# Developmental bias explains the evolutionary trend towards simple leaf shapes

James S. Malone,[1,2] Nora S. Martin,[3] Samuel H. A. von der Dunk,[1] Liliana M. Dávalos,[4] and Ard A. Louis[1]

[1] *Rudolf Peierls Centre for Theoretical Physics, University of Oxford, Oxford, OX1 3NP, United Kingdom*
[2] *Department of Biology, University of Oxford, Oxford, OX1 3SZ, United Kingdom*
[3] *CRG (Barcelona Collaboratorium for Modelling and Predictive Biology), Dr. Aiguader 88, Barcelona 08003, Spain*
[4] *Department of Ecology and Evolution and Consortium for Inter-Disciplinary Environmental Research, Stony Brook University, Stony Brook, NY 11790, United States*

(Dated: August 17, 2025)

The relative influence of developmental bias and natural selection on evolutionary outcomes remains a central, unresolved question in evolutionary biology. Here, we combine large-scale phylogenetic trees with herbarium data to confirm that, across the evolutionary history of angiosperms, transitions from complex to simple leaf shapes have occurred at a significantly higher rate than the reverse. To investigate whether this directional trend reflects an underlying developmental bias, we simulate evolutionary transitions by introducing random mutations into a well-established computational model of leaf development. Despite excluding selection, the transition rates produced by the model closely align with those observed in the phylogenetic data. This agreement indicates that a strong developmental bias, reflecting the relative ease with which gene regulatory networks generate simple over complex leaves, has played a major role in shaping the evolution of leaf morphology. Our findings echo recent work showing that similar biases towards simplicity influence the evolution of molecular structures, including RNA secondary structures and protein quaternary assemblies. Together, these results suggest that a developmental bias favouring simplicity may be a widespread and general feature of evolutionary systems.

## I. INTRODUCTION

Much of the phenotypic diversity across the tree of life can be explained by adaptation to ecological demands via natural selection. However, this process is constrained by the mechanisms that generate phenotypic variation, including development [1, 2], which translates genetic information into form. While this idea is widely acknowledged in evolutionary developmental biology [1–3], only recently have advances in methodology and our understanding of development allowed researchers to examine its predictive power for evolutionary outcomes [4–11]. Yet it remains unclear how widespread and strong this developmental bias is across phylogeny [12, 13], particularly for complex traits like morphology. Here, we test whether developmental bias can account for macroevolutionary patterns in angiosperm leaf shape evolution.

Angiosperm leaves are an attractive system to study developmental bias. They display vast morphological diversity, much of which can can be effectively represented in only two spatial dimensions due to minimal growth along the adaxial-abaxial (upper-lower surface) axis [14]. This diversity arises from variations in growth patterns regulated by gene regulatory networks (GRNs) that orchestrate both local and global growth dynamics during leaf development, allowing complex forms to emerge from anisotropic expansion. These processes include deeply conserved genes and developmental mechanisms [15]. Variation in leaf morphology is driven not so much by changes in gene content, but rather by differences in the spatial and temporal expression of shared developmental regulators. This helps explain why closely related species can display wide variation in leaf shape. For instance, the genus *Pelargonium* includes species with unlobed, lobed, dissected, and compound leaves [16]. Similarly, significant variation can occur within a single organism, as seen in *Morus alba* (mulberry), which produces both unlobed and lobed leaves, often with varying degrees of lobe symmetry [14, 17]. The conservation of basic growth mechanisms also suggests that some developmental biases will be, to first order, shared across many different angiosperm species, enabling large-scale studies of the role of development in determining evolutionary outcomes.

At the same time, there is evidence that leaf shape can be influenced by ecological demands such as thermoregulation [18, 19], hydraulic efficiency [20], light interception [21–23] and herbivory [24, 25]. However, it is still unclear how strong and general these effects are in natural environments [26, 27]. Size is often found as a stronger predictor of ecology than shape *per se* [27–30] (see section S3 for further discussion).

In a notable study, Geeta *et al.* [36] analysed a phylogeny of 561 angiosperm species, showing that most leaf shapes (76% in their sample - table II) are unlobed, and that transition rates between leaf shapes are highly asymmetric. Specifically, transitions from more complex forms - such as lobed, dissected, or compound leaves - to simpler unlobed forms occurred far more frequently than transitions in the opposite direction. What explains this global bias towards simpler leaves? Is it primarily development or is it primarily adaptation?

In order to answer this question, we proceed with two parallel lines of work. In the first, we extend the phylogenetic analysis of Geeta *et al.* [36] with two much larger, recent phylogenies [37, 38] and our own shape classification dataset made from 8,911 herbarium images [31]. These confirm that transitions to unlobed leaves are

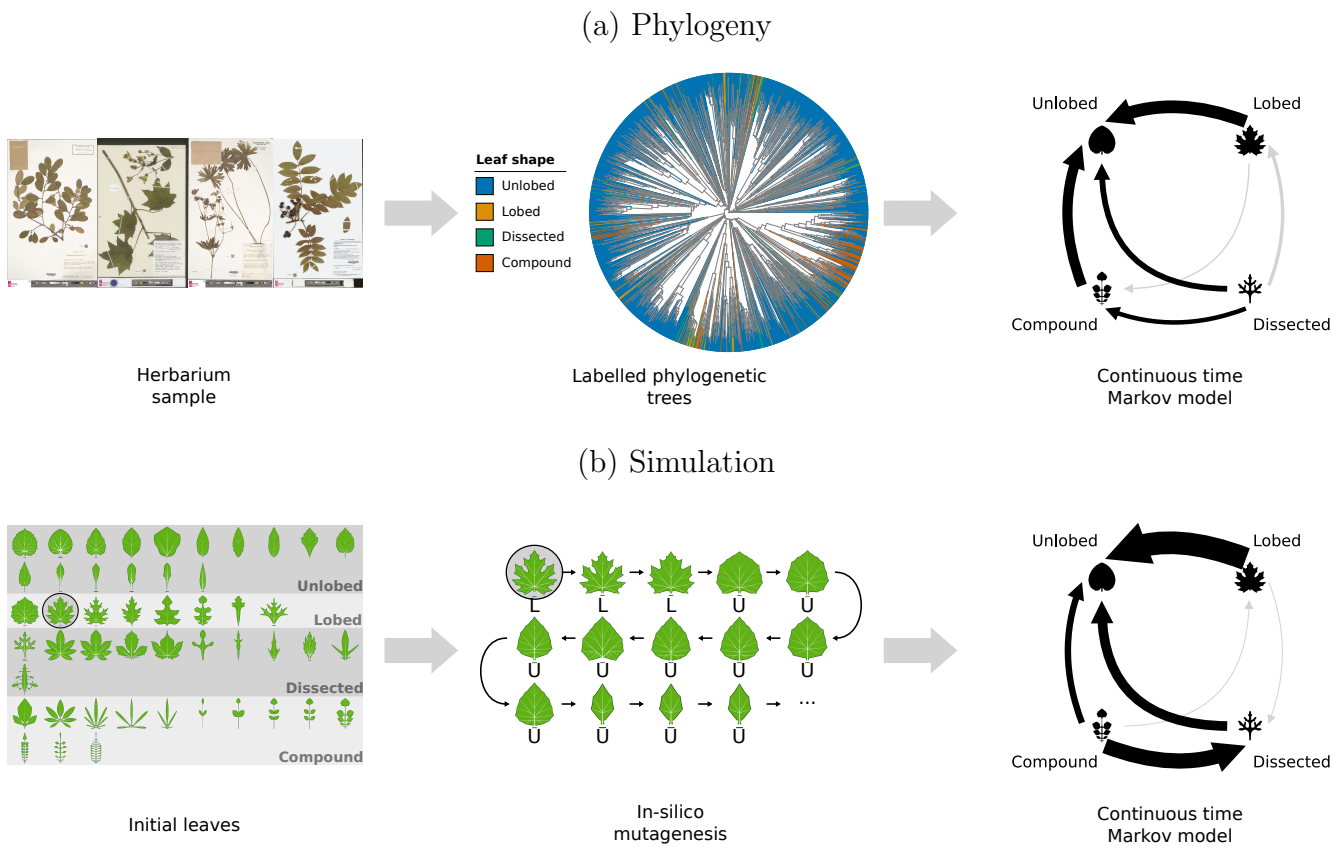



FIG. 1. **A summary of the methods and main results.** We inferred transition rates between leaf shapes in two different ways: **(a)** by fitting a continuous time Markov chain (CTMC) model of trait evolution to phylogenetic trees and **(b)** by simulating mutation in a computational model of leaf development. **(a)** (Left) A sample of images from the Naturalis Biodiversity Center [31][32–35] herbarium are manually classified as unlobed (u), lobed (l), dissected (d) or compound (c) (Methods section IV F). This dataset, as well as an analogous dataset by Geeta *et al.* [36], are used to label the tips of 5 different phylogenetic trees representing angiosperm evolution [36–38]. (Centre) the Zuntini *et al.* [38] species tree is shown in full. Rates are inferred by fitting a CTMC model to the labelled trees using a Markov chain Monte Carlo (MCMC) approach in BayesTraits V4.1.2 [39]. (Right) Posterior net rates for the Zuntini *et al.* [38] species tree are shown as arrows. Black (grey) arrows depict rates for which 0 falls outside (within) a 95% credible interval. **(b)** (Left) An initial sample of 48 leaves representing a wide morphospace is generated using the Runions *et al.* [14] model (Methods sections IV A and IV B). Leaves in this sample are mutated in parallel through a random walk algorithm, with no selection besides rejecting parameter combinations that give physically implausible leaves (table I), generating a sequence of leaves that evolve away from the initial leaf (Methods section IV C). (Middle) Every leaf is classified as unlobed (u), lobed (l), dissected (d) or compound (c) using an automated morphometric approach (Methods section IV D), giving a sequence of character states for each walk. To this trajectory, a CTMC model is fitted using MCMC (Methods section IV E). (Right) Arrows depict posterior net rates for the MUT2 scheme (alg. 1).

much more likely than transitions away from unlobed leaves.

In the second line of work, we examine the role of random mutations using a well-established computational model of angiosperm leaf development by Runions *et al.* [14] (fig. 5). This model captures key conserved developmental mechanisms thought to be broadly shared across angiosperms (Methods section IV A), allowing us to explore the generic effects of neutral evolution in the absence of direct selection on leaf shape. Remarkably, we find that random mutation alone is sufficient to reproduce the relative phylogenetic rates toward unlobed leaf forms. This suggest that developmental bias may be the first order explanation of this macroevolutionary trend.

