryos.

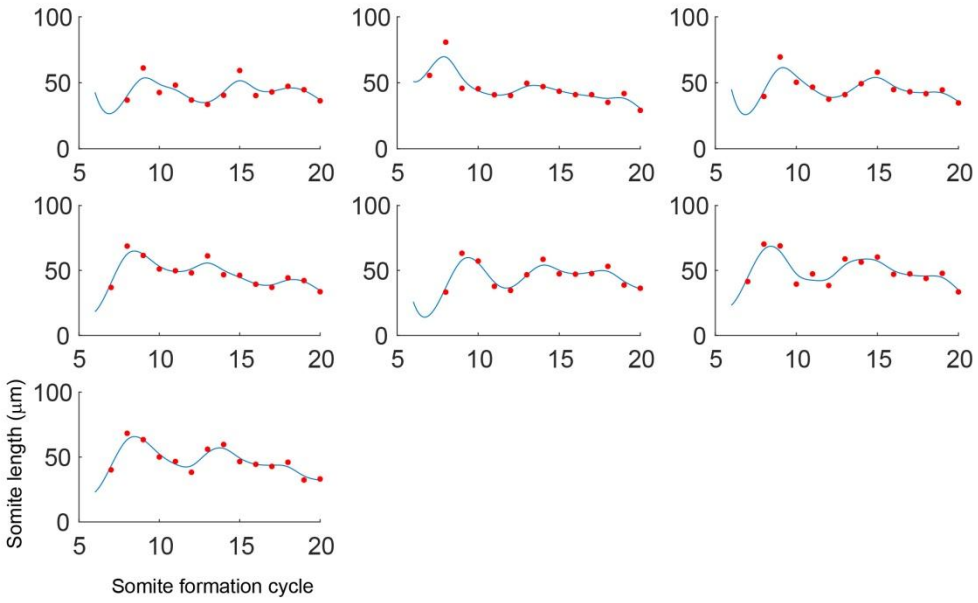


Fig. S13 Somite length change in individual samples in long-term SU5402 treatment. Somite sizes were measured using time-lapse imaging. As predicted in the simulation of the clock and scaled gradient model (Fig. 4K), the echo effect was observed.

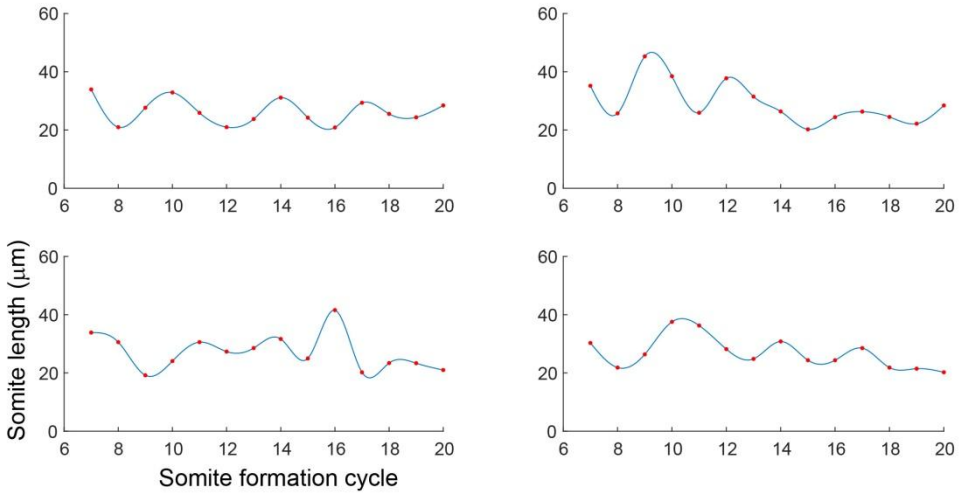


Fig. S14 Somite length change in individual samples in short-term BCI treatment. Somite sizes were measured using time-lapse imaging. Similar to SU5402 treatment, the echo effect was observed.

