# State transitions of a developmental oscillator

**Authors:** Milou W.M. Meeuse[1,2]*, Yannick P. Hauser[1,2]*, Gert-Jan Hendriks[1,2], Jan Eglinger[1], Guy Bogaarts[3], Charisios Tsiairis[1], Helge Großhans[1,2,4]

**Affiliations:** [1] Friedrich Miescher Institute for Biomedical Research (FMI), Maulbeerstrasse 66, CH-4058 Basel.
[2] University of Basel, Petersplatz 1, CH-4001 Basel.
[3] University Hospital, Spitalstrasse 21, CH-4031 Basel.
[4] Correspondence to: helge.grosshans@fmi.ch.
* equal contribution.

**Abstract:** Gene expression oscillators drive various repetitive biological processes. The architecture and properties of an oscillatory system can be inferred from the way it transitions, or bifurcates, between active (oscillatory) and quiescent (stable) states. Here, we have characterized the behavior of a developmental gene expression oscillator in *C. elegans* during naturally occurring and induced bifurcations. We observe a rigid oscillator that appears to operate near a Saddle Node on Invariant Cycle (SNIC) rather than a supercritical Hopf bifurcation, which yields specific system features. Developmental progression and the oscillation period are coupled, and the stable state of the system resembles a specific phase of the oscillator. This phase coincides with the time of transitions between different developmental stages, which are sensitive to nutrition. Hence, we speculate that the system's bifurcation may constitute a checkpoint for progression of *C. elegans* larval development.

Gene expression oscillations occur in many other biological systems as exemplified by the circadian rhythms in metabolism and behavior (*1*), vertebrate somitogenesis (*2*), and plant lateral root branching (*3*), and have thus been of a long-standing interest to both experimentalists and theoreticians. A recently discovered '*C. elegans* oscillator' (*4*, *5*), i.e., a system of genes expressed in an oscillatory manner in larvae, differs from other gene expression oscillators in its unique combination of features (*4*, *5*): an oscillation of thousands of genes is detectable at the whole animal level and occurs with large amplitudes and widely dispersed expression peak times (i.e., peak phases). It also lacks temperature-compensation such that the oscillation period increases as ambient temperature decreases. Thus, a better understanding of this oscillator can provide insights into a striking phenomenon of dynamic gene expression and may help to reveal potential common principles and idiosyncrasies of gene expression oscillators.

As the properties of an oscillatory system are constrained by its architecture (*6*, *7*), examination of oscillator behavior can be used to infer architecture and function. Relevant characteristic behaviors include a system's response to perturbations and the way it transitions, or bifurcates, between stable (quiescent) and oscillatory (active) states. However, for many biological oscillators, including those controlling somitogenesis and circadian rhythms, it has been difficult to access bifurcations.

Here, we have surveyed *C. elegans* oscillator activity at high temporal resolution from embryogenesis to adulthood to observe naturally occurring and induced state transitions of the oscillator during development. This has enabled us to characterize the oscillator at a level where we can identify bifurcation points, and to probe the coupling of the oscillations with development.

**Period increase and abrupt amplitude loss**

Previous analyses of *C. elegans* oscillatory gene expression (*4*, *5*, *8*) were either insufficiently temporally resolved or too short to characterize oscillations quantitatively across *C. elegans* larval development. Hence, we sampled a population of synchronized animals hourly from 5 hours through 48 hours after plating at 25°C; i.e., throughout all four larval stages and until well into the adult stage. A pairwise-correlation plot of gene expression over time showed periodic similarity that was repeated four times (Fig. 1A), presumably reflecting progression through the four larval stages. The larger dataset enabled us to improve on the previous identification of genes with oscillatory expression (*4*), resulting in the classification of 3,739 genes as 'high-confidence oscillating' genes (Fig. S1A, B and Table S1; (*9*)). Individual transcripts peaked at a wide variety of time points (Fig. 1B), as observed for the previous, smaller gene set (*4*), and this peak phase did not appear to correlate with expression amplitude (Fig. S1C).

Strikingly, oscillations disappeared, i.e., the system bifurcated, at ~36 hours, when animals enter adulthood (Fig. 1B). This finding provided us with an opportunity to characterize the oscillator's properties by investigating the dynamics of this transition. Specifically, as amplitudes and periods may change in distinct ways depending on the type of bifurcation involved (*6*, *10*, *11*), we examined these two features. Inspection of the heatmap failed to reveal a global amplitude damping prior to cessation of the oscillation (Fig. 1B). We confirmed this finding, and thus a rather stable amplitude, when we determined the amplitudes of 'high-confidence-oscillating' in the L2, L3 and L4 stages and found them to be in good agreement (Fig. 1C). [We omitted the L1 stage because we had not sampled it from its very beginning.]

Contrasting with the stable amplitude, the period of the last oscillation cycle appeared increased (Fig. 1B, Fig. S2). We quantified the period changes over time with high temporal

resolution using a Hilbert transform (*12*). The mean oscillation period of all 'high-confidence-oscillating' genes, thus calculated at every hour of development, was approximately seven hours during early larval development, but increased after 24 hours after plating (Fig. 1D). This change was also apparent when we reconstructed an oscillation from the mean observed oscillation period and compared it to an oscillation with a constant period of seven hours (Fig. 1E).

A change in period without a noticeable change in amplitude as we have observed here characterizes rigid oscillators (*13*). It also appears incompatible with the function of a simple negative feedback loop (*14*, *15*). Rather, this behavior, and the abrupt loss of amplitude upon transition to adulthood could be explained by the operation of interlinked positive and negative feedback loops. Both behaviors appear also inconsistent with the transition occurring through a supercritical Hopf bifurcation (*6*, *10*, *11*). Instead, the observed dynamics suggest the existence of a bifurcation, such as SNIC (Saddle Node on Invariant Cycle), in which the period changes as a function of a control parameter while amplitude remains unaffected (*6*, *11*).

**Development coupled to oscillatory gene expression**

Although the gene expression oscillations occur in the context of larval development, their role(s) in development have remained speculative (*4*, *5*). As a developmental role would predict coupling to developmental progression, we wondered whether the slowing of oscillations towards the end of larval development was accompanied by a lengthened L4 stage. To quantify larval stage durations, we detected periods of molts, which terminate each larval stage, for individual *C. elegans* animals with high accuracy. We employed a luciferase assay (*16*) that reveals the period of behavioral quiescence, or lethargus, that is associated with the molt (Fig. S3A-C). Using a luciferase-expressing strain that developed with improved synchrony relative to previous transgenic strains (Fig. S3D-J), we examined the durations of developmental stages, molts and intermolts closely. We found that L2 and L3 were very similar to one another in duration (Fig. 2A). By contrast, L4 did take significantly longer than the two preceding stages due to an extension of both the fourth intermolt period and the fourth molt (Fig. 2A).

The similar trends in larval stage durations and oscillation periods, determined by the luciferase assay and RNA sequencing, respectively, are consistent with a coupling of these two processes. To test the coupling hypothesis explicitly, we sought to quantify the synchrony of oscillatory gene expression and developmental progression in individual animals at the same time. To this end, we established a microchamber-based time-lapse microscopy assay by adapting a previous protocol (*17*). In this assay, animals are hatched and grown individually in small chambers where they can be tracked and imaged while moving freely, enabling their progression through molts. Using Mos1-mediated single copy transgene integration (MosSCI) (*18*), we generated transgenic animals that expressed destabilized *gfp* from the promoter of *qua-1*, a highly expressed gene with a large mRNA level amplitude.

Consistent with the RNA sequencing data, we detected oscillations of the reporter with four expression peaks (Fig. 2B). Moreover, when we curated the molts (*9*), we observed similar rates of development as in the luciferase assays (Fig. 2C, Table S2). The stage durations were in good agreement with the averaged oscillation period times for each cycle, obtained through a Hilbert transform of GFP intensities, for all three larval stages, L2 through L4, for which oscillations period lengths could be reliably determined (Fig. 2D). Indeed, molts appeared to occur very uniformly across animals relative to the oscillatory gene expression (Fig. 2B). To quantify this relationship, we determined the gene expression phases at molt entries and exits. We obtained highly similar values across worms within one larval

stage (Fig. 2E), and only a minor drift when comparing phases across larval stages. Two additional reporter transgenes, based on the promoters of *dpy-9* and *F11E6.3*, which differ in peak expression phases from *qua-1* and one another, yielded similar results (Fig. S4).

We considered two possible interpretations of the narrow distributions of oscillation phases at molt entry and exit: first, both oscillations and development could be under independent, but precise temporal control. In this model, certain developmental events would merely coincide with specific phases of oscillations rather than being coupled to them. Therefore, variations in the periods of oscillation and development would add up, non-linearly, to the experimentally observed phase variations. Second, phase-locking of oscillatory gene expression and developmental events might result from the two processes being truly coupled and/or from one driving the other. In this case, the variations in the two periods would partially explain each other, causing a reduction in the expected phase variation relative to the first scenario.

To distinguish between these scenarios, we used error-propagation to calculate the expected error for two independent processes (*9*). Focusing on the L2 and L3 stages, where the oscillation period determination was most accurate, we found that this calculated error consistently exceeded the experimentally observed error (Fig. 2F), for all three reporter genes, for both molt entry and molt exit, and for both larval stages. Thus, our observations agree with the notion that development and oscillatory gene expression are functionally coupled, and potentially causally connected.

**Oscillation initiation with a time lag in L1**

In addition to showing a longer duration of the L4 stage relative to the two preceding stages, both the luciferase assay and the single animal imaging revealed a striking extension of the L1 stage (Fig. 2A–C, Fig. S4A, D). In contrast to the situation in L4, the extended duration of L1 could be attributed almost exclusively to a lengthening of the first intermolt, whereas the first molt was comparable in duration to the subsequent two molts. We can exclude this L1 stage extension reflects the time that starved L1 animals require to recover from starvation (*16*), as larvae hatched directly into food in both luciferase assay and microchamber imaging experiments. Thus, we conclude that the duration of the first intermolt is genuinely different from that of the other intermolts.

This finding prompted us to investigate the early larval gene expression dynamics. We sampled synchronized worms, grown at 25°C, hourly from one hour through 15 hours after plating. A pairwise-correlation plot of the RNA sequencing samples revealed periodic similarities between six hours and 15 hours. By contrast, such periodicity was not detected for the samples collected during the first five hours (Fig. 3A), suggesting a global absence of oscillations during that time. An independent RNA sequencing experiment (Fig. S5A) confirmed this result.

The start of detectable oscillations differed for individual genes (Fig. 3B). Nonetheless, the occurrence of first peaks was globally well correlated to the peak phases calculated from data in Fig. 1 (Fig. 3C). Moreover, the transcript levels of many genes with a late-occurring (11-13 hours) first peak proceeded through a trough before reaching their first peak. We conclude that initiation of larval development after arrest is accompanied by a time lag prior to transition into an oscillatory state, and that oscillations exhibit a structure of phase-locked gene expressing patterns as soon as they become detectable. In other words, it appears that the early stable (pre-oscillatory) state resembles an arrest of the oscillator in a specific phase rather than a distinct state of the system.

**Quiescent state after hatching**

To test directly the possibility that the oscillator was arrested in a specific phase in early L1 stage larvae, we asked whether gene expression at this time resembles any state of the oscillator. Hence, we sought to correlate the expression profiles for these time points to those of the later developmental time points. (In the following, we will use "TPx" to refer to any time point 'x', in hours, after hatching. Technically, this is defined in our experiment as the time after plating synchronized, first larval stage animals on food.) To this end, we fused parts of the early developmental time course (TP1 – 13) to parts of the long developmental time course (TP14 – 48) to generate one long, continuous developmental time course (Fig. 4A, Fig. S6). Validity of this fusion was supported by a high similarity between each of the overlapping time points (TP5 – 15) of the two time courses (Fig. S6B).

For 'high-confidence oscillating' genes, we then examined the correlations of the first five time points, where oscillations do not occur, to all other time points of the fused time course (Fig. 4B; (*9*)). The analysis provided two insights. First, correlation coefficients all exceeded 0.8 with little change over time, confirming the high similarity of samples TP1 – 5 to one another. Hence, the oscillator appears to be quiescent at the beginning of larval development. Second, in addition to one another, TP1 – 5 are particularly highly correlated to a specific time of the oscillatory regime, namely TP13 and the 'repetitive' TP19 and TP27 (Fig. 4B). We confirmed these two key observations on an independent L1 sequencing time course (Fig. S5). We conclude that during the first five hours after plating, 'high-confidence oscillating' genes adopt a stable expression profile that resembles a specific phase of the oscillator. In other words, this phase of the oscillator appears to poise the system for a transition from a stable to an oscillatory regime.

**Initiation of oscillation soon after gastrulation**

Given the behavior of the oscillations in early larvae, we wondered about the expression of 'high-confidence oscillating' genes in embryos. Hence, we examined single embryo gene expression data from a published time series (*19*). When plotting the embryonic expression patterns of 'high-confidence oscillating' genes sorted by their peak phase defined in larvae, we observed a dynamic expression pattern with a striking phase signature (Fig. 4C). To investigate this further, we performed a correlation analysis between embryonic and larval time points focusing on 'high-confidence oscillating' genes (Fig. S7A). Specifically, we determined the peaks of highest correlation for each embryonic time point over larval time (dots in Fig. 4D, E). This revealed that correlations increased rapidly from start of embryogenesis until ~380 min (95%-CI 317.6 min – 444.2 min) (Fig. S7B), with the peak of correlation occurring always at the same time point in each of the four larval oscillation cycles (Fig. 4D). This is particularly obvious for correlation to time points in the second larval cycle: Comprising TP14 (designated 0°) through TP20 (360°), initial correlation always peaks for TP14/0° (Fig. 4D, F, G). By contrast, past ~380 min of embryonic development, correlation levels increased only modestly, but the peaks of correlation moved progressively away from TP14/0° towards TP19/300° (Fig. 4E, F, G).

We conclude that the system adopts two distinct states during embryogenesis. Initially, it approaches the oscillatory regime in a state similar to the oscillator phase seen at TP14/0°. After completion of gastrulation and around the beginning of morphogenesis/organogenesis (*20*), it transitions into the oscillatory state and reaches, at hatching, a phase most similar to larval ~TP19/300° and its cyclical repeats (TP13, TP26/27; TP35). We note that although the oscillations were previously suggested to be linked to cuticle synthesis and molting (Kim, Hendriks), the transition into the active oscillatory state occurs long before the first signs of cuticle synthesis at ~600 min (*21*). Consistent with the wide dispersion of gene expression peak phases in larvae (Fig. S1C), this may indicate

oscillator functions in additional repetitive processes beyond molting. Additionally or alternatively, molting may involve processes, in particular those linked to ECM remodeling, that are executed long before the onset of lethargus (Supplementary Text).