The basic mechanism driving this bias towards simplicity is straightforward [8]: complex leaf shapes need more specificity in their GRNs than the simpler unlobed leaf shape [40]. As a result, random genetic variation is more likely to produce simple shapes rather than complex ones that need more finely tuned parameter combinations.

## II. RESULTS

### A. Transition rates from simple leaves to complex leaves are greater than vice-versa across angiosperm phylogeny

To establish whether there is a phylogenetic bias towards simple, unlobed leaves, we inferred the rate of evolutionary transitions between different leaf shapes across angiosperm phylogeny. We first classified 8,911 herbarium images (fig. 1a) [31] into 4 discrete categories - unlobed, lobed, dissected and compound. We used this data, as well as another shape classification dataset from Geeta *et al.* [36], to label the tips of 5 phylogenetic trees representing angiosperm evolution [36–38]. Of these trees, 3 were resolved at the level of genus (Zuntini *et al.* [38] genus, Janssens *et al.* [37] genus and Geeta *et al.* [36]). The other two were at the level of species (Zuntini *et al.* [38] species and Janssens *et al.* [37] species) enabling us to examine how altering the amount of species-level diversity affects inferred transition rates. In agreement with Geeta *et al.* [36], we found that for these new phylogenies the vast majority have unlobed leaves (table II), already suggesting a strong bias towards simple, unlobed leaves.

To infer the rate of evolution between each of these 4 categories while accounting for tree topology, we fitted a continuous-time Markov chain (CTMC) model to the labelled trees (Methods section IV E), using a Markov chain Monte Carlo (MCMC) approach implemented in BayesTraits V4.1.2 [41]. In agreement with Geeta *et al.* [36], we detected a significantly higher rate of transitions to the simplest shape (unlobed) compared to transitions away to more complex shapes (figs. 1a and 7a).

The bias towards unlobed shapes was consistent across the five phylogenetic trees (figs. 7a and 7b). We found no cases where the direction of two statistically significant (black arrows - where 0 falls outside a 95% credible interval) net rates disagreed between phylogenies (fig. 7a).

### B. A computational model of leaf development recapitulates *in-vivo* experiments demonstrating developmental bias towards unlobed leaf shape.

To investigate how much of the preference for unlobed leaves can be explained by developmental bias, we use a popular model from Runions *et al.* [14] that combines gene-regulatory networks and key growth mechanisms for leaf shape (Methods section IV A). As a first test, we use the model to analyse *in-vivo* experiments by Kierzkowski *et al.* [40] that varied leaf shape in *Arabidopsis thaliana*.

In particular, this study found that expression of the genes *SHOOTMERISTEMLESS* (*STM*) and *REDUCED COMPLEXITY* (*RCO*) together was necessary to generate the compound shape in *Cardamine hirsuta* and sufficient to induce compound shape in *Arabidopsis thaliana* (fig. 2a). *STM* is a member of the *KNOX* (*Knotted1*-like homeobox) gene family, which has a conserved effect of increasing shape complexity across flowering plants [42]. While *RCO* is not widely conserved across phylogeny [43], the general mechanism by which it acts, locally repressing growth on the margin, is. For example, *CUP-SHAPED COTYLEDON2* (*CUC2*) has a similar effect in the sinuses of *A. thaliana* leaves [44] and has been found to be necessary for shape complexity in diverse lineages [45]. These experimental results suggest that compound leaves require more complex and fine-tuned developmental programmes, with the implication that there are more parameter combinations that produce simpler leaf shapes which do not require this fine-tuning.

To reproduce the main results in Kierzkowski *et al.* [40] with the Runions *et al.* [14] model, we selected two model parameters that control similar growth properties to the genes *STM* – total growth duration and *RCO* – morphogen repression strength (fig. 2). In agreement with their experiments, we found that activating both parameters together was sufficient to lead to a compound shape while activating each in isolation produced a lobed shape instead (figs. 2a and 2b). We also systematically varied morphogen repression strength and duration of growth within their biologically plausible ranges (fig. 2c). We found that the majority of this space (20 of 25 parameter combinations) was occupied by simple, unlobed shapes, with only a small fraction representing more complex shapes (1/25 dissected, 4/25 compound - see Methods section IV D). This implies a developmental bias away from complex leaves towards simpler, unlobed shapes.

### C. Mutation in a computational model of leaf development is sufficient to reproduce phylogenetic bias towards simple leaves

Having shown that the model in Runions *et al.* [14] can reproduce some key results from the experiments of Kierzkowski *et al.* [40], even though it was not specifically designed to model *Arabidopsis thaliana*, we next employ it to study the effects of random mutations more generally. We first generated a sample of initial leaves representing a diverse morphospace (figs. S1 and 1b). We mutated the initial sample using a random walk algorithm with two different mutational schemes, MUT2 (alg. 1) and MUT5 (alg. S2), and automatically classified the resulting shapes into the same four shape categories (unlobed, lobed, dissected and compound) using a morphometric classifier (fig. 6). Of the 76,800 total leaves generated per simulation, we found that the proportion of unlobed, lobed, dissected and compound were 73.4%, 9.6%, 11.4% and 5.6% respectively for MUT2 and 76.1%, 7.2%, 11.4% and 5.4% respectively for MUT5. This shows that our results are robust to the two different mutational schemes. Both reflect the fact that unlobed leaves occupy a significantly greater proportion of model parameter space that produces viable leaves than more

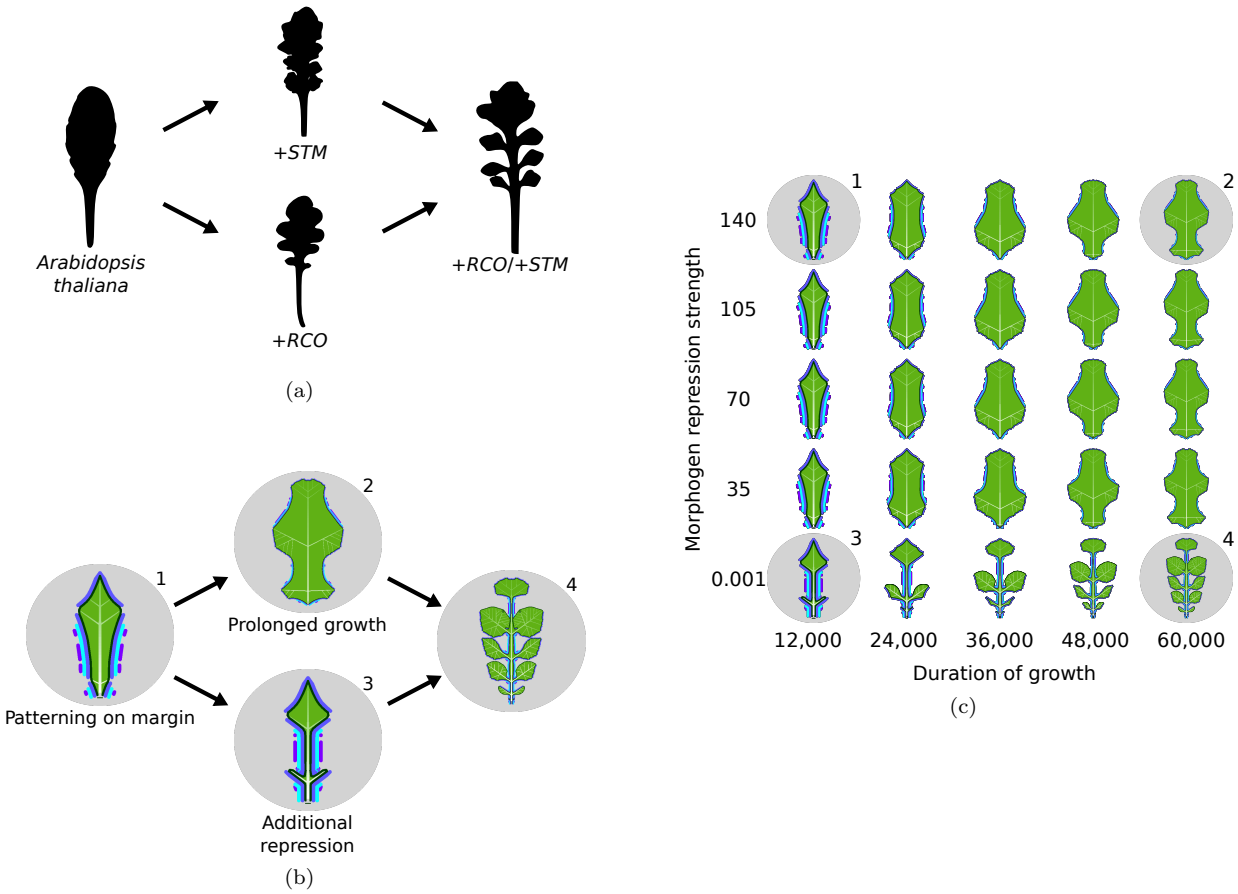



FIG. 2. **An experimentally motivated illustration showing that simple leaves are easier to generate than complex leaves. (a)** Leaf silhouettes of wild-type *A. thaliana* and three mutants, which result in lobed (+*STM*, +*RCO*) and compound shape (+*STM*/+*RCO*) (reproduced from Kierzkowski *et al.* [40] with permission). Ectopic *STM* and *RCO* expression individually create a lobed shape, but together induce compound leaflets. **(b)** A reproduction of the Kierzkowski *et al.* [40] result using the Runions *et al.* [14] computational model of leaf development. +*STM* is approximated by increasing the duration of growth and +*RCO* by increasing the growth repression strength of one morphogen on the margin of the leaf. Morphogen presence is represented by the purple/blue lines outside the margin. **(c)** A morphospace showing the resulting shapes from different levels of the two parameters we use to approximate *STM* and *RCO*. The range shown for *morphogen repression strength* is the range in which biologically plausible shapes can be produced. The range for *duration of growth* is the range in which leaves can be generated in a reasonable timeframe. In the parameter files of the Runions *et al.* [14] model, the parameters we refer to as *morphogen repression strength* and *duration of growth* are given by FAIRING4 and FINALFRAME respectively.

complex shapes do (see also Results section II D).

