



# Reverse engineering morphogenesis through Bayesian optimization of physics-based models



Nilay Kumar<sup>1</sup>, Mayesha Sahir Mim<sup>1,2</sup>, Alexander Dowling<sup>1</sup> & Jeremiah J. Zartman<sup>1,2</sup>

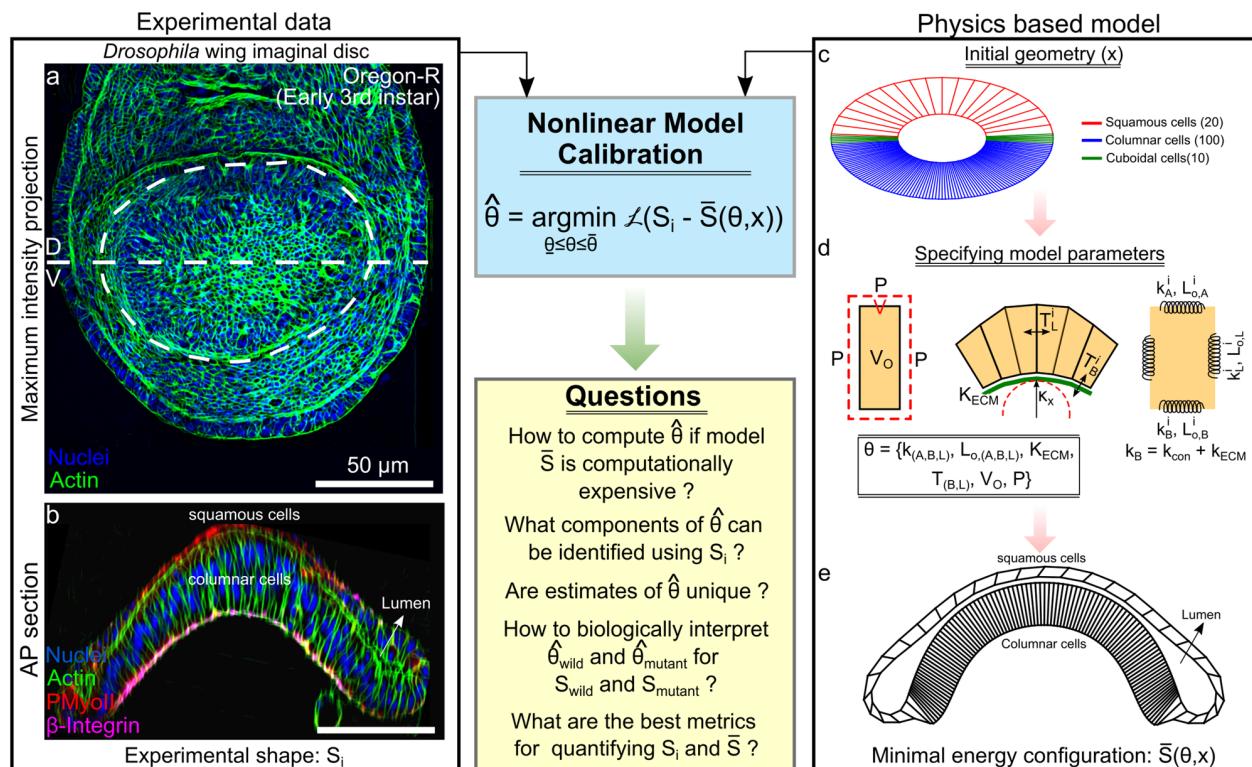
Morphogenetic programs coordinate cell signaling and mechanical interactions to shape organs. In systems and synthetic biology, a key challenge is determining optimal cellular interactions for predicting organ shape, size, and function. Physics-based models defining the subcellular force distribution facilitate this, but it is challenging to calibrate parameters in these models from data. To solve this inverse problem, we created a Bayesian optimization framework to determine the optimal cellular force distribution such that the predicted organ shapes match the experimentally observed organ shapes. This integrative framework employs Gaussian Process Regression, a non-parametric kernel-based probabilistic machine learning modeling paradigm, to learn the mapping functions relating to the morphogenetic programs that maintain the final organ shape. We calibrated and tested the method on *Drosophila* wing imaginal discs to study mechanisms that regulate epithelial processes ranging from development to cancer. The parameter estimation framework successfully infers the underlying changes in core parameters needed to match simulation data with imaging data of wing discs perturbed with collagenase. The computational pipeline identifies distinct parameter sets mimicking wild-type shapes. It enables a global sensitivity analysis to support the regulation of actomyosin contractility and basal ECM stiffness to generate and maintain the curved shape of the wing imaginal disc. The optimization framework, combined with experimental imaging, identified that Piezo, a mechanosensitive ion channel, impacts fold formation by regulating the apical-basal balance of actomyosin contractility and elasticity of ECM. This workflow is extensible toward reverse-engineering morphogenesis across organ systems and for real-time control of complex multicellular systems.

Reverse engineering biological systems requires mathematical tools to infer interactions and mechanisms of a given process to enable forward engineering applications<sup>1,2</sup>. For instance, a detailed understanding of the pathways involved during a wound-healing process can help design new treatment plans or drugs<sup>3,4</sup>. The calibration of computational models of biological systems is challenging due to the large number of interactions whose mathematical description encompasses a very large parameter space. Inevitably, there exists a tradeoff between computational speed and level of detail. For instance, subcellular element models of epithelial morphogenesis can involve hundreds of parameters with an average computational time on the order of days<sup>5</sup>. These models recapitulate a wide variety of biological processes across multiple model organisms with increasing levels of detail.

However, before any model can make new predictions, it must be calibrated and validated against experimental data.

On the other hand, in-vivo biological experiments also are costly and time-consuming. The number of features or variables that can be measured in the lab is often limiting. A wide range of forces that often have nonlinear formulations jointly contribute to the shape of an organ. Identifying parameter space that defines similar tissue shapes is the first step in understanding morphogenetic robustness. Moreover, an ability to parametrize a shape also can allow for a more robust comparison between the control, wild-type condition, and a mutant organ shape yielding physical insights (Fig. 1). For example, a similar methodology of parametrizing signaling data demonstrates that projecting raw  $\text{Ca}^{2+}$  signatures from single cells into a

<sup>1</sup>Department of Chemical and Biomolecular Engineering, University of Notre Dame, Notre Dame, IN 46556, USA. <sup>2</sup>Bioengineering Graduate Program, University of Notre Dame, Notre Dame, IN 46556, USA. e-mail: [jzartman@nd.edu](mailto:jzartman@nd.edu)



**Fig. 1 | Initial model formulation recapitulates the morphological features of the wing disc.** **a** Apical view of a z-projection of a *Drosophila* 3<sup>rd</sup> instar wing imaginal disc. **b** Cross section of the tissue along the anterior-posterior (AP) axis offset from the dorsal-ventral (D, V) boundary. **c** The initial geometry used for Surface Evolver

simulations. **d** Definition of subcellular cytoskeletal interactions used to define the system's total energy. **e** Minimum energy configuration obtained after optimization for parameters in Table 3. Note: The loss function  $\mathcal{L}$  measures the error between the experimental shape  $S_i$  and the predicted shape  $\bar{S}(\theta, x)$  from the model.

more meaningful parameter space of a physics-based model using Approximate Bayesian Computations led to the discovery of four distinct cellular states<sup>6</sup>.

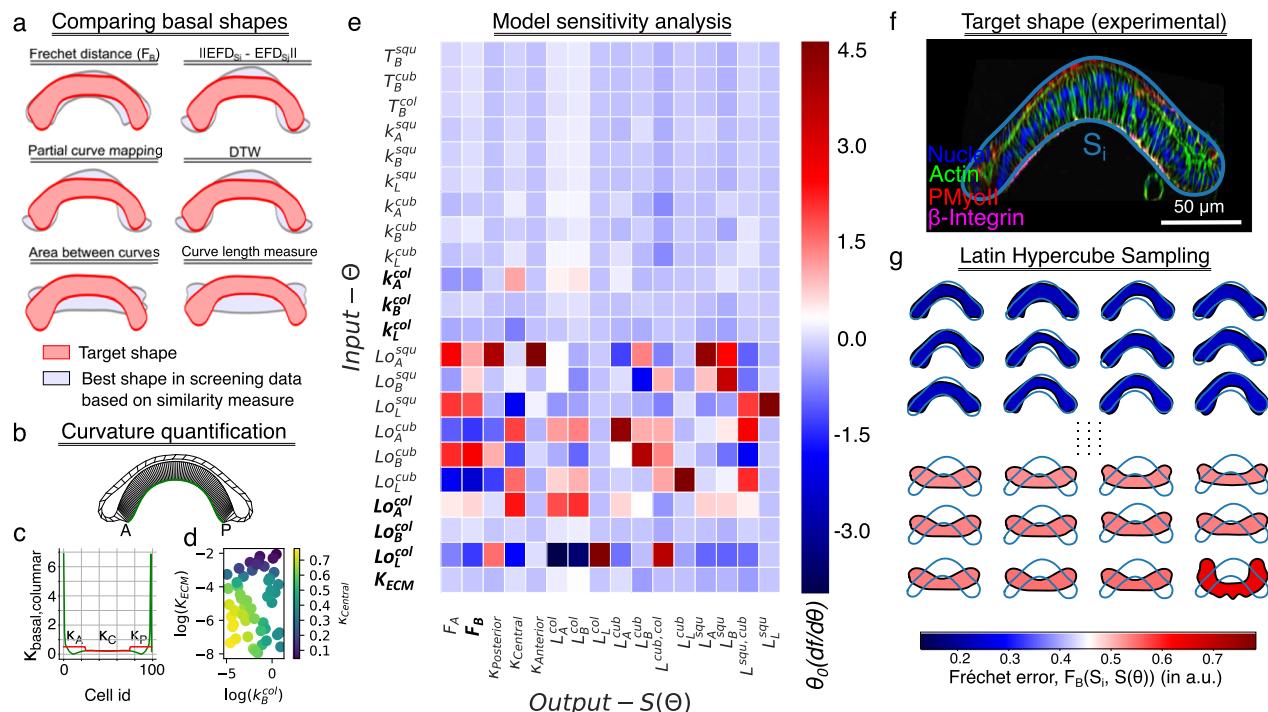
A typical modeling workflow includes the following steps<sup>7</sup>. The overall complex process is conceptually decomposed into a subset of biological processes. These individual subprocesses are first calibrated using a subset of the measurable experimental data before moving towards calibration of the entire process<sup>8</sup>. However, the many interactions between subprocesses often prevent identifying the global optimum. Also, choosing the best error function for comparing experimental data with model output is challenging. Methods of calibrating computational biology models include nonlinear least squares regression<sup>9</sup>, maximum likelihood estimation (MLE), maximum a posteriori (MAP) estimation<sup>10</sup>, Markov chain Monte Carlo (MCMC)<sup>11</sup>, and genetic algorithms<sup>12</sup>, among others<sup>13–16</sup>.

Each of these algorithms has advantages and disadvantages. While least squares and MLE parameter estimation methods typically exhibit fast local convergence, they often get stuck at a local minimum using gradient-based optimization methods. MCMC avoids these (sometimes poor) locally optimal parameter point estimates by inferring a posterior distribution, which often requires at least an order of magnitude more computational effort. For instance, MCMC has been used to estimate parameters of ordinary differential equation-based models in systems biology<sup>11</sup>. Another approach followed is Sequential Monte Carlo Approximate Bayesian Computations (SMC-ABC), where rejection sampling is used to estimate the posterior based on a prior and has been used previously to estimate parameters of models of tissue growth<sup>17</sup> and  $\text{Ca}^{2+}$  signaling<sup>6</sup>. However, such an approach is computationally expensive as it requires calculating a distribution and may not be feasible for calibrating computationally expensive modeling frameworks like a molecular dynamic or a subcellular element (SCE) model. Moreover, these detailed modeling approaches require

calibration based on multiple measured variables within the lab setup. Consequently, developing new or hybrid approaches that leverage the strengths of multiple approaches can lead to more efficient and robust computational methods. As such, new computationally efficient and robust approaches are needed to calibrate complex mechanistic (biological) mathematical models.