**Initiation and termination in a specific oscillator phase**

To complete our understanding of oscillator function across *C. elegans* development, we investigated its mode of termination. Towards this end, we analyzed the correlation of oscillating gene expression for adult time points (TP ≥ 37 h) to all other time points. This revealed correlation peaks, i.e., local maxima, to TP13, TP19 and TP26, and a constant high correlation to TP1–5 (Fig. 5A). Thus, the state of the system in adults again resembles an arrest in a specific oscillator phase. Moreover, this adult state appears highly similar to the quiescent state that the system adopts at the initiation of larval development. Indeed, in both situations, peak values of correlation to TP13/19/26 remain largely unchanged over time, thus contrasting with the increasing correlations during oscillator initiation in early embryogenesis. Whether this reflects a possibility for oscillator reactivation later in adulthood, or whether the system would later in adulthood depart from this towards a more distinct state, resembling that of early embryos, remains to be determined.

In order to explore how the oscillator enters the arrested adult state, we sought to determine whether changes in oscillations over time followed any recognizable trends during larval development. To this end, we investigated the correlation (i.e., similarity) of the second oscillation cycle (C2, starting at TP14) to the other cycles (C1, C3, and C4, starting at TP7, TP20, and TP28, respectively) (Fig. 5B). We use the cycle nomenclature for clarity, because the absence of oscillations at the beginning of L1 and the hourly resolution of sampling mean that experimentally determined cycles can deviate from larval stages. Use of cycle 2 as the reference maintained comparability with the analysis that we performed for embryogenesis above.

To explore similarities of the cycles, we determined for each of the six time points TP14/0° through TP19/300° the peaks of correlation in the other cycles (Fig. 5B). This revealed high and largely invariant correlation values for TP14 through TP19 across each of cycle 1 and cycle 3. By contrast, values declined continuously during cycle 4. This decline in correlations was not a consequence of the extended period, as we employed spline interpolation and local maxima detection to determine correlation peaks (Fig. S8, (*9*)). Hence, in cycle 4, oscillations are progressively deviating from states in the earlier cycles. Indeed, when we plotted the correlations of the peaks over TP14/0° – TP19/300° in a polar plot, correlations to cycle 1 and cycle 3 time points fell on a circle, whereas correlations to cycle 4 spiraled away from this circle (Fig. 5C). We emphasize that, as shown in Fig. 1C, this loss of correlation appears not substantially driven by a loss of oscillation amplitude in cycle 4.

We conclude that the oscillation begins to depart from an invariant cycle already during the last cycle in the L4 stage. Upon approaching the adult stage, the system transitions to a non-oscillatory state, with animals retaining a gene expression state resembling the oscillator state at TP19/300°. This is also the state in which oscillations appear quiescent during the first hours of larval development. Therefore, we wondered whether this oscillator phase was particularly conducive to state transitions of the system in response to changes in the developmental trajectory. To test this, we examined animals that exited from dauer arrest, a diapause stage that animals enter during larval development under conditions of environmental stress such as heat, crowding, or food shortage. Using a published time course of animals released from dauer arrest after starvation (*4*), we found that their oscillating gene expression patterns correlated highly with those of animals initiating larval development in the L1 stage (Fig. 5D). Additionally, gene expression patterns at 1 hour through 5 hours and

at 13 hours post-dauer were highly correlated to those of the repetitive TP13, TP19 and TP26/27 during continuous development. In particular, a period of quiescence during the first five hours after placing animals on food was followed by a transition to an oscillatory state in a phase resembling TP19/300° of the continuous development time course (Fig. 5E).

Taken together, these findings support the notion that transitions of the system between quiescent and oscillatory regime may occur preferentially in a specific oscillator phase (Fig. 5F) that is shared by transitions during normal development (i.e., in newly hatched larvae and in young adults) and in response to perturbation (dauer arrest in response to starvation). [We note that L1 stage sequencing analysis was done on a population of animals synchronized by release from starvation. Although this synchronization procedure does not appear to slow progression through the L1 stage, it remains possible that starvation has contributed to the phase-specific oscillation arrest in this experiment.]

Although dauer diapause is an extreme form of developmental arrest, previous work showed that development can be transiently arrested in any larval stage in response to poor nutritional conditions (*22*). These arrests occur specifically at the beginning of a larval stage, shortly after molt exit, and are thought to involve the function of a developmental checkpoint. Within the limits of our temporal resolution, these times of possible arrest coincide with the oscillator phase that is particularly conducive to bifurcation. Hence, we speculate that food sensing, metabolism, or nutritional state of the animal may contribute to determining the state of the oscillatory system and that, in turn, a bifurcation point of the oscillator may constitute the developmental checkpoint. In this view, the *C. elegans* oscillator appears to provide a striking example of how a system bifurcation can be utilized to provide control over developmental processes, through enabling the integration of environmental inputs into genetic programs.

**References and Notes:**


1. S. Panda, J. B. Hogenesch, S. A. Kay, Circadian rhythms from flies to human. *Nature*. **417**, 329–335 (2002).
2. A. C. Oates, L. G. Morelli, S. Ares, Patterning embryos with oscillations: structure, function and dynamics of the vertebrate segmentation clock. *Development*. **139**, 625–639 (2012).
3. M. A. Moreno-Risueno, J. M. Van Norman, A. Moreno, J. Zhang, S. E. Ahnert, P. N. Benfey, Oscillating gene expression determines competence for periodic Arabidopsis root branching. *Science*. **329**, 1306–1311 (2010).
4. G.-J. Hendriks, D. Gaidatzis, F. Aeschimann, H. Großhans, Extensive oscillatory gene expression during C. elegans larval development. *Mol. Cell*. **53**, 380–392 (2014).
5. D. hyun Kim, D. Grün, A. van Oudenaarden, Dampening of expression oscillations by synchronous regulation of a microRNA and its target. *Nat. Genet.* **45**, 1337–1344 (2013).
6. E. M. Izhikevich, NEURAL EXCITABILITY, SPIKING AND BURSTING. *International Journal of Bifurcation and Chaos*. **10**, 1171–1266 (2000).
7. R. Guantes, J. F. Poyatos, Dynamical principles of two-component genetic oscillators. *PLoS Comput. Biol.* **2**, e30 (2006).
8. D. Grün, M. Kirchner, N. Thierfelder, M. Stoeckius, M. Selbach, N. Rajewsky, Conservation of mRNA and protein expression during development of C. elegans. *Cell Rep*. **6**, 565–577 (2014).
9. Materials and methods are available as supplementary materials.

10. J. D. Salvi, D. Ó Maoiléidigh, A. J. Hudspeth, Identification of Bifurcations from Observations of Noisy Biological Oscillators. *Biophys. J.* **111**, 798–812 (2016).
11. S. H. Strogatz, *Nonlinear Dynamics and Chaos with Student Solutions Manual: Nonlinear Dynamics and Chaos: With applications To Physics, Biology, Chemistry, And Engineering (Studies in Nonlinearity)* (Westview Press, Boulder, CO, ed. 2, 2015).
12. A. Pikovsky, M. Rosenblum, J. Kurths, *Synchronization: A universal concept in nonlinear sciences* (Cambridge University Press, Cambridge, 2001).
13. U. Abraham, A. E. Granada, P. O. Westermark, M. Heine, A. Kramer, H. Herzel, Coupling governs entrainment range of circadian clocks. *Mol Syst Biol*. **6**, 438 (2010).
14. G. Mönke, E. Cristiano, A. Finzel, D. Friedrich, H. Herzel, M. Falcke, A. Loewer, Excitability in the p53 network mediates robust signaling with tunable activation thresholds in single cells. *Scientific Reports*. **7**, 46571 (2017).
15. T. Y.-C. Tsai, Y. S. Choi, W. Ma, J. R. Pomerening, C. Tang, J. E. Ferrell, Robust, Tunable Biological Oscillations from Interlinked Positive and Negative Feedback Loops. *Science*. **321**, 126–129 (2008).
16. M. Olmedo, M. Geibel, M. Artal-Sanz, M. Merrow, A High-Throughput Method for the Analysis of Larval Developmental Phenotypes in Caenorhabditis elegans. *Genetics*. **201**, 443–448 (2015).
17. M. Turek, J. Besseling, H. Bringmann, Agarose Microchambers for Long-term Calcium Imaging of Caenorhabditis elegans. *J Vis Exp*, e52742 (2015).
18. C. Frøkjær-Jensen, M. W. Davis, M. Ailion, E. M. Jorgensen, Improved Mos1-mediated transgenesis in C. elegans. *Nat. Methods*. **9**, 117–118 (2012).
19. T. Hashimshony, M. Feder, M. Levin, B. K. Hall, I. Yanai, Spatiotemporal transcriptomics reveals the evolutionary history of the endoderm germ layer. *Nature*. **519**, 219–222 (2015).
20. Hall, D.H., Herndon, L.A. and Altun, Z., *Introduction to C. elegans Embryo Anatomy. In WormAtlas.*
21. J. E. Sulston, E. Schierenberg, J. G. White, J. N. Thomson, The embryonic cell lineage of the nematode Caenorhabditis elegans. *Dev. Biol.* **100**, 64–119 (1983).
22. A. J. Schindler, L. R. Baugh, D. R. Sherwood, Identification of late larval stage developmental checkpoints in Caenorhabditis elegans regulated by insulin/IGF and steroid hormone signaling pathways. *PLoS Genet.* **10**, e1004426 (2014).
23. R. Edgar, M. Domrachev, A. E. Lash, Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res*. **30**, 207–210 (2002).
24. J. R. Priess, D. I. Hirsh, Caenorhabditis elegans morphogenesis: the role of the cytoskeleton in elongation of the embryo. *Dev. Biol.* **117**, 156–173 (1986).
25. T. T. K. Vuong-Brender, M. Ben Amar, J. Pontabry, M. Labouesse, The interplay of stiffness and force anisotropies drives embryo elongation. *eLife*. **6**, e23866 (2017).
26. V. Lažetić, D. S. Fay, Molting in C. elegans. *Worm*. **6**, e1330246 (2017).
27. C. Lagido, J. Pettitt, A. Flett, L. A. Glover, Bridging the phenotypic gap: Real-time assessment of mitochondrial function and metabolism of the nematode Caenorhabditis elegans. *BMC Physiology*. **8**, 7 (2008).
28. D. G. Gibson, L. Young, R.-Y. Chuang, J. C. Venter, C. A. Hutchison, H. O. Smith, Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods*. **6**, 343–345 (2009).
29. C. Frøkjaer-Jensen, M. W. Davis, C. E. Hopkins, B. J. Newman, J. M. Thummel, S.-P. Olesen, M. Grunnet, E. M. Jorgensen, Single-copy insertion of transgenes in Caenorhabditis elegans. *Nat. Genet.* **40**, 1375–1383 (2008).

30. Stiernagle, Theresa, *Maintenance of C. elegans (February 11, 2006), WormBook, ed. The C. elegans Research Community, WormBook, doi/10.1895/wormbook.1.101.1, http://www.wormbook.org.*
31. R Core Team, R: A language and environment for statistical computing, (available at https://www.R-project.org/).
32. D. Gaidatzis, A. Lerch, F. Hahne, M. B. Stadler, QuasR: quantification and annotation of short reads in R. *Bioinformatics*. **31**, 1130–1132 (2015).
33. K. F. Au, H. Jiang, L. Lin, Y. Xing, W. H. Wong, Detection of splice junctions from paired-end RNA-seq data by SpliceMap. *Nucleic Acids Res*. **38**, 4570–4578 (2010).
34. A.-N. Spiess, *propagate: Propagation of Uncertainty* (2018; https://CRAN.R-project.org/package=propagate).
35. E. Jones, T. Oliphant, P. Peterson, SciPy: Open Source Scientific Tools for Python (2001), (available at http://www.scipy.org/).
36. PANTHER Classification System, (available at http://pantherdb.org/).
37. H. W. Borchers, *pracma: Practical Numerical Math Functions* (2019; https://CRAN.R-project.org/package=pracma).
38. J. Sueur, T. Aubin, C. Simonis, Seewave, a Free Modular Tool for Sound Analysis and Synthesis. *Bioacoustics*. **18**, 213–226 (2008).
39. D. Kim, H.-S. Oh, *EMD: Empirical Mode Decomposition and Hilbert Spectral Analysis* (2018; https://CRAN.R-project.org/package=EMD).
40. D. Bates, M. Maechler, *Matrix: Sparse and Dense Matrix Classes and Methods* (2018; https://CRAN.R-project.org/package=Matrix).
41. R. Gaujoux, C. Seoighe, A flexible R package for nonnegative matrix factorization. *BMC Bioinformatics*. **11**, 367 (2010).
42. K. Soetaert, *plot3D: Plotting Multi-Dimensional Data* (2017; https://CRAN.R-project.org/package=plot3D).
43. N. Gritti, S. Kienle, O. Filina, J. S. van Zon, Long-term time-lapse microscopy of *C. elegans* post-embryonic development. *Nature Communications*. **7**, 12500 (2016).
44. W. McKinney, in *Proceedings of the 9th Python in Science Conference* (2010), pp. 51–56.
45. E. O. Lebigot, Uncertainties: a Python package for calculations with uncertainties, (available at https://pypi.org/project/uncertainties/).

**Acknowledgements:** We thank Stephane Thiry and Kirsten Jacobeit and the FMI Functional Genomics Facility for RNA sequencing, Iskra Katic for help in generating transgenic strain, Maria Olmédo and Henrik Brinkmann for introducing us to the luciferase and the single animal imaging assays, respectively, Dimos Gaidatzis and Michael Stadler for advice on computational analyses, Laurent Gelman for help with imaging, Lucas Morales Moya for comments on the manuscript, and Hanspeter Herzel for discussion. **Funding:** M.W.M.M. is a recipient of a Boehringer Ingelheim Fonds PhD fellowship. This work is part of a project that has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant agreement No. 741269, to H.G.). The FMI is core-funded by the Novartis Research Foundation. **Author contributions:** Gert-Jan Hendriks and Yannick Hauser performed RNA sequencing time courses. Milou Meeuse and Yannick Hauser analyzed RNA sequencing data. Milou Meeuse performed and analyzed luciferase assays. Guy Bogaarts developed the graphical user interface for the luciferase data. Yannick Hauser

acquired and analyzed single worm imaging data. Jan Eglinger wrote the KNIME workflow for the single worm imaging. Charisios Tsiairis conceived parts of the analysis. Helge Großhans, Milou Meeuse and Yannick Hauser conceived the project and wrote the manuscript. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All sequencing data generated for this study have been deposited in NCBI's Gene Expression Omnibus (*23*) and are accessible through GEO SuperSeries accession number GSE133576 https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE133576. The dauer exit time course was published previously (*4*) and is accessible through GEO Series accession number GSE52910 (http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE52910).
All other data are available upon request to the corresponding author. Published research reagents from the FMI are shared with the academic community under a Material Transfer Agreement (MTA) having terms and conditions corresponding to those of the UBMTA (Uniform Biological Material Transfer Agreement).

