Supporting information for: Rotation of sex combs in
*Drosophila melanogaster* requires precise and
coordinated spatio-temporal dynamics from forces
generated by epithelial cells

Ernest C. Y. Ho¹,a, Juan Nicolas Malagón¹,²,a, Abha Ahuja³,4, Rama Singh³, and Ellen Larsen¹

¹Department of Cell and Systems Biology, University of Toronto, Toronto, Ontario, Canada
²Boyer Center for Molecular Medicine, Yale University, New Haven CT, USA
³Department of Biology, McMaster University, Hamilton, Ontario, Canada
⁴College of Natural Sciences, Minerva Schools at KGI, San Francisco, California, USA

aShared first authors

September 8, 2018

Number of Pages: 18 (including this title page)
Data analyses of experimental results of Figure 4

There are 5 experimental data points for each of ♀wt, ♂wt and ♂bab<sup>PR72</sup> (Table A). Each data point consists of 4 coordinates: initial area of EP1, initial area of EP2, final area of EP1 and final area of EP2. We apply a combination of Monte Carlo simulations and bootstrapping to construct artificial data points based on these 5 actual experimental data points for each genotype for statistical analyses. First, we perform statistical bootstrapping: we construct 5000 extra sets of data points based on the 5 experimental data points. Each set of these data points consists of the same number of data points (5) as the original. Each data point of a set comes from random drawing of the 5 data points of the original experimental data set with replacement. Next, we perform Monte Carlo simulations on each of the 5000 bootstrapped data sets for each of ♀wt, ♂wt and ♂bab<sup>PR72</sup>: for each of these bootstrapped data sets, we generate 1 million artificial data points from a multivariate Gaussian distribution with means and covariances identical to the means and covariances of that data set. Therefore, there are 4 sample mean values (each for initial EP1, initial EP2, final EP1, final EP2) and $4 \times 4 = 16$ covariance values arranged in a positive semi-definite symmetric matrix required for the generation of these 1 million artificial data points. Finally, we derive quantities and perform confidence interval estimates based on the $5000 \times 3$ realizations with a total of 15000 million artificial data points from the bootstrapped data sets of all three genotypes (both detailed below).

For each bootstrapped data set with the 1 million Monte Carlo simulated data points, we define the inhomogeneity coefficient $C_{inhomogeneity}$ as,

$$C_{inhomogeneity} = \frac{\text{Final area of EP2/Initial area of EP2}}{\text{Final area of EP1/Initial area of EP1}}.$$

$C_{inhomogeneity}$ is a vector of 1 million elements. In equation 1, $C_{inhomogeneity}$ is formed by element-wise division of two vectors, $\frac{EP2_{Final}}{EP2_{Initial}}$ (numerator) and $\frac{EP1_{Final}}{EP1_{Initial}}$ (denominator), which are themselves formed by element-wise division of the respective individual data points within
the same bootstrapped set. For example, the first element of \( \frac{\text{EP}_2^{\text{Final}}}{\text{EP}_2^{\text{Initial}}} \) is obtained by dividing the value of the final area of EP2 of the first data point with the initial area of EP2 of the first data point. Therefore, an element of \( C_{\text{inhomogeneity}} > 1 \) means that, for that particular simulated data point, EP2 expands more (or contracts less) as a proportion of its initial size than EP1, and \textit{vice versa} for an element of \( C_{\text{inhomogeneity}} < 1 \). Thus, \( C_{\text{inhomogeneity}} \) captures the inhomogeneity of distal cell expansion during SC rotation.

To explore the effects of being \( \varphi \) wt, \( \sigma \) wt or \( \sigma \) bab\(^{PR72} \) has on the inhomogeneity of expansion of distal cells, we define \( \Delta C_{\text{inhomogeneity}} \) as the difference of \( C_{\text{inhomogeneity}} \) between any of the two of \( \varphi \) wt, \( \sigma \) wt and \( \sigma \) bab\(^{PR72} \) (therefore, two bootstrapped data sets, each of 1 million artificial data points, are required for arguments of \( \Delta C_{\text{inhomogeneity}} \)):

\[
\Delta C_{\text{inhomogeneity}}(s1, s2) = C_{\text{inhomogeneity}}(s1) - C_{\text{inhomogeneity}}(s2). \tag{2}
\]