Regardless of the initial leaf shape, over the course of these random walks, the proportion of walks in an unlobed state increased to an equilibrium just below 80% while most other shape categories stabilise to a significantly lower proportion (figs. 3 and 8). These predictions lie within the margins of error of the phylogeny-derived CTMC model predictions (figs. 3 and 8), suggesting that mutation and development alone are, to first order, sufficient to predict the relative frequency of different leaf shapes over evolutionary time.

To enable a more detailed comparison between simulation and phylogeny, we inferred the rates of shape transitions from our simulations. By estimating rates for all 12 possible (non-self) transitions, we gain deeper insight into the dynamics of shape change over time than by examining the proportions of the four shape categories alone. To do this, we fitted the same CTMC model used for the phylogenetic inference to the random walk data, to estimate the evolutionary rates between shape categories in the absence of everything but mutation, using MCMC [46, 47]. In agreement with our phylogenetic inference, we found that transitions from unlobed occurred at a significantly lower rate than transitions to unlobed, giving a consistent net transition rate towards unlobed (figs. 1b and 7a). This robust agreement between relative rates derived independently from simulations and phylogenies suggests that the evolutionary preference for simple leaves observed across angiosperm phylogeny may primarily be driven by developmental bias instead of selection.

By contrast, for transition rates between complex leaf

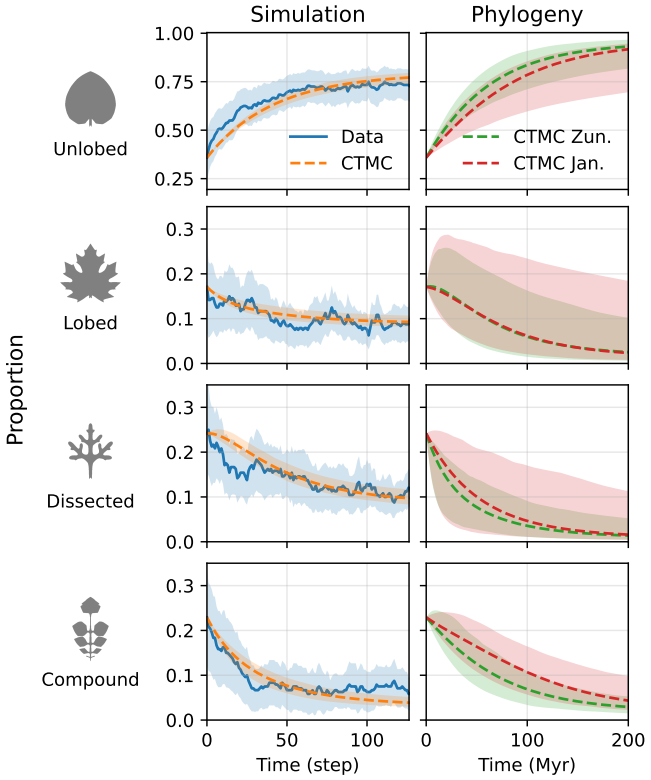



FIG. 3. **The proportions of each leaf shape over time as predicted by *in-silico* mutagenesis and phylogeny. (Simulation)** The solid blue line represents the mean proportion of random walk chains in each state at each time step ($\pm 1.96\times$ standard error) in our simulation (MUT2 alg. 1). To this data, we fit a CTMC via maximum likelihood estimation (MLE) (dashed line) and MCMC with errors inferred from MCMC posteriors via Monte-Carlo error propagation. Since step size is arbitrary, for comparison with the phylogenetic model, we adjust the x-limits for the simulation such that the CTMC predictions for both the simulation and phylogeny reach halfway from the initial distribution to the stationary distribution by approximately the same distance along the axis. **(Phylogeny)** We predict how the proportions of each leaf shape evolve from the same initial proportions used in the simulations, using the CTMC inferred from each phylogeny (green - Zuntini *et al.* [38] genus (fig. S3), red - Janssens *et al.* [37] genus (fig. S5)). Dashed lines represent the prediction using MLE transition rate values. Errors for all CTMC predictions were inferred from the transition rate MCMC posteriors using Monte-Carlo error propagation ($N = 500$), and represent a 95% credible interval for each time step. We show these extrapolations for a time period equal to the approximate age of the angiosperm clade (200 Myr) [38].

shapes (lobed, dissected and compound) (figs. 1 and 7a) there is less agreement between both the different phylogenies, and between the phylogenies and the simulations. These differences could be because selection has played a more prominent role in the transitions among more complex leaf shapes than it has for transitions from complex to simple. The model may also be less reliable at resolving relative rates of complex leaves, for which more detail needs to be specified. Fully resolving the discrepancies for transitions between more complex leaf shapes would require reducing the error associated with the phylogenetic rate estimates, which in turn would necessitate significantly more leaf shape data.

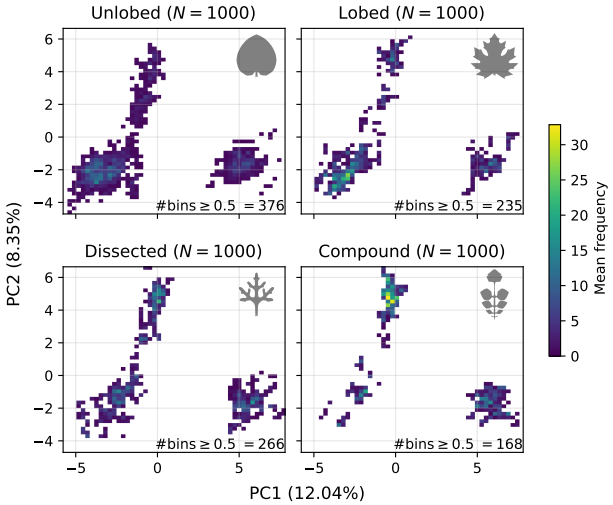
### D. Complex leaves occupy less volume in parameter space and are closer to simple leaves, but not vice versa

To further investigate the mechanisms behind the bias towards simple leaves, we explored – much as we did for the experimentally inspired study in fig. 2c – the relative volumes in parameter space that different shapes occupy. Projecting the parameter values of all random walk outputs into a 2D space via principle component analysis (PCA), reveals several possible factors affecting relative rates (fig. 4). First it shows that each of the four shape categories occupy a different sized area (fig. 4a). The simplest category, unlobed, occupies by far the greatest area (fig. 4b), suggesting that the bias towards simple forms in our simulations can be partially explained because, as argued before (fig. 2), more parameter combinations map to simple leaves than complex leaves. We find that this pattern also holds for PC3-PC4 (fig. S16e) and PC5-PC6 (fig. S16f).

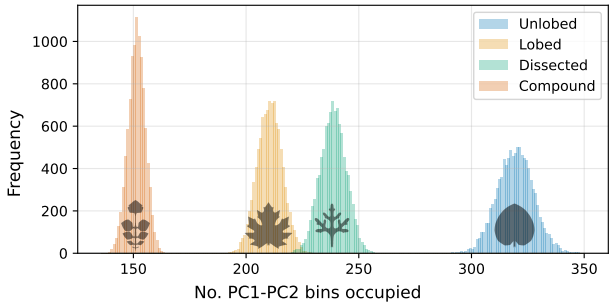
Moreover, the vast majority of lobed, dissected and compound leaves are contained within a broader area occupied by unlobed in the 2D parameter space (fig. 4a), suggesting that more complex leaves are close to simpler leaves in parameter space but not always vice versa. This provides another mechanism making transitions from complex leaves to simple easier than the reverse. We also observe this same pattern for PC3-PC4 (fig. S16b) and PC5-PC6 (fig. S16c).

Finally, the parameter space of unlobed leaves appears to be much more connected than the space of more complex leaves. Constructing a nearest neighbour graph (no. neighbours=6) from the PC1-PC2 parameter space reveals that unlobed separates into far fewer connected components (3) than lobed (131), dissected (96) and compound (235) (fig. 9). This may further contribute to the bias in transition rates towards unlobed, as less connected spaces would be more frequently exited by random walks than more connected spaces.

Our results show that the finding from experiments [40] and manual parameter searches (fig. 2c) – that simpler leaves occupy a larger volume within the parameter space of the developmental programme – extends to the Runions *et al.* [14] model more generally. This volume disparity, in addition to other properties including the degree of overlap between shape categories and the connectedness, may account for much of the bias towards unlobed we see in our simulations,




(a)




(b)

FIG. 4. **Simpler leaves occupy a greater volume than complex leaves in model parameter space. (a)** A 2D histogram of model parameter space given by the 2 principle axes from a principle components anlaysis (PCA) of a sample of 4000 leaves from each shape category from one of the simulations (MUT2). Mean bin occupancy is calculated via a bootstrapping approach, by sub-sampling with replacement $N = 1000$ leaves from each shape category for 10000 iterations. Bins are shaded if the mean occupancy across these iterations is $\geq 0.5$. This reveals that unlobed leaves occupy a larger region of model parameter space than more complex shapes and that this larger space also contains the majority of more complex leaves. Moreover, more complex leaves are concentrated within a much smaller area than unlobed leaves. **(b)** The no. PC1-PC2 2D histogram bins occupied by each leaf shape in 10000 bootstrap samples ($N = 1000$) of a balanced sample of MUT2 (4000 per shape category). Bins with at least 1 leaf present are counted. This reveals that unlobed occupies by far the largest area in this 2D projection of parameter space. We find the same is true for other PCs (fig. S16).

## III. DISCUSSION

We set out to test the hypothesis that mutation alone was sufficient to produce the higher transition rates towards simple leaves compared to complex leaves observed in angiosperm phylogeny [36]. By performing *in-silico* mutagenesis, we found that non-adaptive processes alone produce greater transition rates to simple unlobed leaves than to complex leaf shapes such as lobed or compound (figs. 1b and 3). We also reproduced the observation from angiosperm phylogeny that unlobed leaves have evolved at a faster rate from complex leaves than vice versa [36], using larger and more recent trees [37, 38] and new classification data (figs. 1a and 3). Together, this suggests that a developmental bias towards simple, unlobed leaves can explain this macroevolutionary pattern in flowering plants.