## Figures

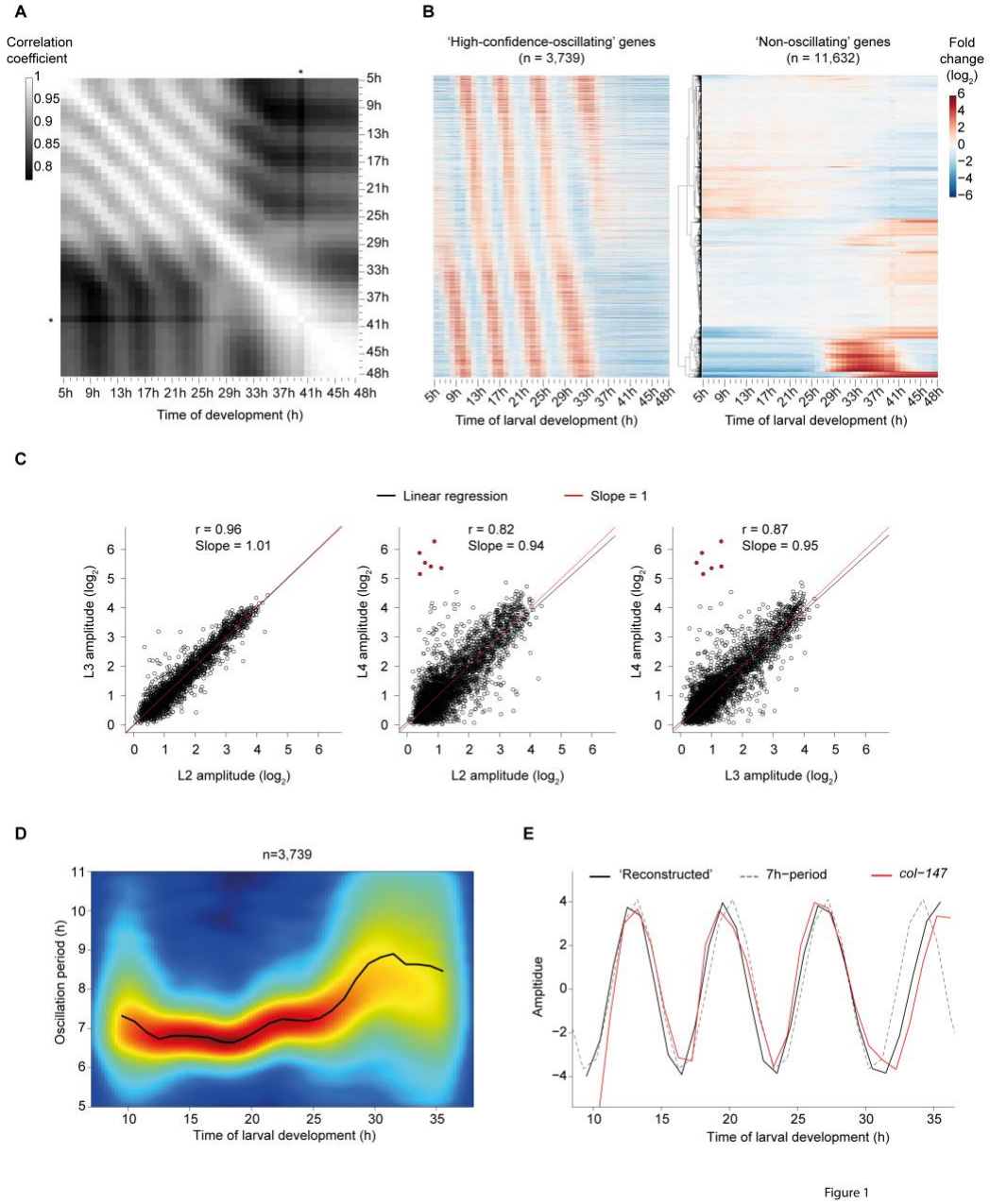



**Fig. 1: Transcript level oscillations with stable amplitudes and variable period during larval development**

(**A**) Pairwise correlation plot of $\log_2$-transformed gene expression patterns obtained from synchronized population of L1 stage larvae sampled and sequenced from t = 5 h until t = 48 h after plating ('L1–YA timecourse'). Asterisk indicates an outlier, time point t = 40 h.
(**B**) Gene expression heatmaps of genes as classified in Fig. S1. 'High-confidence-oscillating' genes were sorted by peak phase and 'non-oscillating' genes were hierarchically clustered.
(**C**) Amplitudes derived from cosine fitting to the individual oscillations of L2, L3 and L4 stage plotted against each other. Pearson correlation coefficient r, slope of the linear regression (black) and the diagonals (slope=1; red) are indicated. Outliers (red circles) are *col-120*, *col-88*, *col-81*, *col-124*, *col-139* and *col-20*; collagen genes whose steep upregulation during the L4 stage causes an artificially inflated calculated amplitude.

(**D**) Density plot showing oscillation period as quantified by Hilbert transform for all 'high-confidence-oscillating' genes (n=3739). Black line: mean oscillation period.
(**E**) Expression changes for an oscillation with a constant 7-h period (dotted line), and an oscillation reconstructed from the mean oscillation period in D (black line), both amplitudes set to four. Red line: a representative gene, *col-147* (mean normalized).

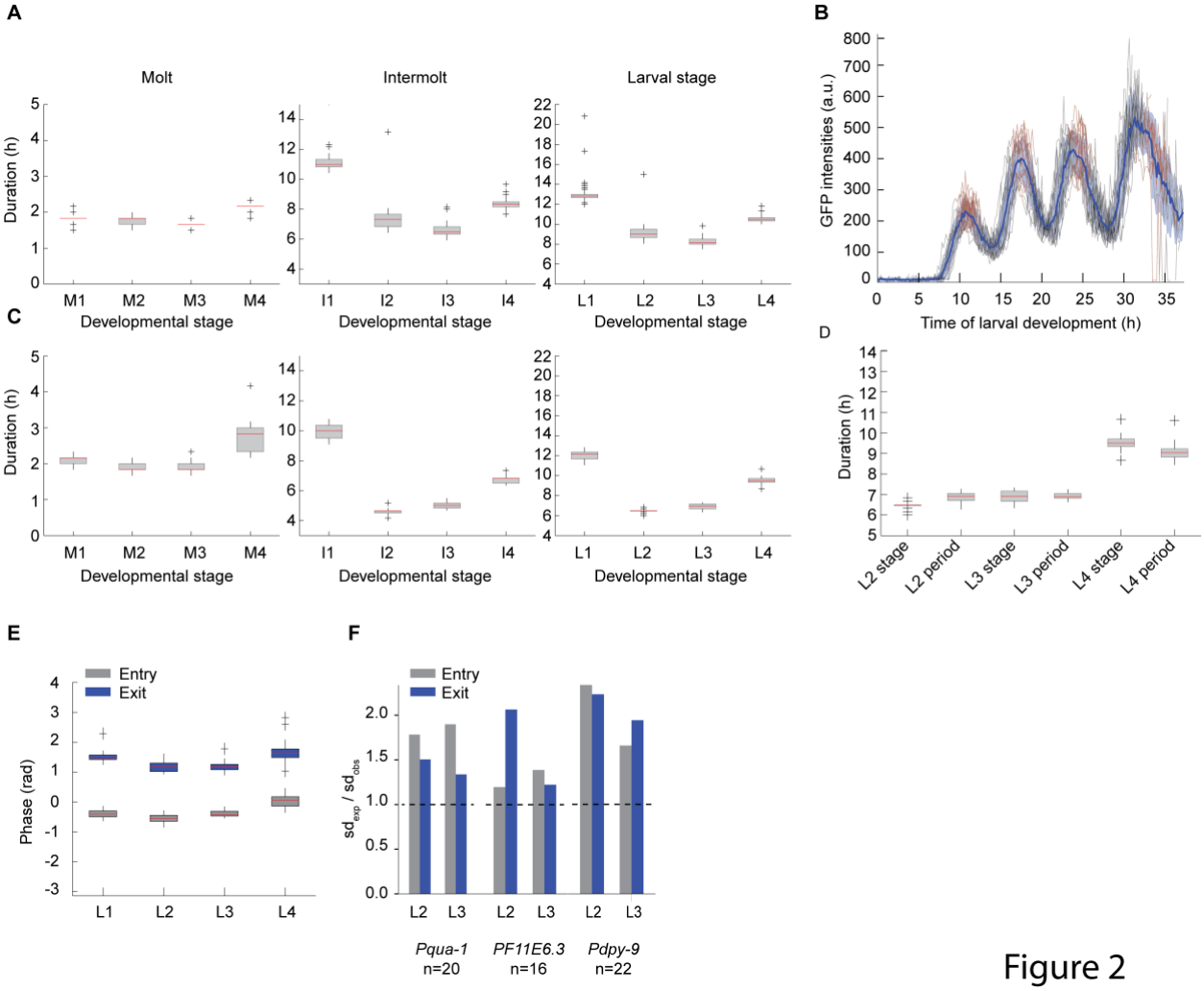



Figure 2

**Fig. 2: Coupling of development and oscillatory gene expression**

(**A**) Boxplots of durations of molts (M), intermolts (I) and larval stages (L) at 20°C, determined through a luciferase assay for single animals (n=86; strain HW1939) (*9*).

(**B**) GFP quantification of HW2523 (*Pqua-1::gfp::pest::h2b::unc-54*$_{3'UTR}$, n=20) single animals over larval development, starting from hatch (t = 0 h). Molts (red), mean intensity (blue line) and standard deviation across population (shading) are indicated.

(**C**) Boxplots of molt, intermolt and larval stage durations of single animals (HW2523) developing in microchambers (n=20).

(**D**) Boxplot of larval stage durations and period times of oscillations in hours for HW2523 animals (n=20). L1 was excluded because of the time lag before oscillations manifest after hatching.

(**E**) Boxplot of phase at molt entry (start of lethargus) and molt exit (end of lethargus) separated by larval stages for strain HW2523 (n=20).

(**F**) Barplots displaying the ratio of expected standard deviation over observed standard deviation for phase calling at either molt entry or molt exit for the indicated reporters. A dashed line indicates parity.

Boxplots extend from first to third quartile with a line at the median, outliers are indicated with a cross, whiskers show 1.5*IQR.

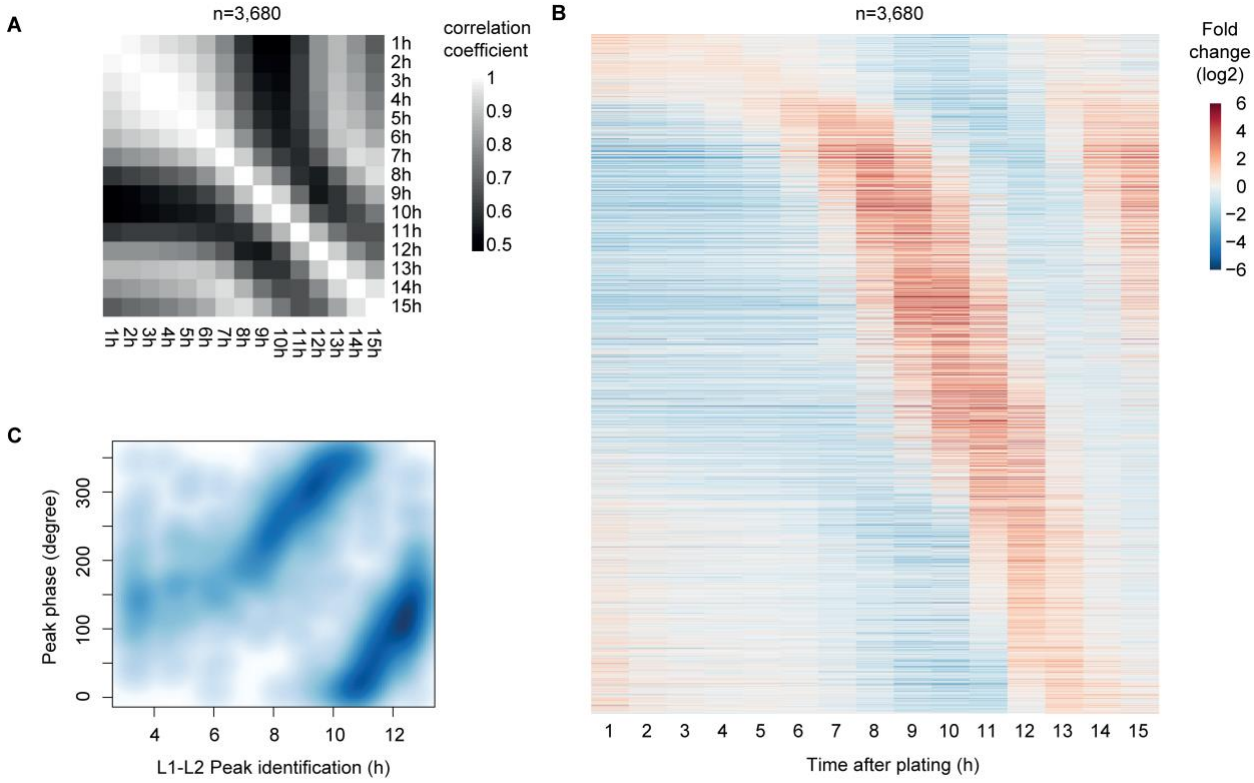



Figure 3

**Fig. 3: Lack of gene expression oscillations in early L1 stage worms**
(**A**) Pairwise correlation plot of $\log_2$-transformed 'high-confidence oscillating' gene expression obtained from a synchronized population of L1 stage larvae placed on food at t = 0 hours at 25°C ('L1–L2 time course'). mRNA was sampled and sequenced from t = 1 h until t = 15 h after plating.
(**B**) Gene expression heatmap of detectably expressed 'high-confidence oscillating' genes. Genes were ranked according to their first peak detected by a spline analysis.
(**C**) Comparison of expression peaks determined for 'high-confidence oscillating' genes in the L1–L2 time course with their peak phase calculated in Fig. 1.
All analyses for 'high-confidence oscillating' genes identified in Fig. 1 with detectable expression (n=3680).