Thus, \( \Delta C_{\text{inhomogeneity}}(s1, s2) \) (a vector of 1 million elements) captures how much such an asymmetry of distal cell expansion varies between \( s1 \) and \( s2 \) where \( s1 \) and \( s2 \) can be any of \( \varphi \) wt, \( \sigma \) wt and \( \sigma \) bab\(^{PR72} \). In obtaining \( \Delta C_{\text{inhomogeneity}}(s1, s2) \), we randomly shuffle the elements of \( C_{\text{inhomogeneity}}(s2) \) before the vector subtraction to avoid potential artifacts resulting from the seeds of the pseudo random number generator. Figure A shows how the values of \( C_{\text{inhomogeneity}} \), \( \Delta C_{\text{inhomogeneity}}(s1, s2) \) and other related quantities based on three example bootstrapped data sets can be visualized as distributions. For statistical analyses, there is a total of \( 5000 \times 3 \) such \( C_{\text{inhomogeneity}} \) distributions. We have also used \( 5000 \times 3 \) such \( \Delta C_{\text{inhomogeneity}}(s1, s2) \) distributions (5000 for each pair of \( (s1, s2), s1 \neq s2 \)), out of a total of \( 3 \times 5000 \times 5000 \) possible combinations, not counting interchange of indices.

In the main text, we used derived quantities such as \( Pr(\frac{\text{EP}_2^{\text{Initial}}}{\text{EP}_1^{\text{Initial}}} < 1) \) or \( Pr(C_{\text{inhomogeneity}} > 1) \). These probability values for each realization of bootstrapped data sets are defined as the proportion of area of the relevant distribution that satisfies the said criterion. These probability values are themselves random variables with a probability distribution that can be visualized by calculating every realization of the probability value from every bootstrapped data set.
Thus, confidence intervals of these probability values and other derived quantities can be estimated in such a manner (Table B).

**Data analyses of ABASCT SD statistics of simulated intact SC without temporal dynamics**

We first perform ANOVA on the ABASCT SD data of simulated intact SC without temporal dynamics (i.e. delay=0 mcs) to test the equality of the means of ABASCT SD values between SC lengths (5, 7, 9 or 11-tooth SCs). The resulting F-test statistic has a value of 120 with \( \text{Pr}(>F) < 2 \times 10^{-16} \), which shows at least one of the means is significantly different from the other three. Post-hoc Tukey’s test is then used to pairwise compare the ABASCT SD data of each group (SC lengths). The results are shown in Table C and Figure E.

**Calculation of p-value in Table D**

In Table D we estimated the p-value between the breaking statistics of two sets of SC simulations. Given that \( 0 \leq A \leq B \leq 48 \), the p-values quoted in Table D are related to the probability that any two sets of simulations with intact ratio of at most \( \frac{A}{48} \) and intact ratio of at least \( \frac{B}{48} \) share a common probability distribution. These values are calculated as follows.

Assuming that each simulation is independent with the rotated SC having a probability \( P \) of being intact, in a set of 48 simulations the probability of getting \( A \) intact SCs is thus,

\[
f(A, 48, P) = \binom{48}{A} \times P^A \times (1 - P)^{48-A}
\]

where

\[
\binom{48}{A} = \frac{48!}{A! \times (48 - A)!}.
\]
Therefore, the probability that a set of simulations has at most $A$ intact SCs is simply,

$$F_1(A, 48, P) = \sum_{i=0}^{A} f(i, 48, P). \quad (5)$$

Similarly, the probability for a set of simulations to get at least $B$ intact SCs is,

$$F_2(B, 48, P) = \sum_{i=B}^{48} f(i, 48, P). \quad (6)$$

Multiplying equations 5 and 6 gives the probability that both events happen together,

$$G(A, B, 48, P) = F_1(A, 48, P) \times F_2(B, 48, P). \quad (7)$$

To derive the p-value, one needs the maximum of $G(A, B, 48, P)$ along the $P$-axis ("worst case scenario"). One method to find such values is to differentiate $G(A, B, 48, P)$ with respect to $P$ and set the expression to zero,

$$\frac{dG(A, B, 48, P)}{dP} = 0 \quad (8)$$

and look for the values of $P$ that solve equation 8 and satisfy the second derivative condition

$$\frac{d^2 G(A, B, 48, P)}{dP^2} < 0.$$  

Once we get these values we can determine whether any of those values of $P$ or the "end-point" values (i.e. $P = 0$ or 1) yields the maximum of equation 7 for $0 \leq P \leq 1$. In our case, we use an alternative numerical approach of scanning the numerical maximal value of $G$ by increasing $0 \leq P \leq 1$ step-wise. The p-value reported in each row of Table D is two times this numerical maximum.
Figure A: Example distributions of Monte Carlo simulated quantities.