In this work, we present a computational pipeline employing Bayesian Optimization<sup>18–20</sup> (BO) to infer the primary biophysical mechanisms driving the shape of an organ. We utilize the *Drosophila* wing imaginal disc cross-section shape ( $S_i$ ) for inferring the parameters ( $\theta$ ) of a biophysical model ( $S$ ) of the wing disc cross-section (Fig. 1). In general, the framework allows projection of the shape of an organ to a more meaningful parameter space describing biophysical mechanisms driving organ shape generation and maintenance. In this work, we used Surface Evolver<sup>21</sup> for simulating the wing disc cross-section. Compared to more detailed biological models of wing disc morphogenesis, such as our collaborative SCE model<sup>12</sup>, the simplified model generated by Surface Evolver provides a testing platform to assess the utility of the approach toward model calibration of multicellular systems.

To increase the computational efficiency, the framework couples the mechanistic model to Gaussian process regression (GPR) surrogate models to map the model parameters to the quantitative objective functions while considering uncertainty in model prediction (Fig. 2). The surrogate model is then used to sample new points based on an acquisition function that guides the sampling of an optimal solution. While the uncertainty with GPR models can be significant, they are routinely used for real-time control and automation<sup>23</sup>. A critical theme of the predictive control literature in the past few decades is that a mediocre model is still valid due to the power of feedback. GPRs are



**Fig. 2 | Selection of the optimal quality of fit metric and input parameters for surrogate model training dataset.** **a** Evaluating similarity measures for computing the objective function. **b, c** Curvature ( $\kappa$ ) along the basal surface of the columnar cells is plotted as green lines. The basal surface is divided into three equal regions (Anterior-A, Central-C, Posterior-P). The average basal curvature is plotted in red for the subregions. **d** A scatter plot visualizes 50 points sampled by varying model parameters  $k_B^{\text{col}}$  and  $K_{\text{ECM}}$  using LHS. Points are color-coded based on the averaged

central curvature ( $\kappa_C$ ) of the model output. **e** Heatmap showing a sensitivity measure represented by the color bar legend. The vertical axis represents the input parameter values of the physics-based model, while the horizontal axis represents different morphological features extracted from an output shape. **f** A representative AP cross-section of early 3<sup>rd</sup> instar wing imaginal disc (~84–90 h AEL). **g** Different shapes generated using the parameter sampling are arranged as per increasing Fréchet errors with respect to the representative experimental cross-section.

especially attractive because they are non-parametric models that “learn” as new data are incorporated. This also prevents the model from getting trapped in a local minimum. BO is well suited to problems with a large parameter space and easily handles constraints. As such, BO has been applied in a wide range of research areas<sup>24</sup> including parameter estimation for computationally expensive scientific models<sup>25</sup>. Moreover, several prior studies have demonstrated GPR surrogate models are computationally efficient emulators of biological processes<sup>16</sup>, including post-transcriptional regulation in *Drosophila*<sup>26</sup>, dynamics of microbial systems<sup>27</sup>, cancer tumor growth<sup>28</sup>, and biopharmaceutical manufacturing<sup>29</sup>.

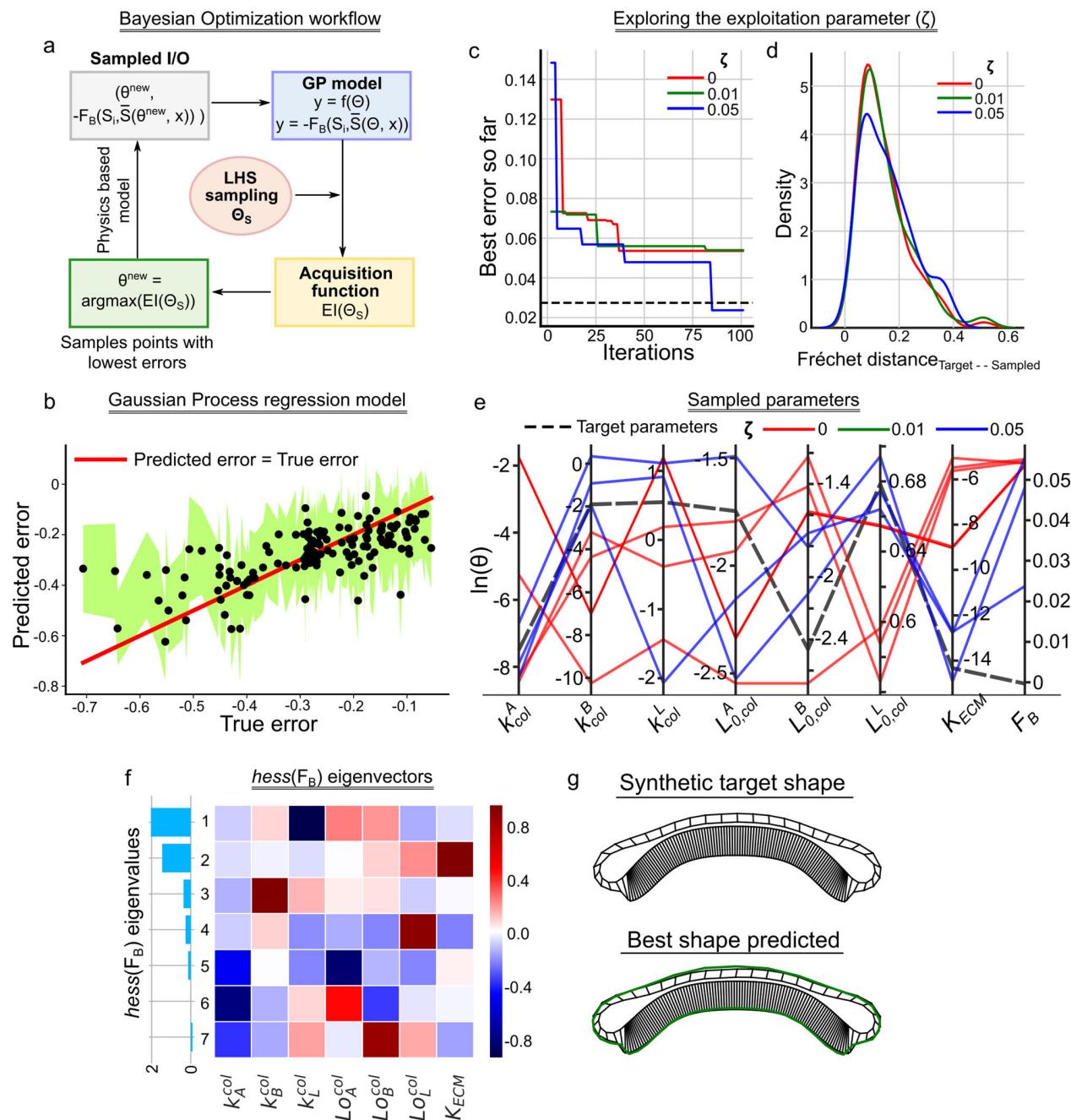
We first benchmark the pipeline using a synthetic tissue shape with known model parameters (Fig. 3). Post benchmarking, we employ the pipeline to predict the morphology of the wing disc undergoing degradation of the extracellular matrix (ECM). To do so, we utilized confocal microscopy data from our earlier investigation<sup>30</sup> (Fig. 4). Lastly, the computational framework, along with fixed tissue imaging data, demonstrated that Piezo, a mechanosensitive ion channel, impacts fold formation within the *Drosophila* wing imaginal disc through the regulation of actomyosin contractility and cell volume (Fig. 5). This paves a direction to discover new mechanisms of mechanotransduction in cytoskeletal regulation during organ growth and morphogenesis. The contributions of this work in the domain of systems biology include (1) A successful application of BO to biophysical models of tissue morphogenesis, (2) a demonstration of Fréchet distance as a useful error metric for calibration of model parameters to define organ shape, and (3) A study highlighting the role of Piezo mediated mechanosensation in fold formation within *Drosophila* wing imaginal disc. Overall, this work provides an efficient pipeline to infer biophysical mechanisms of morphogenesis using morphological data of organs and tissue and identifies a key role of Piezo in regulating tissue shape during wing disc development.

## Results

### Formulation of a flexible mechanistic model that recapitulates the morphology of the wing disc cross-section

We first utilized Surface Evolver<sup>21</sup> to formulate a model of the anterior-posterior (AP) cross-section of the *Drosophila* wing imaginal disc (Fig. 1a, b). This approach enabled us to systematically test the utility of benchmarking the Bayesian optimization framework of nonlinear model calibration. The *Drosophila* wing imaginal disc is an established model system for studying the calibration of models of epithelial morphogenesis<sup>31,32</sup>. At later stages of development, the wing disc consists of a fluid-filled sac with a lumen surrounded by epithelial cells of different subtypes (squamous, cuboidal, and columnar) (Fig. 1c). A thin extracellular matrix (ECM) encloses the basal surface of cells<sup>33</sup>. The central oval-shaped region termed the pouch resembles a dome-like structure. The shape of the pouch along each direction is patterned by highly conserved morphogens<sup>30,34–36</sup>. The geometrical attributes of wing imaginal disc on a cellular basis was estimated based on a literature review<sup>33,37</sup> (Table 1), and cell lengths were normalized to avoid numerical instabilities within the model.

A set of energy functions were defined to incorporate the contributions of known cytoskeletal regulators in the wing imaginal disc (Fig. 1d, Table 2). For instance, phosphorylated non-muscle Myosin II (pMyoII) generates contractile forces by pulling on actin cytoskeleton filaments<sup>38</sup>. This contractility drives shape changes within the tissue that include fold formation<sup>39,40</sup>. In our model, we assumed each cell edge was a Hookean spring with a natural length,  $l_0$ , and a spring constant,  $k$ . Conceptualizing cell lengths as a spring enables the modeling of length changes. Higher cell contractility increases stiffness and causes resistance to size variations, whereas lower contractility enables more deformations. The energy is calculated using