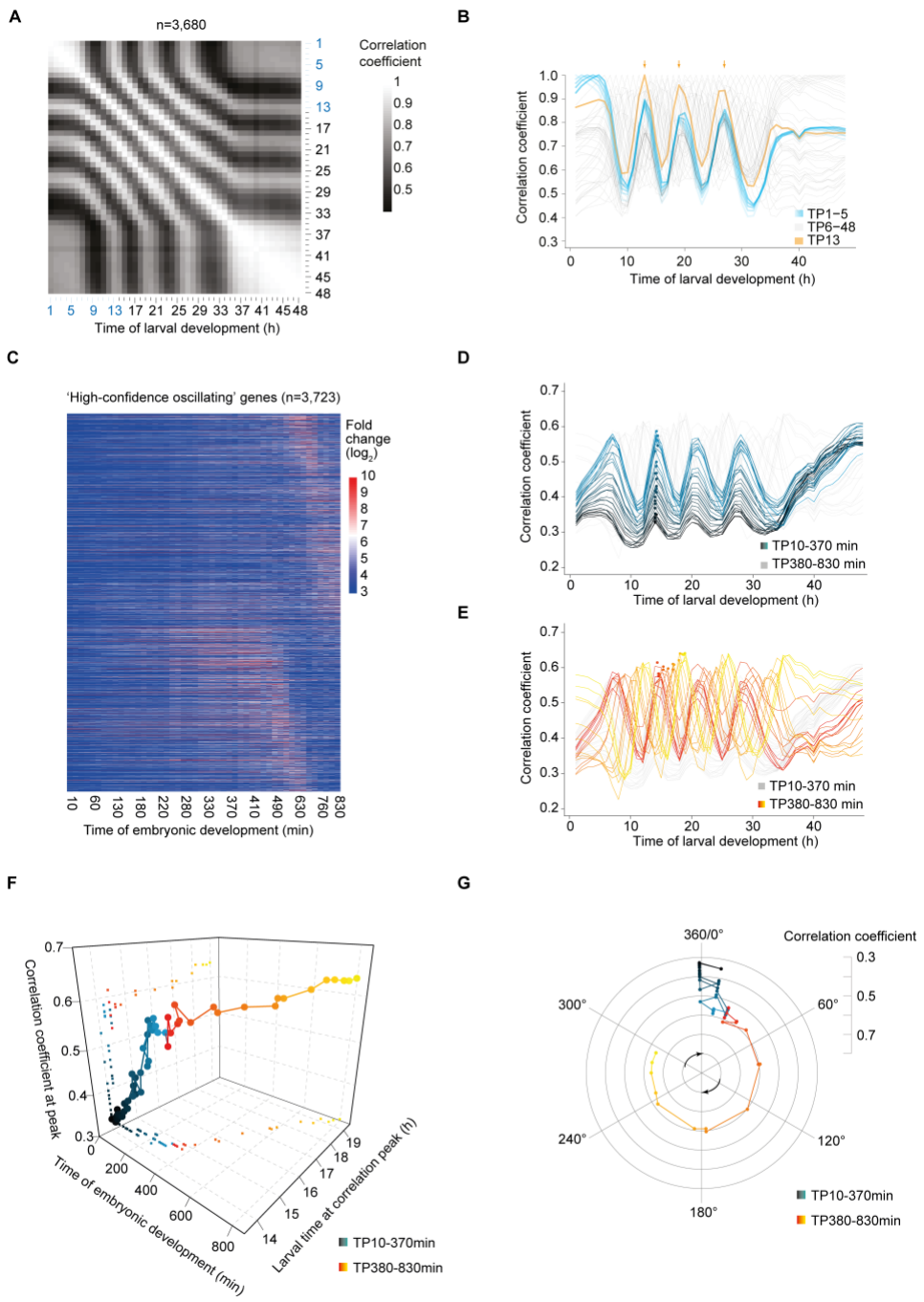



Figure 4

**Fig. 4: Transition to an oscillatory state during embryogenesis**

(**A**) Pairwise correlation plot of $\log_2$-transformed 'high-confidence oscillating' gene expression obtained from a fusion of the L1–L2 time course (blue labels) and the L1–YA time course (black).

(**B**) Correlation of TP1–5 (blue) and TP13 (orange) to all other time points in the fused larval time course. Arrows indicate local correlation maxima at TP13, 19 and 26/27.

(**C**) Heatmap of $\log_2$-transformed embryonic expression of 'high-confidence-oscillating' genes sorted by larval peak phase.

(**D, E**) Correlation of each embryonic time point to larval time points of the fused larval time-course for embryonic TP10-370 min (D, black-blue gradient) and TP380-830 min (E, red-

yellow gradient), respectively. Dots represent peaks of the correlation lines after spline analysis in the second oscillation cycle.
(**F**) 3D-scatter plot of the maximum correlation coefficient for each embryonic time point to the second larval oscillation cycle. Embryonic time is determined by time of sample collection, larval time by spline interpolation.
(**G**) Polar plot of maximum correlation coefficients over the point in the second larval oscillation cycle at which the correlation peak is detected. TP14 is defined as 0° and correlates most highly to TP20, thus defined as 360°. Values are as in F; color scheme as in D and E.
All correlations were determined by Pearson correlation.

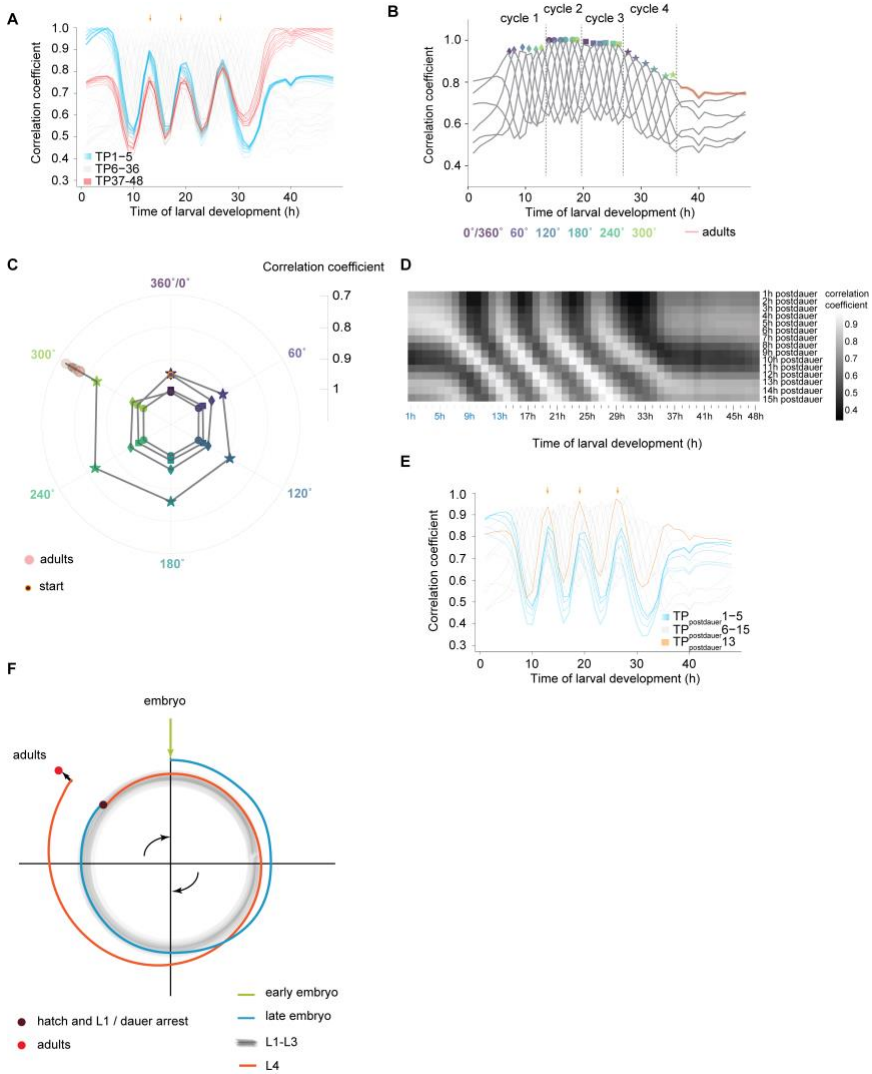



Figure 5

**Fig. 5: Transition towards a quiescent state during the L4 stage**

(**A**) Correlation of TP1–5 (blue lines) and TP37–48 (adult; red) to all other time points of the fused larval time course from Fig. 4A. Orange arrows indicate local correlation maxima to TP13, 19 and 26/27.

(**B**) Correlation of cycle 2 time points (TP14–19; corresponding to 0° to 300°; marked by indicated colors) to all other time points of the fused larval time course. Diamonds (cycle 1), circles (cycle 2), squares (cycle 3) and stars (cycle 4) indicate correlation peak values and peak times determined by spline interpolation.

(**C**) Polar plot displaying correlation of cycle 2 gene expression patterns with those of the corresponding points in the other cycles. Color scheme and symbols as in B. Adult time points (red circles) and start (0° in cycle 1; orange) are placed according to correlations in A and B, respectively.

(**D**) Pairwise-correlation map of oscillatory gene expression of dauer exit samples ("postdauer") and fused larval time course samples.

(**E**) Correlation of the indicated time points after plating dauer-arrested animals on food ($TP_{postdauer}$) to the fused larval time course. Arrows indicate peaks of correlation to TP13/19/26.5 (300°) of the larval time course.

(**F**) Model of the '*C. elegans* oscillator'.

All correlations were determined by Pearson correlation.

# Supplementary Materials for

## State transitions of a developmental oscillator

**Authors:** Milou W.M. Meeuse, Yannick P. Hauser, Gert-Jan Hendriks, Jan Eglinger, Guy Bogaarts, Charisios Tsiairis, Helge Großhans*

*Correspondence to: helge.grosshans@fmi.ch.

### Supplementary Text

At ~380 min into embryo development, oscillations become detectable not only long before the first signs of cuticle synthesis, but also do so specifically at a time when embryo elongation and the formation of an apical ECM structure known as embryonic sheath (ES) (*24*, *25*) occur. Of the five known genes encoding ES components (*25*), four (*sym-1; fbn-1; noah-1; noah-2*) are also detectably expressed in larvae. (The fifth gene has not been mapped in our data.) For all four, the levels oscillate with high amplitudes and peak shortly after cuticle shedding (ecdysis), i.e., after a molt has just been completed. Hence, molting may involve processes, in particular those linked to ECM remodeling, that are executed long before the onset of lethargus. In other words, the traditional definition of molting as comprising lethargus, during which the old cuticle is detached from the epidermis and a new cuticle synthesized, and ecdysis (*26*), may fail to capture the full complexity and duration of this fundamental process of nematode development.

### Materials and Methods

<u>*C. elegans* strains</u>

The Bristol N2 strain was used as wild type. The following transgenic strains were used:
HW1370: *EG8079; xeSi136 [F11E6.3p::gfp::h2b::pest::unc-54 3'UTR; unc-119 +] II* (this study)
HW1939: *EG8079, xeSi296 [Peft-3::luc::gfp::unc-54 3'UTR, unc-119(+)] II* (this study)
HW2523: *EG8079*: *xeSi437 [Pqua-1::gfp::h2b::pest::unc-54 3'UTR; unc-119 +] II* (this study)
HW2526: *EG8079: xeSi440 [Pdpy-9::gfp::h2b::pest::unc-54 3'UTR; unc-119 +] II* (this study)
PE254: *feIs5 [Psur-5::luc::gfp; rol-6(su1006)] V* (ref. (*27*))
PE255: *feIs5 [Psur-5::luc::gfp; rol-6(su1006)] X* (ref. (*27*))

All transcriptional reporters and luciferase constructs produced for this study were generated using Gibson assembly (*28*) and the destination vector pCFJ150 (*29*). First a starting plasmid was generated by combining *Not*I digested pCFJ150, with either *Nhe-1::GFP-Pest-H2B* or *Nhe-1::luciferase::GFP* (adapted from pSLGCV (*27*) and ordered as codon optimized, intron containing gBlocks® Gene Fragment (Integrated DNA Technologies), and *unc-54 3'UTR*

(amplified from genomic DNA) to yield pYPH0.14 and pMM001 respectively. Second, promoters consisting of either 2kb upstream of the ATG or up to the next gene were amplified from *C. elegans* genomic DNA before inserting them into *Nhe*I-digested pYPH0.14 or pMM001. PCR primers and resulting plasmids are listed in the table below. Third, we obtained transgenic worms by single-copy integration into EG8079 worms, containing the universal *ttTi5605* locus on chromosome II by following the published protocol for injection with low DNA concentration (*18*). All MosSCI strains were at least backcrossed two times. All primer sequences can be found in Supplemental Table 3.

Methods luciferase assay

Gravid adults were bleached and single embryos were transferred by pipetting into a well of a white, flat-bottom, 384-well plate (Berthold Technologies, 32505). Embryos hatched and developed in 90 µL volume containing *E. coli* OP50 ($OD_{600}$ = 0.9) diluted in S-Basal medium (*30*) and 100 µM Firefly D-Luciferin (p.j.k., 102111). Plates were sealed with Breathe Easier sealing membrane (Diversified Biotech, BERM-2000). Luminescence was measured using a Luminometer (Berthold Technologies, Centro XS3 LB 960) for 0.5 seconds every 10 minutes for 72 hours at 20°C in a temperature-controlled incubator and is given in arbitrary units. Luminescence data was analyzed using an automated algorithm for molt detection on trend-corrected data as described previously (*16*) but implemented in MATLAB, and with the option to manually annotate molts in a Graphical User Interface. The hatch was identified as the first data point that exceeds the mean + 5*stdev of the raw luminescence of the first 20 time points and also exceeds the raw luminescence by 3.
To quantify the duration of the molts, we subtracted the time point at molt entry from the time point at molt exit. To quantify the duration of larval stages, we subtracted the time point at molt exit of the previous stage (or time point at hatch for L1) from the time point at molt exit of the current stage. The duration of the intermolt was quantified as duration of the molt subtracted from duration of the larval stage. For statistical analysis, we assumed the durations to be normally distributed and used Welch two-sample and two-sided t-test, i.e. the function 't.test' of the package 'stats' (version 3.5.1) (*31*) in R.

RNA sequencing

For RNA sequencing, synchronized L1 worms, obtained by hatching eggs in the absence of food, were cultured at 25°C and collected hourly from 1 hours until 15 hours of larval development, or 5 hours until 48 hours of larval development, for L1–L2 time course and L1–YA time course respectively. A repeat experiment was performed at room temperature from 1 hours until 15 hours. RNA was extracted in Tri Reagent and DNase treated as described previously (*4*). For the L1–YA time course, libraries were prepared using the TruSeq Illumina mRNA-seq (stranded – high input), followed by the Hiseq50 Cycle Single-end reads protocol on HiSeq2500. For the early larval time course, libraries were prepared using the Illumina TruSeq mRNA-Seq Sample Prep Kit (Strand-sequenced: any), followed by the Hiseq50 Cycle Single-end reads protocol on HiSeq2500.

