Pairing of segmentation clock genes drives robust pattern formation

https://doi.org/10.1038/s41586-020-03055-0
Received: 24 September 2019
Accepted: 3 November 2020
Published online: 23 December 2020

Oriana Q. H. Zinani1,2, Kemal Keseroğlu2, Ahmet Ay3,4 & Ertuğrul M. Özbudak1,2,5,6

Gene expression is an inherently stochastic process1,2; however, organismal development and homeostasis require cells to coordinate the spatiotemporal expression of large sets of genes. In metazoans, pairs of co-expressed genes often reside in the same chromosomal neighbourhood, with gene pairs representing 10 to 50% of all genes, depending on the species3–6. Because shared upstream regulators can ensure correlated gene expression, the selective advantage of maintaining adjacent gene pairs remains unknown. Here, using two linked zebrafish segmentation clock genes, her1 and her7, and combining single-cell transcript counting, genetic engineering, real-time imaging and computational modelling, we show that gene pairing boosts correlated transcription and provides phenotypic robustness for the formation of developmental patterns. Our results demonstrate that the prevention of gene pairing disrupts oscillations and segmentation, and the linkage of her1 and her7 is essential for the development of the body axis in zebrafish embryos. We predict that gene pairing may be similarly advantageous in other organisms, and our findings could lead to the engineering of precise synthetic clocks in embryos and organoids.

The subdivision of the anterior–posterior axis into a fixed number of repeating units, known as somites, is a prime example of how coordinated gene expression patterns the vertebrate embryo. During somitogenesis, groups of cells sequentially and synchronously commit to segmentation in a notably short time frame. The pace of segmentation is set by the period of an oscillator, the segmentation clock, which is active in cells of the unsegmented presomitic mesoderm (PSM) (Fig. 1a). Oscillatory expression of the Hes or her clock genes is conserved in vertebrates; disruption in this oscillatory expression results in defects in vertebral segmentation (for example, congenital scoliosis in humans). At the conclusion of each clock cycle (approximately 30 min in zebrafish), a cohort of around 200 cells bud from the unsegmented PSM to form a new somite (Fig. 1a). To form the full-length body axis, segmentation is carried out for a species-specific number of cycles (33 in zebrafish). The segmentation clock relies on a transcriptional negative-feedback loop7,11. In zebrafish, two paired clock genes (her1 and her7) are separated by a 12-kb regulatory sequence (Fig. 1b). Her1 and Her7 function as hetero- or homodimers that repress their own transcription. her1 and her7 have similar transcriptional time delays8,9 and RNA half-lives8; therefore, the transcription of her1 and her7 is mainly concomitant in the PSM. To achieve the rapid tempo and reproducible precision of segmentation, the transcription of her1 and her7 should be tightly coordinated.

To quantify her1 and her7 transcripts in single cells, we performed high-resolution single-molecule fluorescence in situ hybridization (smFISH). Consistent with previous findings10,11, we affirmed that segmentation clock genes display very low transcriptional amplitude (46 ± 3 (mean ± s.e.m.) her1 and 57 ± 2 her7 RNA molecules) (Extended Data Fig. 1a–f, Supplementary Table 1, Methods). As segmentation clock RNAs and proteins are short-lived (t1/2 = 3–5 min)11, temporal averaging cannot be used to reduce gene expression variability, resulting in a highly variable (noisy) biological clock11. To determine the degree of her1 and her7 coexpression in the PSM, we calculated the Spearman correlation and Pearson correlation scores of her1 and her7 transcription in the region where they display characteristic oscillatory kinematic waves (that is, between 40–80% of PSM from the posterior end) (Fig. 1d, h, i). This analysis revealed that her1 and her7 have a high transcriptional correlation (mean Spearman correlation = 0.76) (Extended Data Fig. 1g). These observations raise two important questions. The first relates to the mechanism that drives the highly correlated transcription of the two clock genes. Another key question is whether the correlated transcription of two clock genes is under selection pressure (that is, beneficial for development). In this study, we investigated the effect of two mechanisms on the correlated transcription of her1 and her7: (1) negative feedback by the Her1 and Her7 repressors, and (2) gene pairing (that is, chromosomal adjacency or linkage of the her1 and her7 genes).

Because the Her1 and Her7 proteins autoinhibit their own transcription, one obvious source of correlated transcription of her1 and her7 could be this shared upstream regulation. To test this hypothesis and determine the role of negative-feedback loops in controlling correlated transcription, we used genome editing to generate a double her1 SNAPi her7hmu2526 homozygote mutant by deleting two base pairs of the her1 coding sequence in the her7hmu2526 mutant15 (Extended Data Fig. 1h). Whereas double heterozygous mutants successfully segment all somites, the oscillatory waves of clock transcription and segmentation are disrupted in double homozygous mutants; similar to previously published mutants16–18 (Fig. 1c–e, Extended Data Fig. 2a–c). As our smFISH probes do not distinguish between wild-type and mutant her1 or her7 RNAs, we found that the transcriptional amplitude of oscillations (Fig. 1f, t(18) = 0.7, P = 0.98 for her1 and t(18) = 1.7, P = 0.24 for her7) and the...
distribution of RNA counts (Fig. 1g, mean = 49 versus 51, Supplementary Tables 1, 2) were similar in wild-type and double heterozygous mutants. These results suggest that in heterozygous mutants— even though approximately half of the RNAs are non-functional— a sufficient amount of Her1 and Her7 proteins are translated to maintain clock oscillations and similar RNA levels to wild-type embryos. By contrast, the mean RNA level increased by 74% in double homozygous mutant embryos compared to wild-type embryos, reflecting the loss of negative feedback.
regulation (Fig. 1g, mean = 49 versus 85, Supplementary Tables 1, 3). By comparing the spatial Spearman correlation scores, we assessed the changes in spatial coexpression of her1 and her7 among different genotypes (Fig. 1j). The spatial Spearman correlation score significantly decreased in both double homozygous and heterozygous mutants as compared to wild-type embryos (Fig. 1j). To test this hypothesis, we detected nascent transcription loci in the nucleus of single cells located in stripes in the anterior PSM and measured probabilities of transcriptional co-firing of a her1 gene with its linked her7 gene on one chromosome (P(het7|het1)) versus with the second her1 on its homologous chromosome (P(het1|het1)) at a given time point. We found that the probability of transcriptional co-firing of paired her1 and her7 genes is significantly higher than the two unpaired her1 genes on separate chromosomes (Fig. 1m, U = 212.5, z = −6.9, P = 5.803 × 10^−12). This finding demonstrated negative-feedback loop, which increases correlated transcription of her1 and her7 in the PSM.

Chromosomal adjacency was previously shown to cause correlated expression of synthetic reporters; thus, we hypothesized that adjacency of her1 and her7 causes their transcription to co-fire on the same chromosome, resulting in highly correlated transcript levels in wild-type embryos (Fig. 1j). To test this hypothesis, we detected nascent transcription loci in the nucleus of single cells located in stripes in the anterior PSM and measured probabilities of transcriptional co-firing of a her1 gene with its linked her7 gene on one chromosome (P(het7|het1)) versus with the second her1 on its homologous chromosome (P(het1|het1)) at a given time point. We found that the probability of transcriptional co-firing of paired her1 and her7 genes is significantly higher than the two unpaired her1 genes on separate chromosomes (Fig. 1m, U = 212.5, z = −6.9, P = 5.803 × 10^−12). This finding demonstrated
that gene pairing boosts correlated transcription of the two clock genes by facilitating transcriptional co-firing.