Our finding is consistent with the growing body of evidence that developmental bias can play a significant role in evolutionary outcomes for morphological traits [4–11]. The particular bias towards simplicity is consistent with recent work arguing for a generic bias towards descriptional simplicity in genotype-phenotype maps [8]. Johnston *et al.* [8] provided examples of this bias strongly affecting RNA secondary structure, where simpler structures are much more likely to be observed in nature than complex ones (see also [48]) and for self-assembling protein quaternary structure, where this bias helps explain the prevalence of highly symmetric forms in the protein data bank (PDB). Furthermore, this bias has been predicted to hold more widely for input-output maps [49]. This generic argument implies that the simplicity bias observed in molecular phenotypes may extend to the level of macroscopic phenotypes, which we show here for leaf shape, but also potentially to a much wider set of morphological phenotypes. More specifically, our results suggest that this bias toward simplicity may hold generically when development uses conserved mechanisms and GRNs where complex phenotypes need tighter parameter specification than simpler phenotypes.

Our study also opens up a number of future directions for studying the evolution of leaf shape. First, it would be valuable to repeat similar simulations in alternative models of leaf development [40, 44, 50]. The Runions *et al.* [14] model is uniquely suited to macroevolutionary studies such as this one, as it captures key conserved elements of angiosperm leaf development while generating a large morphospace, but, if the same bias towards simplicity was observed in other models, this would provide further evidence that it is a general property of leaf development. Second, it would be illuminating to conduct similar *in-silico* mutagenesis, but classify the resulting leaves using a continuous classification system instead of discrete categories. For example directly measuring margin complexity using a compression algorithm on the leaf silhouette [51]. This could allow for more subtle biases between different leaf forms to be identified. Third, when inferring transition rates from phylogeny, we were limited by the availability of herbarium data, the resolution of phylogenetic trees and the rate at which herbarium leaves could be manually classified. To more accurately resolve transition rates between more complex shape categories, it would be useful to repeat this study with larger phylogenies. This would likely require an automated method

for herbarium shape classification. Finally, the methodology presented in this study – combining phylogenetic analysis with evolutionary and developmental computational models – could be used to study the evolution of other traits where a reliable developmental model exists, and phylogenetic data is available (e.g.[5]). This would greatly expand our understanding of how well development can predict trait evolution across phylogeny.

## IV. METHODS

### A. Simulating leaf development

We use a well-established model by Runions *et al.* [14] to approximate the developmental programme underpinning many angiosperm leaf shapes. In the model, leaves are represented in 2D with 3 components: a triangular mesh representing the lamina, an open tree structure representing the vasculature and a polygon with morphogen presence or absence at each vertex representing the margin (figs. 5a and 5b).

We selected this model for several reasons. Firstly, unlike similar models which often focus on a limited range of shape diversity e.g. variation within *A. thaliana* [44, 50], Runions *et al.* [14] captures a wide range of angiosperm leaf shape diversity in one parameter space. This is consistent with the known lability and conservation of the leaf developmental programme [14, 16, 52] and enables the inference of transition rates between diverse shapes such as unlobed and compound.

Furthermore, the model integrates key generic properties of angiosperm leaf development such as patterning on the margin and the alignment of growth with the vasculature (fig. 5b). Morphogens affect growth locally and several can be simulated simultaneously on one leaf (fig. 5d). They are abstract and do not directly represent specific genes, however their behaviour can be modified to mimic the effects of real genes. For example, a morphogen can be tuned to repress growth in sinuses like *CUC2* in *A. thaliana* serrations [44, 53], or *Cardamine hirsuta* leaflets [54] (fig. 2). Another key generic property captured in the Runions *et al.* [14] model is that veins determine the axes of growth (fig. 5c). This reflects the way that auxin patterns veins [55, 56] and protrusions [44, 57] in real leaves. In the model, consistent with observed patterns of auxin efflux carriers [58], new veins are inserted between the current vasculature and special points on the margin known as "convergence points". New convergence points form between old convergence points depending on the distance between old points and the effect of any morphogens present (fig. 5c). Growth and form therefore emerge through feedback between growth of the vasculature and patterning on the margin (fig. 5d).

Finally, leaf shapes can be produced from parameter files in a matter of seconds, enabling simulations on a large scale which is necessary for our large-scale, random walk approach.

### B. Generating the initial sample

Since a range of different forms are contained within each shape category, we began our simulations from leaves drawn from an initial sample representing a wide variety of morphologies (16 unlobed, 8 lobed, 11 dissected, 13 compound) (figs. S1 and 1b). This sample was generated using a variety of methods. Some were taken from Runions *et al.* [14], others were generated by tuning parameters by hand and some were taken from the output of random walks on leaves from Runions *et al.* [14].

### C. *In-silico* mutagenesis

To simulate mutation to our sample, we used a constrained random walk algorithm. The algorithm samples the parameter space of biologically plausible leaf shapes surrounding the starting leaf by making a small change to one of the parameters, checking the validity of the resultant leaf (table I), and then moving onto another parameter (see fig. S2 for examples of rejected leaves). We chose this approach over other methods, such as uniform sampling or latin hypercube sampling, as the biologically relevant ranges of each parameter were unknown. Random walks have been shown to cause some biases towards high-robustness regions of parameter space in discrete parameter spaces represented as networks [59], which in this case would result in the under-representation of invalid leaves. We have no reason to assume that this would systematically bias the rates between valid leaves but this cannot be fully excluded. For each of the 48 leaves in our initial sample (fig. S1), we ran 5 random walks for 320 steps, generating a total of $5 \times 48 \times 320 = 76,800$ per simulation. We allow 100/122 model parameters to vary throughout the simulations (table S1). Not all parameters are included, as some are not relevant to leaf shape (e.g. control how the image is rendered).

To test the robustness of our results, we repeated the simulation with 2 different mutational schemes. MUT2 (alg. 1) iterates through the parameters in a random order, and attempts to change the parameter by a value selected at random from an array of numbers randomly generated at 3 different orders of magnitude. MUT5 (alg. S2) is the same as MUT2 except the value each parameter is multiplied by 10% of the range of that parameter within the initial leaves (fig. S1). The aim here was to provide some way of accounting for the sampling range.

### D. Shape classification

To sort the leaves into each of the four shape categories (unlobed, lobed, dissected, compound), we developed a computer-vision program (fig. 6), to reduce error from subjectivity. It finds the coordinates of key landmarks

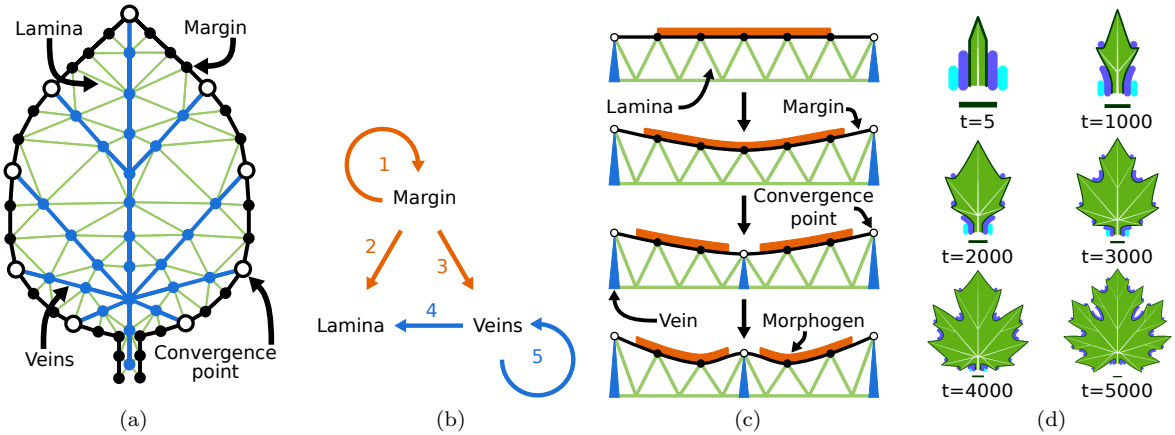



FIG. 5. **A description of the leaf model from Runions *et al.* [14], highlighting the three key components of leaf development and their interactions. (a)** Model components illustrated in leaf form (inspired by Runions *et al.* [14] fig. 3). Margin in black, veins in blue, lamina in green, convergence points in white with black outlines. **(b)** The three components of the model with arrows indicating interactions that affect growth. (1) Existing morphogens and convergence points influence where new ones will form. (2) Convergence points and morphogens influence local growth rate. (3) Convergence points induce new veins. (4) Veins determine local growth direction. (5) Existing veins influence the growth of new veins. **(c)** Zoomed view of the margin, showing how growth of the vasculature and patterning on the margin interact to produce protrusions and sinuses. Veins (in blue) extend, pushing the margin out, however a morphogen (in orange) inhibits webbing between the veins, causing a sinus to form. This increases the distance between the two convergence points, leading to the insertion of another convergence point between the two that excludes the morphogen from its vicinity. Existing vasculature connects to the convergence point and specifies a new axis of growth. **(d)** Different growth stages of a model leaf. The dark blue morphogen inhibits webbing and convergence point formation, the light blue morphogen delineates the petiole. The value of $t$ denotes the iteration of the simulation.

**Algorithm 1:** Random Walk MUT2

$\boldsymbol{\theta} \leftarrow$ initial parameter values
**for** $step \leftarrow 0$ **to** $N_{maxGenerated} - 1$ **do**
  **for** $attempt \leftarrow 0$ **to** $N_{maxAttempts} - 1$ **do**
    $\boldsymbol{\theta}' \leftarrow \boldsymbol{\theta}$  `// Copy parameters for this step`
    $mutation \leftarrow$ random draw from {
      uniform$[0.1, 1)$,
      uniform$[1, 10)$,
      uniform$[10, 100)$
    }
    $sign \leftarrow$ random choice from $\{-1, 1\}$
    $i \leftarrow$ random draw from $\{0, \ldots, \text{length}(\boldsymbol{\theta}) - 1\}$
    $\boldsymbol{\theta}'[i] \leftarrow \boldsymbol{\theta}[i] + (mutation \times sign)$  `// Mutate`
    Run simulation with $\boldsymbol{\theta}'$
    **if** *leaf passes sanity checks* **then**
      $\boldsymbol{\theta} \leftarrow \boldsymbol{\theta}'$  `// Save new parameters`
      **break**

in the leaf image, extracts morphometric measurements and assigns a category based on threshold values of these measurements (eq. (1)). These morphometric measurements were also used to check if leaves were valid (table I) in the random walk algorithm.