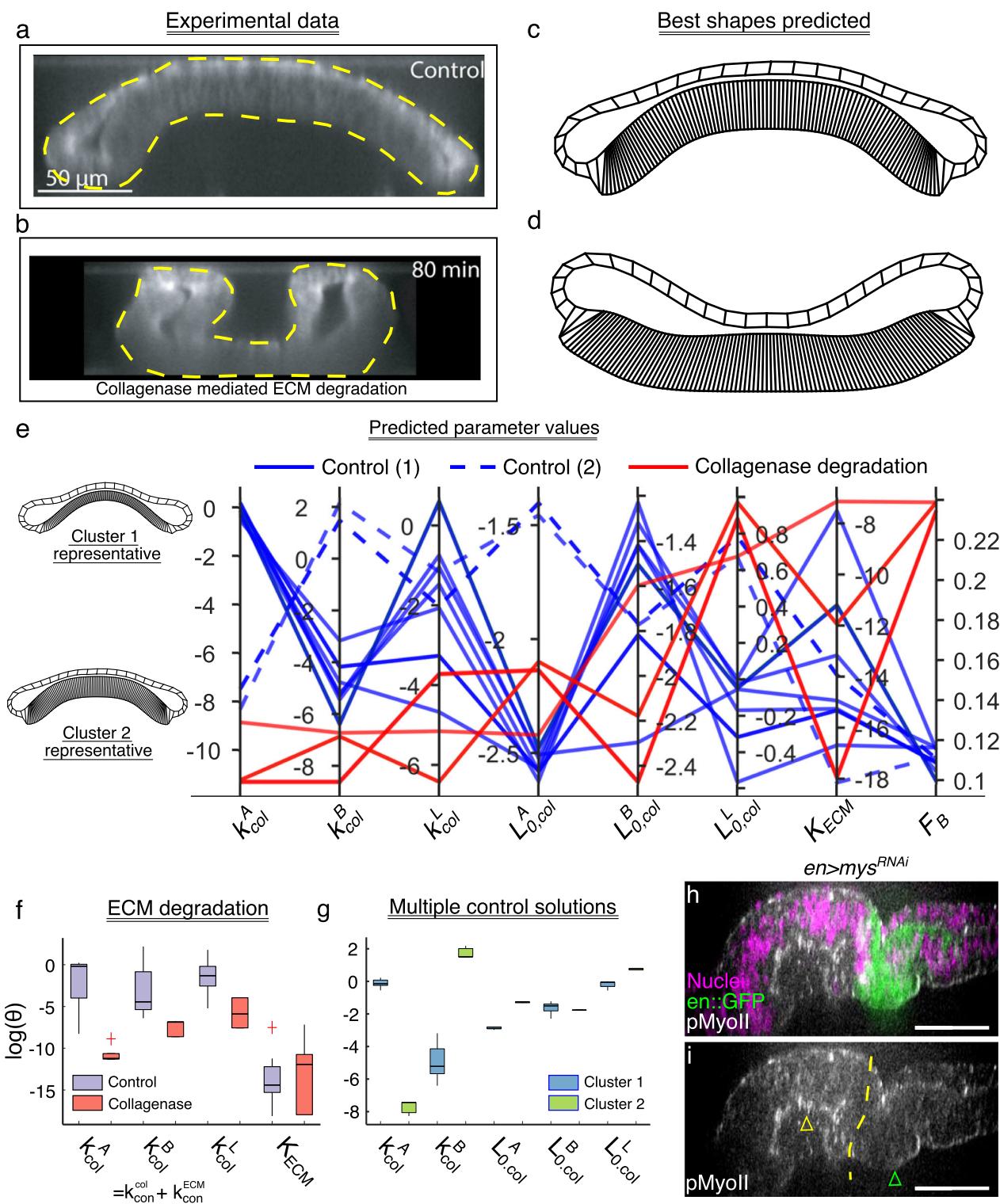


**Fig. 3 | Bayesian optimization using a Gaussian Process regression emulator model enables efficient parameter estimation based on the tissue cross-section shape.** **a** A schematic for the Bayesian optimization framework. **b** A scatter plot of the error predicted by the GPR model against the true error values. The shaded green region represents the lower and upper bound in prediction while the red solid line is a line of parity. **c** Plot showing the best error predicted so far with the number of iterations for three different exploration parameter values. **d** Kernel density plots of the error sampled during the BO for three different exploration parameter values. **e** A parallel coordinate plot showing the best parameters identified by the BO framework. The plot displays each data point as a line spanning multiple parallel vertical lines. Each vertical axis in the plot represents the estimated Surface

Evolver parameter in log scale. The position of the line on each y-axis corresponds to the value of the parameter represented by the particular y-axis. Different colors represent different  $\zeta$  values and have been included as a legend. The last vertical axis represents the  $F_B$  with respect to the target shape. **f** Eigendecomposition analysis of the local Hessian of  $F_B$ . Each row of the heatmap corresponds to the eigenvector of the Hessian matrix. The rows are arranged in decreasing order of the corresponding eigenvalues represented by a bar plot on the left-hand side of the heatmap. **g** Tissue geometry for the synthetic target shape and the best shape obtained by the BO framework. The green contour on top of the best shape predicted represents the basal contour of the synthetic target shape.

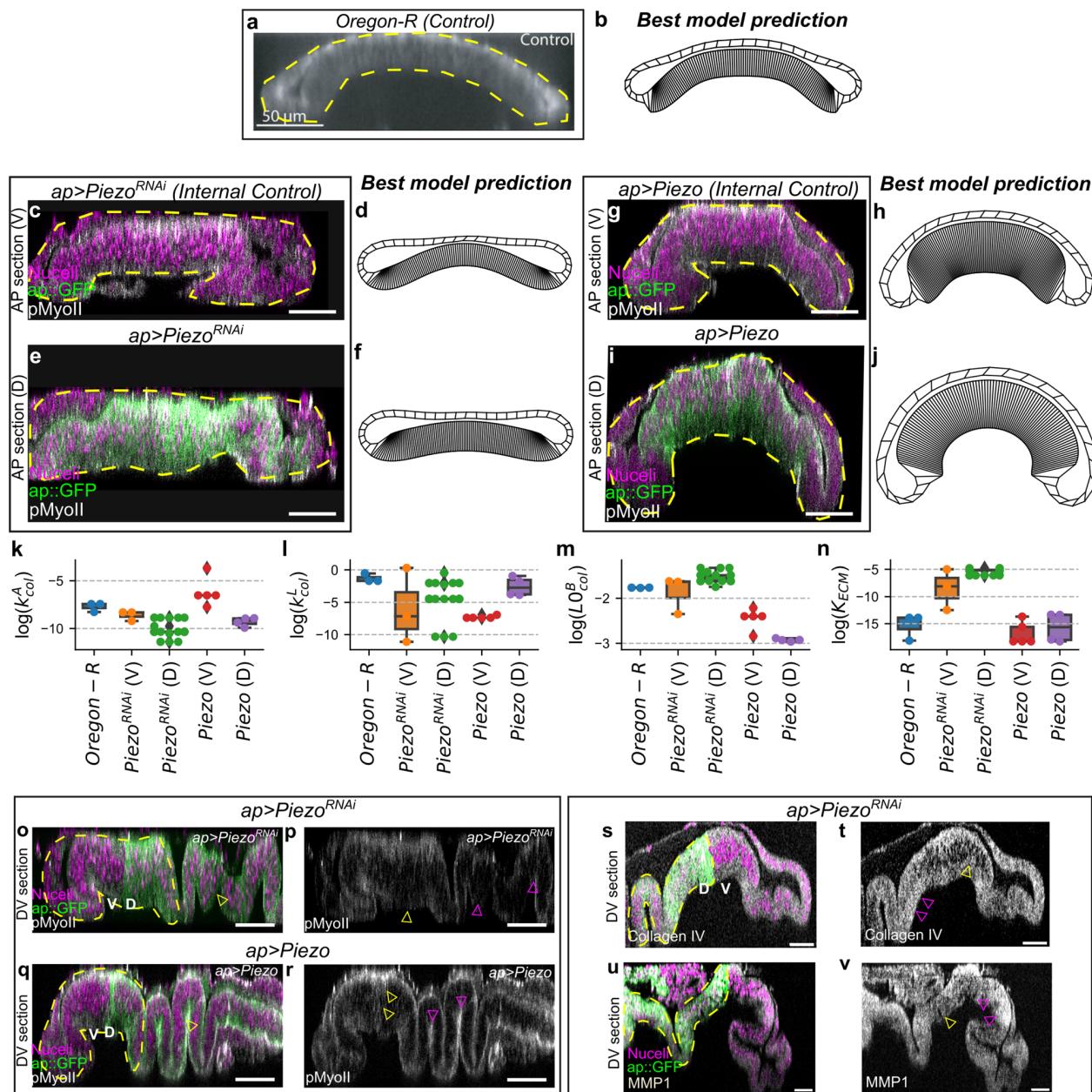
Hooke's law. Energy is also defined for each individual cell and lumen to penalize changes in volume. The target volume of each component is defined, and Surface Evolver calculates the product of a user-defined constant pressure and any changes in the target volume to estimate the

volume energy due to relative compressibility. Apical and lateral adhesion of cells is primarily mediated by apical localization of E-Cadherin and is modeled through the definition of lateral tension<sup>41</sup>. Further, the basal surface of the tissue adheres to the extracellular



**Fig. 4 | Bayesian Optimization successfully recapitulates the experimental perturbation using only the experimental output shape.** **a, b** DV cross-section of wing discs before and after treatment with Collagenase. **c, d** Best shapes predicted by the BO framework in response to the experimental data. **e** A parallel coordinate plot showing the best parameters identified by the BO framework. Different colors represent control and mutant samples and have been included as a legend. Each vertical axis represents one of the parameters represented in a log scale. The last vertical axis represents the  $F_B$  with respect to the target shape. **f** Box plot showing the

variation of parameters for the best shapes predicted by the BO framework. **g** Box plot showing the variation of parameters for the two clusters of parameter sets recapitulating the control “wildtype” shape. The boxes indicate the interquartile range (25th to 75th percentiles), the line inside the box marks for the median, and lines extending from the box represents the range from the 10th to the 90th percentiles. **h, i** Expression of pMyoII in the AP cross sections of discs expressing  $en > mys^{RNAi}$  ( $n = 5$ ).



**Fig. 5 | Piezo regulates cell height and fold formation in the *Drosophila* wing imaginal disc.** An ap-Gal4 driver was used to downregulate and overexpress Piezo in the dorsal compartment of the wing disc using UAS-Piezo<sup>RNAi</sup> and UAS-Piezo transgenic fly lines, respectively. **a** AP cross section of the Oregon-R disc has been used as a global control for comparison. Further, the ventral compartments are also interpreted as an “internal” control, given that the perturbation does not explicitly change Piezo expression levels in that compartment. However, some morphological changes are evident compared to the wild-type strain. It should be noted that the variation of parameters is lower for Oregon-R. Additionally, since tissue mechanics is interconnected, the comparison of the ventral compartment with Oregon-R should not be made directly, as perturbations in one compartment can influence the biomechanics and morphology of the other. **b** The best shape prediction corresponding to the control data. **c, e** AP cross sections along the ventral (V, control) and dorsal (D, mutant) compartments of the disc expressing *ap>Piezo<sup>RNAi</sup>*. Fluorescent

labels are indicated within the plot ( $n = 8$ ). **d, f** Best shapes predicted by the BO framework in response to the experimental data in A-i and A-ii. **g-j** A similar analysis as A for *ap>Piezo*. **k-n** Box plot showing the variation of parameters (x-axis) for the best shapes predicted. Individual data points have been scattered over the boxes. **o-r** Expression of pMyoII in the DV cross sections of mutant discs. The yellow dashed line represents the approximate dorsal-ventral compartment boundary. Yellow arrows indicate changes in pMyoII and fold formation within the dorsal pouch compartment. At the same time, the magenta arrow suggests changes in the hinge and notum portion of the wing imaginal disc. **(e: n = 7, f: n = 8)** **(s-v)** Piezo regulates ECM elasticity by controlling MMP1 and Collagen IV expression levels. This is demonstrated by the expression levels of **(g) Collagen IV** ( $n = 7$ ) and **(h) Matrix metalloproteinases 1 (MMP1)** ( $n = 8$ ) in the DV cross sections of *Piezo* knockdown discs. All scale bar = 50  $\mu$ m.

matrix (ECM) via Integrin adhesion molecules<sup>42,43</sup>. This adhesion is modeled as an additional tension residing in the basal cell edges. Lastly, the ECM is modeled as an elastic string where the energy is evaluated as the integral of the squared curvature over the length of the

string. The formulation is similar to one adopted by Storgel et al.<sup>44</sup>. Additionally, the basal contractility ( $k_B$ ) is defined as the sum of actomyosin-mediated contractility at the basal surface ( $k_{con}$ ) and the ECM ( $k_{ECM}$ ). All the energy functions defined are available as

**Table 1 | Geometrical attributes of a wing imaginal disc**

Attribute	Value	Citations
Columnar cell apical area	10 $\mu\text{m}^2$	<a href="#">37</a>
Columnar cell height	50 $\mu\text{m}$	<a href="#">22,30,37</a>
Columnar cell volume	250 $\mu\text{m}^3$	<a href="#">37</a>
Squamous cell height	8 $\mu\text{m}$	<a href="#">30,33</a>
$L_{\text{squamous.apical}}$	5	<a href="#">30</a>
$N_{\text{cells.columnar}}$	100	<a href="#">22</a>

The quantities reported above are the mean values of the attributes. Note: All cell lengths are normalized against the maximum edge length as surface evolver is known to work best for edges and volumes less than 1. Citations for the numerical values used for the initial condition are included in the table's rightmost column.

