LEARNING A MECHANICAL GROWTH MODEL OF FLOWER MORPHOGENESIS

Argyris Zardilis
Sainsbury Laboratory, University of Cambridge
argyris.zardilis@slcu.cam.ac.uk

Alexandra Budnikova
Sainsbury Laboratory, University of Cambridge
ab3017@cam.ac.uk

Henrik Jönsson
Sainsbury Laboratory, University of Cambridge
henrik.jonsson@slcu.cam.ac.uk

1 INTRODUCTION

Morphogenesis, the process by which organs and organisms acquire their shape, requires the integration of genetic, biochemical and mechanical signals. The data that we collect for this process is therefore multimodal and multi-scale, typically image time series of the morphology of the developing tissue and some measure of the (spatial) expression of genetic and hormonal components. In plants (and beyond) there has been an increasing number of 'atlases' where expression data are connected to their morphological context in space, either through multiplexed imaging techniques (spatial transcriptomics; Nobori et al. (2023)), single-cell sequencing mapped spatially Neumann et al. (2022), or computationally multiplexing low-throughput imaging experiments Oughou et al. (2019); Refahi et al. (2021). As the data grows in size and diversity, integrated analysis of these datasets and incorporation into mechanical theories of growth becomes extremely challenging to do manually Berens et al. (2023). We address this problem in sepals – the first organs of the flower, emerging in the outer whorl of the flower bud and transitioning from outgrowths to folded structures – as an example system of complex organogenesis Roeder (2021). In previous work we have developed an 'atlas' of spatial gene expression for this process at cellular resolution, integrating a number of genetic components in the same reference time series Refahi et al. (2021) (Figure 1a). Here we use this dataset to infer a mechanical model of growth to propose possible physical mechanisms of action of molecular species by (i) computing separate representations for the morphology and genetic views of the dataset and learning a mapping between them to identify genetic components more likely to have contributed to morphological characteristics of the tissue and (ii) inferring a linear mechanical model of growth for this process using the identified genetic components from (i), calculated cell growth rates and computed mechanical state. This representation integrates genetic and mechanical signals to give a more comprehensive picture of the process and is an example of a method to gain a more mechanistic understanding from the increasingly available spatially mapped gene expression datasets.

MEANINGFULNESS STATEMENT

A meaningful representation of life is one that fits with our analogies and models of biology so that it can be used to generate interpretable and testable hypotheses and ideas. Here, we propose learning a mechanical model of growth as a way to integrate large and multimodal datasets in the complex process of organogenesis. This model can be used to propose multi-scale hypotheses, for example the physical actions of genes. These hypotheses are physically interpretable and testable either experimentally or computationally by simulations.

2 RESULTS

2.1 DATA REPRESENTATION

One way to infer the causality of genes in (flower) development is through perturbations, so if we perturb a gene and see an effect on the organ identity or morphology of the tissue, we can link them

