PNAS
www.pnas.org

Supplementary Information for

Temporal scaling in *C. elegans* larval development

Olga Filina[1], Burak Demirbas[1], Rik Haagmans[1] and Jeroen S. van Zon[1*]

[1]AMOLF, Science Park 104, 1098 XG Amsterdam, the Netherlands

*corresponding author: Jeroen S. van Zon.

Email: j.v.zon@amolf.nl

This PDF file includes:

Supplementary Text
Figures S1 to S7

Supplementary text

1. Strains and experimental conditions

To monitor *wrt-2::GFP* expression in *eat-2* and *lin-42* mutants, these mutations were crossed into *heIs63*. The *eat-2* genotype was confirmed by measuring the rate of pharyngeal pumping, which is decreased 5-fold compared to wild-type animals. As *lin-42* constitutes a complex genetic locus encoding multiple isoforms, we chose to use the *lin-42(ox461)* allele that deletes the entire locus of *lin-42.* In addition, the *lin-42(ok2385)* allele that deletes the main isoform *lin-42a* and the overlapping exons of *lin-42b*, was analyzed and showed similar phenotypes as *lin-42(ox461)* (data not shown). For maintenance, all strains were cultured at 20˚C on NGM (Nematode growth medium) agar plates seeded with OP50 strain of *E. coli* bacteria, using standard *C. elegans* techniques. For the *E. coli* HB101 diet experiment in Fig. 4, animals were maintained on HB101 for 5-7 generations prior to the experiment.

2. Time-lapse microscopy

Custom time-lapse microscopy setup was used to monitor the entire development of individual *C. elegans*. Late-stage embryos were placed inside the 250x250x20 µm polyacrylamide microchambers (one embryo per chamber) filled with *E. coli* OP50 bacteria as food source. Nikon Ti-E inverted microscope with a large chip camera (Hamamatsu sCMOS Orca v2) and a 40x magnification objective (Nikon CFI Plan Fluor 40x, NA=1.3, oil immersion) were used for imaging. Transmission and fluorescence images were acquired with an LED light source (CoolLED pE-100 615nm) and a 488 nm laser (Coherent OBIS LS 488-100), respectively. Each chamber containing a single animal was imaged every 20-40 minutes during the entire larval development (40-100 hours depending on the genotype and temperature). A stack of 20-30 images in the Z-direction was acquired using short exposure times (1-10 ms), such that the motion of the animal was insignificant.

3. Temperature control and condition shifts

All experiments were performed in a temperature-controlled room with a constant temperature inside the sample of 23˚C. To perform an experiment at different temperature, an additional temperature control system was used. A thermoelectric chiller (Thermotek T257P) was used to cool the custom-made objective jacket by circulating an antifreeze fluid (a mixture of water and glycerin) between the chiller and the objective jacket. In order to calibrate the system, a thermocouple temperature sensor measuring 0.025 mm in diameter (RS Pro) was placed inside the sample in contact with the polyacrylamide hydrogel and connected to a digital thermometer (RS Pro). The temperature was then varied on the thermoelectric chiller while the resulting temperature inside the sample was being monitored. In this work, experiments were performed at 23, 19 or 15

˚C. Temperature shifts were implemented by changing the temperature of the chiller mid-experiment. As animals in different microchambers hatch at slightly different times, the time after hatching at which the shift in temperature was initiated varied between animals. For shifts in food supply, we loaded each chamber with reduced amounts of OP50. Full depletion of food occurred at different times after hatching for each animal and we only analyzed data for animals that entered L4 arrest.

4. Image analysis

Time-lapse image stacks were processed with custom Python software in order to obtain the precise timing of ecdysis events, seam cell divisions, and peaks in oscillatory *wrt-2* expression. For every animal, the times of hatching and ecdysis events were annotated based on visual inspection of transmitted light images. Hatching was defined as the time larvae first appears out of the egg shell, while ecdysis events were defined as the first appearance of the shed cuticle in the chamber (Fig. 1C). Times of seam cell divisions were annotated based on visual inspection of the *wrt-2p::GFP* fluorescence signal in the nucleus and the membrane of seam cells. Divisions of V1-V6 seam cells occur close together in time. We therefore defined the time of each round of divisions as the average time V1-V6 cells had divided or had started dividing, as determined by the formation of the metaphase plate. We only analyzed seam cells located on the side of the body closest to the objective.

To determine the time of peaks in oscillatory *wrt-2* expression, we first obtained *wrt-2p::GFP* expression profiles as a function of time for individual animals. For this, in every time frame we automatically segmented the region encompassing seam cells using a Watershed algorithm and calculated the average fluorescence intensity inside this region. Finally, to find the time of each peak (μ_i), we fitted the obtained oscillatory profiles with a combination of Gaussian functions and a linear offset using non-linear least-squares minimization (Fig. 1B):

$$f(x, m, b, A_i, \mu_i, \sigma_i) = mx + b + \sum_{i=1}^{n} \left[\frac{A_i}{\sigma_i \sqrt{2\pi}} e^{\left[-\frac{(x-\mu_i)^2}{2\sigma_i^2} \right]} \right] \quad (1)$$

where *n* is the number of peaks, A_i, μ_i, σ_i are the amplitude, center and width of peak. Finally, we fitted the experimentally measured times for pairs of events a and b to a line function of the form $t_b = s_{a,b} \cdot t_a$ using non-linear least-squares minimization (Fig. 1G), using the Linear Model from the *lmfit* package in Python. To visualize *wrt-2p::GFP* expression dynamics, e.g. in Fig. 1E, we applied a Savitzky-Golay filter with window size of 2 h and polynomial order of 1 on the raw segmented data we show in Fig. 1B. To measure the animal's body length as a function of time (Fig. 5A), we manually annotated ~10 points along the anterior-posterior body axis and performed spline interpolation. Body length was defined as the length of the resulting spline curve.

5. Timing models and simulations

We model the progression of development as the evolution of a developmental phase ϕ, that increases from ϕ=0 (start of larval development at hatching) to ϕ=1 (entry into adulthood at the L4 ecdysis). The exact assignment of a phase to a particular developmental event is arbitrary. Here, we define the phase so that for standard conditions (wild-type animals at 23 ˚C) the phase increases linearly with time, $\frac{d\phi}{dt}$=const, and ϕ=1 at time t=T, where T is the total duration of development at standard conditions. As a result, for standard conditions we use the following definition for the developmental phase of event a:

$$\phi_a = \frac{t_a}{T} \tag{2}$$

For other conditions or genotypes, the time evolution $\phi(t)$ of the developmental phase has a different form. As a result, the time of event a, occurring at a developmental phase ϕ_a is given by $t_a = f(\phi_a)$, where f is a monotonically increasing function that is specific for each condition or genotype. Expressions for $f(\phi)$ are discussed further below.