**Table 2 | Formulation of energy terms included in the Surface Evolver simulations**

$E_v$	Compressibility energy	$P * V$
$E_{\text{spring}}$	Hookean energy (actomyosin contractility)	$k * (l(\mathbf{e}) - l_0(\mathbf{e}))^n$
$E_{\text{adhesion}}$	Cell Tension	$T(\mathbf{e}) * l(\mathbf{e})$
$E_{\text{ECM}}$	Energy of extracellular matrix	$K_{\text{ECM}} \int^{\text{basal}} \kappa^2 dl$

In our model,  $P$  and  $V$  represent the system pressure and volume of the individual body within the simulation. Our model defines each cell as a closed object.  $k$  and  $n$  denote the spring constant and Hooke's power for the Hooke's energy term.  $l$  and  $l_0$  represent the actual and set cell lengths for edges ( $\mathbf{e}$ ) of the closed body, respectively.  $T$  denotes the line tension parameter. Lastly  $K_{\text{ECM}}$  and  $\kappa$  are the bending rigidity and the local curvature of the extracellular matrix, respectively.  $dl$  represents the infinitesimal length element within the 2D approximation of the basal surface along which the integral of  $\kappa^2$  is evaluated. This model characterizes the bending elasticity of the ECM.

sub-routines to be called within the Surface Evolver environment. We also define a customized repulsion module to stop the apical edges from crossing each other. After every specified number of iterations (50), the repulsion subroutine is called that tracks the distance of each node in the apical surface with the center of all the other apical edges. Suppose the measured distance is less than a particular threshold. In that case, the vertices are shifted away by a minimal distance in the opposite direction normal to the line joining the vertex and the center of the edge.

Surface Evolver minimizes the system's total energy under shape constraints to obtain a minimal energy configuration<sup>21</sup>. As a qualitative validation, we chose model parameters based on known experimental constraints (Table 3). Over time, the system's total energy decreases and then converges to a stable minimum value (Supplementary Fig. 1). This minimum energy configuration qualitatively resembles the experimental cross-section of the wing imaginal disc (Fig. 1e). The total computational time for convergence of a typical simulation is around 40 minutes using a desktop workstation running Ubuntu 20.04.2 LTS with an Intel® Xeon(R) CPU E5-1603 v3 @ 2.80 GHz × 4 processor and 16 GB of RAM. This is ideal for benchmarking a nonlinear model calibration framework as the model provides sufficient detail to capture salient features of the cross-sectional shape but is not too computationally expensive to preclude analysis of the overall pipeline. In the next sections, we describe the sensitivity of model parameters and a framework to compare the error between the model-generated shape with the experimental cross-section.

#### Defining the input-output data for the surrogate model

Fréchet error is the best metric for comparisons between tissue shapes. Here, we first describe a methodology of comparing

**Table 3 | Model parameters: the italicized parameters are to be defined within the function that writes the initial geometry file**

Parameter name	Definition	Value
$N_{\text{squamous}}$	Number of squamous cells	20
$N_{\text{cuboidal}}$	Number of cuboidal cells	10
$N_{\text{columnar}}$	Number of columnar cells	100
$\text{Pressure}$	Pressure of system	0.001
$V_{\text{Lumen}}$	Volume of lumen	15
$V_{\text{squ.cub.col.0.cells}}$	Target volume of individual cell type	(0.6,0.6,0.36)
$T_{\text{api.bas.lat.squ}}$	Tension of edges of squamous cells	0
$T_{\text{api.bas.lat.cub}}$	Tension of edges of cuboidal cells	0
$T_{\text{api.bas.lat.col}}$	Tension of edges of columnar cells	0
$k_{\text{api.bas.lat.squ}}$	Spring constant of edges of squamous cells	(0.1,0.1,1,10)
$k_{\text{api.bas.lat.cub}}$	Spring constant of edges of cuboidal cells	(0.1,0.1,0.1)
$k_{\text{api.bas.lat.col}}$	Spring constant of edges of columnar cells	(0.1,10,0.1)
$L_{\text{api.bas.lat.0.squ}}$	Hooke length of edges of squamous cells	(1,1,0.6)
$L_{\text{api.bas.lat.cub}}$	Hooke length of edges of cuboidal cells	(0.6,0.6,0.6)
$L_{\text{api.bas.lat.col}}$	Hooke length of edges of columnar cells	(0.2,0.2,0.3)
$k_{\text{squ.cub}}$	Spring constant at cuboidal -squamous interface	0.01
$k_{\text{col.cub}}$	Spring constant at cuboidal -columnar interface	0.01
$L_{\text{squ.cub}}$	Hooke length of cuboidal-squamous interface	0.6
$L_{\text{col.cub}}$	Hooke length of cuboidal-columnar interface	1.8
$n$	Power law constant for spring energy model	2
$K_{\text{ECM.basal}}$	Modulus of squared curvature energy at ECM and basal cells	0.01
$K_{\text{apical.lumen}}$	Modulus	0.001

The other parameters are to be declared within the data file that is used to run Surface Evolver.   
sq squamous, cub cuboidal, col columnar, api apical, bas basal, lat lateral.

the outer (basal) contours of any two-wing disc cross sections. The input to a model of the wing disc cross section is the set of parameters ( $\theta$ ) representing cytoskeletal regulation (Fig. 1d). For model calibration, an objective function comparing the target experimental shape ( $S_i$ ) and Surface Evolver generated cross-section ( $\hat{S}(\theta, x)$  where:  $x$  represents the initial geometry) needs to be defined (Fig. 1). Elliptic Fourier Descriptors (EFD)<sup>45</sup> are used to normalize the basal surface of epithelia against the size and translation given the experimental data and the simulated cross sections are of arbitrary length scales.

To select the best similarity measure for comparing two contours in general, we employed the following strategy: A random simulation was selected from the parameter screening data, and an error was computed based on several similarity measures with respect to the other cross-sections generated within the screen. In particular, we tested the following metrics for comparing two cross-sectional shapes (Fig. 2b):

- Area between curves, which is used to compute the area enclosed between any two curves.
- Curve length measure, which is a metric that is used to compare two curves based on their total arc length.
- Elliptic Fourier Descriptors (EFD), which are the calculated Fourier coefficients of the chain-encoded closed contour. This shape descriptor is both rotation and translation invariant. An L2-norm between the EFD coefficients of two shapes is used to quantify similarity<sup>45</sup>.
- Partial Curve Mapping, which is a method of comparing curves through alignments of the smaller curve regions.

- Dynamic Time Warping (DTW), which compares similarities between two signals. It works by distorting one of the signals to maximize its alignment with the other signal with which it is compared<sup>46</sup>.
- Fréchet distance, which is defined as the shortest distance between any two curves given one, is allowed to traverse along the two curves with different speeds<sup>47</sup>.

The best shapes identified using each of the similarity measures listed above for the randomly selected target shape are plotted in Fig. 2a. Qualitatively, Fréchet distance<sup>47</sup> ( $F_B$ ) performed the best in identifying the basal contour most closely approximating the target shape. One advantage of Fréchet distance is that the measure in itself is an error quantification. This simplifies data-driven modeling as it also serves as the objective function. We also report average apical ( $L_A$ ), basal ( $L_B$ ), and lateral ( $L_L$ ) lengths of each cell subtype along with the average basal curvature for the anterior ( $\kappa_{Anterior}$ ), medial ( $\kappa_{Medial}$ ), and the posterior ( $\kappa_{Posterior}$ ) halves of the wing disc (Fig. 2b).

**Parameter sensitivity analysis identifies cell contractility as a key regulator of tissue shape.** A sensitivity analysis was carried out to study the effect of varying  $\theta$  on overall tissue shape ( $\bar{S}(\theta, x)$ ) (Fig. 2c, Supplementary Fig. 2). Each parameter  $\theta$  was increased and decreased by 70% of its original value. Morphological features within  $\bar{S}(\theta, x)$  were measured, and a central finite difference scheme was used to compute the sensitivity of model parameters. As expected, changing the natural lengths of apical, basal, or lateral edges of the squamous, cuboidal, and columnar cells ( $L_{0,(A,B,L)}^{squ,cub,col}$ ) caused the most changes across any of the measured features. A loss in the natural length of the cell often represents the cell's failure to regulate actin polymerization and depolymerization. Dysregulation of actin polymerization causes severe morphological defects<sup>48</sup>.

A change in the spring constants of either of the apical, basal, and lateral edges of a columnar cell ( $k_A^{col}, k_B^{col}, k_L^{col}$ ) led to changes in the shape of the tissue. It should be noted that the basal contractility in our model is a sum total of contractility generated by the basal actomyosin complex and the ECM. This agrees with our previous work, where we used a more detailed subcellular element model (SCE model) of a wing imaginal disc to show that the tissue shape is mainly generated by actomyosin contractility and maintained by the ECM<sup>30</sup>. We also found that varying the contractility of squamous cells did not impact overall tissue shape (Fig. 2c). This observation confirms a recent report where altering contractility and growth through downregulation of PI3K in the squamous epithelia did not significantly impact the shape of wing imaginal disc<sup>33</sup>. Besides highlighting the significance of basal cell contractility and cellular geometrical properties, our study emphasizes the pivotal role of extracellular matrix (ECM) bending rigidity ( $K_{ECM}$ ). Through simulations, we demonstrate that increasing  $K_{ECM}$  can effectively counteract the impact of basal contractility ( $k_B^{col}$ ) on basal tissue curvature, thus contributing to the maintenance of tissue shape (Fig. 2b). We sampled 50 points in the parameter space by varying  $k_B^{col}$  and  $K_{ECM}$  as shown in Fig. 2c, d. We next computed minimal energy configurations and measured the average central curvature ( $\kappa_C$ ) of the output shapes. Notably, an increase in  $K_{ECM}$  led to a reduction of  $\kappa_C$ . Additionally, we observed an almost linear trend between  $K_{ECM}$  and  $\kappa_C$  at higher limits of  $K_{ECM}$  suggesting that energy due to ECM may dominate that of actomyosin complex generated contractility. Further experimentations are required to validate these findings.

**Gaussian Process Regression surrogate model.** To further benchmark the calibration of the nonlinear model, we selected a small subset of the parameters identified through sensitivity analysis. The selected parameters are highlighted in bold along the vertical axis of the heatmap within Fig. 2e. Latin Hypercube Sampling was used to sample this reduced parameter space uniformly<sup>49</sup>. Surface Evolver was run for all the

sampled points, and the geometrical features were extracted. The sampled parameters and the corresponding morphological features constitute the input-output data for a surrogate model to be used for the optimization task. As an example, the sampled cross-sections were arranged based on an increasing Fréchet distance (top to bottom) (Fig. 2g) with respect to the target experimental cross-section  $S_i$  (Fig. 2f). A low Fréchet error corresponds to a better approximation of the experimental data.

Our work used Gaussian Process Regression (GPR) as a surrogate model<sup>50-52</sup>. GPRs, also known as Kriging models, have a rich history in surrogate-assisted optimization<sup>53</sup>, especially for problems with ten to fifty degrees of freedom. We employed a leave-one-out strategy for training where a model was trained iteratively, leaving exactly one data point out and using everything else for training<sup>52</sup>. The remaining data points were then used for the assessment of the model. With this strategy, we found that the model predictions are in good accordance with the true output.

Based on this result, we utilized all the data points within the parameter screening for the initial training of the GPR model (Fig. 3b). In the following sections, we used the described GPR model to develop a framework for Bayesian optimization of the physics-based model of wing disc morphogenesis.

#### Bayesian optimization of the Gaussian Process Regression (GPR) model of tissue shape enables parameter estimation of the physics-based models

The Bayesian optimization (BO) approach utilizes and updates a prior belief between the inputs and the outputs used for the calculation of the objective function in the form of surrogate models. The surrogate GPR model approximates the functional values for input as a Gaussian distribution allowing quantification of uncertainty in the form of covariances. This is crucial for the computation of acquisition functions, described later in the text, which is used to sample new points within the parameter space, as shown in Fig. 3a.