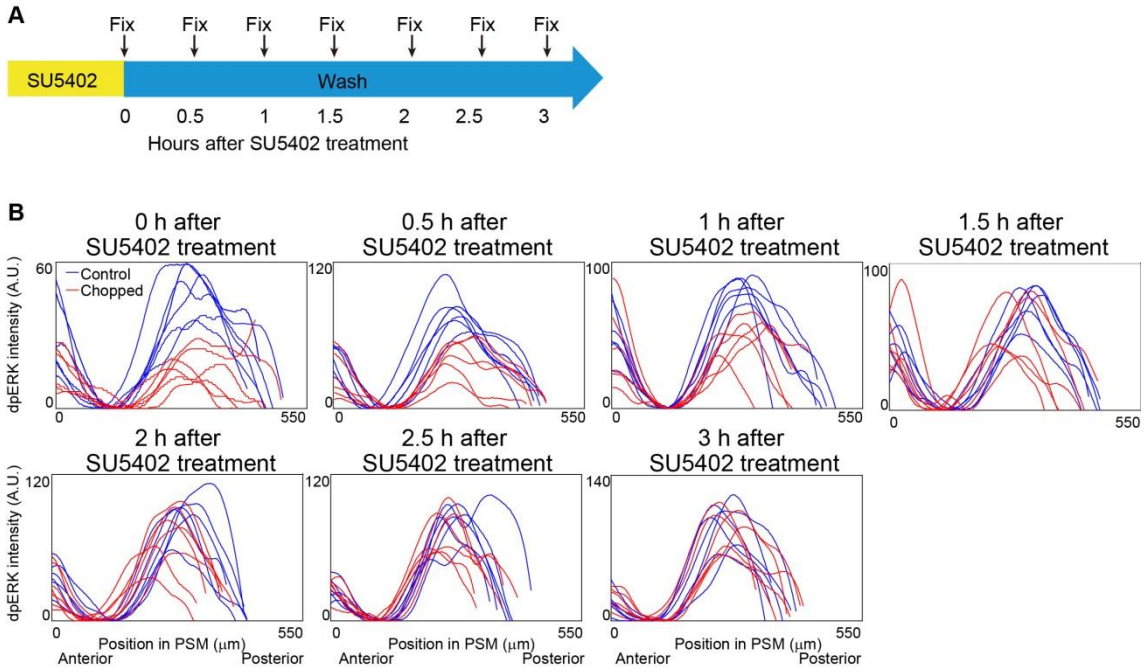


Fig. S15 dpErk intensity change in individual samples for echo experiment. (**A**) Schematic illustration of the experiment. Samples were fixed every 30 min after SU5402 treatment. (**B**) Intensity profiles of dpErk immunostaining for csntrol (blue) and treated embryos (red).