Robustness of a system can be tested by perturbing the environment. Temperature fluctuations are a natural environmental noise source that affects fish embryos. Although wild-type embryos successfully segment somites over a large range of temperatures21, certain zebrafish that affects fish embryos. Although wild-type embryos successfully segment somites over a large range of temperatures21, certain zebrafish that gene pairing boosts correlated transcription of the two clock genes by facilitating transcriptional co-firing.

Robustness of a system can be tested by perturbing the environment. Temperature fluctuations are a natural environmental noise source that affects fish embryos. Although wild-type embryos successfully segment somites over a large range of temperatures21, certain zebrafish that affects fish embryos. Although wild-type embryos successfully segment somites over a large range of temperatures21, certain zebrafish that affects fish embryos. Although wild-type embryos successfully segment somites over a large range of temperatures21, certain zebrafish that affects fish embryos. Although wild-type embryos successfully segment somites over a large range of temperatures21, certain zebrafish that affects fish embryos. Although wild-type embryos successfully segment somites over a large range of temperatures21, certain zebrafish that affects fish embryos. Although wild-type embryos successfully segment somites over a large range of temperatures21, certain zebrafish that affects fish embryos. Although wild-type embryos successfully segment somites over a large range of temperatures21, certain zebrafish that affects fish embryos. Although wild-type embryos successfully segment somites over a large range of temperatures21, certain zebrafish that affects fish embryos. Although wild-type embryos successfully segment somites over a large range of temperatures21, certain zebrafish.
genes are located on the same chromosome than when they are located on different chromosomes at both temperatures (Fig. 1p, U = 266, $z = -6.2, P = 2.489 \times 10^{-6}$ for 28°C and U = 242, $z = -6.4, P = 7.458 \times 10^{-30}$ for 21.5°C). We further found that both $P_{\text{her1:her1-Venus}}$ and $P_{\text{her1:her1-ci302}}$ decreased significantly in embryos raised at 21.5°C compared with those at 28°C (Fig. 1p, $U = 411, z = -5.1, P = 1.745 \times 10^{-6}$ and $U = 736.5, z = -2.5, P = 0.048$), indicating that transcriptional co-firing is temperature-dependent.

We then hypothesized that co-firing of two paired clock genes is advantageous for somite segmentation as that would coordinate transcript levels. To test this, we generated a new her1:her1 mutant carrying the same mutation described above in an otherwise wild-type background (Extended Data Fig. 1h, Methods). We subsequently mated her1:her1 fish with her1:her1 fish to obtain double heterozygote embryos with unpaired functional her1 and her7 genes. These embryos carried a mutant her1 gene adjacent to a wild-type her7 gene on one chromosome, and a mutant her7 gene adjacent to a wild-type her1 gene on the other chromosome. Hence, fish that are compound heterozygous for some, and a mutant gene on one chromosome and a mutant other chromosome. These alleles will have the same functional gene dose as the previously described double heterozygous embryos with paired her1 and her7 genes (Fig. 2a, b). We next raised the gene-paired and gene-unpaired embryos at different temperatures. At 28°C, both groups successfully segmented all their somites (Supplementary Table 8). However, at 21.5°C, gene-unpaired embryos had reduced success in somite segmentation as compared to gene-paired embryos (Fig. 2c–f). This defect could be quantified by an increase in both the percentage of embryos with defects (Fig. 2e, 77% for gene-unpaired and 30% for gene-paired, n > 130, $\chi^2(1) = 60.3, P = 8.043 \times 10^{-15}$) and the severity of defects (Fig. 2f, median = 4 for gene-unpaired and 2 for gene-paired, U = 4837.5, $z = -5.5, P = 4.878 \times 10^{-9}$). These results revealed that maintaining paired genes in the genome is beneficial for successful pattern formation during embryonic development under environmental stress.

We next investigated the mechanism by which gene pairing is beneficial for somite segmentation. Regulatory networks containing negative-feedback loops are widespread among metazoans. However, negative-feedback loops usually do not give rise to oscillations, but instead act as a rheostat to tightly maintain the system output around a steady-state (a stable point in a phase diagram). For a negative-feedback loop to generate oscillations, several important criteria need to be satisfied. Experimental studies previously revealed the importance of time delays and short RNA or protein half-lives to generate sustained oscillations. However, another important criterion needed to generate oscillations is overshooting of the steady-state. The rate of mRNA synthesis needs to be high enough to push the system into an unstable steady-state with a periodic orbit, establishing a limit cycle in the phase diagram. We found that when genes are paired on the same chromosome, co-transcription happens more frequently (Fig. 1p). We hypothesized that frequent co-firing of transcription results in a high rate of RNA production, and thereby overshoots the limit-cycle threshold. When her1 and her7 are on different chromosomes, co-transcription is less frequent, which lowers the RNA accumulation rate, resulting in an inability to overshoot the steady-state needed to establish a limit cycle (that is, oscillations) (Fig. 2j–I).

To test this hypothesis, we first developed a single-cell level stochastic model that incorporates the negative-feedback loop established by two different dimers (Methods). This model, transcription rates of her1 and her7 only occur from genes that are free of repressor dimers. Transcription rates of her1 and her7 were kept equal to each other and between the two chromosomes. We used a random number generator to implement stochasticity in transcriptional firing. In the case of gene-paired embryos, transcription firing was determined simultaneously for both genes (in agreement with the experimental data) (Fig. 1m). By contrast, for gene-unpaired embryos, transcription firing was determined separately for each gene. Alternatively, we simulated other scenarios in which the transcriptional firing of two clock genes occurred at different rates (Methods). Our minimal model recapitulated our observations. In all scenarios, simulations showed that correlated expression and thereby sustained oscillations of her1 and her7 occurred when two genes are paired, and less correlated expression and occasionally failed oscillations occurred when two genes are unpaired (Fig. 2m–o, U = 14, $z = -6.5, P = 1.148 \times 10^{-30}$, Extended Data Fig. 4).

To provide further experimental evidence for this mechanism, we performed smFISH experiments and found that transcription of the two clock genes was less correlated in gene-unpaired than in gene-paired embryos raised at 21.5°C (Fig. 3a–c, U = 993.0, z = -4.9, P = 1.000 × 10^{-9}, Extended Data Fig. 3d, Supplementary Table 9). These results are consistent with the computational modelling showing that less correlation expression of her1 and her7 leads to defects in kinematic clock waves and indicate that the segmentation defects occurring in mutant embryos are due to the failure of sustained clock oscillations.