First, the main visual features are extracted from the leaf image, including silhouette, margin, lamina and veins. The coordinates of each pixel along the edge of the silhouette (the contour) are extracted. A reference point is defined in the centre of the mid-vein at the point where it first branches outwards into lateral veins (fig. 6b). Then, the distance between each point on the contour and this reference point is calculated (fig. 6c). The local maxima and minima in this distance array that pass a certain prominence threshold are determined using `find_peaks` from the scipy.signal package with `distance=10` and `prominence=25`. These pertinent extrema represent estimates for the coordinates of the tips and sinuses of any protrusions in the leaf margin (fig. 6b shown in red and blue). Because of the prominence threshold, small protrusions such as serrations are excluded.

From the coordinates of the local minima, maxima and the reference point, we calculate a number of morphometric measurements which are then used to classify the leaf. 3 measurements are used: $\epsilon$, the total no. pertinent local maxima and minima; $\alpha$, the average angle bound by each local maxima and its two adjacent minima; and $\omega$, the average horizontal distance between each pair of local minima, closest to the midvein, that share the same vertical coordinate (measured in pixels). The definitions

| Check | Condition |
|---|---|
| *Weight difference* | The difference in leaf area either side of the midvein must not exceed 6% of the total leaf area. |
| *Leaf width* | Final leaf must be wider than 110% of the width of the leaf at the first iteration of the growth simulation. |
| *Veins outside lamina* | Less than 3.7% of the total vein area can be non-overlapping with the lamina. |
| *Veins width* | The maximum width of the vasculature must be greater than 5% of the total image width. |
| *Veins area* | The proportion of total leaf area that is occupied by veins must be greater than 5%. |
| *Overlapping margin* | The proportion of the lamina (interior area of the leaf) that is occupied by leaf margin must be less than 0.1%. |
| *Error or timeout* | Simulation exits unexpectedly or lasts longer than 160 seconds. |

TABLE I. **Random walk constraints.** The only type of selection we apply during the walk is stabilising selection for physically plausible leaves. See fig. S2 for examples of rejected leaves. *Weight difference* (fig. S2a) ensures that the midvein and petiole are roughly in line with the centre of mass of the leaf. *Leaf width* (fig. S2b) ensures that leaves which stop growing after the first few stages of the model simulation are rejected. *Veins outside lamina* (fig. S2c) prevents leaves forming whose veins grow outside the lamina. *Veins width* (fig. S2d) and *Veins area* (fig. S2e) ensure that leaves that generate without adequate vasculature are rejected. *Overlapping margin* (fig. S2f) rejects leaves with ingrown margins. *Error or timeout* – some parameter combinations result in the simulation exiting unexpectedly or taking a very long time to complete, posing an additional constraint.

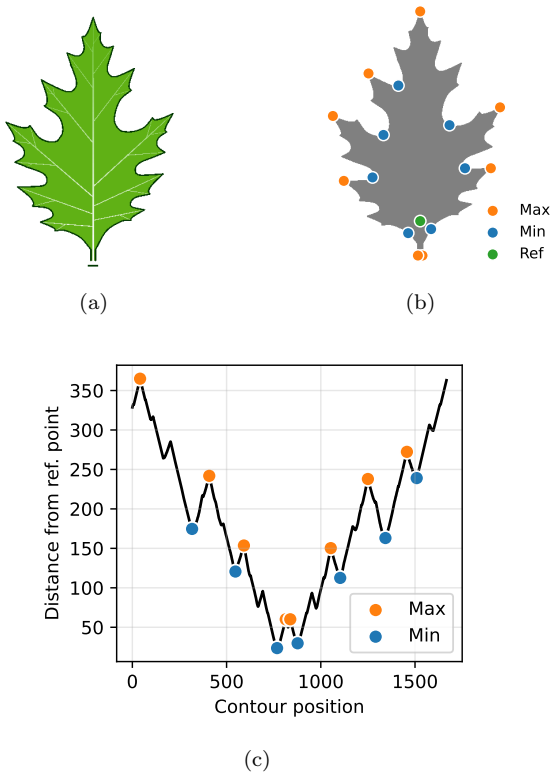



FIG. 6. **A summary of the morphometric classifier. (a)** Example lobed leaf. **(b)** Silhouette with local maxima in orange, minima in blue and reference point in green. **(c)** Graph of distances of each contour point to the reference point. Local maxima shown in orange, minima in blue.

for each of the shape categories are as follows:

$$\left.\begin{array}{r}(\epsilon > 5 \wedge \alpha < 8^\circ)\vee \\ (\epsilon > 5 \wedge \omega < 25 \text{ pixels})\end{array}\right\} \implies \text{Compound (c)},$$
$$\epsilon > 5 \wedge \alpha < 30^\circ \implies \text{Dissected (d)},$$
$$\epsilon > 5 \wedge \alpha < 60^\circ \implies \text{Lobed (l)},$$
$$\epsilon \leq 5 \vee \alpha > 60^\circ \implies \text{Unlobed (u)}. \quad (1)$$

Any leaf that fails to satisfy any of these conditions is labelled as "unclassified", however there were no unclassified leaves in both the MUT2 and MUT5 simulations.

Leaf shape is a continuous trait, thus the boundaries between discrete categories, such as those in eq. (1), are arbitrary. This makes it difficult to assess the accuracy of the classifier algorithm. Nevertheless, we wanted to have some measure of how close it matches classification by eye. Therefore, we manually classified 200 leaves randomly selected from one of the simulations, with an equal number from the output of each initial leaf. 78.5% of the manually ascribed labels matched those assigned by the classifier. For comparison, we fit a machine learning algorithm (Histogram-based Gradient Boosting Regression Tree from the scikit-learn python package) to a sample of 500 manually labelled leaves, sampled in the same way as the set of 200. The algorithm was trained on the distance of 200 equally spaced points along the leaf contour to the reference point for each leaf (fig. 6c). When used to predict shape for the same sample for 200 leaves used for the classical classifier, it returned an accuracy of 73.0%. We then repeated the machine learning approach using morphometric data rather than contour data, which returned an accuracy of 79.5%. This suggests that the threshold classifier is a good enough match with subjective shape classification compared to other classification methods. It also demonstrates that the morphometrics used to construct the original classifier are useful features when classifying shape as unlobed, lobed, dissected or compound. The two machine-learning methods agreed on 77.0% of the labels.

### E. Inferring evolutionary rates from random walk data

To test the model for bias in leaf shape transition rates upon mutation, we ran 5 random walks on each of the 48 initial leaves, for 320 steps, generating a total of 76,800 leaves for both MUT2 and MUT5. Using the classifier (fig. 6), we assigned each leaf the category

unlobed, lobed, dissected or compound (eq. (1)). Evolutionary rates between shape categories were inferred by fitting a continuous-time Markov chain (CTMC) to the labelled walks.

This model is defined by a matrix of transition rates $\boldsymbol{Q}$ between each state - unlobed ($u$), lobed ($l$), dissected ($d$) and compound ($c$)

$$\boldsymbol{Q} = \begin{bmatrix} \sum_{\theta\in\{l,d,c\}} -q_{u\theta} & q_{ul} & q_{ud} & q_{uc} \\ q_{lu} & \sum_{\theta\in\{u,d,c\}} -q_{l\theta} & q_{ld} & q_{lc} \\ q_{du} & q_{dl} & \sum_{\theta\in\{u,l,c\}} -q_{d\theta} & q_{dc} \\ q_{cu} & q_{cl} & q_{cd} & \sum_{\theta\in\{u,l,d\}} -q_{c\theta} \end{bmatrix} \tag{2}$$

- where the diagonal elements must be equal to the negative sum of all other elements in the same row [39].

The model assumes that at any given instant, over a continuous time interval, the system is in one of these four states. The probability $p(i, j, t)$ of transitioning between state $i$ and $j$, depends only on the current state, the transition rate matrix $\boldsymbol{Q}$ and the time $t$ until the end of the transition. The matrix of all possible transition probabilities $\boldsymbol{P}(t)$ relates to $\boldsymbol{Q}$ and $t$ like so [39, 46]

$$\boldsymbol{P}(t) = \exp(\boldsymbol{Q}t). \tag{3}$$

Therefore, the probability of a transition from state $i$ to $j$ over time $t$ is

$$p(i, j, t) = \exp(\boldsymbol{Q}t)_{ij}. \tag{4}$$

The probability of a set of observed walks $x = \{x_0, x_1, ..., x_W\}$, where each walk contains a sequence of labelled steps $x_w = \{x_{w0}, x_{w1}, ..., x_{wN}\}$, given a particular set of rate parameters $\boldsymbol{Q}$ is

$$P(x|\boldsymbol{Q}) = \prod_{w=0}^{W} \prod_{n=0}^{N-1} p(x_{wn}, x_{w(n+1)}, t), \tag{5}$$

where $w$ and $n$ are walk and step index respectively. We assume $t = 1$ for every step, as over the course of the entire random walk, the average step size in parameter space will be approximately constant.

Rates were inferred from the random walk simulation data by finding the maximum of this function, using MCMC with the emcee package for python [60]. We ran 24 chains for 5,000 steps with uniform priors between 0 and 0.1 for all transition rates. Chains were thinned by removing all but every 10th step (figs. S9 and S10) and the first 2,500 were discarded to obtain the posterior distribution. We took the mean of the posterior for each transition as the estimate for each transition rate in fig. 1. As an additional check, we also inferred the rates via maximum likelihood estimation (MLE) and found that they converged with the MCMC estimates (figs. 7a and 7b).

Because the inferred rates depend on the unit of time, we needed to normalise the transition rates to compare with the rates inferred from phylogeny. We did this for each mutational scheme by dividing by the mean of all transition rate estimates $q_{ij}$. We define a normalised rate as

$$\hat{q}_{ij} = \frac{q_{ij}}{N}, \quad N = \frac{1}{12} \sum_{\substack{i,j\in\{u,l,d,c\} \\ i\neq j}} q_{ij}. \tag{6}$$

### F. Inferring evolutionary rates from phylogeny

Although Geeta *et al.* [36] had already inferred rates between these 4 categories with a phylogeny of 561 taxa, we wanted to test how robust these results are. Those results used a phylogeny that is now 25 years old [61], and there are now much larger (e.g. Zuntini *et al.* [38] - 9,506 species) and more reliable phylogenies available. To do this analysis, we combined 2 phylogenetic trees [37, 38] with leaf images from the Naturalis herbarium [31]. We first filtered all digitised records to just those from angiosperm families as per APG IV [62]. We then removed all but the first record for each species.