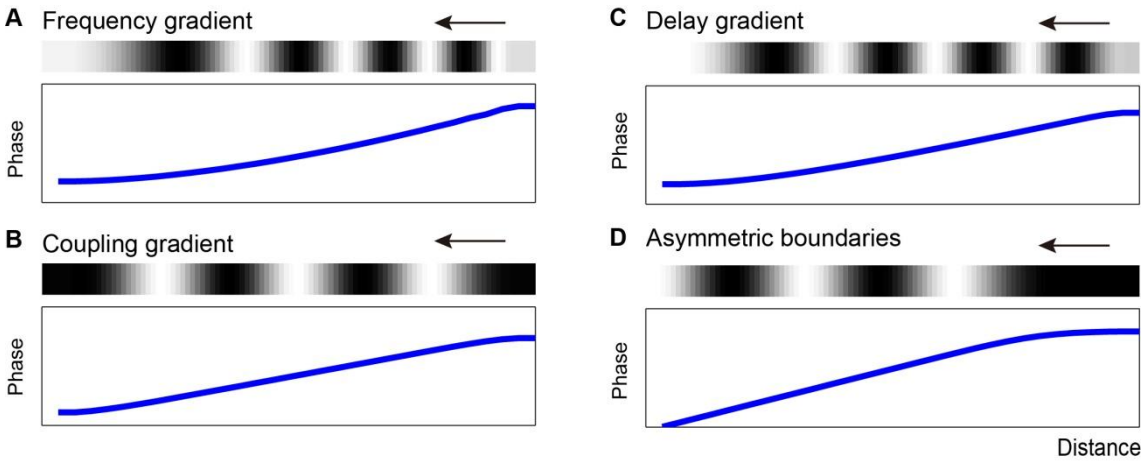
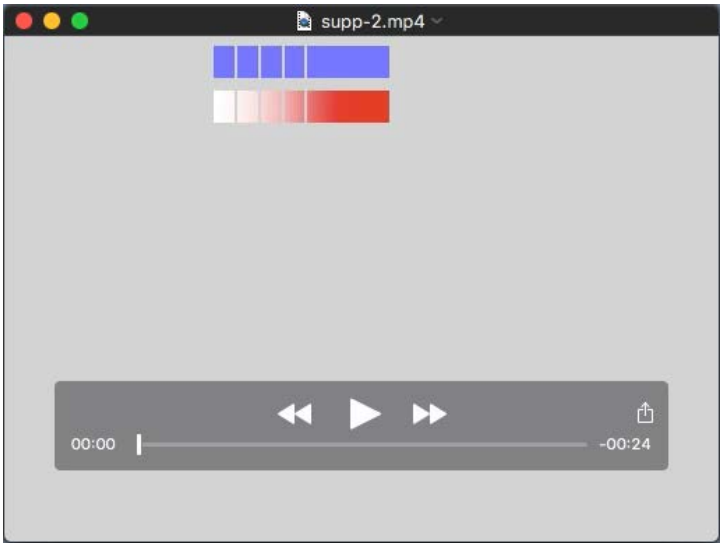


Fig. S16 Traveling wavevs arise under many conditions. Simulation of the Kuramoto model (see supplementary materials and methods for detail) with reflective boundary conditions and linear gradients in frequency (**A**), coupling strength (**B**), or phase delay (**C**) generate phase gradients and hence traveling waves. This is also observed in the absence of any gradients, but with a cut boundary at the anterior (**D**).



Movie 1 Simulation of clock and scaled gradient model

In the clock and scaled gradient model, each cell has a clock and a morphogen that shows a gradient in posterior to anterior direction. For simplicity, we assumed the clock oscillates synchronously along the PSM (blue in upper column) but phase gradients do not affect the results of the model. The morphogen gradient always scales with PSM size by an unspecified mechanism (red in lower column).. The mature somites are indicated by black lines, whereas 4 somites that are not visible but already determined are indicated by gray lines.

Supplementary Materials and Methods

Formulation of the Clock and Scaled Gradient model

As outlined in the main text, we develop a simple model of somitogenesis based on two key features: a 'clock' and a 'scaled gradient'. Note, we do not assign molecular identities to either the clock or the gradient, although, as described in the main text, we have candidates for each. To formulate the model mathematically, we convert our observations of the system into a set of concrete mathematical assumptions, namely:

1. **Clock**: To model an oscillator, we describe its phase by:

$$\frac{\partial \phi}{\partial t} = \frac{2\pi}{T} \tag{S1}$$

   where T is the clock period. Here we assume that the clock oscillates synchronously throughout the PSM (see below for the more general case where ϕ also varies in space).