To assess the function of gene pairing in the segmentation clock in real-time, we next used the transgenic Tg(her1:her1-Venus) zebrafish in which the transcriptional regulatory regions, coding sequences, and untranslated regions (UTRs) of her1 were fused to Venus sequences then inserted into a different chromosome. This transgenic line leaves the endogenous locus intact and has been used as a reporter for the segmentation clock. We imaged embryos carrying heterozygous her1-Venus reporter along the entire PSM from 14 to 25 somite stages (Fig. 3d, g, Methods). From the kymographs, we quantified the amplitude of oscillations in the next presumptive somites at the anterior end of PSM for 11 somite cycles (n = 12 for each background; Fig. 3e, h, k). We found that the average amplitude was highest in otherwise wild-type embryos, was decreased in gene-paired embryos (60%, U = 4423.5, z = -6.9, P = 9.426 × 10{-13}, compared to wild-type embryos), and was lowest in gene-unpaired embryos (20%, U = 2264.5, z = -10.4, P = 5.189 × 10^{-25}, compared to gene-paired embryos, Fig. 3f, i, l, m). We consistently found significantly more defective segment boundaries in live imaging of gene-unpaired embryos than of gene-paired embryos (n = 72 out of 132 versus n = 17 out of 132, respectively, $\chi^2(1) = 51.3, P = 1.606 \times 10^{-12}$, Fig. 3d, g, j, n). Within each genotype, the amplitudes of oscillations preceding disrupted boundaries were significantly lower than the ones preceding the successful ones (65% U = 506.5, z = -3.2, P = 2.748 × 10^{-3} in gene-paired and 36% U = 1067, z = -5.0, P = 1.172 × 10^{-4} in gene-unpaired embryos, Fig. 3o). In summary, gene pairing leads to transcriptional co-firing (Fig. 1p) and correlated transcription (Fig. 3c) of clock genes, which establishes a limit cycle (Fig. 2g–i) that drives oscillatory clock waves in the PSM and results in successful segmentation of somites. Unpaired functional clock genes co-transcribe less frequently, which occasionally let the system fall into a steady-state with damped oscillations (Fig. 2j–l), resulting in segmentation defects.

Although co-regulated gene pairs are often found in proximity to each other, the selective advantage of gene pairing has remained unknown. Although many paired genes encode for housekeeping proteins, some encode for transcription factors and signalling regulators (such as MRF4–MYF5, MESP1–MESP2, SIX1–SIX4–SIX6, and CYP26A1–CYP26C1), which have transcription start sites that are 9–516 kb apart. By combining single-cell transcript counting, genetic engineering, real-time imaging, and computational modelling, we have shown that linkage of two segmentation clock genes causes their correlated expression, which is phenotypically beneficial for robust developmental patterning. We anticipate that gene pairing is also similarly advantageous in other systems, and this advantage could enable engineering of precise synthetic clocks in embryos and organoids.

**Online content**

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-020-03055-0.
1. Ozbudak, E. M., Thattai, M., Kurtser, I., Grossman, A. D. & van Oudenaarden, A. Regulation of noise in the expression of a single gene. Nat. Genet. 31, 69–73 (2002).

2. Elowitz, M. B., Levine, A. J., Siggia, E. D. & Swain, P. S. Stochastic gene expression in a single cell. Science 297, 1183–1186 (2002).

3. Adachi, N. & Lieber, M. R. Bidirectional gene organization: a common architectural feature of the human genome. Cell 109, 807–809 (2002).

4. Yang, L. & Yu, J. A comparative analysis of divergently-paired genes (DPGs) among Drosophila and vertebrate genomes. BMC Evol. Biol. 9, 55 (2009).

5. Arnone, J. T., Robbins-Pianka, A., Arace, J. R., Kass-Gergi, S. & McAlear, M. A. The adjacent positioning of co-regulated gene pairs is widely conserved across eukaryotes. BMC Genomics 13, 546 (2012).

6. Yan, C., Wu, S., Pocetti, C. & Bai, L. Regulation of cell-to-cell variability in divergent gene expression. Nat. Commun. 7, 11099 (2016).

7. Lewis, J. Autoinhibition with transcriptional delay: a simple mechanism for the zebrafish somitogenesis oscillator. Curr. Biol. 13, 1398–1408 (2003).

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

9. Hamma, 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 Rep. 3, 1–7 (2013).

10. Ay, A., Knierer, S., Sperlea, A., Holland, J. & Ozbudak, E. M. Short-lived Her proteins drive robust synchronized oscillations in the zebrafish segmentation clock. Development 140, 3244–3253 (2013).

11. Trofka, A. et al. Topology and dynamics of the zebrafish segmentation clock core circuit. PLoS Biol. 10, e1001364 (2012).

12. Hanisch, A. et al. The elongation rate of RNA polymerase II in zebrafish and its significance in the somite segmentation clock. Development 140, 444–453 (2013).

13. Keskin, S. et al. Noise in the vertebrate segmentation clock is boosted by time delays but tamed by Notch signaling. Cell Rep. 23, 2175–2185 (2018).

14. Chooraopokayil, S., Willems, B., Ströhle, P. & Gajewski, M. Analysis of her1 and her7 mutants reveals a spatio temporal separation of the somite clock module. PLoS ONE 7, e39073 (2012).

15. Henry, C. A. et al. Two linked hairy/Enhancer of split-related zebrafish genes, her1 and her7, function together to refine alternating somite boundaries. Development 129, 3693–3704 (2002).

16. Lleras Forero, L. et al. Segmentation of the zebrafish axial skeleton relies on notochord sheath cells and not on the segmentation clock. eLife 7, e33843 (2018).

17. Novák, B. & Tyson, J. J. Design principles of biochemical oscillators. Nat. Rev. Mol. Cell Biol. 9, 981–991 (2008).

18. Trofka, A. et al. The Her7 node modulates the network topology of the zebrafish segmentation clock via sequestration of the Hes6 hub. Development 139, 940–947 (2012).

19. 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. Dev. Cell 23, 995–1005 (2012).

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2020
The CRISPR–Cas9 editing in zebrafish was done as previously described. sgRNA was designed using CRISPRScan (http://www.crisprscan.org/); three sgRNA with zero off-target sites were chosen. Fill-in PCR was used to generate a DNA template for sgRNA in vitro transcription. AmpliScribe T7-Flash transcription kit (Epitentre) was used to generate sgRNA for injection. Invitrogen mMESSAGE mMACHINE SP6 Transcription Kit was used to generate cas9 mRNA from pCS2-nCas9n (Addgene plasmid 47792; http://n2t.net/; RRID: Addgene_47792). Injections containing 300 pg of cas9 mRNA and 20 pg of each sgRNA were injected into one-cell stage wild-type or her1ci302 her7hu2526 mutant embryos. her7hu2526 has a Lys-to-Stop mutation in the helix of the HLH-domain. DNA was isolated from fins or whole embryos. Primers used for genotyping are listed in Supplementary Table 11. Two base pairs (GC) were deleted in the second exon of her1 (deletion position corresponds to ENSDARG00000014722 1071–1072 bp) (Extended Data Fig. 1h). Mutant her1 mRNA translates into a 39-amino-acid truncated polypeptide, with 30 amino acids matching Her1 (328 aa) and 9 amino acid mismatches before a premature stop code in the bHLH domain. Because this aa allele of her1 was generated in both the wild-type (her1+; and her7hu2526) background, we used it for experiments. Mutant fish were outcrossed at least twice before using in final experiments.