To incorporate animal-to-animal variability, we assumed two different sources of variability. First, there is an intrinsic variability in the stage ϕ_a at which each event a occurs, that is uncorrelated between different events occurring within the same animal. Second, we assumed variability in the total duration of development, T. This corresponds to an animal-to-animal variability in the global rate of development that impacts each event occurring within the same animal equally. Then, the event time $t_{a,i}$ for event a in animal i is given by:

$$t_{a,i} = \frac{f\left(\phi_a + \eta_{\phi,i}\right)}{T}\left(T + \eta_{T,i}\right) \tag{3}$$

where T and ϕ_a correspond to the population average values, while η_T and η_ϕ are Gaussian noise sources with standard deviation σ_T and σ_ϕ, respectively.

The function $f(\phi)$ changes for differing environmental conditions or mutants that perturb the duration of development. In particular, we considered three different models, the ‘Uniform’, ‘Pause’, and ‘Rate change’ models (Fig. 3A). For the ‘Uniform’ model, event times are given by Eq. 2, but now with an increased duration of development T'. For the ‘Pause’ model, development occurs at the same rate as for standard conditions, but with a pause at developmental phase ϕ' that results in a total duration of development T'=$(1+\kappa)T$. This results in:

$$t_a = \begin{cases} T\phi_a & \phi_a < \phi' \\ T(\kappa + \phi_a) & \phi_a \geq \phi' \end{cases} \tag{4}$$

Finally, for the ‘Rate change’ model, the developmental rate differs between events occurring prior to a developmental phase ϕ" and events occurring afterwards. This yields:

$$t_a = \begin{cases} T_1\phi_a & \phi_a < \phi" \\ (T_1 - T_2)\phi" + T_2\phi_a & \phi_a \geq \phi" \end{cases} \tag{5}$$

where $1/T_1$ and $1/T_2$ correspond to the two developmental rates and the total duration of development is given by T'=$T_1\phi$"+$T_2(1-\phi")$.

6. Calculation of deviation from scaling for timing models

For the 'Uniform', 'Pause' and 'Rate change' model, we calculate the deviation from scaling as follows. First, we use that fact that for events a and b that occur in the same animal, the total duration of development, T' , has the same value, to express t_b as function of t_a. For the 'Uniform' model, this yields:

$$t_b = \frac{\phi_b}{\phi_a} t_a \tag{6}$$

meaning that that event pairs lie along the scaling line as measured for individuals under standard conditions, and that the changes in timing can be fully captured by a simple rescaling of event times with the duration of development T' under non-standard conditions. In contrast, for the 'Pause' model, this yields:

$$t_b = \begin{cases} \frac{\phi_b}{\phi_a} t_a & t_a, t_b < t^* \\ \frac{\kappa+\phi_b}{\phi_a} t_a & t_a < t^*, t_b > t^* \\ \frac{\kappa+\phi_b}{\kappa+\phi_a} t_a & t_a, t_b > t^* \end{cases} \tag{7}$$

Here, pairs of time points for events a and b only lie along the scaling line for standard conditions when both events occur before the time of the delay, t^*=$T\phi'$. Otherwise, the slope of the line is different from wild-type conditions and depends explicitly on the delay parameter κ. Finally, for 'Rate change' model, corresponding to the *lin-42(ox461)* mutant, we have:

$$t_b = \begin{cases} \frac{\phi_b}{\phi_a} t_a & t_a, t_b < t^* \\ \left(\frac{\phi"}{\phi_a} + \rho \frac{\phi_b - \phi"}{\phi_a}\right) t_a & t_a < t^*, t_b > t^* \\ \frac{\phi" + \rho(\phi_b - \phi")}{\phi" + \rho(\phi_a - \phi")} t_a & t_a, t_b > t^* \end{cases} \tag{8}$$

where $t^* = T_1\phi"$ and $\rho = \frac{T_2}{T_1}$. Apart from the case when $t_a, t_b < t^*$, this expression depends explicitly on the parameters ϕ" and ρ and does not lie along the same line event pairs for standard conditions. Finally, we calculate the deviation from scaling as:

$$\delta_{ab} = \text{atan}(s_{ab}^S) - \text{atan}\ (s_{ab}^P) \tag{9}$$

where $t_b = s_{ab} \cdot t_a$, with the slope s_{ab} given by Eqs. 6-8, and S and P denote standard and perturbed conditions, respectively.

7. Model fitting

We examined which of the three models, 'Uniform', 'Pause' or 'Rate change', provided the best fit to the experimental data for each mutant or condition examined, using the following procedure. The expression for the deviation from scaling in Eqs. 6-9 either does not depend on model parameters ('Uniform' model) or only on a subset of parameters (κ and ϕ' for the 'Pause' model and ρ and ϕ" for the 'Rate change' model). Hence, we first fitted the models to the experimentally measured deviation from scaling, by minimizing the root mean square error (RMSE) of the difference between experiment and models summed over all event pairs (Supplementary Fig. 4B,C). Using the best-fit values of these parameters, we then obtained best-fit values for the remaining parameters (T for the 'Uniform' and 'Pause' models, T_1 for the 'Rate change' model) by minimizing the RMSE of the difference between the model and the measured average event times as function of developmental phase, using Eqs. 2, 4 and 5 (Supplementary Fig. 4D). For HB101, we excluded the last two events, Peak 4 and Ecdysis 4, that formed the third epoch. Their timing deviated strongly from that of the events directly preceding them, but the long delay between Ecdysis 3 and Peak 4 and the small number of events in this epoch made it challenging to constrain a timing model for this epoch. For least-squares fitting, we used the *lmfit* package in Python.

8. Model parameters

For the model results in Fig. 3A-D, we used the following parameters, chosen to emphasize and clarify the differences between models. For the 'Standard condition' model: T=30 h. For the 'Uniform' model: T'=60 h. For the 'Pause' model: κ=1 and ϕ'=0.4, resulting in a total duration of development of T'=60 h. For the 'Rate Change' model: T_1=25 h, T_2=95 h and ϕ''=0.5, also resulting in T'=60 h. For the stochastic simulations in Fig. 3C-D: σ_T=5 h and σ_ϕ=$7\cdot 10^{-3}$. For Figs. 2, 3E,F, 4 and 5, we used model parameters that were fitted to the experimental data, as outlined above. In Fig. 2, we used the following fitted parameter values: $T^{23^\circ\mathrm{C}}$=39 h, $T^{19^\circ\mathrm{C}}$=57 h, and $T^{15^\circ\mathrm{C}}$=110 h. In Fig. 4 and Fig. 3E, we used the following fitted parameter values: T^{eat-1}=42 h, T^{HB101}=43 h, κ_{HB101}=0.032 and ϕ'_{HB101}=0.42. In Fig. 5 and Fig. 3F, we used the following fitted parameters: T_1^{lin-42}=45 h, T_2^{lin-42}=70 h and ϕ''_{lin-42}=0.37. In Fig. 6C,E, we used T_1=45 h, T_2=70 h, as determined by the experiments in Fig. 2, and ϕ" =0.3, corresponding to the average time after hatching at which the temperature shift was initiated.