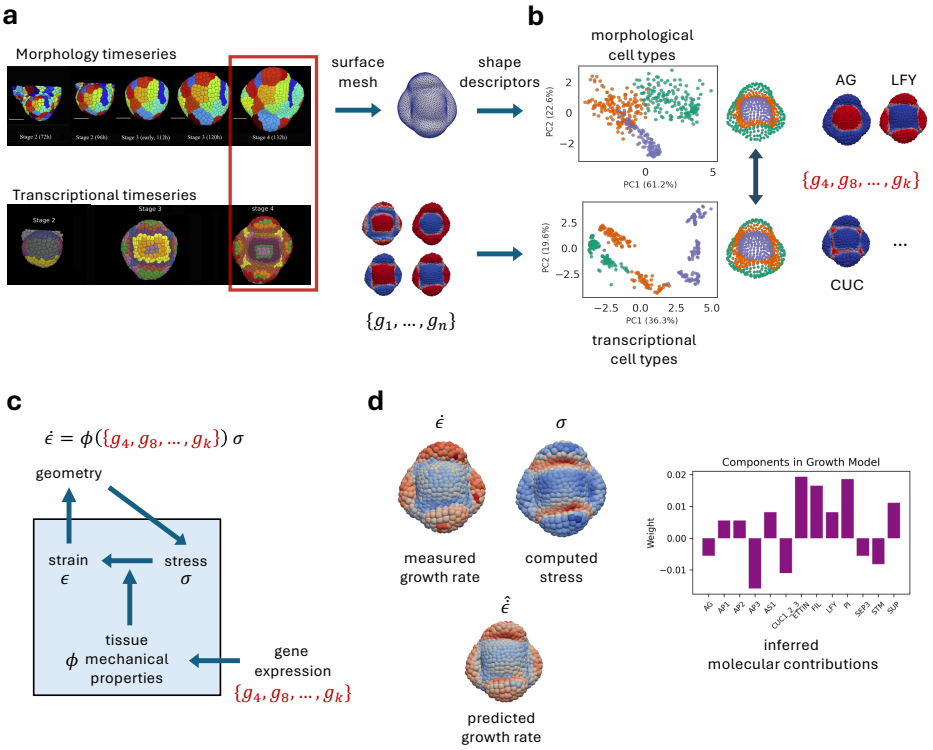

Figure 1: a. Morphological characteristics are computed from the last time point of an imaging time series of the developing organ and are b. mapped to spatial gene expression to get the genes most correlated to morphology c. Schematic of the mechanical growth model including the proposed influence of genetic factors d. Inferred molecular contributions to growth using the computed strain rates and mechanical stress of the tissue. Images of the time series in a. are reproduced from Refahi et al. (2021).

causally. This led to the identification of important developmental genes and models of their action in flower morphogenesis (ABC model; Bowman & Moyroud (2024); Coen & Meyerowitz (1991); Bowman et al. (1989)). Since plant cells are non-migrating, their action should be local so here we will use spatial correlation between morphological characteristics and gene expression as a proxy of causality. From the reference time series we focus on the last time point as it is the first stage where we can see distinct morphological aspects of the sepals (Figure 1a). At that stage, we focus on the first layer of cells (450 cells) as that is considered the determinant of overall shape.

To compute morphological characteristics we first extract a surface mesh from the 3D image stack and then for each point on the mesh compute the principal curvatures at that point, κ_1 and κ_2, and other curvature descriptors based on those; gaussian curvature ($\kappa_1\kappa_2$), mean curvature ($(\kappa_1+\kappa_2)/2$), and deviatoric curvature ($(\kappa_1-\kappa_2)/2$). These curvature metrics are transferred to the tissue cells and combined with cell volume. With the chosen measures we tried to capture the local geometry around a cell but there are more morphological features that could potentially be interesting like the shape of the cell itself and other more series-based descriptions of shape (based e.g. on spherical harmonics as in Pönisch et al. (2024)). On the other hand the spatial gene expressions are defined in the atlas per cell and are spread using a diffusion process to make them continuous as they were initially binary. Using PCA for dimensionality reduction and a Gaussian Mixture model for clustering, these features seem to capture the main regions and cell types at this stage, the central stem cell region, the developing organs at the four poles, and the boundary between them (Figure 1b). We next use logistic regression to learn functions between the morphological characteristics and each of the 25 genes across the cells in the dataset. From the 25 genes, 13 were well predicted from morphology. This was assessed by first binarising the predicted expression with a threshold of 0.5 and computing the balanced accuracy score with the atlas patterns (score > 0.75). The highlighted genes include

some well known markers for cell states, like CUC (boundary), LFY (sepals) and AG (central region) (Figure 1b, Alvarez-Buylla et al. (2010)).

2.2 MODEL INFERENCE

We have identified a statistical relation between morphology and gene expression but how could the genes have caused the morphology? Since plant cells are surrounded by a cell wall, morphogenesis becomes purely a problem of growth and the role of cell wall mechanics is especially apparent. Assuming a linear mechanical model, the irreversible growth (strain) rate of the tissue at any point is given by: $\dot{\epsilon} = \phi\,\sigma$ where σ is a measure of the mechanical stress and ϕ is a measure of the material properties and addition of new wall material (extensibility) Lockhart (1965); Ortega (1985). We can think of growth as the balance between the amount of force and the amount of resistance by the tissue. Here, for simplicity, we omit the 'yield' stress by which growth only occurs after a threshold ($\dot{\epsilon} = \phi\,(\sigma - Y)$). One proposed way that genes affect this process locally is by changing the material properties of the tissue to allow either more or less resistance to the mechanical forces that then gives rises to differential growth, $\dot{\epsilon} = \phi(\{g_i\})\,\sigma$ (Figure 1c).