To see how the rates are affected by looking at the species level vs the genus level, we generated 2 trees for each original tree, one with species at the tips and one with all but one tip per genera removed. For the species level trees, we took the intersect between the filtered herbarium data and the full trees. If more than 3,100 tips remained, we sampled at random 3,100 tips, to allow for manual classification within reasonable time limits. This gave a combined set 5,552 leaf images to classify for both species trees. Notably, the intersect between the herbarium data and the Zuntini *et al.* [38] species tree gave 3057 taxa. Since this was less than 3,100, no further sub-sampling was required, and so unlike all the other trees in this analysis, the Zuntini *et al.* [38] species tree represents the raw intersect between the original tree and the herbarium. For this phylogeny, our genus tree and our species tree should both have similar levels of reliability. For the Janssens *et al.* [37] phylogeny, we trust the genus level tree more than the species level tree, because the overlap with the herbarium data was not as good at the species level. For example, taxa were more heavily concentrated in a smaller number of angiosperm families, compared to the other trees (fig. S8). Moreover, we sub-sampled down to 3,100 taxa from 16,436 in the raw intersect to make manual classification feasible - something not required for the other trees.

For the genus level trees, we further sub-sampled the angiosperm filtered herbarium data to 200 records per family, or the maximum if fewer than 200 records were present. The purpose of this step was to sample more

| Tree | Unlobed (u) | Lobed (l) | Dissected (d) | Compound (c) | ld | lc | ldc | No. tips |
|---|---|---|---|---|---|---|---|---|
| Geeta *et al.* [36] | 428 (76.3%) | 38 (6.8%) | 17 (3.0%) | 68 (12.1%) | 7 (1.2%) | 1 (0.2%) | 2 (0.4%) | 561 |
| Zuntini *et al.* [38] species | 2445 (86.0%) | 102 (3.6%) | 83 (2.9%) | 212 (7.5%) | - | - | - | 2842 |
| Zuntini *et al.* [38] genus | 1910 (88.3%) | 82 (3.8%) | 47 (2.2%) | 123 (5.7%) | - | - | - | 2162 |
| Janssens *et al.* [37] species | 2606 (87.2%) | 125 (4.2%) | 99 (3.3%) | 160 (5.4%) | - | - | - | 2990 |
| Janssens *et al.* [37] genus | 2178 (88.6%) | 97 (3.9%) | 55 (2.2%) | 128 (5.2%) | - | - | - | 2458 |

TABLE II. **The number of taxa in each shape category for each labelled tree used in our inference.** Columns with multiple shape categories (e.g. ld) represent taxa which could fit into any of the listed categories.

evenly over phylogeny, accounting for potential biases in the herbarium. We then took the intersect between genera in this sample and each of our phylogenies and selected at random one species from within each genus that was present in the herbarium to use as the label. This gave a combined set of 3,359 leaf images to classify for both genus trees.

These images were then used to classify each species as one of the 4 categories (unlobed, lobed, dissected or compound) or ambiguous. This was done by eye, relying on a sample of model leaves and their morphometric classifier output for reference (figs. S1 and 1b). After classification, all species labelled as ambiguous were removed, leaving an unambiguously labelled phylogeny (table II and figs. S3 to S7).

We used 3 different phylogenies from the literature to infer the evolutionary rates. First, to compare our predictions directly to the Geeta *et al.* [36] study we used their phylogeny. Second, we used a tree by Janssens *et al.* [37] for a large sample, as this tree contained over of 36,101 species, representing 8,399 genera, 426 families and all orders of flowering plants (available in Supplementary material 5 of Janssens *et al.* [37]). Finally, we used a recent phylogeny by Zuntini *et al.* [38], containing 9,506 species (7,923 genera and all angiosperm families and orders), inferred using 353 nuclear genes and calibrated with 200 fossil records (available at `https://zenodo.org/records/10778207` we used 4-young-tree-smoothing-10-pruned-for-diversification-analyses). We show results from the Geeta *et al.* [36] phylogeny, labelled with their own classification data and the Janssens *et al.* [37] and Zuntini *et al.* [38] trees labelled with our own classification data from Naturalis Biodiversity Center [31].

Rates were inferred using the MultiState method from BayesTraits V4.1.2 [39]. We used MCMC for 1,010,000 iterations with a sample period of 1000 and removed the first 10000 iterations to obtain posterior distributions. We used a uniform prior between 0 and 0.1 for the Zuntini *et al.* [38] trees (figs. S11 and S12) and Janssens *et al.* [37] trees (figs. S13 and S14), however, as the unit of branch length was substitutions per site rather than millions of years, for Geeta *et al.* [36] we instead used a uniform prior between 0 and 100 (fig. S15). We took the mean of each posterior as the transition rate estimate. In addition, we estimated the rates using MLE for comparison (figs. 7a and 7b). As the units of time differ between our phylogenies and the simulations, we needed to normalise the rates to allow for comparison. We did this using the same method as for the simulations (eq. (6)): by dividing each transition rate estimate by the mean transition rate estimate for each tree.

## ACKNOWLEDGMENTS

JSM was supported by funding from the Biotechnology and Biological Sciences Research Council (UKRI-BBSRC) [grant number BB/T008784/1]. LMD was supported, in part, by National Science Foundation OISE 2020577 and a Stony Brook University Presidential Innovation and Excellence. SvdD and NSM were supported by the Issachar fund. NSM acknowledges support of the Spanish Ministry of Science and Innovation through the Centro de Excelencia Severo Ochoa (CEX2020-001049-S, MCIN/AEI/10.13039/501100011033), the EMBL partnership and the Generalitat de Catalunya through the CERCA programme. Research for this publication has been partially carried out in the Barcelona Collaboratorium for Modelling and Predictive Biology. This research is part of grant JDC2022-049526-I to NSM, funded by MCIU/AEI/10.13039/501100011033 and by European Union NextGenerationEU/PRTR.

[1] J. M. Smith, R. Burian, S. Kauffman, P. Alberch, J. Campbell, B. Goodwin, R. Lande, D. Raup, and L. Wolpert, Developmental Constraints and Evolution: A Perspective from the Mountain Lake Conference on Development and Evolution, The Quarterly Review of Biology **60**, 265 (1985).

[2] P. Alberch, Ontogenesis and morphological diversification, Am Zool **20**, 653 (1980).

[3] T. Uller, A. P. Moczek, R. A. Watson, P. M. Brakefield, and K. N. Laland, Developmental Bias and Evolution: A Regulatory Network Perspective, Genetics **209**, 949 (2018).

[4] P. Alberch and E. A. Gale, A developmental analysis of an evolutionary trend: Digital reduction in amphibians, Evolution **39**, 8 (1985).

[5] P. Prusinkiewicz, Y. Erasmus, B. Lane, L. D. Harder,

and E. Coen, Evolution and Development of Inflorescence Architectures, Science **316**, 1452 (2007).

[6] K. D. Kavanagh, O. Shoval, B. B. Winslow, U. Alon, B. P. Leary, A. Kan, and C. J. Tabin, Developmental bias in the evolution of phalanges, Proc. Natl. Acad. Sci. U.S.A. **110**, 18190 (2013).

[7] P. F. Hagolani, R. Zimm, R. Vroomans, and I. Salazar-Ciudad, On the evolution and development of morphological complexity: A view from gene regulatory networks, PLoS Comput Biol **17**, e1008570 (2021).

[8] I. G. Johnston, K. Dingle, S. F. Greenbury, C. Q. Camargo, J. P. K. Doye, S. E. Ahnert, and A. A. Louis, Symmetry and simplicity spontaneously emerge from the algorithmic nature of evolution, Proc. Natl. Acad. Sci. U.S.A. **119**, e2113883119 (2022).

[9] R. Zimm, F. Berio, M. Debiais-Thibaud, and N. Goudemand, A shark-inspired general model of tooth morphogenesis unveils developmental asymmetries in phenotype transitions, Proc. Natl. Acad. Sci. U.S.A. **120**, e2216959120 (2023).

[10] M. Staps, P. W. Miller, C. E. Tarnita, and R. Mallarino, Development shapes the evolutionary diversification of rodent stripe patterns, Proc. Natl. Acad. Sci. U.S.A. **120**, e2312077120 (2023).

[11] P. T. Rohner and D. Berger, Developmental bias predicts 60 million years of wing shape evolution, Proc. Natl. Acad. Sci. U.S.A. **120**, e2211210120 (2023).

[12] K. Laland, T. Uller, M. Feldman, K. Sterelny, G. B. Müller, A. Moczek, E. Jablonka, J. Odling-Smee, G. A. Wray, H. E. Hoekstra, D. J. Futuyma, R. E. Lenski, T. F. C. Mackay, D. Schluter, and J. E. Strassmann, Does evolutionary theory need a rethink?, Nature **514**, 161 (2014).

[13] J. J. Welch, What's wrong with evolutionary biology?, Biol Philos **32**, 263 (2017).

[14] A. Runions, M. Tsiantis, and P. Prusinkiewicz, A common developmental program can produce diverse leaf shapes, New Phytologist **216**, 401 (2017).

[15] L. A. Nikolov, A. Runions, M. Das Gupta, and M. Tsiantis, Leaf development and evolution, in *Current Topics in Developmental Biology*, Vol. 131 (Elsevier, 2019) pp. 109–139.

[16] A. B. Nicotra, A. Leigh, C. K. Boyce, C. S. Jones, K. J. Niklas, D. L. Royer, and H. Tsukaya, The evolution and functional significance of leaf shape in the angiosperms, Functional Plant Biol. **38**, 535 (2011).

[17] C. A. Kidner and S. Umbreen, Why is leaf shape so variable?, International Journal of Plant Developmental Biology **4**, 64 (2010).

[18] S. Vogel, Convective cooling at low airspeeds and the shapes of broad leaves, J Exp Bot **21**, 91 (1970).

[19] T. J. Givnish, Comparative studies of leaf form: Assessing the relative roles of selective pressures and phylogenetic constraints, New Phytologist **106**, 131 (1987).

[20] S. Sisó, J. Camarero, and E. Gil-Pelegrín, Relationship between hydraulic resistance and leaf morphology in broadleaf *Quercus* species: A new interpretation of leaf lobation, Trees **15**, 341 (2001).

[21] J. S. Roden, Modeling the light interception and carbon gain of individual fluttering aspen (*Populus tremuloides* Michx) leaves, Trees **17**, 117 (2003).