9. Event order

We first determined the average time of each event, as measured under standard conditions (wild-type animals, fed OP50 and raised at 23°C). We then established the average event order as the sequence of events, as ordered by their average timing. To detect permutations of event order in individual animals (either raised under standard conditions, under different conditions or in mutants)

as compared to the average event order, we followed the following procedure. We stepped through each event i in the order determined by the average event order and examined the relative timing of event $i+1$. If event $i+1$ precedes event i in this individual, we first swapped the order of these two events in the sequence of events recorder in this animal, then recorded this as a permutation of events i and $i+1$, and then restarted the procedure at the first event, $i=0$. Upon completion, this procedure has sorted the sequence of events recorded in this individual so that it is identical to the average event order, while recording all permutations between subsequent events i and $i+1$ that are required to achieve this. The frequency of such permutations within a population of animals under different conditions is presented in Supplementary Fig. 2. Note that changes in order of events that on average are not adjacent in time, e.g. the normal order (*a,b,c*) is changed into (*b, c,* a), are with this procedure recorded as two permutations, namely first swapping *a* and *b*, and second *b* and *c*. However, such more complex event reorderings occur relatively rarely.

10. Calculating the deviation from scaling for experimental data

We consider two different kinds of deviations from scaling. First, deviations from inter-individual scaling refer to the deviation of event times $t_{a,i}$ and $t_{b,i}$, measured in each animal i, from the scaling line with slope $s_{a,b}$ that is fitted to the data points for the entire population. The deviation from inter-individual scaling would be zero if all data points for event pairs a and b fall exactly on the line with slope $s_{a,b}$. Second, deviations from population-level scaling refer to differences in the fitted slopes $s_{a,b}^S$ and $s_{a,b}^P$, measured for the same event pair a and b under standard (S) and perturbed (P) conditions. The deviation from population-level scaling is zero when the data points for standard and perturbed condition are fitted by lines with the same slope. To quantify the deviation from inter-individual scaling, we calculated $\theta_i = \arctan\left(\frac{t_{b,i}}{t_{a,i}}\right)$. The angle made by the data point of animal i with the scaling line is then given by $\delta\theta_i = \theta_i - \arctan(s_{a,b})$. Next, we calculated the distance λ_i to the scaling line along a circle with radius $\left(t_{a,i}^2 + t_{b,i}^2\right)^{\frac{1}{2}}$ as $\lambda_i = \delta\theta_i\left(t_{a,i}^2 + t_{b,i}^2\right)^{\frac{1}{2}}$. The deviation from inter-individual scaling is then given by the standard deviation $\sigma_\lambda = \sqrt{\langle \lambda_i - \langle \lambda_i \rangle \rangle}$ calculated over all individuals in the population. We calculated the deviation from population-level scaling as $\delta = \langle \theta_i^P \rangle - \langle \theta_i^S \rangle$. However, in particular for event pairs with small average event times $\langle t_{a,i} \rangle$ and $\langle t_{b,i} \rangle$, values of θ_i can vary substantially, leading to non-zero deviation δ for the typical number of animals analyzed for these experiments. Hence, we also estimate the probability that the two series θ_i^P and θ_i^S are sampled from the same distribution, using the two-sample Kolmogorov-Smirnov test (*ks_2samp* from the *scipy.stats* package in Python). We reported the P value, with high P meaning that the distributions of the two samples are likely the same, and, hence, obey the same temporal scaling relationship.

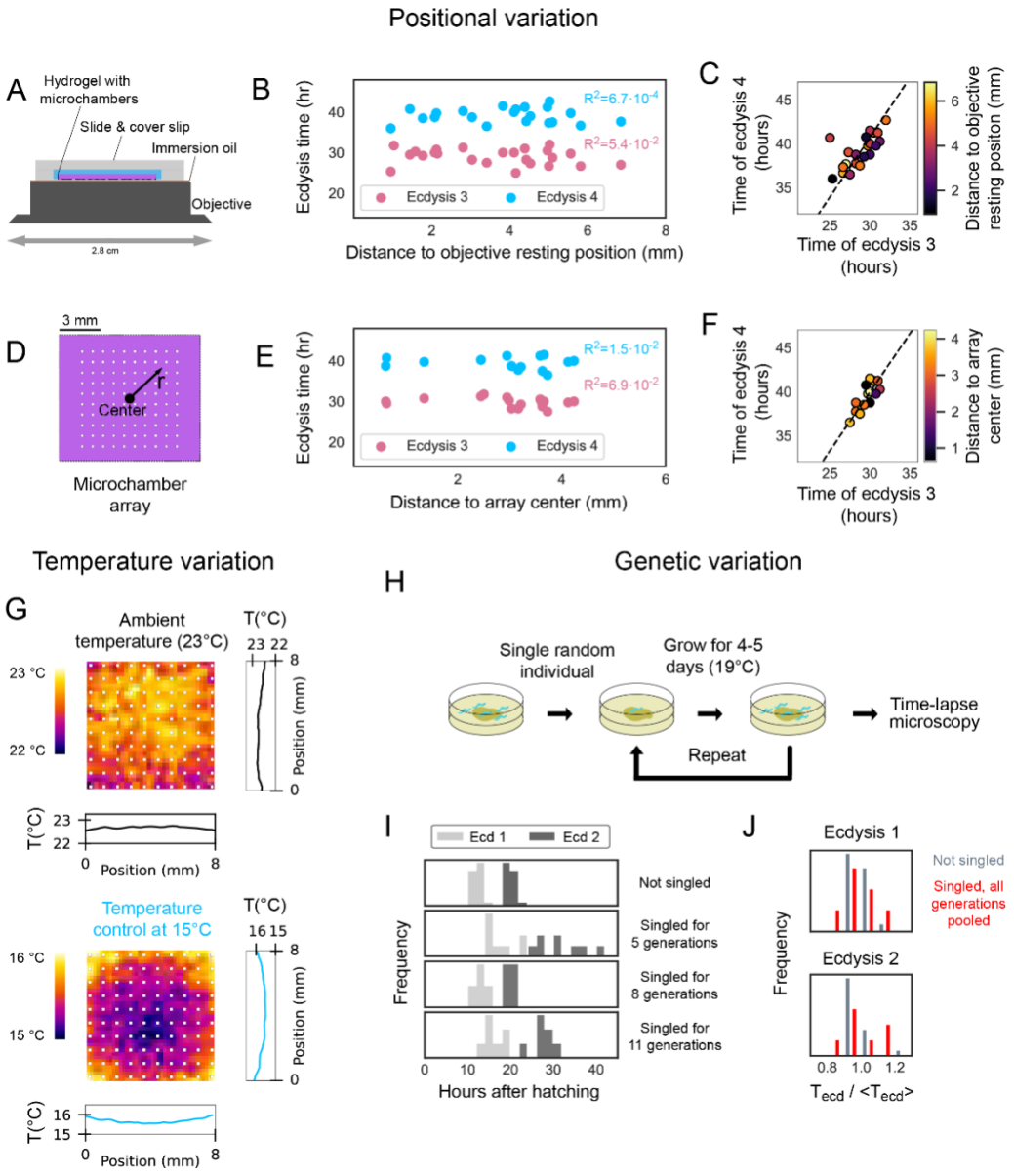

Figure S1 External sources of timing variability.