Since we have live imaging and tracking of the cells, we can compute their volumetric growth rate ($\dot{\epsilon} = \frac{1}{V}\frac{dV}{dt}$, computed as in Refahi et al. (2021), range $[-0.04\,h^{-1}, 0.14\,h^{-1}]$) and we can also computationally infer the mechanical state of the cells (maximal stress, σ, computed as in Bozorg et al. (2014)). The inference problem can then be posed as:

$$\frac{\dot{\epsilon}}{\sigma} = \mathbf{g_i}^T w + \eta_i \tag{1}$$

where we assume a linear effect of the genes on the mechanical properties. The weights to learn w can then be interpreted as the contributions of each gene to the mechanical properties and consequently growth via mechanical stress (Figure 1d). Practically, in solving 1 we also used the l^2-norm of the weights to balance the fitting with model simplicity. The predicted growth captures the main spatial pattern of growth in the organ (mean squared error=7×10^{-4}, Figure 1d) and the learned weights propose new hypotheses as they disambiguate the effects proposed by previous analyses Refahi et al. (2021).

3 DISCUSSION

We have proposed a method to infer a mechanical growth model to propose physical action of genes, connecting molecular, mechanical and morphology data to get a more comprehensive picture of development in a complex organ. As with any model there are simplifications. Plant growth is unlikely to be linear and volumetric growth, like we have assumed, misses important aspects of growth like direction and anisotropy. There have been many other components not included here that have been proposed to influence aspect of growth like the microtubule networks in cells Hamant et al. (2008). Instead of volumetric growth, growth can also be captured tensorially (theory is developed Goriely (2017) and the stress tensor is already available), which would capture its direction, for example, and would allow testing (or learning) a larger range of hypotheses including the effect of the cytoskeleton. Practically, however, this would make the learning problem harder and the growth tensor would have to be computed for every point between successive images.

As for our data, we took advantage of the integrated dataset and despite the increasing number of 'atlases', this might not be the case in general. It is more likely that expression data comes in as, for example, single-cell sequencing data. There are already cartography methods to spatially map single-cell data Nitzan et al. (2019) which would improve integration in those cases. Moreover here we ignored previous time steps in our time series data but computing trajectories in shape space would allow a possible mapping with the (pseudo)time developmental trajectories routinely computed for single-cell RNA-seq data using similar temporal cartography methods (e.g., Lis et al. (2020)). Linking trajectories in different spaces in integrated or unintegrated datasets will be valuable and come closer to causality. We can imagine, for example, even learning the terms of the physical theoretical model to further explore the space of hypotheses Kamienny et al. (2022); Maddu et al. (2021).

In conclusion, our proposed method is a powerful way to integrate data from multiple processes in development and can unlock the mechanistic potential in the increasing number of spatial gene

expression data available in plant systems either through spatial transcriptomics or lower throughput methods. These methods can be developed further to act as a vehicle for integrating diverse hypotheses going beyond genes in developmental biology and should carry to other developmental processes in plants and beyond Özpolat et al. (2025).

REFERENCES

Elena R Alvarez-Buylla, Mariana Benítez, Adriana Corvera-Poiré, Álvaro Chaos Cador, Stefan de Folter, Alicia Gamboa de Buen, Adriana Garay-Arroyo, Berenice García-Ponce, Fabiola Jaimes-Miranda, Rigoberto V Pérez-Ruiz, et al. Flower development. *The Arabidopsis Book/American Society of Plant Biologists*, 8, 2010.

Philipp Berens, Kyle Cranmer, Neil D Lawrence, Ulrike von Luxburg, and Jessica Montgomery. AI for science: an emerging agenda. *arXiv preprint arXiv:2303.04217*, 2023.