**Fish stocks**

The Tg(Ola.Actb:Hsa.HRAS-EGFP) transgenic line expressing membrane-localized-GFP, Tg(her1.1-Verus) transgenic line, AB wild-type line and her7hu2526 and Df(Chr5)her1,her7,ndrg3a;her7hu2526 mutant lines were used. her1ci302 and her1ci301her7hu2526 mutants are generated for this study, as described above. All of the fish experiments were performed under the ethical guideline of Cincinnati Children’s Hospital Medical Center; the animal protocol was reviewed and approved by Cincinnati Children’s Hospital Medical Center Animal Care and Use Committees (protocol 2017-0048).

**smFISH experiments**

The RNAscope Fluorescent Multiplex Detection kit (Advanced Cell Diagnostics, 320831) was used as previously described. We adopted most of the protocol, with minor changes. In brief, 300 pg of membrane-localized GFP RNA was injected into one-cell-stage embryos. Embryos were incubated at 23 °C unless specified otherwise. Embryos were fixed at 10–14 somite stage with 4% paraformaldehyde (PFA) the next day. After dehydration, embryos were stored at −20 °C overnight. To reduce the background staining, we added a pre-hybridization step. After washing off the pre-treatment solution, embryos were hybridized with pre-hybridization (4 M urea, 0.1% Tween-20, 5× SSC, 0.5 mg ml−1 tRNA, 0.05 mg ml−1 heparin salt, 0.0092 M citric acid) at 40 °C for 2 h before probe hybridization overnight. The probes used for this study were listed in Supplementary Table 11. The C3 probe was diluted with the C1 probe solution in a 1:50 ratio. Amp4B was used to quantify her1 and her7 transcripts. Immunohistochemistry was used to amplify the membrane label after the RNAscope assay. Embryos were permeabilized with 1% PBSTX (1% Triton X-100 in PBS) at room temperature for 1 h, blocked with blocking buffer (1% Triton X-100, 2% BSA, 5% goat serum in PBS) at room temperature for 2 h, incubated with primary antibodies (chicken IgY anti-GFP, Life Technologies, A10262, 1:200) diluted with blocking buffer at 4 °C overnight. The next day, embryos were rinsed with 1% PBSTX for 30 min, blocked with blocking buffer for 10 min, incubated with secondary antibodies (goat anti-chicken IgY (H+L); Alexa Fluor 488, A11039, Life Technologies, 1:200) and Hoechst trihydrochloride (Invitrogen, 33342) 1:400 at 4 °C overnight. Before imaging, embryos were washed with 0.2× SSC (0.01% Tween-20 in 0.2× SSC) for 10 min on ice then stored in 4% PFA.

**smFISH imaging**

ProLong Gold antifade reagent (Life Technologies, P36934) was used as mounting medium. Imaging was performed by confocal microscopy on a Nikon A1R HD confocal on TIE microscope with a 100× NA 1.45 Plan Apo objective or 100× NA 1.49 TIRF Apo and resonant scanner, and sampled at Nyquist with a pinhole diameter of 73.55 μm and xy pixel size of 0.13 μm per pixels. Large tiled images were acquired to cover the whole PSM tissue of an embryo with 0.27 μm z-stack. Images were stitched with Nikon NIS-Elements software. In total, we imaged 25 wild-type, 18 her1ci301 her7hu2526, 28 her7hu2526, 24 her1ci301 her7hu2526, 14 wild-type sibling of her1ci301 her7hu2526, 23 wild-type grown at 21.5 °C, 23 wild-type grown at 28 °C, 27 her1ci301 her7hu2526 gene-paired mutant (grown at 21.5 °C), 37 her1ci301 her7hu2526 gene-unpaired mutant (grown at 21.5 °C) embryos. Original microscopy image files are provided at the BioStudies Database.

**In situ hybridization**

In situ hybridization was performed according to standard protocols. xip24 (cb1045) probe was used to label already formed somite. DIG-labelled RNA probes were prepared by in vitro transcription and anti-digoxigenin (DIG)-AP Fab fragments (Roche, 1093274) were used. Embryos were fixed with 4% PFA in PBS at room temperature for 2 h. NBT/BCIP (Roche, 1168145100) stained in situ hybridization samples were imaged with a Nikon SMZ1500 stereomicroscope (HR Plan Apo IX WD 54), Nikon DS-Ri1 digital camera with reflected light at 23 °C room temperature. Fragmented (broken) or incomplete segment boundaries were scored as disrupted.

**Counting transcripts in single cells**

Imaris 9 Cell module was used for cell segmentation and RNA counting. The tissues adjacent to the PSM (skin, notochord, neural tube and lateral plate mesoderm) were masked using the Imaris surface tool as previously described. Manual surface creation tool was selected to mark the regions surrounding the PSM. The voxels inside the surface were set to zero to completely remove undesired tissues. To detect nuclei, nucleus smooth was enabled with filter width 0.3 μm, background subtraction was turned on with width 1.2 μm and split nuclei by seed point was enabled with 3.0 μm seed diameter. The number of voxels filter was used to select the seed points; the nucleus number of voxels was set between 8,543.1 and 67,759 to detect a single nucleus. For cell detection, cell smallest detail was set to 0.25 μm, and cell volume was set between 150 μm3 and 450 μm3. To detect single RNA molecules, a spots detection tool was used with a diameter of 0.5 μm, background subtraction enabled, and quality score was used to pick up true mRNA signal. The RNA concentration is calculated by dividing RNA counts of each cell by its volume. The total number of analysed cells and slices were 50,819 and 2,832 in wild-type; 47,973 and 3,956 in her1ci301 her7hu2526; 62,952 and 3,956 in her1ci301 her7hu2526; 41,921 and 2,743 in her1ci301 her7hu2526; 20,852 and 1,625 wild-type sibling of her1ci301 her7hu2526; and 45,283 and 2,618 in wild-type embryos grown at 21.5 °C; 42,868 and 2,630 in wild-type embryos grown at 28 °C; 59,674 and 3,032 in same-chromosome mutants; 92,255 and 4,641 in wild-type embryos grown at 21.5 °C; and 62,952 and 3,956 in her1ci301 her7hu2526; 41,921 and 2,743 in her1ci301 her7hu2526; 20,852 and 1,625 wild-type sibling of her1ci301 her7hu2526; and 45,283 and 2,618 in wild-type embryos grown at 21.5 °C; 42,868 and 2,630 in wild-type embryos grown at 28 °C; 59,674 and 3,032 in same-chromosome mutants; 92,255 and 4,641 in different-chromosome mutant backgrounds. False RNA counts due to background staining are measured separately for each gene and in each embryo. The background RNA counts for her1 and her7 genes are measured by counting spots detected in more than 40 cells located in somites, where their expression is turned off. The image analysis
pipeline performed in Python (see ‘Code availability’) as previously described. Most of the cells have low RNA numbers (fewer than 100). Cytoplasmic RNAs are well separated even for cells with 100–200 total RNA levels.