(A) Schematic cross-section of a sample (hydrogel between cover slip and microscopy slide) and objective during a typical experiment. All proportions are true to scale. Note the relatively large size of the microscope objective compared to the microchamber array. **(B)** Impact of sample position. Time of ecdysis measured in an individual animal, as function of the distance of its microchamber to center of the objective's resting position. The resting position is the location where the center of the objective is positioned between each imaging cycle, and varies between experimental batches. For both Ecdysis 3 and 4, the Pearson correlation R is small, indicating a poor correlation with distance to the objective's resting position. **(C)** Times of Ecdysis 3 and 4 plotted as event pairs, with marker colors indicating the distance to the objective's resting position. While event pairs

exhibit scaling, no clear correlation is apparent between distance and early or late ecdysis timing. **(D)** Top view of the microchamber array. **(E),(F)** Same as (B),(C) but comparing to the distance *r* of each microchamber to the center of the microchamber array, as indicated in (D). Overall, the correlation of ecdysis timing with *r* is also weak (E), and does not explain the difference between early or late event timing (F). **(G)** Images of temperature distribution, measured by Forward Imaging InfraRed (FLIR), of a microchamber array at ambient conditions (top, corresponding to the standard condition of 23°C) and with temperature control at 15°C (bottom). Graphs show the average temperature profile measured along the two axes. Temperature differences within the microchamber array are <1°C. **(H),(I),(J)** Impact of genetic variation. We propagated wild-type *C. elegans* populations as single animals (H), for either 5, 8 or 11 generations, to reduce any genetic variability, and measured ecdysis times (I). Animals that were propagated by singling showed delayed ecdysis times (dark grey) compared to wild-type animals (light grey), consistent with impact of accumulated deleterious mutations (Vassilieva et al., Evolution 2000). However, variability in ecdysis times was not reduced compared to wild-type animals (J), as singled animals showed a similar distribution of ecdysis times T_{ecd} relative to the population average $\langle T_{\mathrm{ecd}} \rangle$ compared to non-singled animals. Here, times measured for animals singled 5, 8, and 11 generations were pooled together. This result shows that variability in timing is not due to genetic variability within the population.

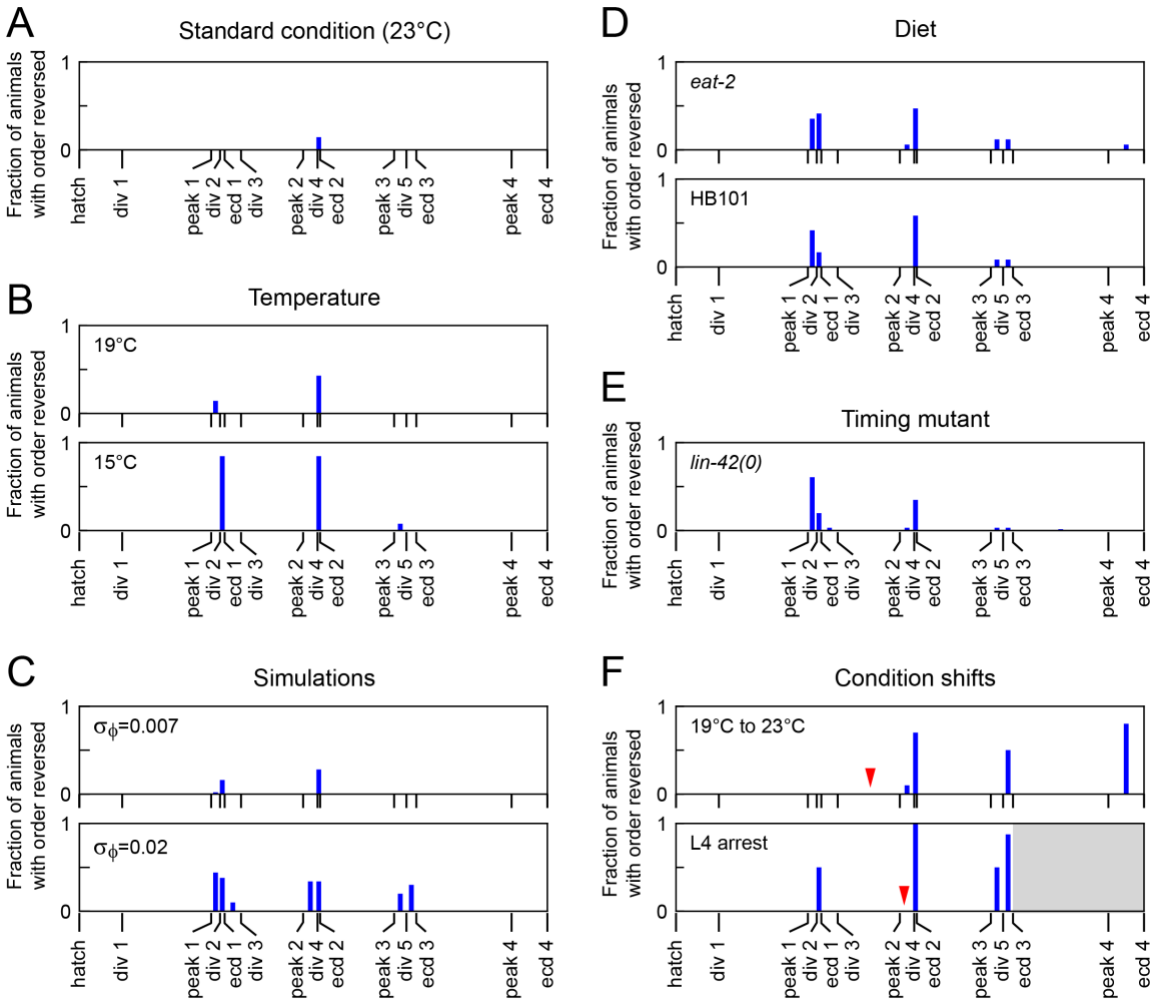

Figure S2. Event order.

(A) Fraction of animals with a measured event order that is permuted compared to the average event order, under standard conditions. Tick marks indicate the average time of each event, as measured under standard conditions. When the order of two events is reversed, e.g. Division 4 and Ecdysis 2, the fraction of animals that exhibited such a reversal is indicated by a bar placed between the corresponding tick marks. Overall, few animals have permutations in order compared to the average event order. **(B)** Permutations in event order upon changes in temperature. Tick marks correspond to average times measured under standard conditions. In particular for development at 15°C, reversals in order between seam cell divisions and ecdysis are frequent in the L1 and L2 larval stages. In general, reversals in event order are only seen in events that, on average, occur close together in time. **(C)** Permutation in event order in stochastic simulations of the timing model in Eq. 2 and 3 in the Methods, for increasing σ_ϕ, i.e. the variability in the developmental phase at which events are executed. For larger variability in phase, permutations of event order become more frequent, but impact more strongly those events that occur close together in time. **(D)**, **(E)**

Permutations in event order upon changes in diet (D) and in *lin-42(0)* mutants (E). Tick marks correspond to average times measured under standard conditions. These animals occasionally show changes in event order that impact events that are non-adjacent in time, e.g. resulting in the permutated event order Division 5 - Ecdysis 3 - Peak 3. In our analysis, this is recorded as two order permutations, of Peak 3 and Division 5 and of Division 5 and Ecdysis 3. For *lin-42(0)* mutants, the decrease in permutations towards the end of larval development reflects the observation that many animals arrest before that stage. **(F)** Permutations in event order for animals exposed to a shift in temperature (top, 19°C to 23°C) or a shift to starvation (bottom, L4 arrest). Markers correspond to the average observed timing of the shift. The grey area indicates events that do not take place as the starved animals arrest prior to Ecdysis 3 (L4 arrest).