[22] R. P. A. Perez, J. Dauzat, B. Pallas, J. Lamour, P. Verley, J.-P. Caliman, E. Costes, and R. Faivre, Designing oil palm architectural ideotypes for optimal light interception and carbon assimilation through a sensitivity analysis of leaf traits, Annals of Botany **121**, 909 (2018).

[23] D. Otto, S. Munz, E. Memic, J. Hartung, and S. Graeff-Hönninger, A computer vision approach for quantifying leaf shape of maize (*Zea mays* L.) and simulating its impact on light interception, Front. Plant Sci. **16**, 1521242 (2025).

[24] D. D. Dell'Aglio, M. E. Losada, and C. D. Jiggins, Butterfly learning and the diversification of plant leaf shape, Front. Ecol. Evol. **4**, 81 (2016).

[25] Y. Higuchi and A. Kawakita, Leaf shape deters plant processing by an herbivorous weevil, Nat. Plants **5**, 959 (2019).

[26] B. E. Campitelli and J. R. Stinchcombe, Testing potential selective agents acting on leaf shape in *Ipomoea hederacea*: Predictions based on an adaptive leaf shape cline, Ecology and Evolution **3**, 2409 (2013).

[27] A. Leigh, S. Sevanto, J. Close, and A. Nicotra, The influence of leaf size and shape on leaf thermal dynamics: Does theory hold up under natural conditions?, Plant Cell & Environment **40**, 237 (2017).

[28] L. Sack and K. Frole, Leaf structural diversity is related to hydraulic capacity in tropical rain forest trees, Ecology **87**, 483 (2006).

[29] C. Scoffoni, M. Rawls, A. McKown, H. Cochard, and L. Sack, Decline of leaf hydraulic conductance with dehydration: Relationship to leaf size and venation architecture, Plant Physiology **156**, 832 (2011).

[30] M. J. Donoghue, D. A. R. Eaton, C. A. Maya-Lastra, M. J. Landis, P. W. Sweeney, M. E. Olson, N. I. Cacho, M. K. Moeglein, J. R. Gardner, N. M. Heaphy, M. Castorena, A. S. Rivas, W. L. Clement, and E. J. Edwards, Replicated radiation of a plant clade along a cloud forest archipelago, Nat Ecol Evol **6**, 1318 (2022).

[31] Naturalis Biodiversity Center, Naturalis Bioportal (2024).

[32] TC. Plowman, Erythroxylum confusum Britton, `https://data.biodiversitydata.nl/naturalis/specimen/U.1284618` (1981).

[33] J. de Bruijn, Kirengeshoma palmata Yatabe, `https://data.biodiversitydata.nl/naturalis/specimen/WAG.1457187` (1981).

[34] P. van Royen, Geranium caespitosum E.James, `https://data.biodiversitydata.nl/naturalis/specimen/L.2011772` (1973).

[35] CR. Bhikhi, Crepidospermum goudotianum (Tul.) Triana & Planch., `https://data.biodiversitydata.nl/naturalis/specimen/U.1615275` (2011).

[36] R. Geeta, L. M. Dávalos, A. Levy, L. Bohs, M. Lavin, K. Mummenhoff, N. Sinha, and M. F. Wojciechowski, Keeping it simple: Flowering plants tend to retain, and revert to, simple leaves, New Phytologist **193**, 481 (2012).

[37] S. B. Janssens, T. L. Couvreur, A. Mertens, G. Dauby, L.-P. M. Dagallier, S. Vanden Abeele, F. Vandelook, M. Mascarello, H. Beeckman, M. Sosef, *et al.*, A large-scale species level dated angiosperm phylogeny for evolutionary and ecological analyses, BDJ **8**, e39677 (2020).

[38] A. R. Zuntini, T. Carruthers, O. Maurin, P. C. Bailey, K. Leempoel, G. E. Brewer, N. Epitawalage, E. Françoso, B. Gallego-Paramo, C. McGinnie, *et al.*, Phylogenomics and the rise of the angiosperms, Nature **629**, 843 (2024).

[39] M. Pagel, Detecting correlated evolution on phylogenies: A general method for the comparative analysis of discrete characters, Proc. R. Soc. Lond. B **255**, 37 (1994).

[40] D. Kierzkowski, A. Runions, F. Vuolo, S. Strauss, R. Lymbouridou, A.-L. Routier-Kierzkowska, D. Wilson-Sánchez, H. Jenke, C. Galinha, G. Mosca, Z. Zhang, C. Canales, R. Dello Ioio, P. Huijser, R. S. Smith, and M. Tsiantis, A Growth-Based Framework for Leaf Shape Development and Diversity, Cell **177**, 1405 (2019).

[41] M. Pagel, A. Meade, and D. Barker, Bayesian Estimation of Ancestral Character States on Phylogenies, Systematic Biology **53**, 673 (2004).

[42] H. Nakayama, A. R. Leichty, and N. R. Sinha, Molecular mechanisms underlying leaf development, morphological diversification, and beyond, The Plant Cell **34**, 2534 (2022).

[43] D. Vlad, D. Kierzkowski, M. I. Rast, F. Vuolo, R. Dello Ioio, C. Galinha, X. Gan, M. Hajheidari, A. Hay, R. S. Smith, P. Huijser, C. D. Bailey, and M. Tsiantis, Leaf shape evolution through duplication, regulatory diversification, and loss of a homeobox gene, Science **343**, 780 (2014).

[44] G. D. Bilsborough, A. Runions, M. Barkoulas, H. W. Jenkins, A. Hasson, C. Galinha, P. Laufs, A. Hay, P. Prusinkiewicz, and M. Tsiantis, Model for the regulation of *Arabidopsis thaliana* leaf margin development, Proc. Natl. Acad. Sci. U.S.A. **108**, 3424 (2011).

[45] T. Blein, A. Pulido, A. Vialette-Guiraud, K. Nikovics, H. Morin, A. Hay, I. E. Johansen, M. Tsiantis, and P. Laufs, A Conserved Molecular Framework for Compound Leaf Development, Science **322**, 1835 (2008).

[46] J. D. Kalbfleisch and J. F. Lawless, The Analysis of Panel Data under a Markov Assumption, Journal of the American Statistical Association **80**, 863 (1985).

[47] R. T. McGibbon and V. S. Pande, Efficient maximum likelihood parameterization of continuous-time Markov processes, The Journal of Chemical Physics **143**, 034109 (2015).

[48] S. H. A. von der Dunk, N. S. Martin, K. Dingle, and A. A. Louis, RNA secondary structures are conserved but random, in prep.

[49] K. Dingle, C. Q. Camargo, and A. A. Louis, Input–output maps are strongly biased towards simple outputs, Nat Commun **9**, 761 (2018).

[50] E. E. Kuchen, S. Fox, P. Barbier De Reuille, R. Kennaway, S. Bensmihen, J. Avondo, G. M. Calder, P. Southam, S. Robinson, A. Bangham, and E. Coen, Generation of Leaf Shape Through Early Patterns of Growth and Tissue Polarity, Science **335**, 1092 (2012).

[51] N. S. Martin, C. Q. Camargo, and A. A. Louis, Bias in the arrival of variation can dominate over natural selection in Richard Dawkins's biomorphs, PLoS Comput Biol **20**, e1011893 (2024).

[52] M. Bar and N. Ori, Leaf development and morphogenesis, Development **141**, 4219 (2014).

[53] K. Nikovics, T. Blein, A. Peaucelle, T. Ishida, H. Morin, M. Aida, and P. Laufs, The Balance between the *MIR164A* and *CUC2* Genes Controls Leaf Margin Serration in *Arabidopsis*, The Plant Cell **18**, 2929 (2006).

[54] N. Bhatia, D. Wilson-Sánchez, S. Strauss, F. Vuolo, B. Pieper, Z. Hu, L. Rambaud-Lavigne, and M. Tsiantis, Interspersed expression of CUP-SHAPED COTYLEDON2 and REDUCED COMPLEXITY shapes Cardamine hirsuta complex leaf form, Current Biology **33**, 2977 (2023).

[55] E. Scarpella, M. Barkoulas, and M. Tsiantis, Control of Leaf and Vein Development by Auxin, Cold Spring Harbor Perspectives in Biology **2**, a001511 (2010).

[56] E. Scarpella, Leaf Vein Patterning, Annual Review of Plant Biology **75**, 377 (2024).

[57] A. Hay, M. Barkoulas, and M. Tsiantis, ASYMMETRIC LEAVES1 and auxin activities converge to repress *BREVIPEDICELLUS* expression and promote leaf development in *Arabidopsis*, Development **133**, 3955 (2006).

[58] E. Scarpella, D. Marcos, J. Friml, and T. Berleth, Control of leaf vascular patterning by polar auxin transport, Genes Dev. **20**, 1015 (2006).

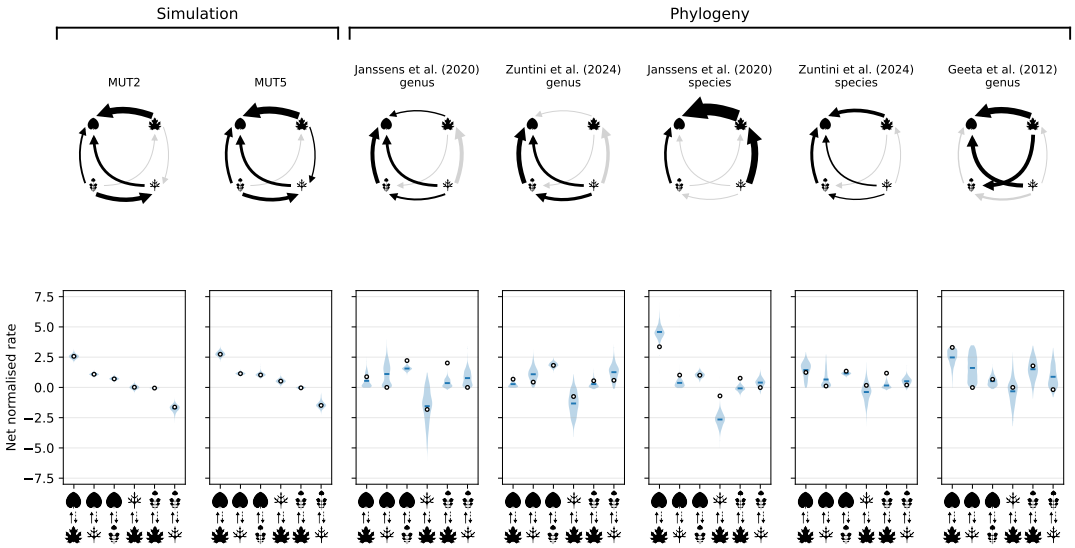
[59] E. Van Nimwegen, J. P. Crutchfield, and M. Huynen, Neutral evolution of mutational robustness, Proc. Natl. Acad. Sci. U.S.A. **96**, 9716 (1999).