To benchmark the pipeline of parameter estimation, we generated a synthetic target shape with known parameters, which was not included in the training data. For training a GPR model, Fréchet distance ( $F_B$ ) is first computed for all the samples in the screening dataset with respect to the synthetic target shape. The parameters and the corresponding negative value of Fréchet distances ( $-F_B$ ) act as input and output to our GPR model. BO uses an acquisition function guided by the GPR model to draw a new parameter set that maximizes the response value of the GPR. Since we chose  $-F_B$  as our response variable, maximizing it corresponds to finding  $\Theta$  leading to the shape closest to the selected target. We use Expected Improvement as our acquisition function. This allows the sampling of new points around the region that maximizes the output of the surrogate model. The exploration parameter ( $\zeta$ ) within the Expected Improvement defines the amount of exploration during the sampling process. Higher exploration parameters tend to sample points from the regions where the GPR model uncertainty is low instead of sampling guided mainly by the mean value of the surrogate model. The pseudo-code for full pipeline is provided as Algorithm 1. This pipeline was implemented in Python (v 3.8.13) using Surface Evolver (v 2.40), Surrogate Modeling Toolbox (v 1.0.0), GPyTorch (v 1.5.0) and scikit-learn (v 0.24.2).

Three different values of the exploration parameter between 0 and 0.05 were selected to run the BO framework for a synthetic tissue cross-section whose parameter values were already known. We plotted the distribution of Fréchet errors for parameters sampled by the BO framework for the different  $\zeta$  values. A kernel density was calculated to approximate the distributions (Fig. 3d). The plots reveal that increasing  $\zeta$  tends to flatten out the distributions of errors between the tissue shapes generated by the sampled parameter and the synthetic target shape. This also suggests that the newly sampled points are farther away from the mean of the distribution as compared to points sampled using lower  $\zeta$  values.

**Algorithm 1.** Bayesian optimization using GPR surrogate models.

```

Input: Input target shape or an experimental data,  $S_i$ 
Output: Find  $\hat{\theta} = \operatorname{argmin} F_B(S_i, \bar{S}(\theta, x)) \equiv \operatorname{argmax} (-F_B(S_i, \bar{S}(\theta, x)))$ 

1 Use Latin Hypercube Sampling (LHS) to sample  $O(10^2)$  parameters and store in  $x_{train}$ 
2 Run Surface Evolver (SE)  $\forall \theta \in x_{train}$  and evaluate  $-F_B(S_i, \bar{S}(\theta, x))$ . Store the results in  $y_{train}$ 
3 Train initial GPR model with input  $x_{train}$  and  $y_{train}$ 
4 Systematically sample many points in the parameter space,  $\theta_s$ 
5 for  $i = 0, \dots, N_{iterations}$  do
6    $\theta^{best} = \operatorname{argmax} (\{-F_B(S_i, \bar{S}(\theta_j, x)): \forall \theta_j \in \theta_s\})$ 
7   for  $j = 0, \dots, N_{sampled}$  do
8     Evaluate  $\mu(\theta_j)$  and  $\sigma(\theta_j)$   $\forall \theta_j \in \theta_s$  using the trained GPR model
9     if  $\sigma(\theta_j) > 0$  do
10     $Z = (\mu(\theta_j) + F_B^{best}(S_i, \bar{S}(\theta_j, x)) - \zeta) / \sigma(\theta_j)$ 
11     $EI(\theta_j) = (\mu(\theta_j) + F_B^{best}(S_i, \bar{S}(\theta_j, x)) - \zeta) \phi(Z) + \sigma(\theta_j) \phi'(Z)$ 
12    else do
13     $EI(\theta_j) = 0$ 
14  end
15  Sample new point  $\theta^{new} = \operatorname{argmax} (\{EI(\theta_k)\}: \forall \theta_k \in \theta_s)$ 
16  Evaluate  $F_B(S_i, \bar{S}(\theta^{new}, x))$ 
17  Add  $\{\theta^{new}, -F_B(S_i, \bar{S}(\theta^{new}, x))\}$  to the training data  $\{x_{train}, y_{train}\}$  and re-train the GPR model
18 end

```

El denotes expected Improvement.  $\mu$  and  $\sigma$  are the mean and variance in GPR predictions.  $\phi$  and  $\phi'$  represent the pdf (probability distribution function) and cdf (cumulative distribution function) for a standard normal distribution.

We next plotted the best error sampled so far with each iteration in BO. The best error so far for an  $i^{\text{th}}$  iteration is defined as the minimal Fréchet distance reported till the particular step during the parameter sampling process. With increasing values of  $\zeta$ , the points sampled have a lower  $F_B$  as compared to  $F_B$  from the best shape in training data. An increase in  $\zeta$  also allows for a faster convergence (Fig. 3c). For each  $\zeta$ , we also plotted the best parameter values that lead to an approximation of the target shape closer or lesser to the best error in the training data. Each vertical axis of the parallel coordinate plot represents the log of parameter values indicated in the horizontal axis label. The dashed black lines represent the true parameter values of the synthetic target shape. The parameters of the best shapes extracted using the pipeline are in good accordance with the synthetic target parameters (Fig. 3e, g).

We next analyzed the curvature (Hessian) of the objective function, i.e.,  $F_B$ , in the vicinity of the parameters of the target shape to assess the parameters that can be approximated using  $F_B$  alone (Fig. 3f). The eigenvectors corresponding to the two largest eigenvalues predominantly point in the direction of parameters  $k_L^{\text{col}}$  and  $K_{ECM}$ , respectively. This means that the predicted organ shape, as determined by  $F_B$ , is most sensitive to these parameters. Thus,  $k_L^{\text{col}}$  and  $K_{ECM}$  can be estimated with the least uncertainty. Previous studies demonstrate that the elasticity of ECM (modeled as  $K_{ECM}$ ) plays a significant role in determining bulk organ shape<sup>22,23</sup>. A change in lateral contractility (modeled as  $k_L^{\text{col}}$ ) is also known to cause severe morphogenetic defects as shown by the shape of the cross-section<sup>22</sup>.

The eigenvectors of the third and fourth largest eigenvalues predominantly are in the direction of  $k_B^{\text{col}}$  and  $I_{0,L}^{\text{col}}$ , respectively, which means these parameters can be inferred with moderate uncertainty. Finally, the three smallest eigenvalues are near zero, and their eigenvectors predominantly point in the direction of  $I_{0,A}^{\text{col}}$ ,  $k_A^{\text{col}}$ , and  $L_{0,B}^{\text{col}}$ . Thus, these

parameters are (near) non-estimable using only  $F_B$ , and the corresponding parameter estimates are highly uncertain.

A parallel coordinate plot showing the best parameters identified by the BO framework. Different colors represent different values and have been included as a legend.

Previous studies have highlighted the role of basal contractility (modeled as  $k_B^{\text{col}}$ ) in generating and maintaining the dome shape of wing imaginal disc<sup>30</sup>. A genetic loss of apical-basal contractility through the expression of the dominant negative form of Rho, an upstream regulator of actomyosin contractility, causes the tissue to flatten out<sup>22</sup>. However, a loss of actomyosin contractility upon pharmacological treatment with ROCK inhibitors did not cause severe changes in the shape of a late development stage 3<sup>rd</sup> instar wing imaginal disc<sup>30</sup>. Genetic perturbation, achieved by expressing Rho<sup>DN</sup>, was performed in the earlier stages of morphogenesis (2<sup>nd</sup> instar larval stage). In contrast, pharmacological perturbations utilizing ROCK inhibitors were administered during the later larval stage (early 3<sup>rd</sup> instar larval stage). This suggests that once the wing disc acquires its bent shape, it is less sensitive to changes made in apical-basal actomyosin contractility. Measurements of more variables like local cell lengths can also help better approximate these parameters, which is an open avenue for future investigations.

Through these steps, we conclude that our pipeline employing BO can successfully recover a subset of parameters of the computational model from the representation of the basal surface alone. The proposed methodology allows the transformation of the shape of an organ into a more meaningful parameter space, where each parameter in this space indicates distinct cytoskeletal regulators of wing disc morphogenesis. In the following sections, we describe how to transform mutant organ shapes into the parameter space to identify changes in cytoskeletal regulation that cause changes in the

overall tissue shape. Such a pipeline can serve to identify new functions for specific genes and gene products.

### Bayesian optimization framework predicts loss in contractility of columnar cells upon removal of ECM

As a first test of the framework, enzymatic degradation of the extracellular matrix (ECM) with Collagenase<sup>30</sup> was performed to test if the BO framework can recapitulate specific perturbations to cell mechanics, in agreement with our earlier work, the removal of ECM led to a striking inversion or flipping of the curvature within the tissue (Fig. 4a, b) and a loss of inwards bending towards the basal surface of the columnar cell was observed. External contours along the basal surface of the experimental cross-sections are fed into the BO framework as inputs. Our pipeline next estimates the parameters of the Surface Evolver model that would best represent the shape changes. BO was carried out to minimize the Fréchet distance between the experimental cross-section and the simulated shape. The pipeline successfully captured qualitative shape changes as observed within the experimental data, as shown in Fig. 4c, d. Significantly, the estimated parameters using the basal contours of the wing disc shape after collagenase treatment indicate a significant decrease in contractility compared to the control (Fig. 4e). All apical, basal, and lateral contractility ( $k_A^{col}$ ,  $k_B^{col}$ ,  $k_L^{col}$ ) levels decreased (Fig. 4f).

The basal contractility ( $k_B^{col}$ ) in our model is the sum of contractility imposed by the actomyosin complex ( $k_{con}^{col}$ ) and the extracellular matrix contractility ( $k_{con}^{ECM}$ ). Modeling these two independently within the Surface Evolver would increase the model complexity. No significant variations were found in the  $K_{ECM}$  parameter, a multiplier of the curvature-based energy used to describe the ECM. The parameter  $K_{ECM}$  contributes towards the elastic energy of the ECM. A higher  $K_{ECM}$  penalizes for higher elastic energy, restricting the wing disc from folding, while a lower  $K_{ECM}$  allows wing imaginal disc tissue to make dramatic curvature changes.

We also examined the curvature of the objective function ( $F_B$ ) around the average parameter value obtained from the set of sampled parameters generated by the BO framework (Supplementary Table 1) to assess the identifiability of the model parameters. The Surface Evolver output of the average parameter value yielded a shape similar to one observed when ECM is degraded by collagenase (Supplementary Fig. 3A). Notably, the eigenvector corresponding to the dominant eigenvalue primarily aligns with the parameter  $k_L^{col}$  indicating that the predicted organ shape, as described by  $F_B$ , is most sensitive to variations in the specific parameter (Supplementary Fig. 3B). Consequently, the estimation of  $k_L^{col}$  can be carried out with minimal uncertainty. Conversely, the three subsequent eigenvalues exhibit significantly smaller magnitudes than the first one, and their predominant directions align with  $K_{ECM}$ ,  $k_B^{col}$  and  $I_{0,L}^{col}$  respectively. This observation suggests that these parameters can be inferred with moderate uncertainty. Our framework suggests that tissue can still regulate its shape even with varying elasticity by modulating other subcellular features like cell edge lengths and actomyosin contractility. These predictions align with previous findings where a subcellular element model recapitulated the changes upon ECM removal by removing basal contractility and the ECM stiffness from the model<sup>30</sup>.

The control group in our pipeline resulted in two sets (Cluster 1, Cluster 2) of parameters that produced similar tissue shapes (Fig. 4c). Examination of the parameters reveals that the differences in shape were due to differences in columnar cell height ( $Lo_L^{col}$ ) (Fig. 4e). Cluster 1 corresponds to tissue shapes with lower lateral cell height but higher apical contractility. In contrast, Cluster 2 corresponds to a higher lateral cell height with higher basal contractility (Supplementary Table 1). This suggests that maintaining the inwards doming shape with stretching of the pouch cells during growth requires increased basal contractility. It can be achieved either through increasing the ECM stiffness or increasing actomyosin-mediated contractility. Our previous work demonstrated that both ECM stiffness and basal contractility play a crucial role in generating and maintaining the dome shape of the *Drosophila* wing imaginal disc<sup>30</sup>. Other recent work shows increased ECM stiffness as the *Drosophila* wing disc grows in size<sup>33</sup>.