The oscillation period of segmentation clock genes increases incrementally from the posterior-to-anterior (tail-to-head) along the PSM. The slowdown of oscillations along the posterior-to-anterior axis causes a phase delay in cells located in the anterior PSM compared to those located in the posterior PSM. As a result, different phases of the oscillator cycle are observed in space along the PSM. Therefore, we quantified average transcript numbers among cells located at the same posterior–anterior position in a two-dimensional, single-cell-wide cross-section and in the same phase of oscillations. Therefore, we quantified average transcript numbers among cells located in single-cell-wide cross-sections along the PSM. The position of each cell and the number of her1 and her7 RNA molecules in each cell were measured in each embryo. Each embryo was divided into left and right halves (Extended Data Fig. 1a). Within half of each PSM, cells were grouped (sliced) on the basis of their oscillation phases along the axis. The slice width was set to 8 μm, which corresponds to the diameter of cells in the PSM. A cell was assigned to a slice when the centre of the cell is located within a spatial slice. The angle of expression stripes of the segmentation clock gene changes incrementally in a posterior–anterior direction along the PSM. We first measured the changes in stripe angles along the axis, fitted an equation to the data, and applied to samples as previously described.

Calculating the amplitude of clock genes

All embryos were aligned from their posterior ends. mRNA data from slices corresponding to the same anterior–posterior positions were grouped (blue dotted line in Extended Data Fig. 1c). The spatial amplitudes of oscillations were measured by subtracting the lowest 10% mRNA mean from the highest 90% mRNA mean in slices corresponding to the same position (Extended Data Fig. 1d). The spatial amplitudes of her1 and her7 RNA were averaged over all positions in the PSM (Extended Data Fig. 1e).

Calculating spatial Spearman and Pearson correlation coefficients

Approximately 40–80% of PSM corresponds to regions where one can detect characteristic kinematic clock waves. We quantified how much transcription of her1 and her7 co-vary in space by Spearman rank and Pearson correlations of slice averages along 40–80% of PSM from the posterior end.

Identification of the heterozygote her1b567/+/her7b567/+ embryos

Homozygous her1b567/+/her1b567 embryos are lethal owing to large chromosomal deletion. her1b567/+/her7b567 fish were incrossed, and embryos that did not have segregation defects (containing a mixture of heterozygous and wild-type siblings) were selected to perform smFISH analysis. The chromosomal junction that covers the deficiency deletion has not been mapped. Thus, the heterozygous embryos cannot be easily separated from their wild-type siblings by PCR-based sequencing. We used the number of nascent transcription loci in smFISH images to sort out the two populations. The Imaris 9 spot detection tool was used with a 1.0-μm diameter to identify bright transcription loci inside the nucleus. We selected spots detected jointly by the her1 and her7 channels that colocalize within 1 μm from each other; such spots should correspond to endogenous transcription loci owing to adjacent positioning of the two genes. The closest spot-to-spot distance was calculated with Imaris XTensions. The histogram of the spot-to-spot closest distances was plotted for each embryo. Wild-type embryos have a notably higher number of cells containing two transcription loci than their heterozygote siblings. The histogram of distances peaks around 2–3 μm and 7–8 μm for wild type and her1b567/+/her7b567 embryos, respectively (Extended Data Fig. 2g–i).

Detecting nascent transcription loci in single cells

We focus on the anterior PSM where we can see expression stripes of clock genes. Because her1 is three times longer than her7, its nascent transcription loci are larger than her7 loci and are more easily differentiated from cytoplasmic dots. Therefore, we first identified the her1 loci. The Imaris 9 Spots module was used to identify her1 transcription loci with diameters of 1.0 μm on the x–y axis and 1.5 μm on the z axis, background subtraction enabled, and quality score was used to pick up transcription loci signal. The same parameters were used to identify her7 transcription loci. Two loci on the same chromosome are colocalized. Therefore, Bitplane XTensions Spots colocalize were used to identify her1 loci colocalize with her7 loci within 1 μm. To detect nuclei, the Imaris Surface module was used. Region growing was enabled, nucleus smooth was enabled with filter width 0.3 μm, background subtraction was turned on with 1.2-μm width and split nuclei by seed point was enabled with 3.0-μm seed diameter. BitplaneXTension Spots Split Into Surface Objects were used to split colocalized loci into nuclei. Nuclei with more than two loci were excluded from the analysis. The probabilities of co-firing of loci were calculated with the following equations:

\[ P(\text{her1}^+ \mid \text{her1}^-) = \frac{2 \times \text{number of nuclei with two her1 loci}}{\text{total number of her1 loci}} \]

\[ P(\text{her7}^+ \mid \text{her1}^-) = \frac{\text{number of her1 loci that colocalized with her7 loci}}{\text{total number of her1 loci}} \]

Time-lapse imaging and image analysis

We outcross homozygous Tg(her1:her1Venus) with wild-type fish to generate heterozygous Tg(her1:her1Venus) as a control. her1b567/+:Tg(her1:her1Venus) homozygous fish were outcrossed with her1b567/+:Tg(her1:her1Venus). Homozygous Tg(her1:her1Venus) fish were outcrossed with her1b567/+ her7b567/+. Fish to generate gene-paired embryos with heterozygous Tg(her1:her1Venus). Embryos were collected within 15 min, incubated at 28 °C till the tailbud stage, then transferred to 21.5 °C until imaging. Before imaging, embryos were dechorionated and laterally aligned in holes (600 μm in diameter that fits yolk) in 1% agarose gel in E3 medium with 4% tricaine that was cast in 35-mm with 15-mm glass-bottom dish. The dish contained E3 medium with 4% tricaine to prevent muscle twitching at later stages and allowed to equilibrate at 21.5 °C before imaging. Images were captured under Nikon Ti-E SpectraX Widefield Microscope (Plan Apo λ 10X) with an Andor Zyla 4.2 megapixel 16-bit sCMOS monochromatic camera, at a 512 × 512 pixel resolution. Four wild-type, four gene-paired, and four gene-unpaired embryos were imaged at a time; the experiment was repeated three times (12 embryos in total per genotype). The samples were excited by 508-nm solid-state diode laser and images were recorded in 2 channels, YFP (535/30 filter) and brightfield, 5 × slices in 30-μm steps with a range of 120 μm, at a time interval of 5 min for 11 h. The imaging temperature of the medium around the embryos was maintained at 21.4 ± 0.1 °C. Kymographs were generated from the YFP images in Fiji using the LOI interpolator from 14 to 25 somite stage for 11 h. Kymographs were aligned at the last formed somite boundary (on the right side). The width of the
segmented line was set to 20-pixel (52 μm), and the line was drawn along the PSM at the formation of each segment during somitogenesis. The intensity profiles were plotted at the anterior end of PSM from the kymographs, using the plot profile with a 4-pixel line width, after subtracting the minimum intensity along the line for each kymograph. The anterior amplitude of each somite from 15 to 25 was calculated by subtracting the trough intensity from the peak intensity before the somite formation. Amplitudes of gene-paired or gene-unpaired embryos were normalized to the mean amplitude of wild-type embryos for each somite stage within the same experiment. To compare the amplitudes preceding normal and disrupted segments, the amplitudes were normalized to the average amplitude of clock cycles preceding normal segments in gene-paired or gene-unpaired embryos. The error bars show s.e.m. of amplitude percentages.