[60] D. Foreman-Mackey, D. W. Hogg, D. Lang, and J. Goodman, Emcee: The MCMC Hammer, Publications of the Astronomical Society of the Pacific **125**, 306 (2013).

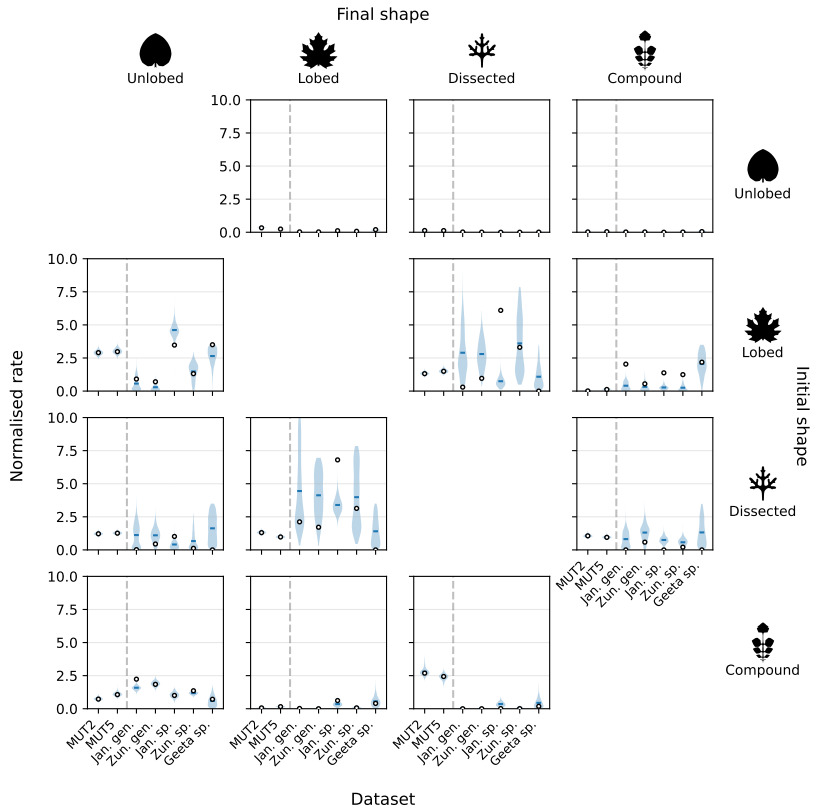
[61] D. Soltis, Angiosperm phylogeny inferred from 18S rDNA, rbcL, and atpB sequences, Botanical Journal of the Linnean Society **133**, 381 (2000).

[62] The Angiosperm Phylogeny Group, An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV, Bot. J. Linn. Soc. **181**, 1 (2016).

## V. EXTENDED DATA




(a)




(b)

FIG. 7. **Further details of transition rates from phylogenies and simulations. (a)** Net normalsied transition rates between the 4 leaf shape categories, inferred from both simulation and phylogeny. (Top row) Arrow thickness and direction represents the mean net normalised transition rate between each pair of leaf shape categories. Black arrows represent transitions where the 0 does not fall within an equal-tailed 95% credible interval for the net normalised posterior, and grey where it does. (Bottom row) Posterior distributions for the net normalised rate of all transitions. Solid bars indicate the mean. The net rate is calculated by subtracting the transition given by the dashed arrow from the transition given by the solid arrow in the x-axis labels. **(b)** Normalised transition rates between different leaf shapes, inferred from both *in-silico* mutagenesis (left of dashed line) and phylogeny (right of dashed line). Each subplot represents a different transition, from the shape shown on the vertical figure axis to the shape shown on the horizontal figure axis. These rates were normalised by dividing by the mean transition rate estimate within each dataset (eq. (6)).

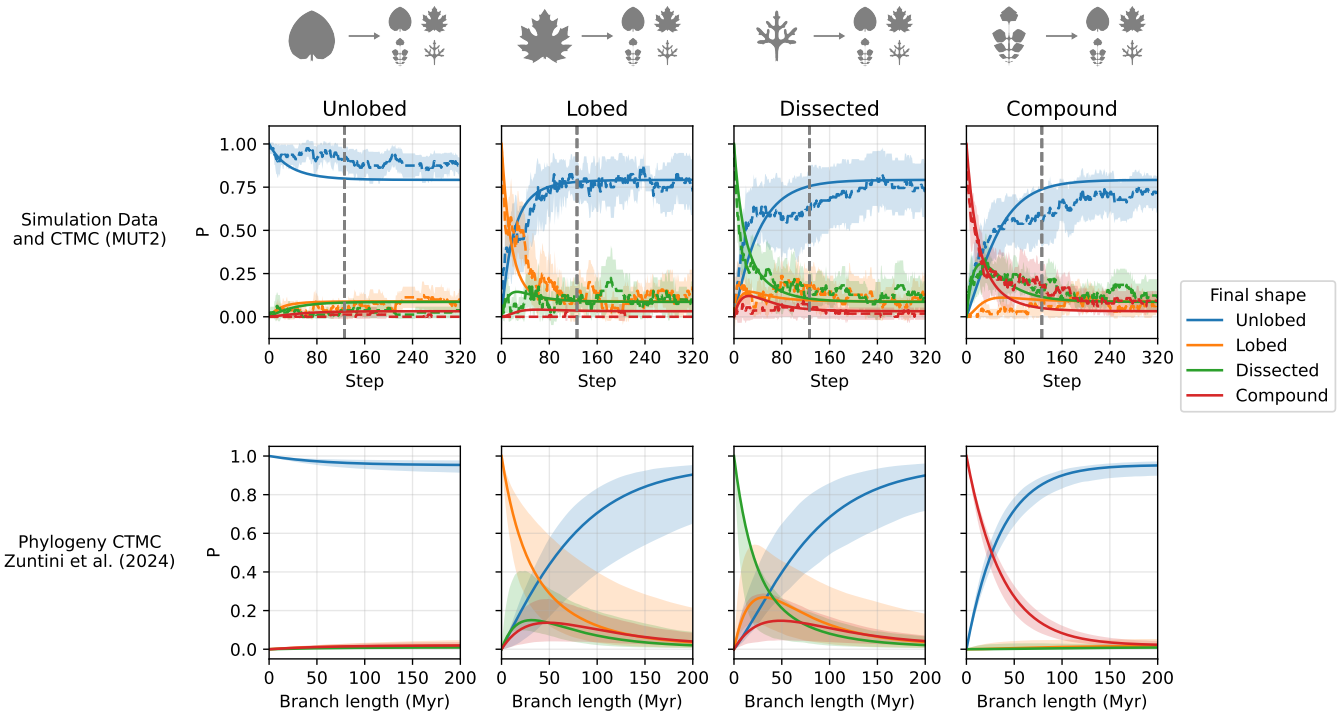



FIG. 8. **Comparing *in-silico* mutagenesis and phylogeny for the probability of each shape transition over time, separated out by starting shape.** The column header gives the initial shape and the line colour gives the final shape. **(Top row)** The dashed line represents the proportion of walks from each unique initial leaf (fig. S1), averaged over all initial leaves in each state. The error bands represents $\pm 1.96\times$standard error for this average. The solid lines represents the CTMC model fit (MLE) to the output of the random walks. The vertical, grey, dashed line represents the estimated simulation step that is equivalent to $t = 200$ Myr for the phylogeny CTMC. This is calculated as described in fig. 3, and is also the simulation x-limit in fig. 3. We show all 320 simulation steps here for a more detailed summary of the simulation trajectories. **(Bottom row)** The predicted transition probabilities given by the CTMC model fit to the labelled Zuntini *et al.* [38] genus tree (fig. S3). The solid line represents the MLE fit and the errors are inferred from MCMC posteriors via Monte-Carlo error propagation ($N = 500$) as per fig. 3. Note the good qualitative agreement between the dynamics between the phylogenies and the simulations, especially for the (blue) data related to transitions to/from unlobed leaves.

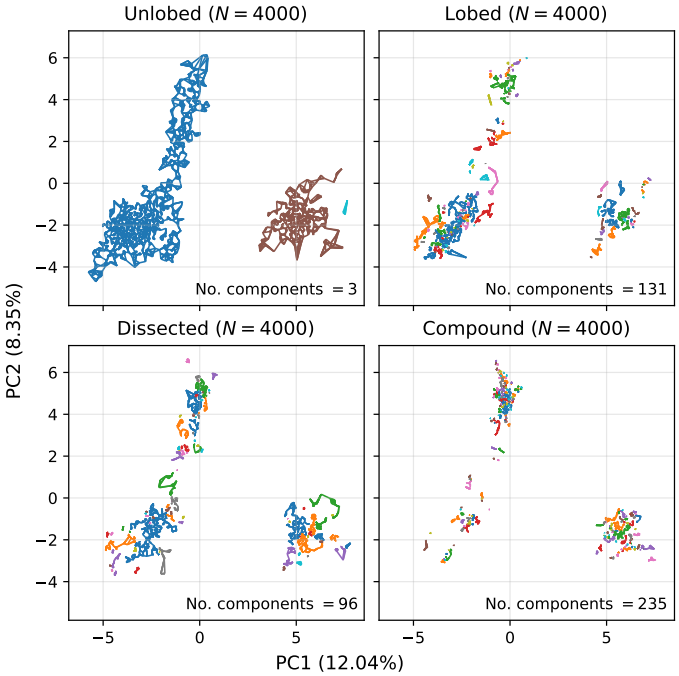



FIG. 9. **The nearest neighbour graph of parameter PCA space for each shape category highlights how the unlobed space is the most connected, followed by dissected, lobed and then compound.** Complex leaves subdivide into many more connected components than unlobed. The number of nearest neighbours used is 6. Colours highlight different connected components, however, since the number of unique colours is only 10, this under represents the true number of connected components for lobed, dissected and compound. Thus, the colours here are intended only as a visual guide.