Further, the sum of apical and basal contractility is also higher for Cluster 2 than for Cluster 1. Cluster 2 also exhibits cells with lower apical cell area. As the pouch grows in size, the apical cell area decreases in size<sup>37</sup>. The reduction in this apical cell area also contributes to the increase in cell height. We also compared the ratio of columnar cell height ( $L_{Lateral}$ ) to apical cell diameter ( $D_{Apical}$ ), represented by the wing disc model parameters  $L_{col}^{lat}$  and  $L_{col}^{api}$  respectively. Our analysis, in line with Breen et al.'s<sup>54</sup> findings, demonstrates a comparable ratio (Literature:  $L_{Lateral} = 21.5 \mu\text{m}$ ,  $D_{Apical} = 3.49 \mu\text{m}$ ,  $L_{Lateral} / D_{Apical} = 6.16$ , Model Predictions:  $L_{col}^{lat} / L_{col}^{api} = 10.03$ ). Furthermore, we illustrate that tension in the basal columnar epithelium exceeds that in the apical columnar epithelium ( $k_{col}^{bas} > k_{col}^{api}$ ,  $k_{col}^{lat}$ ) in agreement with Sui et al.<sup>55</sup>.

Our framework also predicts a loss of contractility ( $k_{col}^{api,bas,lat}$ ) within the tissue upon Collagenase mediated degradation of the Extracellular Matrix (ECM) (Fig. 4e, f). To test if inhibition of cell-ECM adhesion as a result of ECM degradation, we downregulated *mys*, a subunit of  $\beta$ -integrin, within the posterior compartment of the wing imaginal disc (Fig. 4h).  $\beta$ -integrin plays a crucial role in cell-ECM adhesion and transduction of extracellular signals into the cells<sup>42,56</sup>. Subsequently, we performed an immunohistochemistry assay using a phosphorylated non-muscle myosin (pMyoII) antibody to study the impact of the loss of cell-ECM adhesion on pMyoII expression. pMyoII, along with Actin, regulates the generation of contractile forces within the tissue<sup>30</sup>. Our findings indicate that the loss of  $\beta$ -integrin through *en>mys<sup>RNAi</sup>* results in decreased basal pMyoII levels in the posterior compartment (Fig. 4h, i). Furthermore, the knockdown of  $\beta$ -integrin qualitatively reproduces a similar phenotypic profile, characterized by the generation of folds in the opposite direction compared to the wild type.

In summary, by applying BO, we show that enzymatic degradation of ECM causes a reduction in global tissue actomyosin contractility. It also causes an inversion in the tissue curvature. Interestingly, our pipeline also reveals two distinct mechanisms of cytoskeletal regulation leading to similar tissue shapes defined by its outer surface. Analysis of the parameters of the two groups further describes biophysical mechanisms driving the thickening of the columnar cells as the tissue grows in size. Our model predicts that an increase in basal contractility and apical cell area is required to sustain the bent shape of the wing imaginal disc as it grows in size.

### Piezo regulates basal epithelial curvature through actomyosin contractility and ECM elasticity

Next, we explored the utility of the BO framework for inferring mechanisms that define new morphological shapes downstream of specific genetic perturbations (Fig. 5a, b). To do this, we used the Gal4-UAS system to either knockdown or over-express *Piezo* in the dorsal compartment of the wing imaginal disc using an apterous-Gal4 driver. Piezo proteins are a class of mechanosensitive ion channels involved in regulating multiple biophysical processes<sup>57,58</sup>. However, its role in regulating overall organ shape is poorly understood. Cross-sections parallel to the AP axis were taken in the dorsal side of the pouch guided by the apterous::GFP expressions. The cross-section in the ventral side was used as an internal control. Knockdown of *Piezo* caused loss of doming within the pouch (Fig. 5e) along with a reduction in curvature of folds in the notum region of the tissue compared to the control (Supplementary Fig. 4). On the other hand, overexpression of *Piezo* caused the tissue to increase in bending as compared to the internal control (Fig. 5i).

We used the definition of the basal surface of the contours, as shown in Fig. 5c, e, g, i), as our inputs to the BO framework. Our framework identified shapes qualitatively matching the mutant cross-section (Fig. 5d, f, h, j). A comparative analysis between the model parameters of the best cases identified revealed a decrease in both apical and lateral contractility (defined by  $k_A^{col}$ ,  $k_L^{col}$ ) upon *Piezo* knockdown as compared to prediction for the internal control and the global Oregon-R control.  $k_L^{col}$  also increased upon overexpression of *Piezo* in a compartment-specific manner (Fig. 5k-l). As predicted through model analysis, quantifying the expression of pMyoII within the *Piezo* mutants also shows an increase in both apical and lateral pMyoII upon overexpression of *Piezo* (Fig. 5q, r, Supplementary Fig. 4). We

also report a decrease in pMyoII in a compartment-specific manner upon *Piezo* knockdown (Fig. 5o, p, Supplementary Fig. 4).

Apart from changes in actomyosin contractility ( $k_i^{col}$ ), our model also predicts a compartment specific increase in  $LO_B^{col}$  upon a knockdown of *Piezo* (Fig. 5m). It further predicts a compartment-specific decrease but a global increase in  $LO_L^{col}$  on *Piezo* overexpression as compared to *Piezo*<sup>RNAi</sup>-expressing discs. An increase in pMyoII levels upon *Piezo* overexpression can cause the lateral pouch to contract, decreasing its length (Fig. 5n, Supplementary Fig. 4). Interestingly, our pipeline also predicts an overall decrease in  $K_{ECM}$  upon *Piezo* overexpression (Fig. 5g, i). A previous study of cell dissemination conducted in *Drosophila* midgut has shown that *Piezo* is required for the degradation of the ECM through Ras<sup>V12</sup> [59]. Further experiments are needed to measure changes in ECM elasticity upon *Piezo* overexpression. Changes in the  $K_{ECM}$  suggest that there will be experimentally observable changes in the composition of the ECM, including regulation of remodeling. Furthermore, *Piezo* plays a crucial role in the degradation of the ECM through its interaction with matrix metalloproteinase 1 (MMP1), an endopeptidase involved in ECM remodeling, and Collagen IV, a fundamental component of the basal ECM<sup>59–62</sup>. To test the implications of this interplay in the wing imaginal disc, we evaluated the expressions of Collagen IV and MMP1 in *Piezo* downregulated (*ap>Piezo*<sup>RNAi</sup>) tissues. Consistent with the general trends predicted by the model, we observed a downregulation of MMP1 (Fig. 5s, t) and an upregulation of Collagen IV (Fig. 5u, v) for *Piezo* knockdown, consistent with the interpretation that *Piezo* regulates ECM elasticity in the wing discs.

Overall, this section used a combination of quantitative image analysis of fixed tissues and the BO platform to infer potential mechanisms of fold formation regulation by *Piezo* mechanosensitive ion channels. Our work highlights that *Piezo* regulates bending in *Drosophila* wing imaginal disc by regulating actomyosin contractility, ECM elasticity, or regulation of single cell volume. A more detailed experimental analysis is required to establish mechanisms of cytoskeletal regulation through *Piezo*.

## Discussion

Morphogenesis involves dynamic shape changes within the organ until it reaches a target size and shape<sup>63,64</sup>. Changes in the global shape of an organ arise from the integration of changes at the cellular level over time. A key challenge for quantitative systems biology is to elucidate how the complex interplay of chemical signals at the single-cell level contributes to the overall organ shape. However, the large diversity of proteins that regulate shape changes and the difficulty of measuring them all at once necessitate the formulation of computational models<sup>65</sup>. A major challenge in calibrating multiscale models of morphogenesis is that they are often highly nonlinear and computationally expensive. Both conventional gradient-based optimization methods (e.g., gradient descent, quasi-Newton methods) and MCMC for Bayesian calibrations require too many evaluations for computationally expensive large-scale models, including subcellular element simulations, to be practical.

In this work, we developed and validated a computational framework, which is robust to initial conditions (Supplementary Fig. S5), to efficiently infer parameter sets with Bayesian optimization using GPR surrogate models that match experimentally obtained shapes of organs. We found that the optimization process works best using Fréchet distance as the error metric, i.e., the objective function (Fig. 2). A multi-faceted Bayesian optimization and local sensitivity analysis revealed that actomyosin contractility and extracellular matrix stiffness are the primary contributors to basal curvature and shape control (Fig. 3). This confirms previous reports that relied on experiments and scenarios of more complex subcellular element simulations<sup>30</sup>, demonstrating the robustness of the conclusions.

The current model calculates errors based on overall changes in tissue shape using distances. Often, problems in complex biology are nonlinear and sloppy, leading to multiple parameters within the model that can generate similar shapes<sup>66</sup>. Our pipeline suggests that tissues with different patterns of apical-basal contractility can generate similar shapes of basal curvature while also revealing variable regulation of columnar cell height

(Fig. 3c). The identification of multiple solutions depends on the number of morphological features used in generating the optimization cost function. For example, including additional components within the objective function in the form of columnar cell height can reduce the number of solutions. The selection rationale of morphological features to include in the cost function depends on the form of experimental data and the choice of the physics-based model. In future work, expected improvement should be compared against alternative acquisition functions such as lower/upper confidence bounds, probability of improvement, and Thompson sampling, which may result in improved performance.

At this stage, our model estimates errors based on the overall geometry of the organ, as quantified by a single metric, the Fréchet error. As a limitation, it does not provide predictions on the geometry of individual cells or the shape of the internal lumen. This limitation is mitigated as features, including tissue height, can be inferred from the cross sections. Future development of this modeling framework can incorporate additional optimization criteria for matching individual cell shapes and the shape of the lumen. Further, a fully 3D modeling framework remains a natural future step. To do so, one can incorporate multi-objective optimization principles to identify the Pareto optimal trade-offs between alternate error metrics solutions<sup>23</sup>. Such enhancements will lead to greater predictive capabilities and enable deeper mechanistic insights into the cellular contributions to an organ's shape. The computational cost will grow to achieve this, and the surrogate modeling must account for relative accuracy for a given computational cost.

The BO framework, in conjunction with immunohistochemistry assays, further reveal the role of *Piezo* mechanosensitive ion channels in regulating fold formation during wing imaginal disc morphogenesis (Fig. 5). It does so through a combined regulation of patterning in apical-basal contractility and cell volume. Even though significant efforts have investigated the roles of *Piezo* in regulating single-cell processes like proliferation, apoptosis, and cytoskeletal regulation<sup>67–71</sup>, emergent functions at the next, multiscale hierarchy of overall organ function remain poorly understood. Previous studies related to *Piezo* have shown its impact on regulating RhoA<sup>72</sup>. Further, RhoA is upstream of pMyoII, suggesting a potential mechanism. Since a knockdown of *Piezo* significantly decreased the accumulation of basal pMyoII (Fig. 5o, p), we also hypothesize that it may be through Integrin clustering mediated activation of mechanosensation. Increased basal contractility reduces the basal cell area, bringing integrin molecules sufficiently close to form clusters<sup>73,74</sup>. Once initiated, integrin clusters can stabilize the formation of focal adhesion complexes to increase tension further<sup>75</sup>. The final step can be the activation of *Piezo* to regulate Rho through  $Ca^{2+}$  to further promote pMyoII and consequently contractility<sup>76</sup>. This is equivalent to the autoregulation of pMyoII, which is supported by the experimental evidence that a loss of Integrin in the *Drosophila* wing imaginal disc is also known to reduce basal pMyoII levels<sup>22</sup>. This work thus motivates future experiments to map the exact mechanisms of regulation of pMyoII by *Piezo*.

Our computational pipeline also predicts a decrease in ECM elasticity (modeled as  $K_{ECM}$ ) upon overexpression of *Piezo* (Fig. 5g, i). Previous studies have proposed the degradation of ECM through *Piezo* in *Drosophila* midgut<sup>59</sup>. In accordance with the literature, our experimental data confirms loss of *Piezo* reduced the levels of Mmp2, an enzyme known to degrade ECM. The work further proposed that *Piezo* can also contribute towards ECM degradation through  $Ca^{2+}$ -mediated activation of Calpains. Further experimental measurements of ECM elasticity are required to confirm the predicted function of *Piezo* during organ development.