2. **Gradient** We model the gradient, $g(x)$, by the function:

$$g(x) = g_0 G\left(x/l_{\mathrm{PSM}}\right) \tag{S2}$$

   where x is the distance to the tailbud, and l_{PSM} is the length of the PSM (from tailbud to mature somite boundary). The functional form $G(x/l_{\mathrm{PSM}})$ embodies the scaling of the gradient with PSM size, with the parameter g_0 denoting the overall gradient amplitude. We choose G to be of the form: $G(u) = 1 + \tanh\left(\beta\left(0.5 - u\right)\right)$, with $\beta = 3$, a gradient chosen to qualitatively reflect the dpErk gradient from Fig. 3 in the main text. For completeness, we also include noise in the gradient by adding a random number at each position, normally distributed with standard deviation σ. However, we emphasize that we see similar qualitative somitogenesis dynamics regardless of the precise functional form of G - even a simple linear gradient ($G(u) = 1 - u$) can recapitulate the *in vivo* behaviour rather closely (Fig. 4).

3. **Somite determination** We assume that a somite boundary is placed when the clock reaches a certain value (we take $\mathrm{mod}(\phi, 2\pi) = 0$) and at the position where the gradient exceeds a certain threshold, α i.e. the i^{th} somite boundary position, b_i, is given by $g(t_i - b_i) = \alpha$, where t_i is the position of the tail at this timepoint (and thus $t_i - b_i$ is the distance from b_i to the tail). We must also include the 4-cycle delay as outlined in Fig. 4A - when the i^{th} somite boundary is specified, b_i, it takes a further 4 cycles (i.e. 8π phase) for it to fully mature.

4. **Changes in PSM size**. We assume that PSM size reduces as somites mature at the anterior end. We also incorporate tail elongation to increase PSM size. For the time-window we are studying, our data suggests that tail elongation speed is approximately constant; therefore we set it to be a constant value, v.

5. **Initial conditions** The key initial condition we must set is the initial PSM size , which we label as as $l^{(0)}_{\mathrm{PSM}}$.

6. **First four somites** In the above, we have assumed (i) the PSM size is defined from tailbud to mature somite boundary and (ii) somites take 4 cycles to fully form. These assumptions break down for the earliest somites, which form (i) at a faster rate than the rest, (ii) before the tailbud has formed and gastrulation is occurring, and (iii) before there is a clear 'mature somite' boundary to set the PSM size in our model. Due to this complexity (and the fact that our data does not cover the earliest somites), we do not consider the first 4 cycles explicitly. Instead, we set the widths of the first 4 somites as equal to some arbitrary constant, l_{initial}, which is chosen such that somite size changes continously from the 4th to the 5th somite.

Together, these assumptions form the basis for our model. The model is simple, containing only 4 free parameters $\{T, g_0, \alpha, v\}$ (and during unperturbed development, we can nondimensionalize the equations leaving only a **single free parameter,** α), as well as the initial conditions. The model's simplicity allows us to obtain simple, qualitative insights into the phenomena and thus build an intuition for how somite sizes are controlled.

Model Results

Analytical solution

As per Fig. 4A, we write the position of somite boundary i as it is specified as b_i, and the tailbud position at this time as t_i. Without loss of generality (WLOG), the threshold condition $\alpha = g_0 G(x_i/l_{\mathrm{PSM}})$ can be rewritten as $x_i/l_{\mathrm{PSM}} = (1-f)$ where f is the fraction along the PSM at which the boundary is placed. Then, the Clock and Scaled Gradient model gives:

$$b_i = f\,(t_i - b_{i-4}) + b_{i-4} \tag{S3}$$

$$b_{i+1} = f\,(t_{i+1} - b_{i-3}) + b_{i-3} \tag{S4}$$

Subtracting gives an expression for the newly specified somite size $l_i \equiv b_{i+1} - b_i$:

$$l_i = fvT + (1-f)l_{i-4} \tag{S5}$$

where we have used $t_{i+1} - t_i = vT$ due to tail elongation. For the steady state, we have $l_i = l_{i-4}$ and thus:

$$l = vT \tag{S6}$$

i.e. we predict that somite size increases with clock period, and with tail speed and, at steady state, is independent of the specifics of the gradient (see Fig. 4C-F).