Processing of RNA-seq data

RNA-seq data were mapped to the *C. elegans* genome using the qAlign function (splicedAlignment=TRUE) from the QuasR package (*32*, *33*) in R. Gene expression was quantified using qCount function from the QuasR package in R. For L1-YA and Dauer exit time

course, QuasR version 1.8.4 was used, and data was aligned to the ce10 genome using Rbowtie aligner version 1.8.0. For L1-L2 time course, QuasR version 1.2.2 was used, and data was aligned to the ce6 genome using Rbowtie aligner version 1.2.0. For L1-L2 replicate time course (Supplemental Figure 4), QuasR version 1.10.1 was used, and data was aligned to the ce10 genome using Rbowtie aligner version 1.10.0.
Counts were scaled by total mapped library size for each sample. A pseudocount of 8 was added and counts were $\log_2$-transformed. For the L1–YA time course, lowly expressed genes were excluded (maximum $\log_2$-transformed gene expression - ($\log_2$(gene width)-mean($\log_2$(gene width))) ≤ 6). This step was omitted in the early time course because many genes start robust expressing only after 5-6 hours. Expression data of the dauer exit time course was obtained from ref. (*4*).

Classification of genes

To classify genes, we applied cosine fitting to the $\log_2$-transformed gene expression levels from t=10 hours until t=25 hours of developmental time (mid L1 until late L3), when the oscillation period is most stable (Figure 1D). During this time the oscillation period is approximately 7 hours, which we used as fixed period for the cosine fitting. We built a linear model as described (*4*) using cos(ωt) and –sin(ωt) as regressors (with 13 degrees of freedom). In short, a cosine curve can be represented as:

$$C * \cos(\omega t + \varphi) = A * \cos(\omega t) - B * \sin(\omega t)$$

With $A = C * \cos(\varphi)$

and $B = C * \sin(\varphi)$

From the linear regression ('lm' function of the package 'stats' in R) we obtained the coefficients A and B, and their standard errors. A and B represent the phase and the amplitude of the oscillation:

$$phase = \arctan(A, B)$$

$$amplitude = \sqrt{A^2 + B^2}$$

As the density of the genes strongly decreased around 0.5 (Supplemental Figure 1A, B) we used amplitude ≥ 0.5 as a first classifier. We propagated the standard error of the coefficients A and B to the amplitude using Taylor expansion in the 'propagate' function (expr=expression(sqrt(((A^2)+(B^2))), ntype = 'stat', do.sim=FALSE, alpha=0.01) from the package 'propagate' (version 1.0-6) (*34*) in R. We obtained a 99% confidence interval (99%-CI) for each gene. As 99%-CI that does not include 0 is significant (p-value=0.01), we used the lower boundary (0.5%) of the CI as a second classifier. Thus, we classified genes with an amplitude ≥ 0.5 and lower CI-boundary ≥ 0 as 'high-confidence-oscillating' and genes with an amplitude < 0.5 or a lower CI-boundary < 0 were classified as 'non-oscillating' (Supplemental Figure 1A, B). Note that the 'non-oscillating' genes should not be considered as genes that were each demonstrated to lack oscillatory expression. Rather, this class may also contain genes for which strong evidence for high-amplitude oscillations is lacking, but that still exhibit low-amplitude oscillations. While we cannot, in the absence of further data, comment on the biological relevance of such potential low-level oscillations, the classification of only genes with more robust and extensive oscillations into the 'high-confidence oscillating genes' group will be beneficial to functional dissection of the oscillator. Compared to our previous annotation of 2,718 oscillating genes (18.9% of total expressed genes) in mRNA expression data of L3 and L4 animals (*4*), we can now detect an additional 1,240 genes as oscillating. We also fail to confirm oscillations for 219 of the previously annotated genes, which we consider to be most likely false

positives from the earlier analysis, resulting from the fact that some genes behave substantially different during L4 compared to previous stages (Figure 1D, E).
To quantify the oscillation amplitude for each larval stage, we split the time course in 4 separate parts, roughly corresponding to the developmental stages, i.e. L1: 7h-14h, L2: 14h-21h, L3: 21h-28h and L4: 28h-36h developmental time. We applied cosine fitting as described above to the 'high-confidence-oscillating' genes in the truncated time courses, with a fixed period of 7 h for L1-L3 and 8.5 h for L4 as determined by quantification of the oscillation period (Figure 1D, E). We applied a linear regression using the function 'lm' of the package 'stats' in R to find the relationship between the amplitudes across different stages, i.e. the slope. The correlation coefficient, r, was determined using the 'cor' function (method=pearson) of the package 'stats' in R.
To compare the peak phases of the L1-YA time course with those of the previously published L3-YA time course (*4*), we calculated the phase difference (L1–YA time course – L3–YA time course) (Supplemental Figure 1D, E). We added 360° to the difference and used the modulus operator (%%360), to maintain the circularity within the data. The coefficient of determination, $R^2$, was calculated by 1-(SSres/SStot), in which the SStot (total sum of squares) is the sum of squares in peak phase of the L1-YA time course. SSres (response sum of squares) is the sum of squares of the phase difference.

Time course fusion

In order to obtain an RNAseq time course spanning the complete larval development, we fused the L1–L2 time course (TP1 – TP15) with the L1–YA course (TP5 – TP48). To decide on which time points to choose from the individual time courses, we correlated both time courses against each other. In general, we saw good correlation between the two time courses, e.g. $TP1_{(L1–L2)}$ correlated well with $TP1_{(L1–YA)}$ etc. Additionally, we could see comparable correlation values of $TP13_{(L1–YA)}$ and $TP13_{(L1-L2)}$ with $TP1–5_{(L1–L2)}$. We concluded that $TP13_{(L1–L2)}$ and $TP13_{(L1–YA)}$ are highly similar and thus fused at this time point, i.e., combined $TP1–TP13_{(L1–L2)}$ with $TP14–TP48_{(L1–YA)}$.

Correlation analyses of RNAseq data

$Log_2$-transformed data was filtered for oscillating genes and then plotted in a correlation matrix using the R command cor(data, method="pearson"). The correlation line plots represent the correlations of selected time points to the fused full developmental time course (Supplemental Figure 6) and are specified in the line plot.
To reveal the highest correlations to chosen time points, we analyzed the correlation line of the chosen time points between TP7 and TP36 (the time in which oscillations occur) using a spline analysis from Scipy (*35*) in python ("from scipy.interpolate import InterpolatedUnivariateSpline" with k=4) and stored the splines as variable "spline". We identified peaks of the correlation line by finding the zeros of the derivative of the spline (cr_points = spline.derivative().roots()). The highest correlations of the respective correlation line were thus the value of the spline at the time point where the spline derivative was zero and the value was above the mean of the correlation line (cr_vals = spline(cr_pts) followed by pos_index = np.argwhere(cr_vals>np.mean(data.iloc[i])) and peak_val = cr_vals[pos_index]). Thus, we identified the correlation of particular time points (e.g. TP14–TP19) with their corresponding time points in the next oscillation cycle. Thereby, we were able to identify cycle time points as described in the results section. We defined the first cycle time point, e.g. TP14 of cycle 2, as 0°,

and the last unique one, TP19, as 300°. TP20 (360° of cycle 2) is also 0° of cycle 3. Note that a sampling interval of 1 hour can mean that a TP in one cycle correlates equally well to two TPs in another cycle, as seen for instance in the correlation of TP13 to TP26 and TP27. The spline interpolation places the peak of correlation in the middle of these time points at ~TP26.5. The spline analysis thus annotates cycle points correctly even in C4 which has an extended period. We performed correlation analyses without mean normalization of expression data, hence correlation values cannot be negative but remain between 0 and 1. We made this decision because a correlation analysis using mean-centered data, where correlations can vary between -1 and +1, requires specific assumptions on which time points to include or exclude for mean normalization, and because it is sensitive to gene expression trends. However, we confirmed, as a proof of principle, the expected negative correlation of time points that are in antiphase when using mean-centered data (Supplemental Figure 9).

GO-term analysis

GO-term analysis was performed using the GO biological process complete option (GO ontology database, release 2019-02-02) from the online tool PANTHER (*36*) (overrepresentation test, release 2019-03-08, standard settings).

Quantification of oscillation period

We quantified the oscillation period using mean-normalized $\log_2$-transformed gene expression levels of 'high-confidence-oscillating' genes from t = 5 h until t = 38 h of developmental time. We excluded t = 39 h until t = 48 h from the analysis, because oscillations cease in the last part of the developmental time course. We used the 'findpeaks' function (with nups=2, ndowns=2) of the package 'pracma' (version 2.1.5) (*37*) in R to call the time point of the expression peaks and selected only those genes for which 4 peaks were found. Peak-to-peak distance (peak 1-2, peak 2-3 and peak 3-4) was calculated by subtracting two subsequent peak times (Supplemental Figure 2).

For a temporally resolved quantification of the oscillation period, we filtered the mean-normalized $\log_2$ transformed gene expression levels of 'high-confidence-oscillating' genes (t=5h until t=38h) using a Butterworth filter ('bwfilter' function of the package 'seewave' (version 2.1.0) (*38*) in R, to remove noise and trend-correct the data. The following command was used to perform the filtering: bwfilter(data, f = 1 , n = 1, from = 0.1, to = 0.2, bandpass = TRUE, listen = FALSE, output = "matrix"). The bandpass frequency from 0.1 to 0.2 (corresponding to 10 hour and 5 hour period respectively) was selected based on the fourier spectrum obtained after fourier transform ('fft' function with standard parameters of the package 'stats'). As an input for the Hilbert transform we used the butterworth-filtered gene expression. The 'ifreq' function (with standard parameters from the package 'seewave') was used to calculate the instantaneous phase and frequency based on the Hilbert transform. To determine the phase progression over time we unwrapped the instantaneous phase (ranging from 0 to $2\pi$ for each oscillation) using the 'unwrap' function of the package 'EMD' (version 1.5.7) (*39*) in R. To avoid edge effects, we removed the first 4 and last 3 data points of the unwrapped phase. The angular velocity is defined as the rate of phase change, which we calculated by taking the derivative of the unwrapped phase. The instantaneous period was determined by $2\pi$/angular velocity and was plotted for each gene individually and as mean in a density plot. The mean of the instantaneous period over all 'high confidence' oscillation genes was used to reconstruct a 'global' oscillation by taking the

following command: sin(cumsum(mean angular velocity)) and plotted together with a 7h-period oscillation and the mean normalized expression of a representative gene, *col-147*.

Identification of first gene expression peaks in L1 larvae

To identify the first peak of oscillating genes, we used a spline analysis in Python ("from scipy.interpolate import InterpolatedUnivariateSpline") from TP3 – TP13. We chose these time points to remove false positives in the beginning due to slightly higher noise for the first 2 time points as well as not to identify the second peak which occurred at ≥TP14 for some very early genes. The function used was "InterpolatedUnivariateSpline" with k=4. After constructing the spline, we identified the zeros of the derivative and chose the time point value with the highest expression value and a zero derivative as the first peak time point.

Embryonic gene expression time course

Embryonic gene expression data was obtained from (*19*), and represented precisely staged single embryos at 10 min intervals from the 4-cell stage up to muscle movement and every 10-70 min thereafter until 830 minutes. We obtained the gene count data from the Gene Expression Omnibus data base under the accession number GSE50548, for which sequencing reads were mapped to WBCel215 genome and counted against WS230 annotation.
We normalized the gene counts to the total mapped library size per sample, added a pseudocount of 8, and $\log_2$-transformed the data. We selected genes according to their being larval 'high-confidence-oscillating' genes, and plotted their embryonic expression patterns according to peak phase in larvae. The embryonic time course was correlated to the fused larval time course using the 'cor' function (method='pearson') of the package 'stats' in R. Correlation line plots were generated by plotting the correlation coefficients for each embryonic time point over larval time. To identify the peaks of the correlation lines with a resolution higher than the sampling frequency, we interpolated the correlation lines using the 'spline' function (n=240, method='fmm') of the package 'stats' in R. To call the peaks of the interpolated correlation lines, we applied the 'findpeaks' function (with nups=5, ndowns=5) of the package 'pracma' on the time points on the interpolated time points 10-185, that cover the four cycles. To find the embryonic time point at which oscillations initiate, we plotted the larval TP in cycle 2 at which the correlation peak occurred over embryonic time (Supplemental Figure 7B) and determined the intersection of the two linear fits, using the 'solve' function of the package 'Matrix' (version 1.2-15) (*40*) and the 'lm' function of the package 'stats' in R respectively. To determine the 95%-CI of the x-coordinate of the intersect, the standard error of the slope a and the intercept b of the two linear fits was propagated using Taylor expansion in the 'propagate' function (expr = expression((b1-b2)/(a2-a1)), ntype = "stat",do.sim = FALSE, alpha=0.05) from the package 'propagate' in R. The pairwise correlation map was generated with the 'aheatmap' function of the package 'NMF' (version 0.21.0) (*41*) and the 3D plot was generated with the '3Dscatter' function of the package 'plot3D' (version 1.1.1) (*42*) in R.

Time-lapse imaging of single animals

Single worm imaging was done by adapting a previous protocol (*17*), and is similar to the method reported in (*43*). Specifically, we replaced the previous 3.5-cm dishes with a "sandwich-like" system in which consists of 4.5% agarose in S-basal with chambers and worms on top of a glass cover slip, which covered the chambers. Two silicon-Isolators (GRACE Bio-Labs, SKU: 666103) with a hole in the middle were then placed on top of each other and glued on the glass

cover slip to surround the agarose with chambers. Low melt agarose (3% in S-basal) was used to seal the agarose chamber which prevented drifts of the agarose chambers during imaging. The sandwich-like system was then covered with a glass slide. We used a 2x sCMOS camera model (T2) CSU_W1 Yokogawa microscope with 20x air objective, NA = 0.8 in combination with a 50µm disk unit to obtain images of single worms. For a high throughput, we motorized the stage positioning and the exchange between confocal and brightfield (a red LED light used to combine with fluorescence without closing shutter). Additionally, we used a motorized z-drive with 2 µm step size and 23 images per z-stack. The 488nm laser power for GFP imaging was set to 70% and a binning of 2 was used.