Future work can also further define morphogenesis as a multi-objective optimization problem based on multiple outputs of the model and corresponding measurable features in experimental data. The pipeline can be extended for organ-level drug screening, allowing for the study of new mechanical functions of genes due to genetic or pharmacological perturbations. Further, this framework can be extended to studying other physical and biochemical processes such as embryogenesis<sup>77,78</sup> and models of plant development<sup>79</sup>. Of note, it can also be used to study any models of

organogenesis as it is independent of the modeling framework or package used in the physics-based simulation, which makes it attractive for more complex computational models that incorporate subcellular elements<sup>22,30,80,81</sup>. In summary, this computational framework enables the systematic elucidation of generalizable biological rules of the morphogenesis of multicellular systems.

## Methods

### Experimental and image analysis methods

**Fly stocks and culture.** *Drosophila* was grown within an incubator maintained at 25°C. The flies were maintained on 12-hour darkness/light cycle. Virgins from the Gal4 driver colonies were collected twice daily. For the first collection, the bottles are emptied before 6 h of collection. Female virgins with Gal4 drivers were crossed with UAS-transgene male flies in a 10:15:4 ratio. Early 3<sup>rd</sup> instar wandering larvae were collected to dissect wing imaginal disc tissue. The wildtype Oregon-R fly line is a long-standing stock in our group originally acquired from the N. Yakoby lab. The following other transgenic stocks and their source, include UAS-Piezo<sup>RNAi</sup>, VDRC # 105132, and UAS-Piezo, BDRC #58772/58773.

**Immunohistochemistry.** Wing imaginal discs were dissected in a phosphate-buffered saline (PBS) solution before fixation in a 4% paraformaldehyde in PBS solution. Quantification included all samples in the reported sample size. Fixation was done by placing the PCR tubes containing wing disc samples in an ice bath for an hour. Post-fixation, the samples were rinsed with a fresh PBT solution (PBS with 0.03% v/v Triton X-100). Three quick rinses were followed by three 10-minute-long washes. PBT within the tubes was next replaced with 250 µL of 5% normal goat serum (NGS) in PBS and agitated for an hour. Following this, the NGS solution was replaced with the primary antibody solution, and the tubes were left in a rotating platform placed in a temperature of 4°C overnight. The following primary antibodies were used: i) Phospho-Myosin Light Chain 2 (Ser19) (1:50, Rabbit, Cell Signaling Technology #3671S) ii) Integrin βPS (myospheroid) (1:5, Mouse, Developmental Studies Hybridoma Bank CF.6G11) iii) Anti-MMP1 (1:1000; Mouse; Developmental Studies Hybridoma Bank, 3B8D12), iv) α-Collagen IV (1:5; Rabbit; Abcam ab6586). The next day, the primary antibody was replaced with PBT. After three quick rinses and three 15-minute-long washes, the PBT in the tube is replaced with a secondary antibody solution. The following secondary antibody and dyes were used in our studies: α-Rabbit Alexa Fluor™ 647 (1:500, Goat, Thermo Fisher Scientific A32733), α-Mouse Alexa Fluor™ 568 (1:500, Goat, Thermo Fisher Scientific A-11031), DAPI (1:500, Sigma Aldrich D9542) and Fluorescein Phalloidin (1:500, Thermo Fisher Scientific F432). After two hours of secondary antibody and dye incubation at room temperature, avoiding light exposure, three rinses of PBT were carried out. After two additional long washes of 15 min, the samples were left for overnight incubation and agitation in PBT at 4°C. The next day, the samples were mounted in a coverslip with spacers to avoid squishing the samples. Spacers were designed using two layers of bio-compatible tapes to create a well to place the samples. Vectashield mounting medium and a cover slip was placed atop, aligned with the spacers.

**Confocal microscopy.** Imaging of wing imaginal disc samples was done with three different microscopes: Nikon Eclipse Ti confocal microscope with a Yokogawa spinning disc, Nikon A1R-MP laser scanning confocal microscope, and Leica Stellaris 8 DIVE Point Scanning Confocal Microscope. For the two confocal microscopes, image data were collected on an IXonEM+colled CCD camera (Andor Technology, South Windsor, CT) using MetaMorph v7.7.9 software (Molecular Devices, Sunnyvale, CA), NIS-Elements software, and LAS X microscope software respectively. The step size for acquiring 3D data was kept between 0.5–1 µm, depending on sample thickness. Imaging was done using 40× and 60× oil objectives with 200 ms exposure time, and 50 nW, 405 nm, 488 nm, 561 nm, and 640 nm laser exposure.

**Image analysis.** All the raw data presented within the manuscript was analyzed using FIJI/ImageJ. Quantification of fluorescence intensity was carried out using an in-house MATLAB code whose details can be found in the supplementary information of the text. CSBDeep, an ImageJ plugin was used for deconvolution and denoising of the Actin channel. A rolling ball background subtraction was also used to remove background noise. QuickStitch<sup>82</sup> was used for stitching individual tiles while imaging the entire volume of the imaginal disc.

### Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this article.

### Data availability

Due to data size constraints, the 3D confocal microscopy data is available on request from the corresponding authors. This includes all of the representative source data and analysis files for experiments within this manuscript.

### Code availability

The optimization framework was developed in Python and is available at <https://github.com/MulticellularSystemsLab/ReverseEngineeringMorphogenesis>.

Received: 21 August 2023; Accepted: 17 April 2024;

Published online: 07 May 2024

### References

1. Friedel, S., Usadel, B., Von Wirén, N. & Sreenivasulu, N. Reverse Engineering: A Key Component of Systems Biology to Unravel Global Abiotic Stress Cross-Talk. *Front. Plant Sci.* **3**, 294 (2012).
2. Narciso, C. & Zartman, J. Reverse-engineering organogenesis through feedback loops between model systems. *Curr. Opin. Biotechnol.* **52**, 1–8 (2018).
3. Vodovotz, Y. Reverse Engineering the Inflammatory “Clock”: From Computational Modeling to Rational Resetting. *Drug Discov. Today Dis. Models* **22**, 57–63 (2016).
4. Shannon, E. K. et al. Multiple Mechanisms Drive Calcium Signal Dynamics around Laser-Induced Epithelial Wounds. *Biophys. J.* **113**, 1623–1635 (2017).
5. Newman, T. J. Modeling multicellular systems using subcellular elements. *Math. Biosci. Eng. MBE* **2**, 613–624 (2005).
6. Yao, J., Pilko, A. & Wollman, R. Distinct cellular states determine calcium signaling response. *Mol. Syst. Biol.* **12**, 894 (2016).
7. Brodland, G. W. How computational models can help unlock biological systems. *Semin. Cell Dev. Biol.* **47–48**, 62–73 (2015).
8. Anderson, A. E., Ellis, B. J. & Weiss, J. A. Verification, validation and sensitivity studies in computational biomechanics. *Comput. Methods Biomed. Engin.* **10**, 171–184 (2007).
9. Reali, F., Priami, C. & Marchetti, L. Optimization Algorithms for Computational Systems Biology. *Front. Appl. Math. Stat.* **3**, 6 (2017).
10. Linden, N. J., Kramer, B. & Rangamani, P. Bayesian parameter estimation for dynamical models in systems biology. *PLOS Comput. Biol.* **18**, e1010651 (2022).
11. Valderrama-Bahamón, G. I. & Fröhlich, H. MCMC Techniques for Parameter Estimation of ODE Based Models in Systems Biology. *Front. Appl. Math. Stat.* **5**, 55 (2019).
12. Kumar, M., Husain, D. M., Upreti, N. & Gupta, D. Genetic Algorithm: Review and Application. SSRN Scholarly Paper at <https://doi.org/10.2139/ssrn.3529843> (2010).
13. Sun, J., Garibaldi, J. M. & Hodgman, C. Parameter estimation using meta-heuristics in systems biology: a comprehensive review. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **9**, 185–202 (2012).
14. Warne, D. J., Prescott, T. P., Baker, R. E. & Simpson, M. J. Multifidelity multilevel Monte Carlo to accelerate approximate Bayesian parameter

inference for partially observed stochastic processes. *J. Comput. Phys.* **469**, 111543 (2022).