Equipment and settings
Figure 1d, k, Extended Data Figures 1a, 2a were acquired with Nikon AIR HD confocal on TIE microscope with a 100× NA 1.45 plan objective and resonant scanner, and sampled at Nyquist (x-y pixel size of 0.13 μm per pixel) with pinhole diameter of 75.35 μm that yielded an optical section thickness of 0.54 μm. Data were acquired as a series of overlapping multichannel z-stacks forming a tiled large image that would cover the whole PSM tissue of an embryo with 0.27 μm z-step. Tiled images were stitched with Nikon NIS-Elements software. Embryos were flat mount with ProLong Gold antifade reagent (Life Technologies, P36934) imaged at 22 °C. Images were acquired as 12 bits per channel. A 561-nm laser with 685/70 and 638-nm laser with 600/50 filters were used to quantify her7 and her1 transcripts. A 405-nm laser with 450/50 filter was used for nuclear imaging. Figure 3a was acquired with Nikon AIR HD confocal on TIE microscope with a 100× NA 1.49 TIRF Apo objective; the rest of the conditions were kept the same as Fig. 1d, k, Extended Data Figs. 1a, 2a images. Figures 1e, 2c, d, Extended Data Figure 2c were acquired using a Nikon SMZ1500 stereomicroscope (HR Plan Apo 1X WD 54), Nikon DS-Ri1 digital camera with reflected light at 23 °C room temperature. Images were acquired as 16 bits. Figure 3d–1 was acquired by a Nikon Ti-E SpectraX Widefield Microscope (Plan Apo λ 10×) with an Andor Zyla 4.2 megapixel 16-bit sCMOS monochromatic camera. The imaging format was set to 512 × 512 pixel resolution (binning 4 × 4). The samples were excited by 508-nm solid-state diode laser, and images were recorded in 2 channels, YFP (535/30 filter) for 300 ms and brightfield for 1 ms, along 120-μm range of 5 z-slices with a 30-μm step-size, at a time interval of 5 min for 11 h. The imaging temperature of the medium around the embryos was maintained at 21.4 ± 0.1 °C.

Computational modelling
We developed a time-delayed stochastic model consisting of 16 reactions with 15 parameters. The variables (Supplementary Table 12): m1 and m7 represent the number of her1 and her7 RNA molecules, respectively; p1l, ph7 and ph6 represent the number of Her1, Her7 and Hes6 proteins, respectively. The variables c1, c1h1l and c1h6l represent free, Her1–Her1-bound and Hes6–Her7-bound states of chromosome 1, respectively. Similarly, the variables c2, c2h1l and c2h6l represent free, Her1–Her1-bound and Hes6–Her7-bound states of chromosome 2, respectively. Each reaction describes a change in the state of DNA or the number of mRNA or protein molecules (Supplementary Tables 13, 14). Protein synthesis and degradation rates are ps1h and pd1h for Her1, and ps7h and pd7h for Her7. mRNA synthesis and degradation rates are described by msh1 and mdh1 for her1, and msh7 and mdh7 for her7, respectively. DNA association rates of Her1–Her1 and Her7–Hes6 dimers to both chromosomes are represented by ah1h1 and ah6h7, respectively. DNA dissociation rates of Her1–Her1 and Her7–Hes6 dimers to both chromosomes are represented by dh1h1 and dh6h7, respectively. Transcriptional and translational delays are represented by tmh1 and tph1 for her1, and tmh7 and tph7 for her7, respectively.

In our model, we have 16 reactions (Supplementary Table 14). We assumed that her1 or her7 mRNA is synthesized only if the chromosome is at the free state (that is, c1 and c2). Only Her1–Her1 and Her7–Hes6 dimers can bind to sites on the promoter regions of her1 and her7 genes23,24, to repress their transcription. Her1–Her1 and Her7–Hes6-bound states of chromosomes 1 and 2 (that is, c1h1l, c2h1l, c1h6h7 and c2h6h7) are the inactive states of her1 and her7 genes. We modelled gene-unpaired heterozygous mutant as chromosomes 1 and 2 leading to transcription of her1 and her7 genes, respectively. Similarly, we modelled gene-paired heterozygous mutant as only chromosome 1 leading to transcription of both her1 and her7 genes. her1 and her7 mRNAs are assumed to be degraded at a linear rate. Similarly, Her1 and Her7 protein synthesis and degradation are modelled with a linear rate. In contrast to Her1 and Her7 proteins, Hes6 protein is assumed to be constant. The time needed to transcribe and translate mRNA is accounted for by the transcriptional and translational delay terms3. The propensity equations for the model are provided in Supplementary Table 14. The model was simulated in Matlab (see ‘Code availability’) for three different scenarios: (1) transcription firing rates of her1 and her7 were kept constant (Fig. 2m–o). (2) The average transcription firing rates of her1 and her7 were kept constant. But, at each incidence of firing, the firing rate for each gene was separately and randomly chosen from a distribution that has the same average rate (Extended Data Fig. 4a, b). (3) Transcription and RNA degradation rates of her7 were set to 50% higher than those of her1, which led to similar average RNA numbers of two genes (Extended Data Fig. 4c, d).

Stochastic simulations
Probabilistically determined propensities and reaction times are used to decide which reaction fires at each iteration. Reactions with higher propensities are more likely to fire. As was done in an earlier study, we performed the stochastic simulations using the next reaction method11. A delayed reaction queue is incorporated into the standard next reaction method to accommodate time delays, as previously suggested15. In total, we ran 30 different 240 min simulations for each genetic background. We generated 30 different random number seeds and used them equally for the gene-paired and -unpaired genotypes so that their results are directly comparable. In the case of paired-gene transcription, their transcription rate was compared to a single random number, transcription firing was determined simultaneously for both genes. By contrast, for unpaired functional her1 and her7 genes, transcription rates were compared to two different random numbers and transcription firing was determined separately for each gene.