To facilitate detection of transgene expression and oscillation, we generated reporters using the promoters of genes that exhibited high transcript levels and amplitudes, and where GFP was concentrated in the nucleus and destabilized through fusion to PEST::H2B (see strain list above). We placed embryos into chambers containing food (concentrated bacteria HT115 with L4440 vector) and imaged every worm with a z-stack in time intervals of 10 min during larval development in a room kept at ~21°C, using a double camera setting to acquire brightfield images in parallel with the fluorescent images. We exploited the availability of matching fluorescent and brightfield images to identify worms by machine learning. After identification, we flattened the worm at each time point to a single pixel line and stacked all time points from left to right, resulting in one kymograph image per worm. We then plotted background-subtracted GFP intensity values from the time of hatch (t = 0 h), which we identified by visual inspection of the brightfield images as the first time point when the worm exited the egg shell. Time lapse images were analyzed using a customized KNIME workflow (Supplemental File 1). We analyzed every worm over time using the same algorithm. First, we identified the brightest focal planes per time point by calculating the mean intensity from all focal planes per time point and selecting the focal planes that had a higher intensity than the mean. Then we maximum-projected the GFP images over Z per time point and blurred the DIC image and also max projected over Z (blurring the DIC improved the machine learning process later on). All images per worm over time were analyzed by Ilastik machine learning in order to identify the worm in the image. The probability map from Ilastik was used to select a threshold that selected worms of a particular experiment best. (The threshold might change slightly as DIC images can look slightly different due to differences in the sample prep amongst experiments.) Using a customized ImageJ plugin, we straightened the worm. The straightened GFP worm image was then max projected over Y which resulted in a single pixel line representing the GFP intensities in a worm and after stacking up all the single pixel lines in Y direction, we obtained the kymographs. In order to remove noise coming from the head and tail regions of the worm due to inaccuracy of the machine learning, we measured mean GFP intensities per time point ranging from 20% until 80% of the worms anterior – posterior axis. For background subtraction we exploited the fact that only the nuclei were GFP positive and thus subtracted the minimum intensity value between GFP nuclei from their intensity values.

After the KNIME workflow, we imported the measured GFP intensities into Python and analyzed the traces using a butterworth filter and Hilbert transform analysis (both from Scipy (*35*)). We used the butterworth bandpass filter using b, a = butter(order =1, [low,high], btype="band") with low=1/14 and high=1/5, corresponding to 14 hour and 5 hour periods respectively. We then filtered using filtfilt(b, a, data, padtype='constant') to linearly filter backwards and forwards.

For individual time points where the worm could not be identified by the Ilastik machine learning algorithm, we linearly interpolated (using interpolation from pandas (*44*)) using "pandas.series.interpolate(method = 'linear', axis =0, limit = 60, limit_direction = 'backward'", between the neighboring time points to obtain a continuous time series needed for the Hilbert transform analysis. Using Hilbert transform, we extracted the phase of the oscillating traces for each time point and specifically investigated the phase at molt entry and molt exit for our different reporter strains.

In order to determine time points in which worms are in lethargus, we investigated pumping behavior. As the Z-stack of individual time points give a short representation of a moving worm, it is possible to determine whether animals pump (feeding, corresponds to intermolt) or not (lethargus / molt). Additionally to the pumping behavior, we also used two additional requirements that needed to be true in order to assign the lethargus time span: First, worms needed to be quiescent (not moving, and straight line) and second, a cuticle needed to be shed at the end of lethargus. Usually worms start pumping one to two time points before they shed the cuticle. This analysis was done manually with the software ImageJ, and results were recorded in an excel file, where for every time point, the worms' behavior was denoted as 1 for pumping and as 0 for non-pumping.

To determine a possible connection between oscillations and development, we applied error propagation, assuming normal distribution of the measured phases and larval stage durations. Thereby, we exploited the inherent variation of the oscillation periods and developmental rates among worms, rather than experimental perturbation, to probe for such a connection. We define the phase $\theta$ at either molt exit or entry as $\theta \equiv \frac{2\pi}{T_o} * T_d \sim (\mu, \sigma^2)$ with $T_o \sim (\mu_o, \sigma_o^2)$ being the period of oscillation and $T_d \sim (\mu_d, \sigma_d^2)$ the intermolt duration (for phase at molt entry) or larval stage duration (for phase at molt exit), resulting in a phase with mean $\mu$ and a standard deviation $\sigma$. Should the two processes be coupled as in scenario 2, we would expect

$\sigma_{observed} < \sigma_{calculated}$.

To calculate the phase at molt entry and molt exit with error propagation we used the "uncertainties" package (*45*) in python. The larval stage duration as well as intermolt duration and period were treated as ufloat numbers, representing the distributions coming from our measurement (e.g. 7.5 +/- 0.2). These distributions were then used to calculate the expected phase at molt entry (using the intermolt duration) and molt exit (using the larval stage duration) using: $phase(err.prop) = \frac{2*pi}{period} * larval\ stage$ . This resulted in the phase being represented by a ufloat number and thus a distribution which we used for plotting after normalizing for the mean to compare the variation of the data.

<u>Synthetic oscillation with a long "L4" period</u>

We generated oscillatory signals using chirp form Python's "scipy.signal": "chirp(t2, f0=1/7, f1=1/9, t1=5, method='linear')" for mimicking the experimental data and "chirp(t2, f0=1/7, f1=1/12, t1=7, method='linear')" for the more extreme last oscillation lengthening. We processed the 1000 generated traces in the same way as the experimental data before using spline interpolation and peak calling in python and plotted the results.

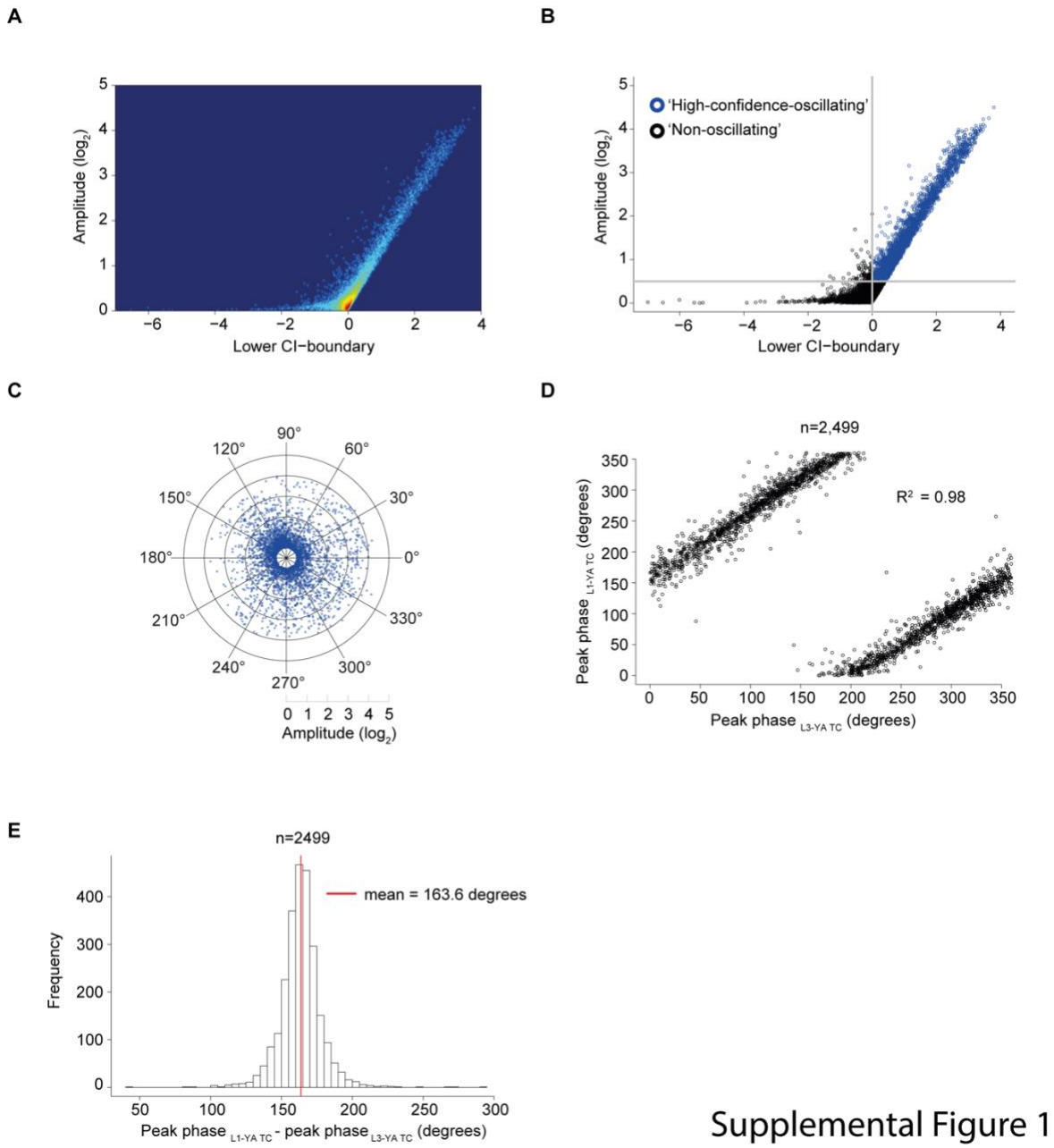


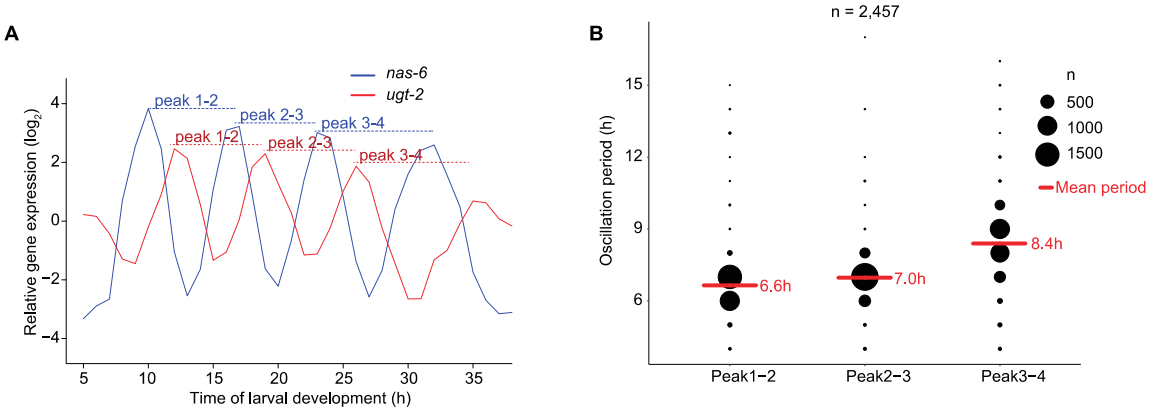

**Fig. S1. Identification of 3,739 'high-confidence-oscillating' genes**

To classify genes as 'oscillating', we used cosine fitting with a fixed period of seven hours (the stable period from mid-L1 through late L3 stage, Figure 1D, E) and two parameters that describe a sine and a cosine respectively to characterize oscillating genes of any peak phase and amplitude (*9*). To identify 'high-confidence-oscillating genes', we determined the oscillating amplitude with a 99% confidence interval (99%-CI).

**(A)** Smooth scatter of amplitude over lower boundary of 99% confidence interval of the amplitude.

**(B)** A lower CI-boundary $\geq 0$, i.e. p-value $\leq 0.01$, and a $\log_2$(amplitude) $\geq 0.5$, which corresponds to a 2-fold change from peak to trough, were used as cut-offs to identify 3,739 'high-confidence-

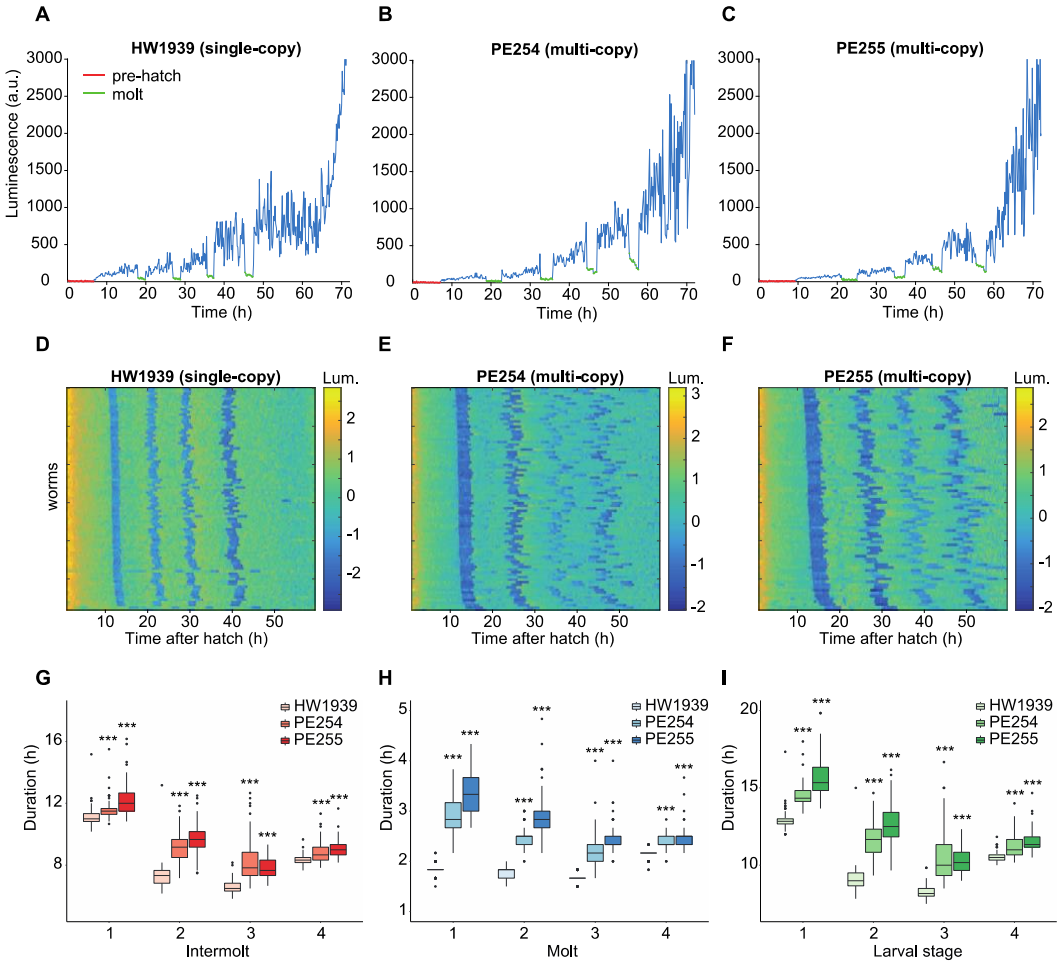
oscillating' genes (24% of total expressed genes), indicated in blue. Other genes for which we were less confident that they oscillate (lower CI-boundary < 0, not significant) or for which we determined a low amplitude (<0.5) were classified as 'non-oscillating' (black).
**(C)** Radar chart plotting amplitude (radial axis, in $\log_2$) over peak phase (circular axis, in degrees) as determined by cosine fitting.
**(D, E)** Scatterplot (D) of the peak phase of the L1 – young adult (YA) developmental time course described here over the previously published L3-YA time course (*4*). Genes that were identified as 'oscillating' in both time courses (n = 2,499) are shown. Peak phases correlate well as confirmed by the coefficient of determination, $R^2$, as indicated. However, they differ systematically (E) because a peak phase of 0° is arbitrarily chosen. A red vertical line indicates the mean phase difference (L1-YA time course – L3-YA time course; corrected for circularity as described (*9*)). Note that the gene-specific peak phase calculated here and previously both also differ from the arbitrarily assigned cycle phases in Figures 4 and 5 and their discussion.