Now, if instead we set $v = 0$, then somites still form, according to:

$$l_i = (1-f)l_{i-4} \tag{S7}$$

i.e. there is a perfect geometric progression of somite sizes, as seen in an *in vitro* model of somitogenesis in which there is no tail elongation (see Fig. 4J)(Lauschke et al., 2013) .

To consider the effects of perturbing the gradient, we allow steady state to be reached (i.e. $l_j \equiv b_{j+1} - b_j = vT$ for $j < i$), and then perturb f. Firstly, we consider the case outlined in Fig. 6 in which the gradient is perturbed transiently for a single cycle. Here we have:

$$b_i = f\,(t_i - b_{i-4}) + b_{i-4} \tag{S8}$$

$$b_{i+1} = (f - \Delta f)\,(t_{i+1} - b_{i-3}) + b_{i-3} \tag{S9}$$

$$b_{i+2} = f\,(t_{i+2} - b_{i-2}) + b_{i-2} \tag{S10}$$

Computing $l_i \equiv b_{i+1} - b_i$ gives:

$$l_i = vT + \Delta f l_{\mathrm{PSM}}^{(i)} \tag{S11}$$

i.e. the i^{th} somite is larger. However, if we now examine the $(i+1)^{th}$ somite, we find:

$$l_{i+1} = vT - \Delta f l_{\mathrm{PSM}}^{(i+1)} \tag{S12}$$

i.e. that the $(i+1)^{th}$ somite is smaller. Furthermore, considering the fundamental 4-cycle periodicity of the equations, we predict (and show through simulation Fig. 6C) that this pattern of big somite followed by small somite is repeated every 4 cycles.

Secondly, we consider the case shown in Fig. 4 K-M, in which Fgf is continuously perturbed. Now we have:

$$b_i = f\,(t_i - b_{i-4}) + b_{i-4} \tag{S13}$$

$$b_{i+1} = (f + \Delta f)\,(t_{i+1} - b_{i-3}) + b_{i-3} \tag{S14}$$

$$b_{i+2} = (f + \Delta f)\,(t_{i+2} - b_{i-2}) + b_{i-2} \tag{S15}$$

and so on. In this case, computing $l_i \equiv b_{i+1} - b_i$ gives:

$$l_i = vT + \Delta f l_{\mathrm{PSM}}^{(i)} \tag{S16}$$

i.e. this somite is altered in size. However, when we compute the size of the subsequent somites, e.g. $l_{i+1} \equiv b_{i+2} - b_{i+1}$

$$l_{i+1} = vT \tag{S17}$$

we find that they return to their unperturbed value, exactly as seen *in vivo* (see Fig. 4L and M)*.

*Note also that our model predicts small somite size changes with 4-cycle periodicity (Fig. 4K), which was seen in Fig. S13

Simulations

We implement the equations described above using a custom MATLAB script[†]. To generate the figures in the main text, we used the parameters: $\{v = 1.5, \alpha = 1, g_0 = 1, T = 1\}$ and the initial conditions $\{l_{\text{initial}} = 2.9, l^{(0)}_{\text{PSM}} = 20\}$. For the chopped embryos we used the same parameters, but different initial conditions $\{l_{\text{initial}} = 0.7, l^{(0)}_{\text{PSM}} = 5\}$. For Fig. 4 C-F, we added noise to the gradient ($\sigma = 0.01$) and noise in the PSM measurement (i.e. normally distributed scatter with $\sigma_{measure} = 0.3$) to reflect the data in Fig. 1 and 2.