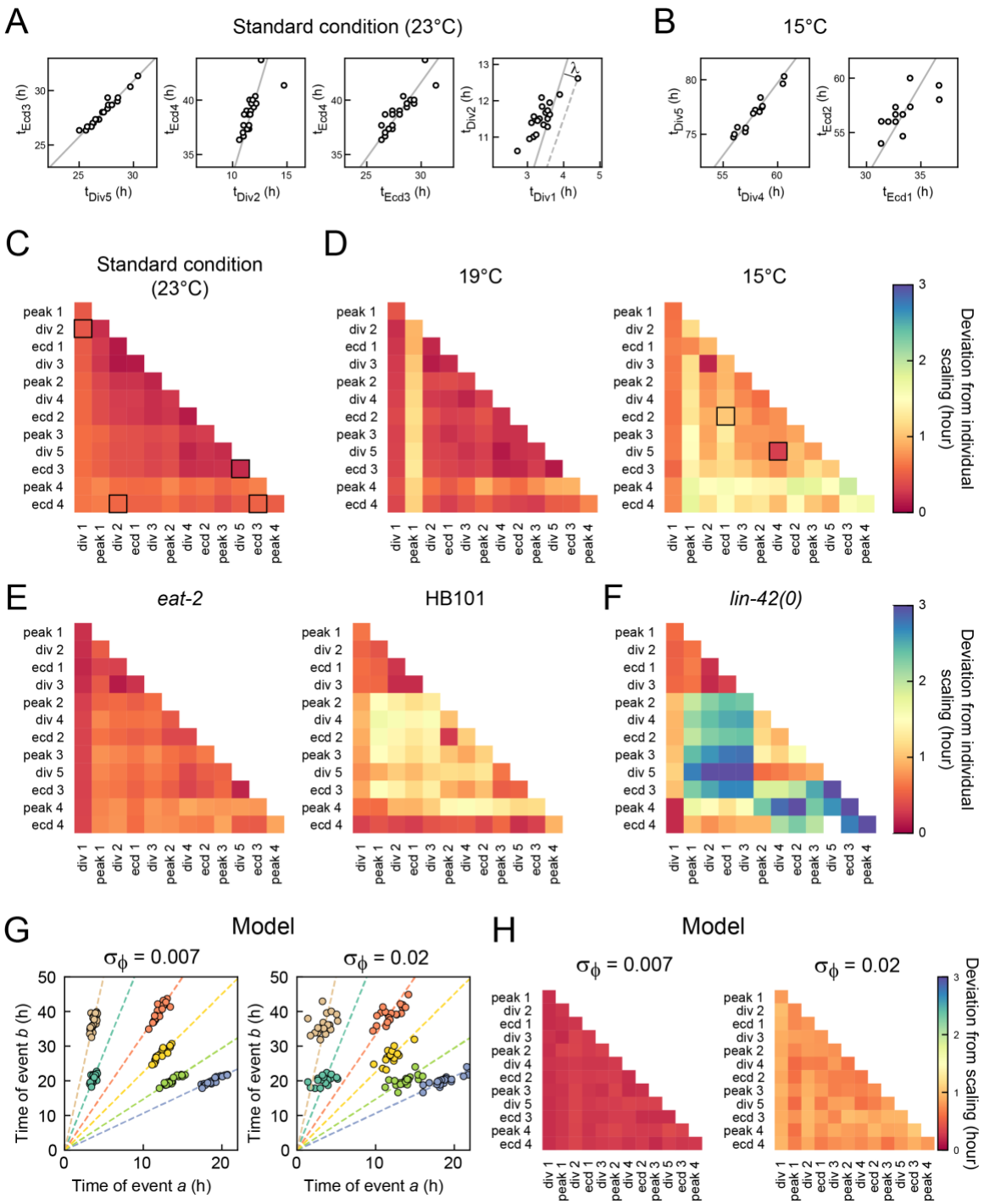

Figure S3. Quality of temporal scaling.

(A) Measured times for three different event pairs a, b under standard conditions (wild-type animals, on an *E. coli* OP50 diet at 23˚C). Markers correspond to times measured in a single animal. Solid grey lines are fits of the form t_b=$s_{a,b} \cdot t_a$ to the data points for individual animals. The four event pairs are ordered by decreasing quality of scaling, i.e. larger variation of single-animal data around the fitted scaling lines. To quantify the scaling quality, we defined $\lambda_i^{a,b} = \left(\mathrm{atan}\left(t_{b,i}/t_{a,i}\right) -\right.$

$\mathrm{atan}\,(s_{a,b})\,\big)\big(t_{a,i}^2 + t_{b,i}^2\big)^{\frac{1}{2}}$, the distance between the data point of individual i and the scaling line, as measured along a circle. **(B)** Times for event pairs measured at 15˚C, ordered by decreasing scaling quality. **(C)** Scaling quality, measured as the standard deviation $\sigma_\lambda^{a,b} = \sqrt{\langle \lambda_i^{a,b} - \langle \lambda_i^{a,b} \rangle \rangle}$ of the distance to the scaling line over all individuals in the population, shown here for all event pairs measured under standard conditions. Squares highlight the event pairs shown in (A). See Materials and Methods for full details. **(D), (E), (F)** Scaling quality for all events pairs measured (D) under different temperatures, (E) upon changes in diet and (F) in *lin-42(0)* mutants. Squares in (D) indicate event pairs shown in (B). **(G)** Simulated times of event pairs a, b in the stochastic timing model (Eqs. 2 and 3 in Methods). Data is shown for events with the same developmental phase as the events in Fig. 1G in the main text. Otherwise, simulation parameters were chosen to resemble the experimental data: we used standard deviations σ_T=3h and σ_ϕ=$7{\cdot}10^{-3}$, meaning that common variability in the rate of development, $1/\langle T \rangle$, is stronger than variability in timing of each individual event. The simulated data closely resemble the experimental data in Fig. 1G, with times for event pairs a, b scattered along lines of $t_b = \frac{\phi_b}{\phi_a}\, t_a$ (dashed lines). Increased variability in developmental phase**,** σ_ϕ=$2{\cdot}10^{-2}$, resulted in weaker clustering along the scaling lines. **(H)** Deviation from inter-individual scaling for the parameters in (G). Data is shown for all events pairs in the experiments, with each event's average phase based on the experimentally measured value.