- **a** Example distributions of Monte Carlo simulated initial area of EP2 for ♀ _wt_, ♂ _wt_ and ♂ _bab^PR72_.
- **b** Example distributions of Monte Carlo simulated final area of EP2 for ♀ _wt_, ♂ _wt_ and ♂ _bab^PR72_.
- **c** Example distributions of Monte Carlo simulated inhomogeneity for ♀ _wt_, ♂ _wt_ and ♂ _bab^PR72_.
- **d** Example distributions of Monte Carlo simulated \( \Delta C_{\text{inhomogeneity}}(s1, s2) \) for ♀ _wt_, ♂ _wt_ and ♂ _bab^PR72_.

![](image.png)
Figure B: Histograms of angle of rotation $\alpha$ of all simulated intact SCs without temporal dynamics. Each histogram shows counts of a single SC length (5, 7, 9 or 11-tooth SC).
Figure C: Histograms of ABASCT SD of all simulated intact SCs without temporal dynamics. Each histogram shows counts of a single SC length (5, 7, 9 or 11-tooth SC).
Figure D: Graph of aggregate ABASCT SD statistics of simulated and intact rotated SCs without temporal dynamics, grouped by adhesion parameters. This graph is based on the same data as Fig 7B, except here the horizontal centre line of each box is the mean. Height of each box from the horizontal centre line represents ± 1SEM (standard error of the mean). The upper (lower) extent of each top (bottom) whisker represents the maximum (minimum) ABASCT SD value of that category.
Figure E: Linear model fit of treatment (number of SC teeth) effects on ABASCT SDs of simulated intact SCs without temporal dynamics.

95% family-wise confidence level

![Graph showing the difference in ABASCT SDs with 95% family-wise confidence level.](image-url)
Figure F: Box plot of ABASCT SD of intact SC simulations without temporal dynamics, but with regular 5-tooth simulations replaced by the “slowed down” version, grouped by adhesion between SC teeth.
Figure G: Histograms of angle of rotation $\alpha$ of all simulated intact SCs grouped by magnitude of delayed expansion of EP1 cells (delay=0 means absence of temporal dynamics). Each histogram shows counts of a single SC length (5, 7, 9 or 11-tooth SC).
Figure H: Histograms of ABASCT SD of all simulated intact SCs grouped by magnitude of delayed expansion of EP1 cells (delay=0 means absence of temporal dynamics). Each histogram shows counts of a single SC length (5, 7, 9 or 11-tooth SC).
Figure I: An example 9-tooth SC simulation which uses 4 times as many pixels as the original ones.

**a** Start of simulation ($t=0$ mcs).

**b** End of simulation ($t=2000$ mcs).
Figure J: Confocal image of an example fly leg showing the results of individual cell tracking according to two cell labelling definitions.

**a** The demarcation line is drawn perpendicular to the SC at the halfway length mark of the SC at 23 hours AP.

**b** The demarcation line is drawn horizontally at the halfway length mark of the SC at 23 hours AP.

Regardless of definitions, the EP2 cells (magenta) on average expand much more than EP1 cells (blue). For example, in **a**, the increase in area of EP2 cells is 187% vs. EP1 cells at 36%. Each scale bar of **a** and **b**: 20 μm.
## Tables

| Fly type       | Pupa Number | Initial EP1 [\(\mu m^2\)] | Initial EP2 [\(\mu m^2\)] | Final EP1 [\(\mu m^2\)] | Final EP2 [\(\mu m^2\)] |
|----------------|-------------|-----------------------------|-----------------------------|---------------------------|---------------------------|
| \(\vartheta\text{wt}\) | 1           | 6.4                         | 6.3                         | 6.9                       | 5.0                       |
| \(\vartheta\text{wt}\) | 2           | 5.8                         | 5.9                         | 6.7                       | 5.3                       |
| \(\vartheta\text{wt}\) | 3           | 6.6                         | 3.7                         | 5.7                       | 3.8                       |
| \(\vartheta\text{wt}\) | 4           | 3.5                         | 5.7                         | 4.2                       | 3.1                       |
| \(\vartheta\text{wt}\) | 5           | 5.9                         | 5.8                         | 7.1                       | 3.6                       |
| \(\sigma\text{bab}^{PR72}\) | 1           | 4.1                         | 2.3                         | 6.6                       | 6.8                       |
| \(\sigma\text{bab}^{PR72}\) | 2           | 5.3                         | 2.1                         | 11.4                      | 7.5                       |
| \(\sigma\text{bab}^{PR72}\) | 3           | 4.7                         | 3.5                         | 9.3                       | 8.6                       |
| \(\sigma\text{bab}^{PR72}\) | 4           | 5.1                         | 4.0                         | 7.2                       | 8.1                       |
| \(\sigma\text{bab}^{PR72}\) | 5           | 5.6                         | 2.6                         | 6.7                       | 7.3                       |
| \(\sigma\text{wt}\) | 1           | 3.6                         | 1.6                         | 9.1                       | 8.2                       |
| \(\sigma\text{wt}\) | 2           | 2.9                         | 1.4                         | 9.0                       | 6.3                       |
| \(\sigma\text{wt}\) | 3           | 3.3                         | 1.1                         | 9.9                       | 7.7                       |
| \(\sigma\text{wt}\) | 4           | 4.8                         | 2.0                         | 9.8                       | 8.1                       |
| \(\sigma\text{wt}\) | 5           | 4.2                         | 1.9                         | 10.5                      | 9.3                       |