For the perturbations, we used the same parameters as wildtype but without noise (to more easily visualize differences between conditions) and:

- For transient Fgf activation (Fig. 4I), we set $g_0 = 1.1$ for the 10th cycle
- For sustained Fgf inhibition (Fig. 4I), we set $g_0 = 0.8$ for all times after the 9th cycle
- For the Fgf bead (Fig. 4I), we add a localized increase in the gradient, $g_{\text{bead}} \exp\left[-\left(\frac{x-x_{\text{bead}}}{w_{\text{bead}}}\right)^2\right]$, with $g_{\text{bead}} = 0.5, x_{\text{bead}} = 29, w_{\text{bead}} = 2$.
- For the slower clock (Fig. 4I), we set $T = 2$.
- For the slower tail (Fig. 4I), we set $v = 1$.
- For sustained Fgf inhibition for Fig. 4K, we set $\sigma = 0.02$, and $g = 0.9$ for all times after the 24th cycle (which we relabel as the 0th cycle on the plot)
- For no axis elongation (Fig. 4J), we set $v = 0$.
- For Fig. 6C, we set $g_0 = 0.92$ for the 10th cycle, and plot the mean and standard deviation for 10 independent simulations with noise $\sigma = 0.03$, parameters chosen to reflect the *in vivo* data.
- For Fig. 6B, we repeated the above but with $g_0 = 0.85$ and no noise term to emphasize the effect for visualization purposes.
- For Fig. 6L-N, we used a higher number of repeats (1000 instead of 10) and a smaller noise ($\sigma = 0.02$) to better compare to the other models.

[†]https://wiki.med.harvard.edu/SysBio/Megason/MegasonSoftware

Origin of *her1* travelling waves

In this study, we found that the wavelength of *her1* waves does not play a central role in determining somite size. Thus we suspect that the spatiotemporal pattern of these waves may not be actively controlled, but is merely a byproduct of imperfect synchronization between oscillators. The emergence of travelling waves has been observed in other oscillatory systems with imperfect synchronization (e.g. due to delays), and can be predicted by fairly general models of coupled oscillators (Ermentrout, 2010). These theoretical tools have also been used to understand the spatiotemporal dynamics of *her1* expression, e.g. (Ares et al.,2012; Morelli et al., 2009). Here we revisit some of these arguments, by modifying equation S1 to a more general form that explicitly considers oscillator synchronization. Specifically, we consider a 1D array of cells ($i = 1, 2, ..., N$, with $i = 1$ anterior and $i = N$ posterior) and describe the phase dynamics by:

$$\frac{\partial \phi_i}{\partial t} = \omega_i + H_i\left(\phi_{i+1} - \phi_i\right) + H_i\left(\phi_{i-1} - \phi_i\right) \tag{S18}$$

commonly referred to as the Kuramoto model[‡] (Kuramoto 1984) . Here, $\omega_i \equiv 2\pi/T_i$ describes the oscillator frequencies, and $H_i(\phi)$ the coupling between oscillators. Following (Ermentrout, 2010), a natural choice for the coupling term is $H_i(\phi) = k_i \sin(\phi + \Phi_i)$, where k_i is the coupling strength and Φ_i the effective coupling delay[§]. Given this equation, we can ask - under what conditions do the oscillators successfully synchronize; or, the complementary question - under what conditions do travelling waves emerge as a result of imperfect synchronization?

Interestingly, we found that many conditions resulted in the formation of travelling waves. As has been hypothesized elsewhere (Giudicelli et al., 2007), a gradient in frequencies along the anterior-posterior (AP) axis, $\omega_i \neq$ const, results in travelling waves along a single direction. However, we also found that if either the coupling strength, k_i, or the coupling phase delay, Φ_i, varied along the AP axis, then this too resulted in travelling waves. Given that there are multiple signaling gradients with the PSM, which could feasibly affect the oscillator frequency, coupling or time delay (Ay et al., 2014), then it is unsurprising that travelling waves will emerge.