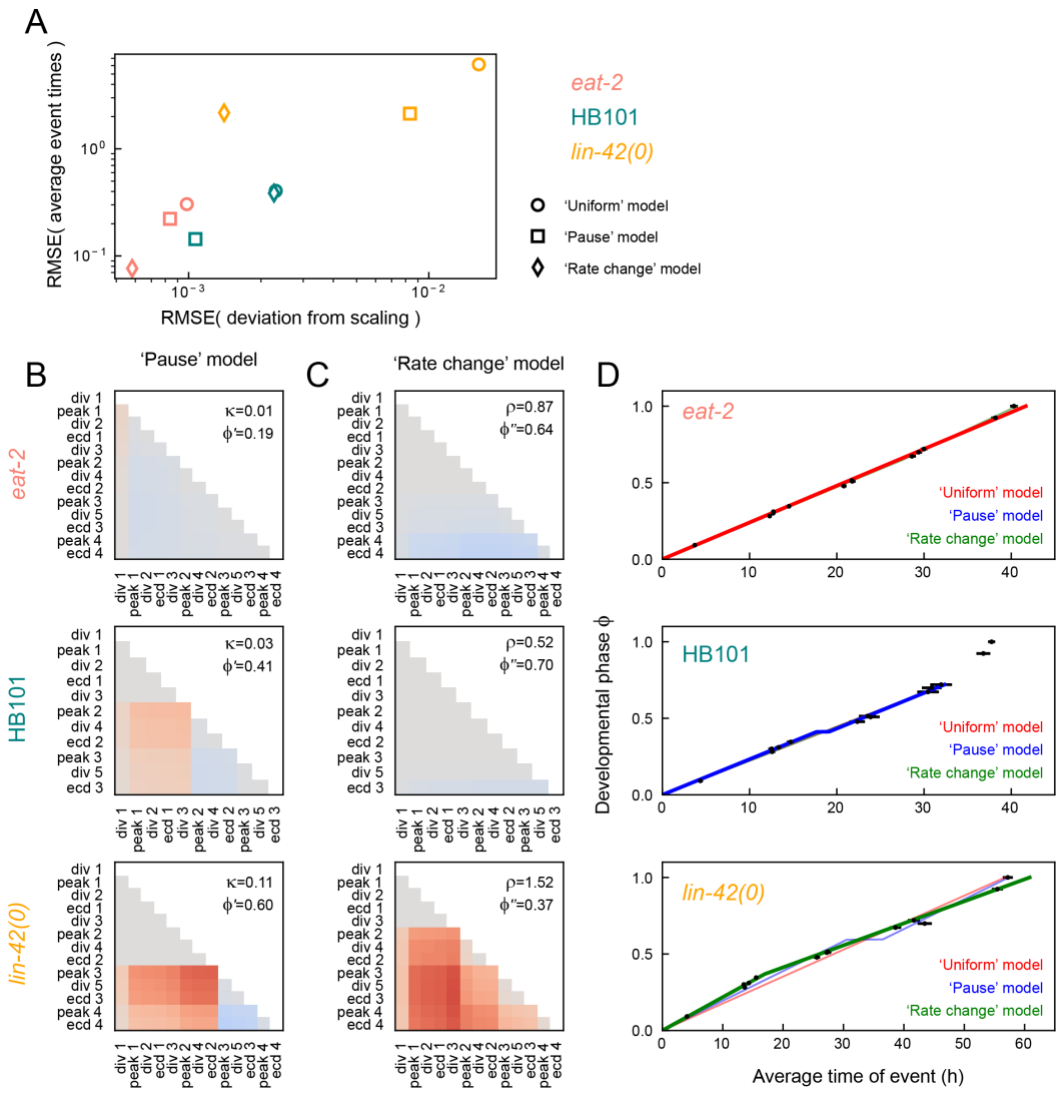

Figure S4. Model fits.

(A) Quality of fits to 'Uniform' (circle), 'Pause' (square) and 'Rate change' (diamond) models for the experimental timing data for *eat-2* (pink) and *lin-42(0)* (orange) mutants and wild-type animals fed an HB101 diet (teal). For each mutant or condition, the root mean square error (RMSE) of the best fit of each model to the experimental deviation from scaling data is plotted on the x-axis and the RMSE of the best fit to the average event times is plotted on the y-axis. For HB101 and *lin-42(0)*, a single model provided the best overall fit, i.e. the 'Pause' model for HB101 and the 'Rate change' model for *lin-42*(0). For *eat-2*, all three models fit the experimental data well. See Materials and Methods for details of the fitting procedure. **(B),(C)** Deviation of scaling predicted by the 'Pause' (B) and 'Rate change' (C) models for the best fit to the experimental data for *eat-2*, HB101 and *lin-42(0)*. The 'Uniform' model predicts zero deviation of scaling for all event pairs. For each model and mutant/condition, the values of the fitting parameters corresponding to the best fit are given. **(D)** Average event times. Experimentally measured average event times are shown as black

markers, with error bars indicating the standard error. Lines indicate the predicted event times for the ‘Uniform’ (red), ‘Pause’ (blue) and ‘Rate change’ (green) models. Thick lines correspond to the model and parameter values used for the best-fitting model in the main text. For *eat-2*, we chose the ‘Uniform’ model as best-fitting model, as it performs approximately as well as the other two models but with fewer fitting parameters. This is supported by the observation that the best-fit of ‘Pause’ and ‘Rate change’ models exhibit either a small delay, κ=0.01, or a weak change in developmental rate, ρ=0.87, resulting in an evolution of the developmental phase that is close to that of the ‘Uniform’ model.