John L Bowman and Edwige Moyroud. Reflections on the ABC model of flower development. *The Plant Cell*, 36(5):1334–1357, 2024.

John L Bowman, David R Smyth, and Elliot M Meyerowitz. Genes directing flower development in arabidopsis. *The Plant Cell*, 1(1):37–52, 1989.

Behruz Bozorg, Pawel Krupinski, and Henrik Jönsson. Stress and strain provide positional and directional cues in development. *PLoS computational biology*, 10(1):e1003410, 2014.

Enrico S Coen and Elliot M Meyerowitz. The war of the whorls: genetic interactions controlling flower development. *Nature*, 353(6339):31–37, 1991.

Alain Goriely. *The mathematics and mechanics of biological growth*, volume 45. Springer, 2017.

Olivier Hamant, Marcus G Heisler, Henrik Jonsson, Pawel Krupinski, Magalie Uyttewaal, Plamen Bokov, Francis Corson, Patrik Sahlin, Arezki Boudaoud, Elliot M Meyerowitz, et al. Developmental patterning by mechanical signals in arabidopsis. *science*, 322(5908):1650–1655, 2008.

Pierre-Alexandre Kamienny, Stéphane d'Ascoli, Guillaume Lample, and François Charton. End-to-end symbolic regression with transformers. *Advances in Neural Information Processing Systems*, 35:10269–10281, 2022.

Julie Pinol Lis, Paul Villoutreix, and Thierry Artières. Towards a general framework for spatio-temporal transcriptomics. In *LMRL Workshop-NeurIPS 2020*, 2020.

James A Lockhart. An analysis of irreversible plant cell elongation. *Journal of theoretical biology*, 8(2):264–275, 1965.

Suryanarayana Maddu, Bevan L Cheeseman, Christian L Müller, and Ivo F Sbalzarini. Learning physically consistent differential equation models from data using group sparsity. *Physical Review E*, 103(4):042310, 2021.

Manuel Neumann, Xiaocai Xu, Cezary Smaczniak, Julia Schumacher, Wenhao Yan, Nils Blüthgen, Thomas Greb, Henrik Jönsson, Jan Traas, Kerstin Kaufmann, et al. A 3d gene expression atlas of the floral meristem based on spatial reconstruction of single nucleus rna sequencing data. *Nature Communications*, 13(1):2838, 2022.

Mor Nitzan, Nikos Karaiskos, Nir Friedman, and Nikolaus Rajewsky. Gene expression cartography. *Nature*, 576(7785):132–137, 2019.

Tatsuya Nobori, Marina Oliva, Ryan Lister, and Joseph R Ecker. Multiplexed single-cell 3d spatial gene expression analysis in plant tissue using phytomap. *Nature Plants*, 9(7):1026–1033, 2023.

Joseph KE Ortega. Augmented growth equation for cell wall expansion. *Plant physiology*, 79(1): 318–320, 1985.

Mohamed Oughou, Eric Cunha, Eric Biot, Nicolas Arnaud, Patrick Laufs, Philippe Andrey, and Jasmine Burguet. Dynamic atlasing of gene expression domains from individual images. In *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*, pp. 662–665. IEEE, 2019.

B Duygu Özpolat, Swathi Arur, and Mansi Srivastava. A case for broadening our view of mechanism in developmental biology. *Development*, 152(2), 2025.

Wolfram Pönisch, Iskra Yanakieva, Guillaume Salbreux, and Ewa K Paluch. Cell shape noise strength regulates the rate of shape change during emt-associated cell spreading. *bioRxiv*, pp. 2024–10, 2024.

Yassin Refahi, Argyris Zardilis, Gaël Michelin, Raymond Wightman, Bruno Leggio, Jonathan Legrand, Emmanuel Faure, Laetitia Vachez, Alessia Armezzani, Anne-Evodie Risson, et al. A multiscale analysis of early flower development in arabidopsis provides an integrated view of molecular regulation and growth control. *Developmental Cell*, 56(4):540–556, 2021.

Adrienne HK Roeder. Arabidopsis sepals: A model system for the emergent process of morphogenesis. *Quantitative Plant Biology*, 2:e14, 2021.