Supplemental Figure 2

**Fig. S2. Oscillation period is constant in L1-L3 stage and increases in L4 stage**

**(A)** Mean-normalized expression (in $\log_2$) of two oscillating genes illustrates peak calling. We identified the peaks of individual 'high-confidence-oscillating' genes by the 'findpeaks' algorithm and used the peak-to-peak distance as a proxy for the oscillation period.
**(B)** Bubble chart showing oscillation period as determined by peak calling. Oscillating genes for which 4 peaks were called are shown (n = 2,457). Red line indicates mean period. Although limited by the sampling frequency and the ability to call four peaks, an increase in the period of the last oscillation cycle relative to earlier cycles is readily observed.




J

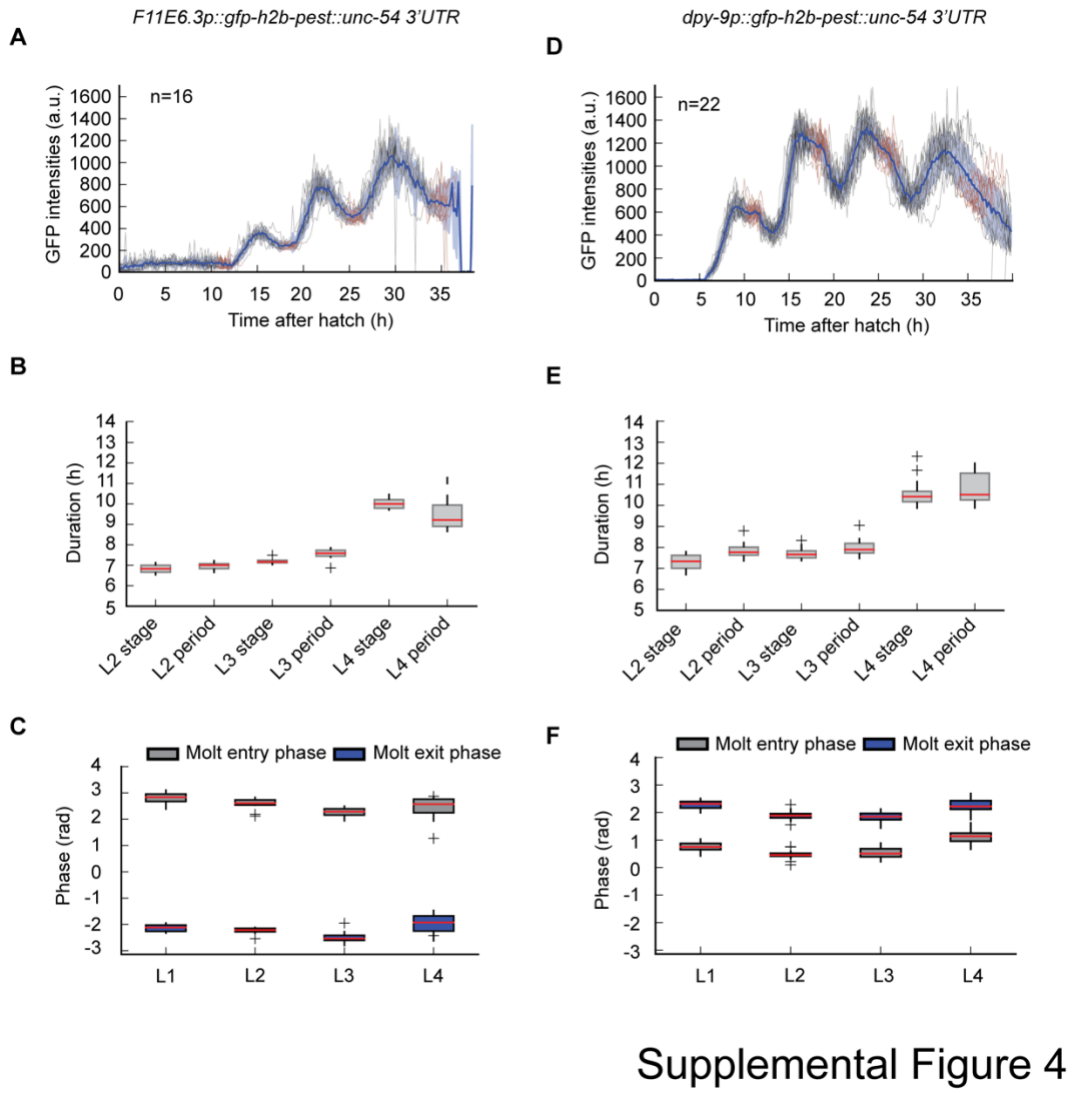
| | stage 1 | stage 2 | stage 3 | stage 4 |
|---|---|---|---|---|
| Intermolt PE254 | 1.04 (p=3.28E-03) | 1.26 (p=4.51E-31) | 1.26 (p=3.07E-19) | 1.05 (p=1.94E-08) |
| Intermolt PE255 | 1.10 (p=1.98E-10) | 1.32 (p=1.39E-35) | 1.19 (p=3.72E-30) | 1.08 (p=2.10E-16) |
| Molt PE254 | 1.60 (p=4.92E-51) | 1.40 (p=9.60E-54) | 1.33 (p=2.11E-29) | 1.11 (p=8.68E-23) |
| Molt PE255 | 1.84 (p=2.49E-50) | 1.65 (p=3.53E-37) | 1.46 (p=2.33E-38) | 1.15 (p=9.74E-21) |
| Larval stage PE254 | 1.11 (p=5.20E-20) | 1.28 (p=2.35E-40) | 1.27 (p=1.99E-22) | 1.07 (p=4.32E-14) |
| Larval stage PE255 | 1.20 (p=6.93E-31) | 1.38 (p=2.97E-42) | 1.25 (p=3.97E-40) | 1.10 (p=8.24E-22) |

Supplemental Figure 3

**Fig. S3. A strain with single-copy integrated luciferase transgene develops rapidly and synchronously**

**(A-C)** Representative raw luminescence traces of individual animals grown at 20°C. As the egg-shell is impenetrable to luciferin, a sudden increase in luminescence at the beginning of the time course indicates hatch (pre-hatch in red). Abrupt drops and subsequent rises in luminescence specify molts (in green). The previously published strains (B, PE254; C, PE255) express luciferase from randomly integrated multi-copy transgene arrays that carry a semi-dominant version of the cuticular collagen *rol-6* as a marker (*27*). To exclude that this genetic make-up could interfere with our quantification, we integrated a luciferase transgene, driven by the strong,

ubiquitous and constitutive *eft-3* promoter, into the genome through Mos1-mediated single copy integration (MosSCI) (A, HW1939).
**(D-F)** Heatmap per strain showing trend-corrected luminescence (Lum.) trace for one animal per horizontal line (D, Single-copy integrated HW1939 (n=86). E, Multi-copy integrated PE254 (n=88). F, Multi-copy integrated PE255 (n=79)). Hatch is set to t = 0 h and traces are sorted by time of entry into first molt. Blue indicates low luminescence and corresponds to the molts.
**(G-I)** Quantification of the duration of each intermolt (G), molt (H), larval stage I) for indicated strains in hours. The newly generated strain developed more rapidly and with less variability with regard to the duration of individual stages. Although the general trend in larval stage durations was shared between the different strains, i.e. L1>L4>L2&L3 (I), animals carrying the *rol-6*-marked multi-copy luciferase arrays also exhibited an extended M1 molt (H) as reported previously (*16*). This effect disappeared when using the single-copy transgene strain. Hence, the duration of molt M1 became comparable to that of M2 and M3 and lengthening of L1 is explained by lengthening of intermolt 1. Significant differences between single-copy integrated (n=86) and multi-copy integrated strains (PE254 (n=88) and PE255 (n=79)) is indicated (*** P<<0.001, Welch two sample, two-sided t-test). Boxplots extend from first to third quartile with a line at the median, outliers are indicated with a cross, whiskers show 1.5*IQR.
**(J)** Table showing fold changes of mean durations of indicated stages for PE254 and PE255 compared to HW1939 for data shown in G-I. P-values are indicated in brackets (Welch two-sample, two-sided t-test).




Supplemental Figure 4

**Fig. S4. Single worm imaging with two additional reporter strains confirms phase-locking of oscillations to molts independently of peak phases**

**(A, D)** GFP quantification of single worm kymographs for the *F11E6.3* (HW1370, n=16) and the *dpy-9* (HW2526, n=22) transcriptional reporters respectively. All traces were aligned to the time of hatching, which was set to t = 0 h. Segments in red indicate lethargus while the blue shading indicates the standard deviation at each time point with the blue line representing the mean across worm. Only three peaks are visible for the *F11E6.3* reporter, because the assay terminated before the final rise in expression seen with RNA sequencing.

**(B, E)** Comparison of larval stage duration and period times of oscillations in hours for L2-L4 larval stages for *F11E6.3* and *dpy-9* transcriptional reporters respectively.

**(C, F)** Boxplot of expression phases at molt entry (start of lethargus) and molt exit (end of lethargus) separated by larval stages; n = 16 for *F11E6.3* (D) and n = 22 for *dpy-9* (I) transcriptional reporters.

Boxplots extend from first to third quartile with a line at the median, outliers are indicated with a cross, whiskers show 1.5*IQR.

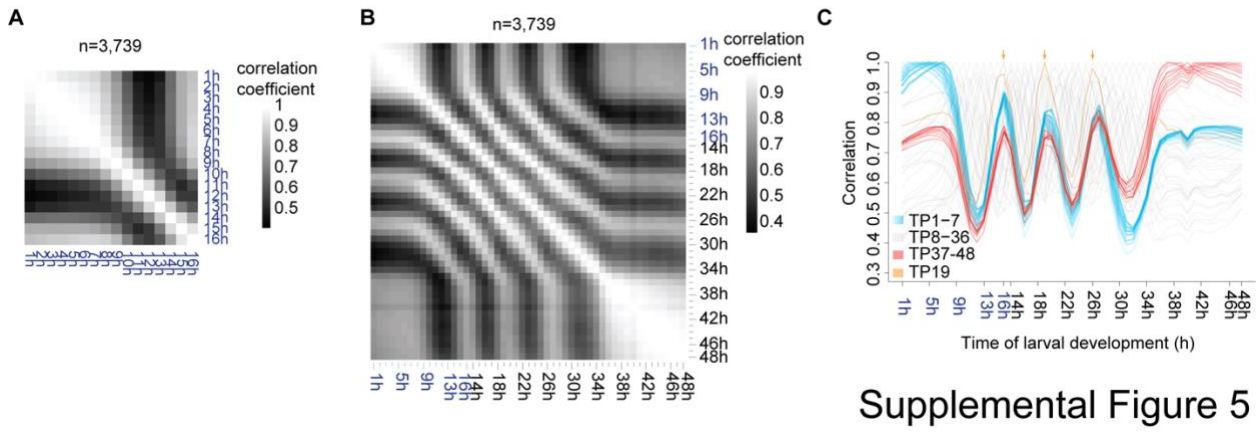



**Fig. S5. Time lag and oscillation initiation phase reproduced in an independent sequencing experiment**

**(A)** Pairwise correlation plot of $\log_2$-transformed expression patterns of 'high-confidence oscillating' genes (n=3,739; identified in Figure 1) following resumption of development after L1 arrest reveals a similar time lag before oscillations start as in L1 in Figure 2.
**(B)** Pairwise-correlation plot of oscillatory gene expression patterns obtained from synchronized population of the replicate L1 stage larvae at (TP1–16 hrs, blue label) combined with samples obtained in Figure 1 (TP14–48 hrs, black label), as in Figure 4A. A lower incubation temperature for this particular experiment resulted in slowed development as seen from an extended time-lag at the beginning and an equivalence of TP16 to TP13 in the fused time course in Figure 4A.
**(C)** Correlations of expression patterns for the indicated time points to all other time points of the fused time course shown in B; label colors as in B. Consistent with the observations in B, TP19 correlates highly with TP16 and TP26 (arrows), and with early time points of the replicate experiments (blue lines, TP1–7) time points. Hence, oscillations are arrested in the same phase at the beginning of the two replicate time courses shown here and in Fig. 5.