Statistical analysis
An independent samples t-test was run to determine whether there were differences in the spatial amplitudes of her1 and her7 in Fig. If the effect size was calculated by Cohen’s d: d = 0.25 for her1 and d = 0.30 for her7 (a small effect) and in Extended Data Fig. 2d (the effect size was calculated by Cohen’s d: d = 1.2 for her1 and d = 2.4 for her7 (a large effect). There were no outliers in the data as assessed by inspection of a box plot in Extended Data Fig. 2d. One outlier is removed in Fig. 1f, although keeping the outlier did not change the statistical significance of our result. Spatial amplitudes for each genetic background were normally distributed, as assessed by Shapiro–Wilks’s test. Homogeneity of variances assumption, as assessed by Levene’s test for equality of variances, only failed for Extended Data Fig. 2d for her1 amplitude data. We have reported the Welch’s t-test for her1 amplitude comparison in this figure. A chi-square test for association was conducted between genotype and segmentation defect in Figs. 2e, 3n (between gene-paired and gene-unpaired embryos). All expected cell frequencies were greater than five. In Fig. 3n, Fisher’s exact test of independence was conducted between wild-type and gene-paired embryos; we could not use chi-squared test owing to the zero disrupted frequency.
in wild-type embryos. A Mann–Whitney U test was run to determine whether there were differences in the number of disrupted boundaries between gene-paired and gene-unpaired in Fig. 2f. Distributions of the number of disrupted boundaries were not similar, as assessed by visual inspection, thus we compared the mean rank of the number of disrupted boundaries between gene-paired and gene-unpaired embryos. A Mann–Whitney U test was run to determine whether there were differences in the co-firing probability between her1–her1 and her1–her7 for embryos raised at different temperatures. Distributions of the co-firing probability were not similar, as assessed by visual inspection (Fig. 1m, p), thus we compared the mean rank of the co-firing probability. A Mann–Whitney U test was run to determine whether there were differences in the normalized amplitude levels between wild-type and gene-paired or gene-paired and gene-unpaired embryos (Fig. 3m) or whether there were differences in the normalized amplitude preceding normal or disrupted segmentation both in gene-paired and gene-unpaired embryos (Fig. 3o). Distributions of the amplitude were not similar, as assessed by visual inspection, thus we compared the mean rank of the normalized amplitude. A Spearman’s rank-order correlation was run to assess the correlation of her1 and her7 gene expressions. Pearson correlation was also run to confirm the results. Our analysis showed the relationship to be monotonic, as assessed by visual inspection of a scatterplot. However, it also showed our data, for some cases, failed assumptions of the Pearson correlation such as both variables being normally distributed, as assessed by Shapiro Wilk’s test, and no outliers. Thus, we reported the Spearman rank correlation in main figures. Pearson correlation scores were reported in Extended Data Fig. 3 to show that the choice of the correlation metric did not change our main findings. A Mann–Whitney U test was run to determine whether there were differences in the Spearman rank and Pearson correlations in Figs. 1j, o, 2o, 3c, Extended Data Figs. 2f, 3, 4b, d. This choice is made owing to the failure of the assumptions of the t-test of independent samples such as existence of outliers and normality in parts of the dataset. Distributions of the Spearman rank and Pearson correlations were sometimes not similar, as assessed by visual inspection, thus we compared the mean rank of the correlation scores. All hypotheses testing has been done by using SPSS 26.0 (IBM). For the cases where multiple tests were done on the same data set, Bonferroni correction is applied by multiplying the P value by the number of tests. The corrected P values are reported. A priori statistical power analysis was performed for sample size estimation, based on preliminary data for each experiment. The effect sizes for our preliminary data-excluding Fig. 1o, p (Pher7/her1 temperature comparison), Extended Data Fig. 3a (wild type vs heterozygous), and Extended Data Fig. 3c varied between 0.3 and 2.4, considered to be a medium-large effect using Cohen’s criteria. Using G*Power software, for each experiment, we determined the minimum sample size to collect the data for using the preliminary effect sizes, α = 0.05 and power = 0.85. For each experimental data, we also conducted a post hoc power analysis to determine whether our design had enough power. We had enough power (power >0.8) for all our experiments except Fig. 1o (power = 0.6), Extended Data Fig. 3c (power = 0.7) and 1p (power = 0.3 only for Pher7/her1 temperature comparison which had P < 0.05), which does not change main conclusions.
Extended Data Fig. 1 | Single RNA molecules are quantified in single cells in the zebrafish PSM. **a.** Top, a single z-section of PSM of a wild-type embryo. *her7* mRNAs and nuclei are coloured in red and blue, respectively. Scale bar, 30 μm. *n* = 24, *N* = 2. Bottom, the PSM is divided into single-cell-wide slices. Cells containing higher or lower RNA than an arbitrary threshold are plotted as red or grey circles, respectively. Left and right halves of the PSM are located at the top and bottom portions of the image, respectively. Three oscillatory waves of *her7* are visible. **b.** *her7* RNA counts are plotted along the posterior-to-anterior direction at the left half of PSM. Each dot corresponds to the average RNA number in a spatial cell population (slice). **c.** All embryos are aligned from their posterior ends and slices corresponding to the same anterior-posterior positions are grouped (blue dashed lines). **d.** The spatial amplitudes of oscillations of total *her* (*her1 + her7*) RNA. Data are mean and s.e.m. **e.** The spatial amplitudes are averaged over all positions in the PSM. Comparison of new wild-type (silver, *n* = 24, *N* = 2) data obtained with a Nikon confocal microscope versus previously published data (dark grey, *n* = 18, *N* = 4) obtained by a Zeiss Apotome®. The box spans the interquartile range, line labels median, the whiskers extend to maximal and minimal observations. Difference assessed by two-sided independent *t*-test with Bonferroni correction. *her1* *P* = 0.48; *her7* *P* = 0.106. **f.** Comparison of histograms of total *her* RNA obtained by two different microscopes. **g.** Spatial Spearman correlation scores for wild-type embryos. Thick line denotes the median; thin black lines denote the 25th and 75th percentiles. **h.** Sequencing showing two base pairs deletion in the *her1* coding sequence in *her1 ci301 her7hu2526* and *her1 ci302* fish. *n* is the number of embryos; *N* is the number of independent experiments.
Extended Data Fig. 2 | Gene pairing boosts correlated transcription. a, A her1<sup>ABC</sup>/her7<sup>ABC</sup> embryo with oscillatory waves of her7 transcription. Scale bar, 30 μm. b, One of the chromosomes has a large deletion including the her1–her7 locus. c, The boundaries of somite segments are marked by xirp2a in situ hybridization staining in sibling wild-type or heterozygous her1<sup>ABC</sup>/her7<sup>ABC</sup> (top) and homozygous her1<sup>ABC</sup>/her1<sup>ABC</sup> (bottom) embryos. Scale bar, 100 μm. d, her1<sup>ABC</sup>/her7<sup>ABC</sup> embryos (n = 24, N = 2) have reduced spatial amplitude from wild-type (n = 14, N = 2) as assessed by two-sided Welch’s t-test with Bonferroni correction for her1 and the independent samples two-sided t-test for her7 (28% her1 amplitude t(13.6) = 2.6, *P = 0.04, 28% her7 amplitude t(18) = 5.3, ***P = 9.800 × 10<sup>−5</sup>). The box spans the interquartile range, line labels median, the whiskers extend to maximal and minimal observations. e, The histogram of total her<sup>+</sup> RNA per cell is plotted in wild-type (grey) and her1<sup>ABC</sup>/her7<sup>ABC</sup> (blue) embryos. her1<sup>ABC</sup>/her7<sup>ABC</sup> embryos have 38% less total her mRNA than wild-type. f, Spatial Spearman correlation scores reflecting correlated expression of her1 and her7 in wild-type (grey) and her1<sup>ABC</sup>/her7<sup>ABC</sup> (blue) embryos as assessed by the two-sided Mann–Whitney U-test (U = 392.5, z = −3.0, **P = 0.003). Median is the thick line, and 25% and 75% are thin black lines. g, The nascent transcription loci (dots) are detected in nuclei (blue) of cells located in a stripe-region in the anterior PSM of a her1<sup>ABC</sup>/her7<sup>ABC</sup> embryo. Scale bar, 5 μm. h, The histogram of the distance between two-closest loci in wild-type embryos. i, The histogram of the distance between two-closest loci in her1<sup>ABC</sup>/her7<sup>ABC</sup> embryos. n is the number of embryos; N is the number of independent experiments.
Extended Data Fig. 3 | Spatial Pearson correlation scores for all genotypes.