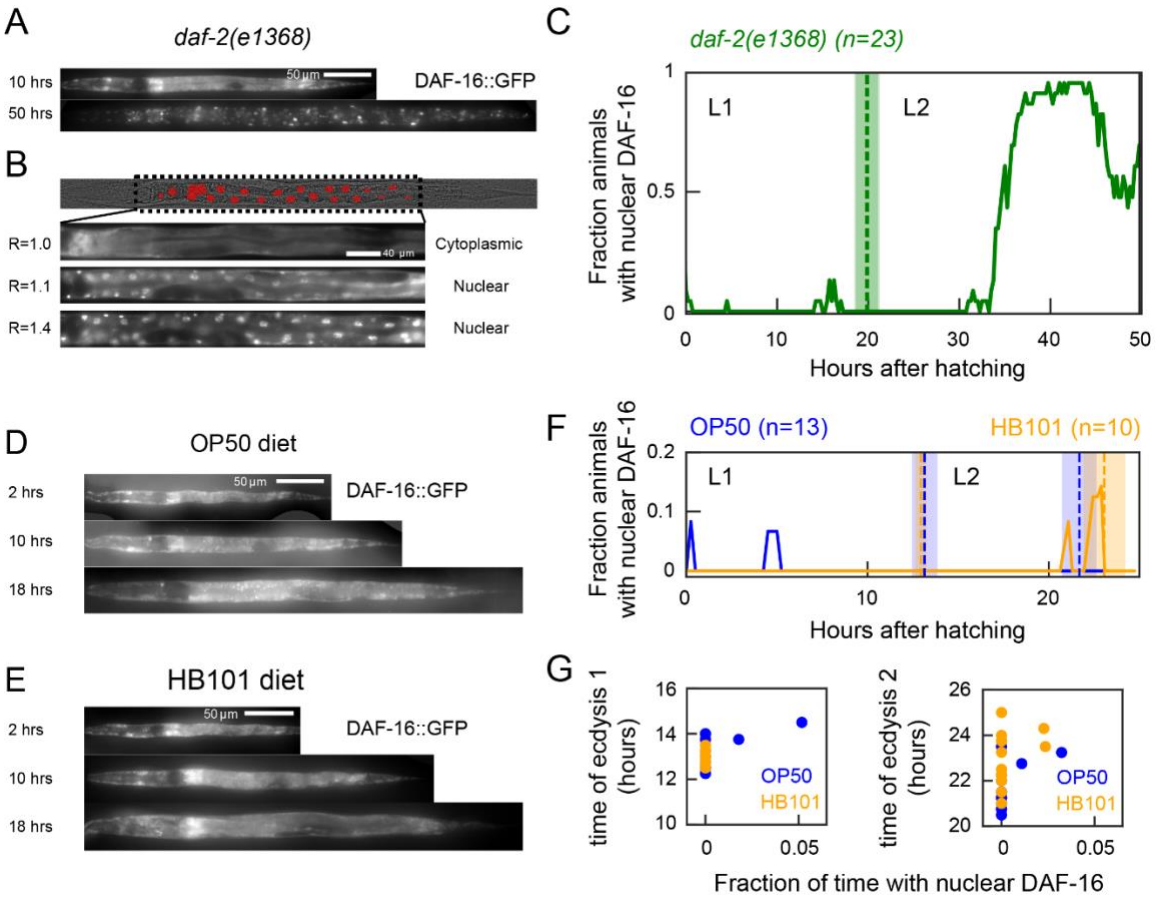

Fig. S5. Insulin signaling and developmental timing.

(A) Computationally straightened *daf-2(e1368)* animal carrying a DAF-16::GFP reporter, at different times after hatching. DAF-16 undergoes nuclear translocation, indicative of a change in insulin signaling, as *daf-2* animals enter dauer. **(B)** Quantification of DAF-16 nuclear localization. Top image shows a transmitted light image of a single, computationally straightened animal, with a fluorescent intestinal nuclear marker (*elt-7p::H1-wCherry*) superimposed in red. The nuclear marker is used to calculate the DAF-16 nuclear to cytoplasmic fluorescence ratio R. The bottom three images show DAF-16::GFP animals with increasing nuclear localization and the corresponding, measured values of R. Animals with clear cytoplasmic localization of DAF-16 have a ratio of R≈1. At a R=1.1, DAF-16 first shows visible nuclear localization, as seen by eye. Therefore, we classify animals with R≥1.1 as showing nuclear DAF-16 localization. **(C)** Fraction of *daf-2(e1368)* animals with nuclear DAF-16 (R≥1.1) during L1 and L2. Dashed lines and shaded area indicate the average and standard deviation of ecdysis time. As expected, DAF-16 becomes nuclear in almost all animals, as they enter dauer. **(D),(E)** DAF-16::GFP individuals fed (D) OP50 and (E) HB101 at different times after hatching. **(F)** Fraction of animals fed OP50 (blue) or HB101 (orange) with nuclear DAF-16 (R≥1.1) during L1 and L2. Dashed lines and shaded area are the average and standard deviation of ecdysis time. The late L2 ecdysis in animals fed HB101 indicates

that the HB101-induced pause is also seen in DAF-16::GFP animals. Note the different scale of the y-axis compared to (C). In most animals, DAF-16 remains cytoplasmic at all times. Infrequent periods of nuclear localization typically reflect values close to the threshold of R=1.1. **(G)** Time of L1 (left) and L2 (right) ecdysis versus the fraction of time DAF-16 was nuclear from hatching to L1 or L2 ecdysis, respectively. Variability in ecdysis time between individuals was not explained by concomitant variation in insulin signaling, as measured by DAF-16 nuclear localization.

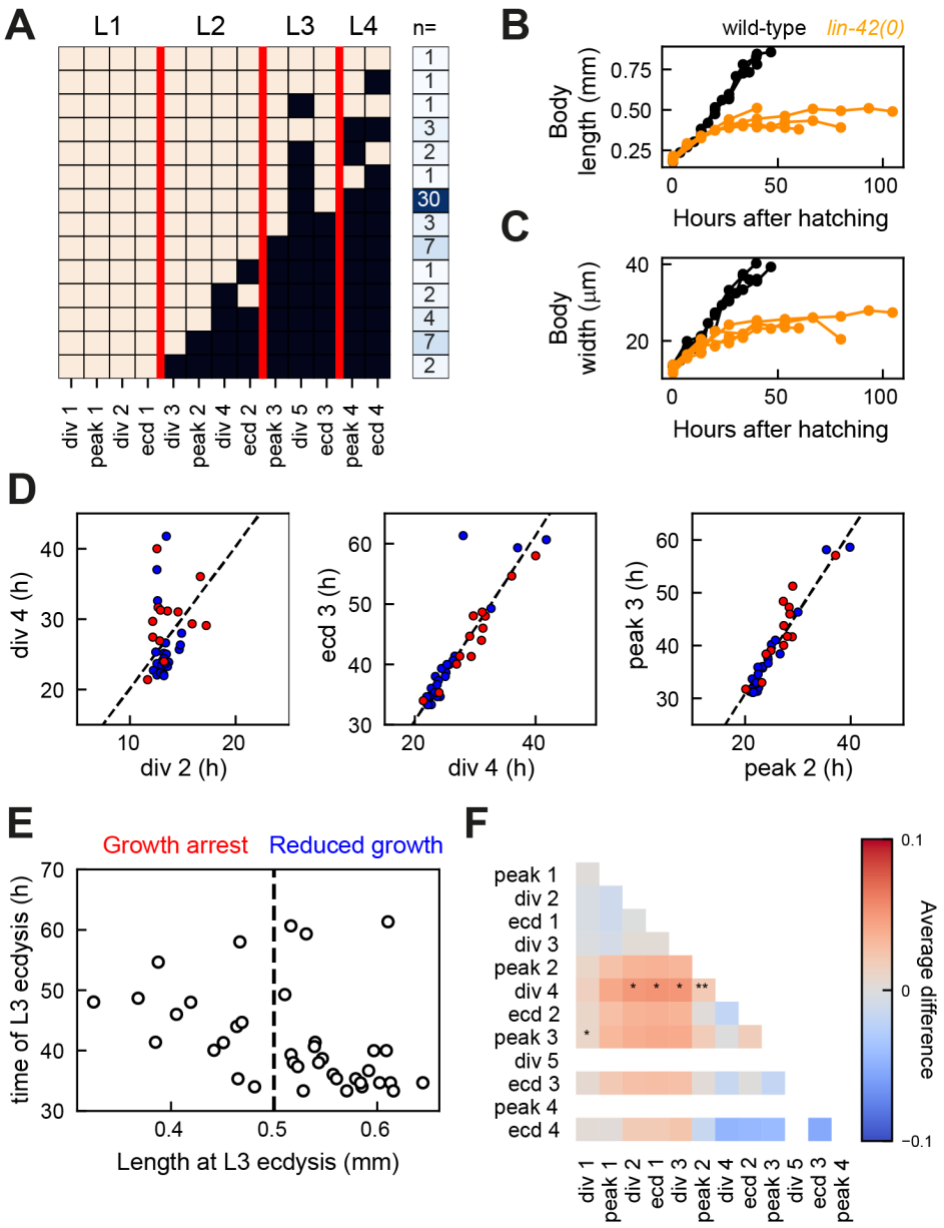

Fig. S6. Developmental progression and growth arrest in *lin-42(0)* mutants.