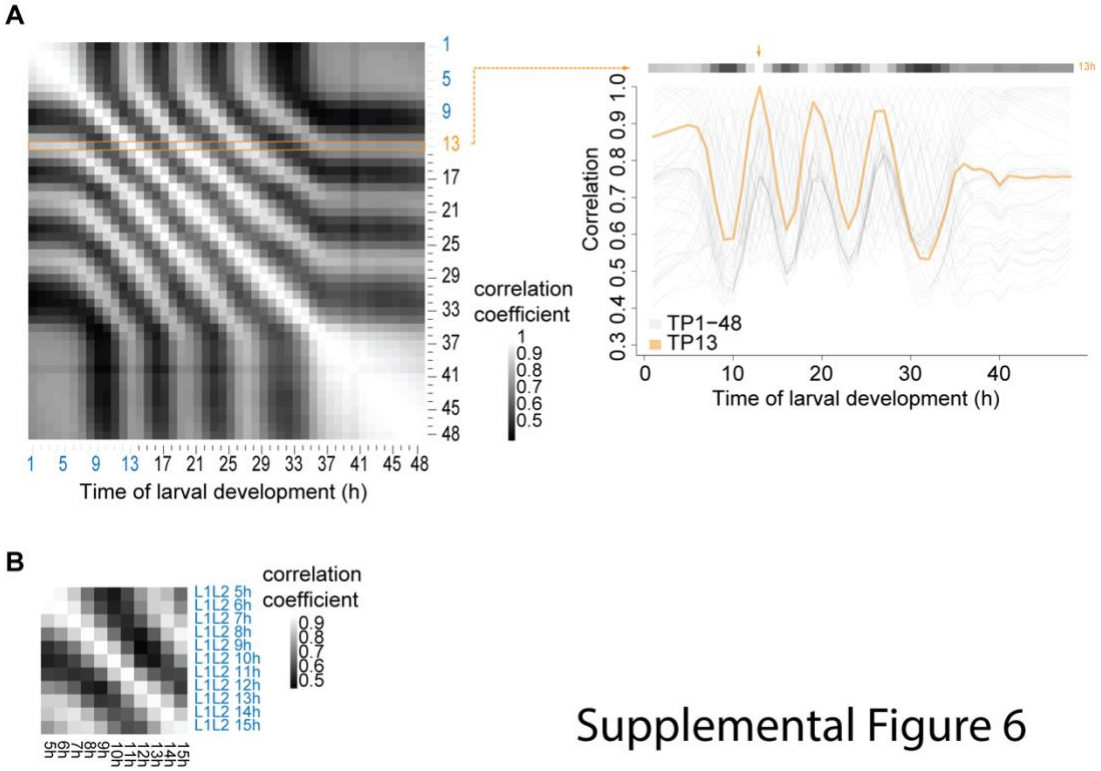



**Fig. S6. Correlation line explanation and validity of time course fusion**

**(A)** Left: Pairwise-correlation plot of oscillatory gene expression patterns obtained from synchronized population of L1 stage larvae at 25°C (TP1–13) combined with samples obtained in Figure 1 (TP14–48), as in Figure 4A. Right: The correlation of TP13 versus all other time points is plotted as a line (orange), correlation with itself at 13 hrs is 1.0 (orange arrow).
**(B)** Pairwise-correlation plot of oscillatory gene expression patterns for L1–L2 time course (TP5-15) and the overlapping time points of the L1–YA time course (Figure 1) confirms extensive similarity of equivalent time points across the two time courses.

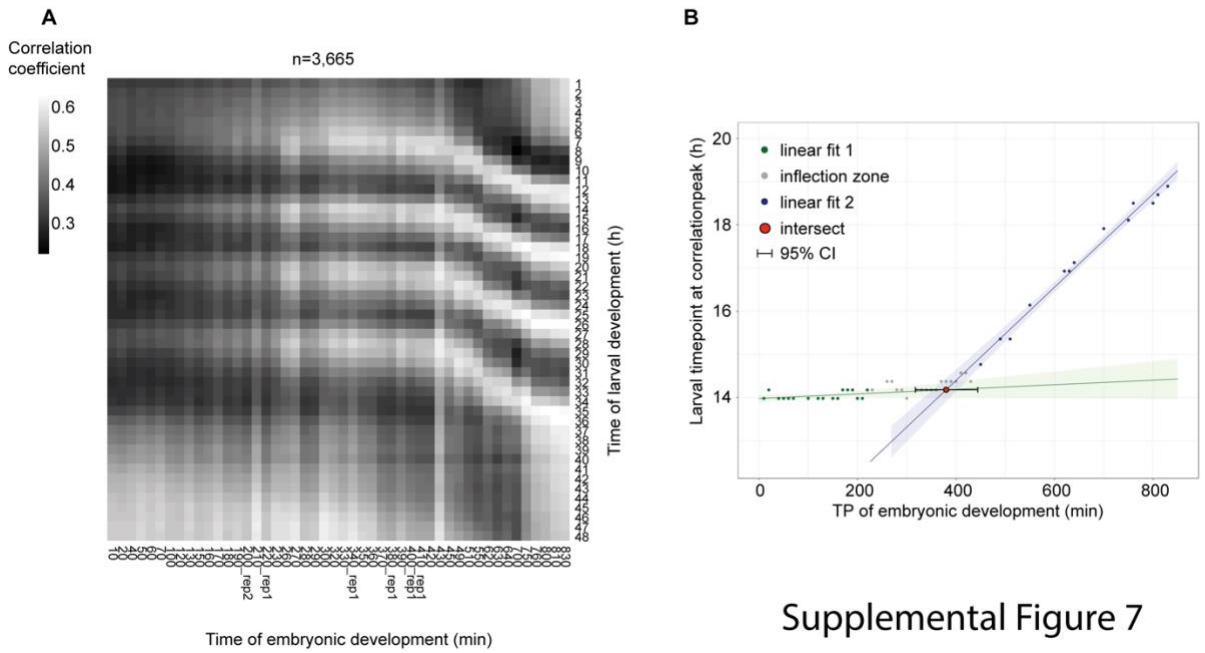
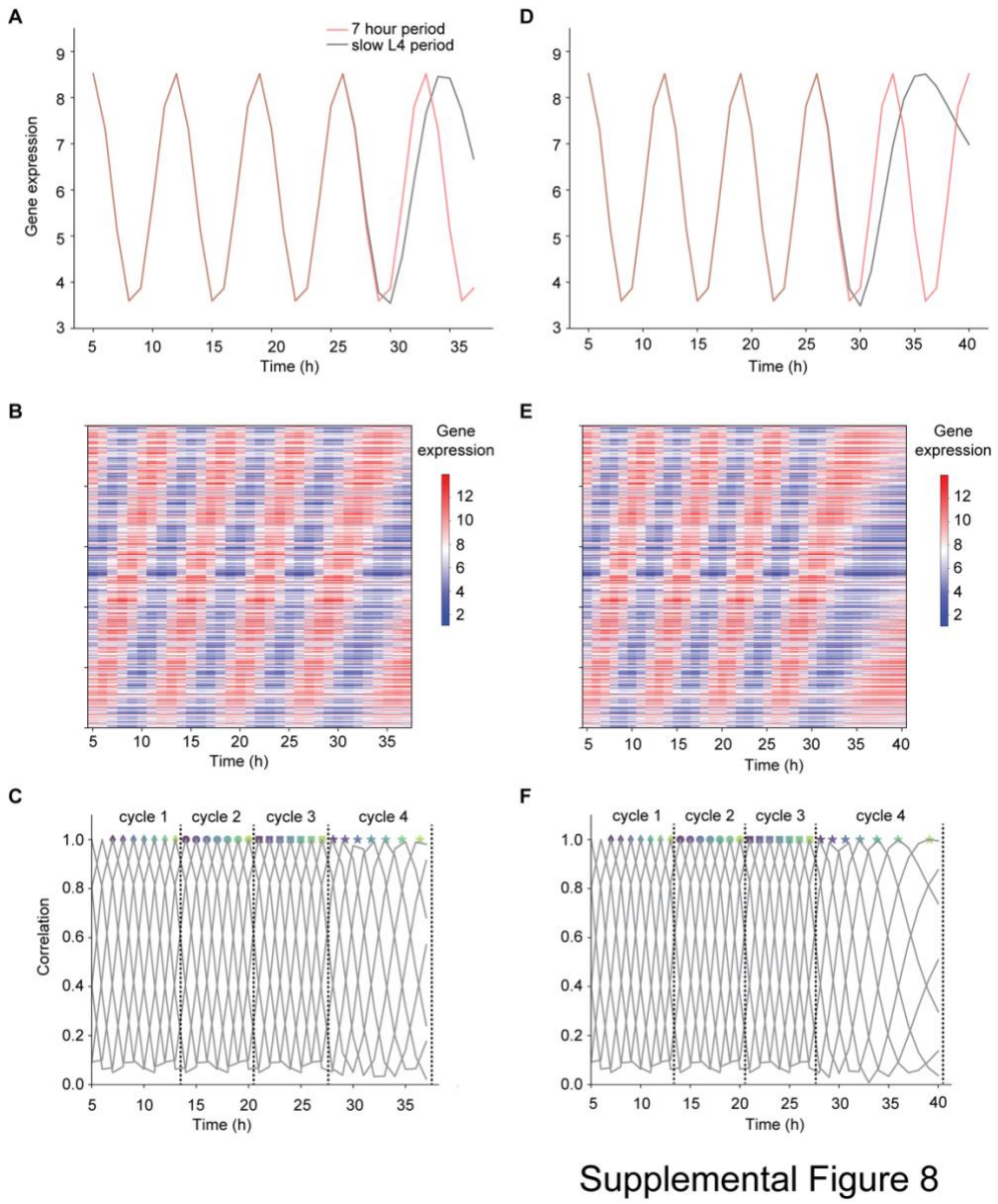



**Fig. S7. Initiation of oscillations during mid-embryogenesis**

**(A)** Pairwise correlation map of $\log_2$-transformed gene counts of fused larval time course to embryonic time course (*19*).

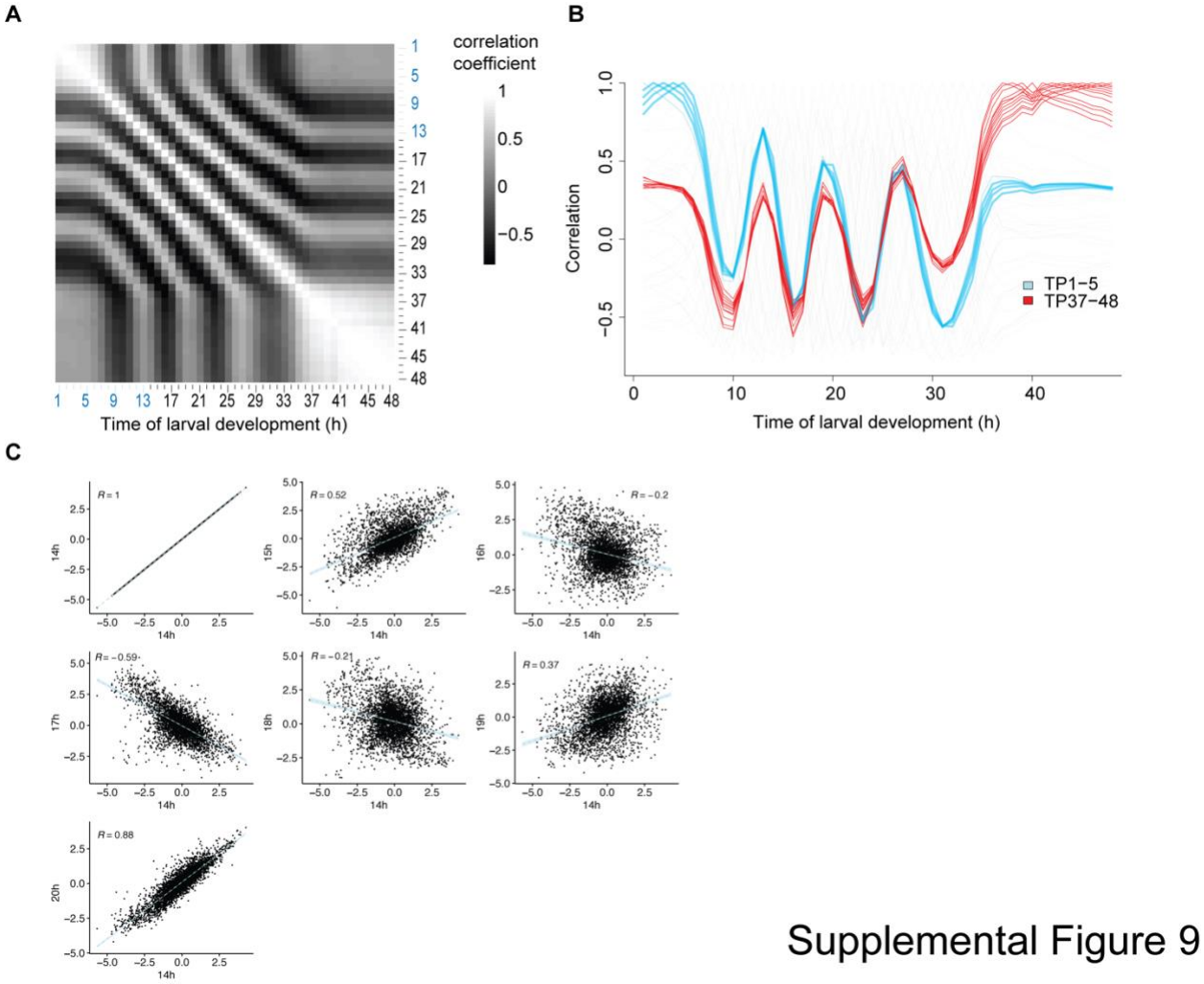
**(B)** Scatter plot showing the larval time point of the larval oscillation cycle 2 (Figure 4D,E) for each embryonic time point. The larval time point of the peak was determined after spline interpolation (*9*). Linear model 1 ($y = 5.312e\text{-}04*x + 13.98$, $p=0.098$, $R^2 = 0.162$, 16 degrees of freedom) was fitted to the data of embryonic TP10-TP230 min (in green) and linear model 2 ($y = 0.0108*x + 10.07$, $p=2.08e\text{-}11$, $R^2 = 0.985$, 11 degrees of freedom) was fitted to the data of embryonic TP450-TP830min (in blue). The embryonic time at the intersection (in red, 380.0 min (95%-CI 317.6 min – 444.2 min)) of the linear models was determined in the inflection zone, i.e. points (in grey) not used for model 1 or model 2 fit, and the 95% CI was determined by propagating the standard errors of the coefficients of the linear models (*9*).




Supplemental Figure 8

**Fig. S8. Period extension does not lead to decreasing correlation**

Correlation analysis with synthetic data. Gene expression oscillation traces were synthetically generated to show oscillations with a 7-h period followed by one oscillation with period increased to 9 hours (A-C), similar to experimental data, or yet further to 12 hours (D–F).
**(A, D)** Example traces of synthetically designed gene expression with increased period length for the last cycle (equivalent to L4). To compare with an unchanged period, the trace with continuous 7-h period is shown in orange.
**(B, E)** Heatmap of 1’000 oscillating traces that show increased period in the last oscillation. Amplitudes, phases and mean expression levels were varied for individual traces.
**(C, F)** Correlation analysis of time points 7 – 13 with all subsequent time points using spline interpolation and peak calling. The peak of correlation of corresponding cycle points has a value close to 1 even when period length increases.




Supplemental Figure 9

**Fig. S9. Correlation analysis with mean normalized data yields qualitatively similar results to an analysis without mean normalization**

**(A)** Pairwise correlation plot of $\log_2$-transformed, mean normalized oscillatory gene expression patterns for the fused developmental time course (Figure 4).
**(B)** Correlation line plots reveal repetitive similarity of TP1-5 to TP13, TP19 and TP26/27. Due to mean normalization, the correlation lines oscillates around 0.
**(C)** Scatter plot comparing $\log_2$-transformed, mean-normalized data of individual time points. The Spearman correlation coefficient is indicated on the left corner and can range from anti-correlation (-1) to full correlation (+1).