Even in the complete absence of spatial gradients (i.e. ω_i, k_i, Φ_i are constant), travelling waves can still form. Consider the case of a 'cut' boundary condition at the anterior end $i = 1$ - here, 'cut' means that the most anterior cells are coupled only on their posterior side. This could be achieved, for example, if cells in mature somites stopped being coupled to the oscillations in the PSM - either by failing to signal to the PSM, or by terminating *her1* oscillations completely.[¶] Given the assumption of a cut boundary condition, the system is described by:

$$\frac{\partial \phi_1}{\partial t} = \omega + H\left(\phi_2 - \phi_1\right) \tag{S19}$$

i.e. the 1st oscillator takes no input from its anterior neighbour. Then, for the remaining oscillators we have:

$$\frac{\partial \phi_2}{\partial t} = \omega + H\left(\phi_3 - \phi_2\right) + H\left(\phi_1 - \phi_2\right) \tag{S20}$$

etc. Assuming solutions of the form $\phi_i = \Omega t + Ai$, then:

$$\Omega = \omega + H(A) \tag{S21}$$

$$\Omega = \omega + H(A) + H(-A) \tag{S22}$$

etc., which has a stable solution of phase-locked oscillations with overall frequency Ω, and a phase gradient defined by $H(-A) = 0$. For the choice of $H(\phi) = k\sin(\phi + \Phi)$, we have $A = \Phi$ and $\Omega = \omega + k\sin(2\Phi)$. Thus, in the complete absence of spatial gradients, but with a cut anterior boundary condition, we have a linear phase gradient $\phi_i = \Omega t + \Phi i$ i.e. waves that travel from posterior to anterior (provided $\Phi > 0$). We confirm this with simulations, using a cut boundary condition in the anterior and a reflective boundary condition in the posterio.

Taken together, given a set of coupled oscillators operating with phase delays, having spatially varying inputs and unknown boundary conditions, it is not surprising to see travelling waves (consistent with previous work e.g (Ares et al.,2012; Morelli et al., 2009)). Therefore, without any clear spatiotemporal perturbations of the *her1* travelling waves, it is difficult to determine to what extent these travelling waves are functional, or could simply be the result of imperfect synchronization. In particular, as discussed in the main text, the presence of a ≈4-cycle delay between somite size determination and the appearance of morphological boundaries suggests that the *her1* dynamics are likely unimportant in the most anterior region of the PSM.

[‡]We emphasize that our model does not aim to describe all features of the waves (e.g. the anterior increase in oscillator amplitude), but merely to show that a very simple model does quite well in capturing the qualitative wave dynamics.

[§]this is not an explicit time delay, for details see (Ermentrout, 2010)

[¶]Note that previous studies support the notion of somite maturation forming an anterior boundary condition e.g. in ripply morphant embryos(Kawamura et al., 2005), in which somite maturation is impaired, *her1* expression persists more anteriorly.

A small effect of *her1* travelling waves on the Clock and Scaled Gradient Model

In our model, we have assumed that the entire PSM oscillates synchronously (i.e. there is no phase gradient) and found that this could qualitatively explain our data. However, measurements of *her1* dynamics *in vivo* show that there is a phase gradient (i.e. travelling waves) within the PSM (Soroldoni et al., 2014). We therefore wished to ask to what extent such a phase gradient would affect the results of our model.

To do this, we assumed a simple linear phase gradient:

$$\phi(x,t) = \frac{2\pi t}{T} + kx \tag{S23}$$

where k is a constant that controls the magnitude of the gradient (with the sign determining the direction of the travelling waves). We then repeated the wildtype simulations with this modified phase profile, using $|k| = 0.5$, corresponding to an initial phase difference of 10 rad across the PSM). The results are shown in Fig. 7 A and B.

Supplementary References

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

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

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

Ermentrout, G. B., Terman, David H. (2010). Mathematical foundations of neuroscience: Springer Science & Business Media.

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

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

Kawamura, A., Koshida, S., Hijikata, H., Ohbayashi, A., Kondoh, H. and Takada, S. (2005). Groucho-associated transcriptional repressor ripply1 is required for proper transition from the presomitic mesoderm to somites. Dev Cell 9, 735-44.

Kuramoto, Y. (1984). Chemical Oscillation Waves, and Turbulence. Mineola, New York: Dover Publications, Inc.

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

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

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