(A) Developmental progression in *lin-42(0)* animals. Bright squares indicate events that were executed, whereas dark squares are events that did not take place. Red lines separate the different larval stages. For each unique sequence of events recorded, we give the number of individuals n in which this sequence is observed. Of 65 animals, 4 reached adulthood (Ecdysis 4), 33 animals arrest in L3, and 21 animals arrest in L2. **(B)** Body length and **(C)** body width measured in the same individuals. Black and orange markers are wild-type animals and *lin-42(0)* animals, respectively. Animals exhibit a simultaneous arrest in extension of both body length and width. **(D)** Measured times for representative event pairs measured in *lin-42(0)* animals, with markers colored according

to the growth phenotype as defined in panel (E). Lines are a linear fit to the data points. The deviation from scaling does not differ strongly between growth-arrested animals (red) and animals with reduced growth (blue). **(E)** Length at L3 ecdysis compared to time of L3 ecdysis in *lin-42(0)* animals. Based on this, we separated the population in growth-arrested animals (length < 0.5mm at L3 ecdysis) and animals with reduced growth (as compared to wild-type animals). Growth-arrested animals developed more slowly than animals with reduced growth, but a small number of animals with reduced growth also displayed very slow development (L3 ecdysis later than 50 h after hatching). We scored the growth phenotype based on L3 characteristics, because most animals skip the L4 larval stage. **(F)** Difference in scaling between growth-arrested (GA) and reduce-growth (RG) animals. Color indicates the difference $\langle\theta^{GA}\rangle - \langle\theta^{RG}\rangle$ between the two populations, where, for each event pair a and b measured in an individual animal, the angle $\theta = \arctan\frac{t_a}{t_b}$. Stars indicate the probability that the distribution of θ is the same for growth-arrested and reduced-growth animals: **:P<0.001, *:P<0.01, and N.S. otherwise (K-S test). Overall, no significant differences in scaling were observed between growth-arrested and reduced-growth animals, indicating that growth-arrested animals do not display stronger breakdown of scaling. Empty squares reflect event pairs for which at least one of the two events did not occur in either growth-arrested or reduced-growth animals.

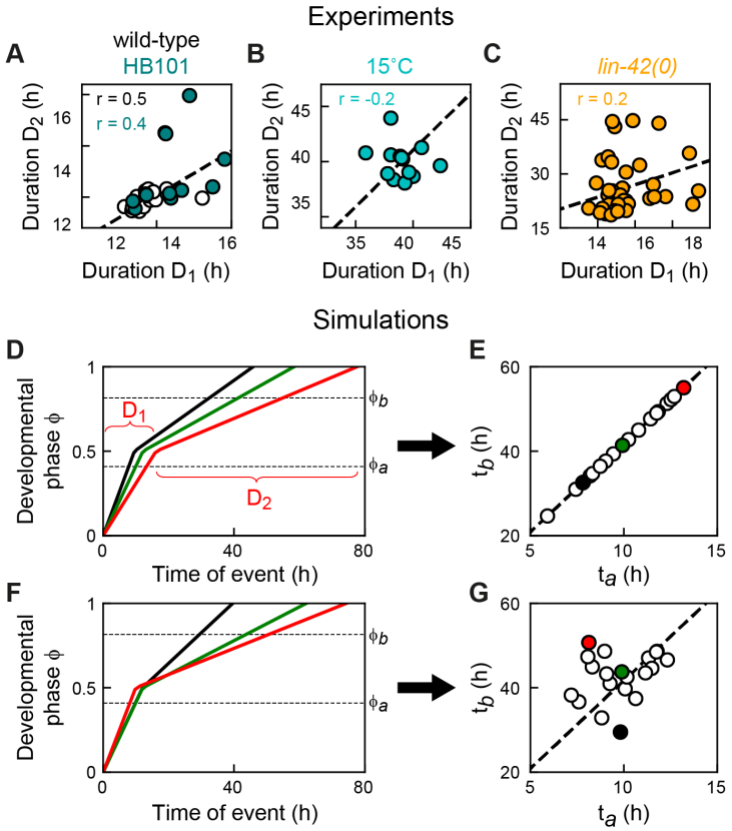

Fig. S7. Breakdown of inter-individual scaling.

(A),(B),(C) Experimentally measured correlation between the developmental duration D_1 and D_2 of development before and after the change in developmental rate, for development under standard conditions (A, wild-type at 23˚C), animals fed HB101 (A), at 15˚C (B) and in *lin-42(0)* mutants (C). The duration D_1 is defined at the time from hatching to the third seam cell division, while D_2 is given by the time from the third seam cell division to the third ecdysis. For D_2, we do not include the time from the third to fourth ecdysis, as this last part of larval development is often skipped in *lin-42* mutants. For each condition and genotype, *r* is the Pearson correlation coefficient. The correlation between D_1 and D_2 is much weaker for animals at 15˚C and in *lin-42* mutants. **(D),(E)** Inter-individual variability in the 'Rate change' model. In this model, all individuals go through the same evolution of developmental phase ϕ but at a rate that differs between individuals (D). Hence, variability in the duration D_1 and D_2 of development before and after the change in developmental rate is strongly correlated within the same individual. As a consequence, times of events t_a and t_b, occurring at phase ϕ_a and ϕ_b, are clustered tightly along a line, indicating temporal scaling (E). The colored markers in (E) correspond to the individuals whose phase evolution is shown in (D). Model is given by Eqs. 3 and 5 in Methods, with T_1=25 h, T_2=95 h, ϕ"=0.5, $\sigma_\phi = 7 \cdot 10^{-3}$ and σ_T=15 h. Dashed line is Eq. 8. **(F),(G)** Breakdown of inter-individual scaling, when variability in durations D_1 and D_2 is not

correlated within the same individual. As a result, individual animals show strong deviations from temporal scaling. Model is given by Eqs. 3 and 5, but T_1 and T_2 vary independently within the same individual, with σ_{T_1}=3 h and σ_{T_2}=15 h. This model result indicates that the lack of correlation between durations D_1 and D_2 observed in lin-42(0) mutants can explain the lack of inter-individual scaling seen in *lin-42(0)* mutants when comparing timing of events occurring before and after the change in developmental rate in Fig. 5E.