15. Warne, D. J., Baker, R. E. & Simpson, M. J. Rapid Bayesian Inference for Expensive Stochastic Models. *J. Comput. Graph. Stat.* **31**, 512–528 (2022).
16. Vernon, I. et al. Bayesian uncertainty analysis for complex systems biology models: emulation, global parameter searches and evaluation of gene functions. *BMC Syst. Biol.* **12**, 1 (2018).
17. Kursawe, J., Baker, R. E. & Fletcher, A. G. Approximate Bayesian computation reveals the importance of repeated measurements for parameterising cell-based models of growing tissues. *J. Theor. Biol.* **443**, 66–81 (2018).
18. Shahriari, B., Swersky, K., Wang, Z., Adams, R. P. & de Freitas, N. Taking the Human Out of the Loop: A Review of Bayesian Optimization. *Proc. IEEE* **104**, 148–175 (2016).
19. Greenhill, S., Rana, S., Gupta, S., Vellanki, P. & Venkatesh, S. Bayesian Optimization for Adaptive Experimental Design: A Review. *IEEE Access* **8**, 13937–13948 (2020).
20. Ulmasov, D., Baroukh, C., Chachuat, B., Deisenroth, M. P. & Misener, R. Bayesian Optimization with Dimension Scheduling: Application to Biological Systems. in *Computer Aided Chemical Engineering* (eds. Kravanja, Z. & Bogataj, M.) 38 1051–1056 (Elsevier, 2016).
21. Brakke, K. A. The Surface Evolver. *Exp. Math.* **1**, 141–165 (1992).
22. Kumar, N. et al. Balancing competing effects of tissue growth and cytoskeletal regulation during Drosophila wing disc development. *Nat. Commun.* **15**, 2477 (2024).
23. Optimal Performance Objectives in the Highly Conserved Bone Morphogenetic Protein Signaling Pathway. *bioRxiv*. <https://www.biorxiv.org/content/10.1101/2024.02.01.578451v2>.
24. Wang, K. & Dowling, A. W. Bayesian optimization for chemical products and functional materials. *Curr. Opin. Chem. Eng.* **36**, 100728 (2022).
25. Befort, B. J., DeFever, R. S., Tow, G. M., Dowling, A. W. & Maginn, E. J. Machine Learning Directed Optimization of Classical Molecular Modeling Force Fields. *J. Chem. Inf. Model.* **61**, 4400–4414 (2021).
26. López-Lopera, A. F., Durrande, N. & Álvarez, M. A. Physically-Inspired Gaussian Process Models for Post-Transcriptional Regulation in Drosophila. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **18**, 656–666 (2021).
27. Oyebamiji, O. K. et al. Gaussian process emulation of an individual-based model simulation of microbial communities. *J. Comput. Sci.* **22**, 69–84 (2017).
28. Rocha, H. L. et al. Bayesian inference using Gaussian process surrogates in cancer modeling. *Comput. Methods Appl. Mech. Eng.* **399**, 115412 (2022).
29. Tulisan, A. et al. Spectroscopic models for real-time monitoring of cell culture processes using spatiotemporal just-in-time Gaussian processes. *AIChE J* **67**, e17210 (2021).
30. Nematbakhsh, A. et al. Epithelial organ shape is generated by patterned actomyosin contractility and maintained by the extracellular matrix. *PLOS Comput. Biol.* **16**, e1008105 (2020).
31. Beira, J. V. & Paro, R. The legacy of Drosophila imaginal discs. *Chromosoma* **125**, 573–592 (2016).
32. Smith-Bolton, R. Drosophila Imaginal Discs as a Model of Epithelial Wound Repair and Regeneration. *Adv. Wound Care* **5**, 251–261 (2016).
33. Harmansa, S., Erlich, A., Eloy, C., Zurlo, G. & Lecuit, T. Growth anisotropy of the extracellular matrix shapes a developing organ. *Nat. Commun.* **14**, 1220 (2023).
34. Parker, J. & Struhl, G. Control of Drosophila wing size by morphogen range and hormonal gating. *Proc. Natl. Acad. Sci.* **117**, 31935–31944 (2020).
35. Bejsovec, A. Wingless Signaling: A Genetic Journey from Morphogenesis to Metastasis. *Genetics* **208**, 1311–1336 (2018).
36. Teleman, A. A. & Cohen, S. M. Dpp Gradient Formation in the Drosophila Wing Imaginal Disc. *Cell* **103**, 971–980 (2000).
37. Breen, D. E., Sui, L., Bai, L., Jülicher, F. & Dahmann, C. Cell-level 3D reconstruction and quantification of the Drosophila wing imaginal disc. *Int. J. Bioinforma. Res. Appl.* **15**, 174–189 (2019).
38. Betapudi, V. Life without double-headed non-muscle myosin II motor proteins. *Front. Chem.* **2**, 45 (2014).
39. Guo, H., Swan, M. & He, B. Optogenetic inhibition of actomyosin reveals mechanical bistability of the mesoderm epithelium during Drosophila mesoderm invagination. *eLife* **11**, e69082 (2022).
40. Heer, N. C. et al. Actomyosin-based tissue folding requires a multicellular myosin gradient. *Dev. Camb. Engl.* **144**, 1876–1886 (2017).
41. Wodarz, A., Stewart, D. B., Nelson, W. J. & Nusse, R. Wingless signaling modulates cadherin-mediated cell adhesion in Drosophila imaginal disc cells. *J. Cell Sci.* **119**, 2425–2434 (2006).
42. Domínguez-Giménez, P., Brown, N. H. & Martín-Bermudo, M. D. Integrin-ECM interactions regulate the changes in cell shape driving the morphogenesis of the Drosophila wing epithelium. *J. Cell Sci.* **120**, 1061–1071 (2007).
43. Fristrom, D., Wilcox, M. & Fristrom, J. The distribution of PS integrins, laminin A and F-actin during key stages in Drosophila wing development. *Dev. Camb. Engl.* **117**, 509–523 (1993).
44. Štorgel, N., Krajnc, M., Mrak, P., Štrus, J. & Ziherl, P. Quantitative Morphology of Epithelial Folds. *Biophys. J.* **110**, 269–277 (2016).
45. Kuhl, F. P. & Giardina, C. R. Elliptic Fourier features of a closed contour. *Comput. Graph. Image Process.* **18**, 236–258 (1982).
46. Dynamic Time Warping. In *Information Retrieval for Music and Motion* (ed. Müller, M.) 69–84 (Springer, Berlin, Heidelberg.). [https://doi.org/10.1007/978-3-540-74048-3\\_4](https://doi.org/10.1007/978-3-540-74048-3_4) (2007).
47. Alt, H. & Godau, M. Computing the Fréchet distance between two polygonal curves. *Int. J. Comput. Geom. Appl.* **05**, 75–91 (1995).
48. Eaton, S., Auvilinen, P., Luo, L., Jan, Y. N. & Simons, K. CDC42 and Rac1 control different actin-dependent processes in the Drosophila wing disc epithelium. *J. Cell Biol.* **131**, 151–164 (1995).
49. Tang, B. Orthogonal Array-Based Latin Hypercubes. *J. Am. Stat. Assoc.* **88**, 1392–1397 (1993).
50. Tonner, P. D., Darnell, C. L., Engelhardt, B. E. & Schmid, A. K. Detecting differential growth of microbial populations with Gaussian process regression. *Genome Res.* **27**, 320–333 (2017).
51. Schulz, E., Speekenbrink, M. & Krause, A. A tutorial on Gaussian process regression: Modelling, exploring, and exploiting functions. *J. Math. Psychol.* **85**, 1–16 (2018).
52. Rasmussen, C. E. & Williams, C. K. I. Gaussian Processes for Machine Learning. (MIT, 2006). <https://doi.org/10.7551/mitpress/3206.001.0001>.
53. Fuhg, J., Fau, A. & Nackenhorst, U. State-of-the-Art and Comparative Review of Adaptive Sampling Methods for Kriging. *Arch. Comput. Methods Eng.* **28**, 2689–2747 (2020).
54. Breen, D. E., Widmann, T., Bai, L., Jülicher, F. & Dahmann, C. Epithelial cell reconstruction and visualization of the developing Drosophila wing imaginal disc. In *2012 IEEE Symposium on Biological Data Visualization (BioVis)* 77–84. <https://doi.org/10.1109/BioVis.2012.6378596> (2012).
55. Sui, L. et al. Differential lateral and basal tension drive folding of Drosophila wing discs through two distinct mechanisms. *Nat. Commun.* **9**, 4620 (2018).
56. Harburger, D. S. & Calderwood, D. A. Integrin signalling at a glance. *J. Cell Sci.* **122**, 159–163 (2009).
57. Wu, J., Lewis, A. H. & Grandl, J. Touch, Tension, and Transduction – The Function and Regulation of Piezo Ion Channels. *Trends Biochem. Sci.* **42**, 57–71 (2017).
58. Honoré, E., et al. The piezo mechanosensitive ion channels: may the force be with you!. *Reviews of Physiology, Biochemistry and Pharmacology*, Vol. 169, pp 25–41 (Springer International Publishing, Cham). [https://doi.org/10.1007/112\\_2015\\_26](https://doi.org/10.1007/112_2015_26). (2015).
59. Lee, J., Cabrera, A. J. H., Nguyen, C. M. T. & Kwon, Y. V. Dissemination of RasV12-transformed cells requires the mechanosensitive channel Piezo. *Nat. Commun.* **11**, 3568 (2020).

60. Kang, H. et al. Piezo1 mediates angiogenesis through activation of MT1-MMP signaling. *Am. J. Physiol. Cell Physiol.* **316**, C92–C103 (2019).

61. Stricker, A. M., Hutson, M. S. & Page-McCaw, A. Piezo initiates transient production of collagen IV to repair damaged basement membranes. Preprint at <https://doi.org/10.1101/2023.12.22.573147> (2023).

62. Gaub, B. M. & Müller, D. J. Mechanical Stimulation of Piezo1 Receptors Depends on Extracellular Matrix Proteins and Directionality of Force. *Nano Lett* **17**, 2064–2072 (2017).

63. Heisenberg, C.-P. & Bellaïche, Y. Forces in Tissue Morphogenesis and Patterning. *Cell* **153**, 948–962 (2013).

64. Buchmann, A., Alber, M. & Zartman, J. J. Sizing it up: the mechanical feedback hypothesis of organ growth regulation. *Semin. Cell Dev. Biol.* **35**, 73–81 (2014).

65. Wu, Q., Kumar, N., Velagala, V. & Zartman, J. J. Tools to reverse-engineer multicellular systems: case studies using the fruit fly. *J. Biol. Eng.* **13**, 33 (2019).

66. Goodwin, B. C., Kauffman, S. & Murray, J. D. Is Morphogenesis an Intrinsically Robust Process? *J. Theor. Biol.* **163**, 135–144 (1993).

67. Nourse, J. L. & Pathak, M. M. How cells channel their stress: Interplay between Piezo1 and the cytoskeleton. *Dev. Biol.* **71**, 3–12 (2017).

68. Tsai, F.-C., Kuo, G.-H., Chang, S.-W. & Tsai, P.-J. Ca<sup>2+</sup> signaling in cytoskeletal reorganization, cell migration, and cancer metastasis. *BioMed Res. Int.* **2015**, 409245 (2015).

69. Song, Y. et al. Mechanosensitive channel Piezo1 induces cell apoptosis in pancreatic cancer by ultrasound with microbubbles. *iScience* **25**, 103733 (2022).

70. Gudipaty, S. A. et al. Mechanical stretch triggers rapid epithelial cell division through Piezo1. *Nature* **543**, 118–121 (2017).

71. Kumar, N. et al. Piezo regulates epithelial topology and promotes precision in organ size control. Preprint at <https://doi.org/10.1101/2023.08.16.553584> (2023).

72. Pardo-Pastor, C. et al. Piezo2 channel regulates RhoA and actin cytoskeleton to promote cell mechanobiological responses. *Proc. Natl. Acad. Sci.* **115**, 1925–1930 (2018).

73. Schwartzman, M. et al. Nanolithographic Control of the Spatial Organization of Cellular Adhesion Receptors at the Single-Molecule Level. *Nano Lett* **11**, 1306–1312 (2011).

74. Coyer, S. R. et al. Nanopatterning reveals an ECM area threshold for focal adhesion assembly and force transmission that is regulated by integrin activation and cytoskeleton tension. *J. Cell Sci.* **125**, 5110–5123 (2012).

75. Roca-Cusachs, P., Iskratsch, T. & Sheetz, M. P. Finding the weakest link: exploring integrin-mediated mechanical molecular pathways. *J. Cell Sci.* **125**, 3025–3038 (2012).

76. Takuwa, Y. et al. Calcium-dependent regulation of Rho and myosin phosphatase in vascular smooth muscle. *Biomed. Rev.* **16**, 13–21 (2005).

77. Velagala, V. & Zartman, J. J. Pinching and pushing: fold formation in the *Drosophila* dorsal epidermis. *Biophys. J.* **120**, 4202–4213 (2021).

78. Velagala, V., Chen, W., Alber, M. & Zartman, J. J. Chapter 4.1 - Multiscale Models Coupling Chemical Signaling and Mechanical Properties for Studying Tissue Growth. in *Mechanobiology* (ed. Niebur, G. L.) 173–195 (Elsevier,), <https://doi.org/10.1016/B978-0-12-817931-4.00010-8>. (2020).

79. Banwarth-Kuhn, M. et al. Combined computational modeling and experimental analysis integrating chemical and mechanical signals suggests possible mechanism of shoot meristem maintenance. *PLOS Comput. Biol.* **18**, e1010199 (2022).

80. Nematbakhsh, A. et al. Multi-scale computational study of the mechanical regulation of cell mitotic rounding in epithelia. *PLOS Comput. Biol.* **13**, e1005533 (2017).

81. Ramezani, A. et al. A multiscale chemical-mechanical model predicts impact of morphogen spreading on tissue growth. *Npj Syst. Biol. Appl.* **9**, 1–12 (2023).

82. Brodskiy, P. A. et al. QuickStitch for Seamless Stitching of Confocal Mosaics through High-Pass Filtering and Recursive Normalization. Preprint at <https://www.biorxiv.org/content/10.1101/075440v1> (2016).

## Acknowledgements

N.K., M.S.M., and J.J.Z. were supported in part by the NIH grant R35GM124935. N.K., A.D., and J.J.Z. were supported by NSF-Simons Center for Quantitative Biology Pilot program and NSF award 2029814. J.J.Z. and M.S.M. were supported in part by EMBRIO Institute, contract #2120200, a National Science Foundation (NSF) Biology Integration Institute, and NSF CBET-2225601 (RECODE). A.D. received partial support from NSF Award CBET-1941596. J.J.Z. and A.D. received partial support from NSF Award 2225601. We thank Dr. Sara Cole from Notre Dame Integrated Imaging Facility Optical Microscopy Core for the assistance with confocal imaging. The authors thank P. Brodskiy, A. Madhwan, M. Unger, V. Velagala, and members of the Multicellular Systems Engineering Lab for their helpful comments and suggestions.

## Author contributions

N.K. built the model framework, performed all preliminary experiments, analyzed the data, and wrote the manuscript. M.S.M. performed additional experiments for the revision and helped write and finalize the revised manuscript. J.J.Z. and A.D. conceived, designed, analyzed, and interpreted results, supervised this study, and wrote the manuscript. All authors read and agreed to the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41540-024-00375-z>.

**Correspondence** and requests for materials should be addressed to Jeremiah J. Zartman.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>

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

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024