a, Spatial Pearson correlation scores of her1 and her7 in wild-type (dark grey), her1ci301/+ her7hu2526/+ (silver), her1ci301 her7hu2526 (red) embryos. Differences are assessed by the two-sided Mann–Whitney U test with Bonferroni correction (wild-type, her1ci301/+ her7hu2526, \(U = 615\), \(z = -2.3\), \(P = 0.072\); wild-type, her1ci301 her7hu2526, \(U = 576\), \(z = -5.0\), *** \(P = 1.652 \times 10^{-6}\); her1ci301/+ her7hu2526, her1ci301 her7hu2526, \(U = 841\), \(z = -1.3\), \(P = 0.546\)).

b, Spatial Pearson correlation scores of her1 and her7 in wild-type (grey) and her1b567/+ her7b567/+ (blue) embryos (\(U = 410\), \(z = -2.8\), ** \(P = 0.005\)).

c, Spatial Pearson correlation scores of her1 and her7 in wild-type embryos raised at 21.5 °C or 28 °C (\(U = 772\), \(z = -2.23\), * \(P = 0.026\)).

d, Spatial Pearson correlation scores for gene-paired and gene-unpaired embryos raised at 21.5 °C (\(U = 1039\), \(z = -4.6\), *** \(P = 4.000 \times 10^{-6}\)).

a–d, Differences in Pearson correlation scores are assessed by the two-sided Mann–Whitney \(U\) test. Thick line denotes the median; thin black lines denote the 25th and 75th percentiles.
Extended Data Fig. 4 | Simulation of alternative scenarios. a, b, Scenario 2. A. Average transcription firing rates of her1 and her7 were kept constant. But, at each incidence of firing, the firing rate for each gene was separately and randomly chosen from a distribution that has the same average rate. b. Spatial Spearman correlation score of her1 and her7 expression over time (U = 3, z = −6.6, ***P = 3.879 × 10−11). Differences in Spearman correlation scores are assessed by the two-sided Mann–Whitney U test. c, d, Scenario 3. c Transcription and RNA degradation rates of her7 were set to 50% higher than those of her1 which led to similar average RNA numbers of two genes. d. Spatial Spearman correlation score of her1 and her7 expression over time (U = 0, z = −6.7, ***P = 2.872 × 10−11). Differences in Spearman correlation scores are assessed by the two-sided Mann–Whitney U test. Thick line denotes the median; thin black lines denote the 25th and 75th percentiles.
Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

- n/a
- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
- Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted. Give P values as exact values wherever suitable.
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen’s d, Pearson’s r), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection
Nikon NIS Elements was used to acquire images at Nikon microscopes.

Data analysis
Python 3, BitPlane Imaris 9, MATLAB 2019a, G* POWER 3.1.9.4, SPSS 26.0, GraphPad Prism 8.4.3, and Fiji (ImageJ 1.58c) were used.

BitPlane Imaris 9 was used to segment tissue into single cell and count transcripts within each cell.

Python 3 was used to analyze data. MATLAB 2019a was used to perform stochastic simulations.

Custom Python and MATLAB codes are provided at GitHub (https://github.com/ozbudak/zinani_genepairing).

Fiji (ImageJ 1.58c) was used to analyze live imaging data.

Power analysis were carried out by using G* POWER 3.1.9.4.

GraphPad Prism 8.4.3 was used to plot graphs.

All hypotheses testing have been done by using SPSS 26.0 (IBM).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Data sets containing RNA counts in each cell for each embryo are provided as Excel files in Supplementary Tables 1-7, 9, 10. Original microscopy image files are provided at the BioStudies (https://www.ebi.ac.uk/biostudies/studies/) (accession number S-BSST434). Source data for Figures 1-3 and Extended Data Figures 1-4 are provided within the manuscript files. There are no restrictions on data availability.
Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

- [ ] Life sciences
- [ ] Behavioural & social sciences
- [ ] Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

**Sample size**
Sample-size calculation was performed using G*Power software.

**Data exclusions**
If a tissue is severely cut or bent during flat mounting, we excluded it from further analysis.

**Replication**
Embryos were collected from at least 10 pairs of fish breedings. Data come from at least two independent experiments. All replication experiments gave similar results. All samples were pulled together during analysis.

**Randomization**
Not applicable for our study. Samples were separated based on genotype or temperature only.

**Blinding**
Blinding was mostly not possible. Embryos with certain genotypes always have visible defects in oscillation waves. So, it is a very easy guess for an experimentalist to discern their genotype (even if blinded). For boundary defect counts reported in Figure 3, we have repeated the experiment three times, the data from different batches were obtained separately by two experimentalists. We have used more than 130 embryos for this single experiment.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

### Materials & experimental systems

| n/a | Involved in the study |
|-----|-----------------------|
|     | Antibodies            |
|     | Eukaryotic cell lines |
|     | Palaeontology         |
|     | Animals and other organisms |
|     | Human research participants |
|     | Clinical data         |

### Antibodies

**Antibodies used**
Chicken IgY anti-GFP, Life Technologies, A10262; Goat anti-Chicken IgY (H+L) Alexa Fluor 488, A11039, Life Technologies; anti-digoxigenin (DIG)-AP Fab fragments, Roche, 1093274. All information is provided in the methods in manuscript.

**Validation**
Chicken IgY anti-GFP, Life Technologies, A10262 is the only primary antibody used in this manuscript. anti-GFP was not used to quantify any molecules, but was only used to mark cell membranes and segment tissue into single cells. Nonetheless, we have validated it: The antibody stained cell membranes only in the GFP RNA-injected fish embryos but not in wild-type embryos.

### Animals and other organisms

**Policy information about** studies involving animals. ARRIVE guidelines recommended for reporting animal research

### Laboratory animals

Zebrafish AB strain embryos are used at 10-14 somite stages or 2 days old. Sex is not determined at this stage. We used Tg(Ola.Actb:Hsa.HRAS-EGFP) transgenic line expressing membrane-localized-GFP, AB wild-type line, Tg(her1:her1-Venus) transgenic line, and her7hu2526 and Df(chr05:her1,her7,ndrg3a)b567 mutant lines. her1ci302 and her1ci301 her7hu2526 mutants were generated for this study.

**Wild animals**
None

**Field-collected samples**
None

**Ethics oversight**
All of the fish experiments were performed under the ethical guideline of Cincinnati Children’s Hospital Medical Center, and the
Ethics oversight: The animal protocol was reviewed and approved by Cincinnati Children’s Hospital Medical Center Animal Care and Use Committees (Protocol #2017-0048).

Note that full information on the approval of the study protocol must also be provided in the manuscript.