Table A: Raw experimental data of area measurement in Figure 4.
### Table B: Monte Carlo simulated quantities on bootstrapped data sets.

| Quantity (↓); Quantile (→) | 0%  | 25%  | 50%  | 75%  | 100% |
|---------------------------|-----|------|------|------|------|
| $Pr(\frac{EP\_{2\text{Initial}}}{EP\_{1\text{Initial}}} < 1)$ |   |      |      |      |      |
| $Pr(\frac{EP\_{2\text{Initial}}}{EP\_{1\text{Initial}}} < 1)$ | 0.03 | 0.36 | 0.54 | 0.69 | 1.00 |
| $Pr(\frac{EP\_{2\text{Initial}}}{EP\_{1\text{Initial}}} < 1)$ | 0.95 | 0.98 | 0.99 | 1.00 | 1.00 |
| $Pr(\frac{EP\_{2\text{Initial}}}{EP\_{1\text{Initial}}} < 1)$ | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |

### Table C: Results of Tukey post-hoc hypothesis testing on the equality of treatments (number of SC teeth) on ABASCT SDs of simulated intact SCs without temporal dynamics. The (null) hypothesis $H_0(A,B)$ tested in each row is that there is "no treatment effect on ABASCT SDs between A-tooth SCs and B-tooth SCs (informally, A-B=0)."

| A | B | Estimate [°] | Error [°] | t value | Pr(>|t|)   |
|---|---|--------------|----------|---------|-----------|
| 7 | 5 | 2.13         | 0.48     | 4.43    | <0.001*** |
| 9 | 5 | 8.62         | 0.52     | 16.69   | <0.001*** |
| 11| 5 | 7.00         | 0.54     | 13.07   | <0.001*** |
| 9 | 7 | 6.49         | 0.53     | 12.2    | <0.001*** |
| 11| 7 | 4.88         | 0.55     | 8.85    | <0.001*** |
| 11| 9 | -1.62        | 0.59     | -2.77   | 0.029*    |
| A  | Delay parameter used in set with A intact SCs [mcs] | B  | Delay parameter used in set with B intact SCs [mcs] | Length | $J(\text{SC T}, \text{SCT})$ | p-value      |
|----|--------------------------------------------------|----|--------------------------------------------------|--------|-----------------------------|-------------|
| 43 | 0                                                | 48 | 120                                              | 11-tooth | 0                           | $\leq 0.0082^{**}$ |
| 41 | 0                                                | 47 | 120                                              | 11-tooth | 2000                        | $\leq 0.0082^{**}$ |
| 38 | 0                                                | 42 | 120                                              | 11-tooth | 4000                        | $\leq 0.077$   |
| 38 | 0                                                | 43 | 200                                              | 11-tooth | 4000                        | $\leq 0.046^{*}$ |
| 32 | 0                                                | 37 | 120                                              | 11-tooth | 6000                        | $\leq 0.068$   |
| 32 | 0                                                | 41 | 200                                              | 11-tooth | 6000                        | $\leq 0.0076^{**}$ |
| 6  | 0                                                | 17 | 120                                              | 11-tooth | 8000                        | $\leq 0.0019^{**}$ |

Table D: Results of hypothesis testing on the equality of SC breaking probabilities between two sets of simulations. The (null) hypothesis $H_0(A, B)$ tested in each row is “a simulation set that has at most A intact SCs and another simulation set that has at least B intact SCs share a common probability distribution of SC breakage”.

