PROCEEDINGS A

rspa.royalsocietypublishing.org

Research

Cite this article: Choi GPT, Mahadevan L. 2018 Planar morphometrics using Teichmüller maps. *Proc. R. Soc. A* **474**: 20170905. http://dx.doi.org/10.1098/rspa.2017.0905

Received: 10 January 2018
Accepted: 31 August 2018

Subject Areas:
computational biology, computational mathematics, applied mathematics

Keywords:
geometric morphometrics, Teichmüller maps, shape classification, community detection

Author for correspondence:
L. Mahadevan
e-mail: lmahadev@g.harvard.edu

Electronic supplementary material is available online at https://dx.doi.org/10.6084/m9.figshare.c.4238162.

THE ROYAL SOCIETY PUBLISHING

Planar morphometrics using Teichmüller maps

Gary P. T. Choi[1] and L. Mahadevan[1,2,3]

[1]John A. Paulson School of Engineering and Applied Sciences, [2]Departments of Physics, and Organismic and Evolutionary Biology, and [3]Kavli Institute for Bionano Science and Technology, Harvard University, Cambridge, MA 02138, USA

GPTC, 0000-0001-5407-9111; LM, 0000-0002-5114-0519

Inspired by the question of quantifying wing shape, we propose a computational approach for analysing planar shapes. We first establish a correspondence between the boundaries of two planar shapes with boundary landmarks using geometric functional data analysis and then compute a landmark-matching curvature-guided Teichmüller mapping with uniform quasi-conformal distortion in the bulk. This allows us to analyse the pair-wise difference between the planar shapes and construct a similarity matrix on which we deploy methods from network analysis to cluster shapes. We deploy our method to study a variety of *Drosophila* wings across species to highlight the phenotypic variation between them, and Lepidoptera wings over time to study the developmental progression of wings. Our approach of combining complex analysis, computation and statistics to quantify, compare and classify planar shapes may be usefully deployed in other biological and physical systems.

1. Introduction

Biological form both constrains and enables biological function. It is thus important to characterize biological shape, and determine methods for its quantification, comparison and classification. Geometric morphometrics focuses on the use of landmark coordinates for shape quantification [1]. For instance, one may quantify a planar shape by its boundary, or by the location of specific features in its interior. Furthermore, one needs to find a way to compare different shapes based on the locations of features along the boundaries and in the interior; this may be done in terms of maps that transform one shape onto another, using modern variants [2,3] of

the theory of transformations [4]. Finally, one needs a shape classification scheme based on these comparisons, such as those based on machine learning and neural networks [5].

Here, we propose a computational morphometric method for planar shapes, a problem of interest in many different fields, from computer vision to biological perception, from pathology to palaeontology, from development to evolution. Our aim is to account for the differences in both the boundaries and landmarks in the interior of compact planar domains by providing an accurate 1-1 correspondence that facilitates meaningful comparison between two shapes, and to furthermore enable the classification of multiple shapes. Current morphometric methods typically consider either the boundary or the interior of shapes but not both, and do not attempt to exactly match landmarks. For instance, Procrustes superimposition [6], which has been widely used in planar morphometrics [7,8], matches landmarks via rigid transformations up to scaling but does not allow for exact landmark or boundary matching. The Thin Plate Spline method [9] allows one to compute non-rigid landmark-based transformations for measuring local variation in shape [10] but does not allow for exact boundary matching or guarantee bijectivity. A promising approach is the Large Deformation Diffeomorphic Method [11–13], which allows for the computation of diffeomorphic mappings of shapes with landmarks; however, the computation is expensive and thus hinders the pairwise comparison between a large set of shapes. Recently, there has been a resurgence of an old method, the use of conformal maps, which are angle-preserving functions to describe shape changes [14–16]. However, in general, maps cannot handle landmark constraints because of the rigidity of these maps imposed by the Riemann mapping theorem. Quasi-conformal maps, a generalization of conformal maps which allow a certain degree of angular distortions with the presence of landmark constraints get around this, and recent attempts [17,18] use quasi-conformal maps for morphometry based on variational principles.

We build on advancements in quasi-conformal theory and functional data analysis and develop a landmark-matching, curvature-guided Teichmüller mapping technique, which overcomes all the above-mentioned problems. We then deploy the technique for comparison and classification of planar wing shapes across developmental and evolutionary time scales.

2. Landmark-matching, curvature-guided Teichmüller maps for comparing shapes

We start by considering S_1, S_2 as two planar shapes that we want to compare, as shown in figure 1*a*. Let $\{l_k^{\text{int}_1}\}_{k=1}^m$ and $\{l_k^{\text{int}_2}\}_{k=1}^m$ be two sets of landmarks at the interior of S_1 and S_2, respectively, and $\{l_k^{\text{bdy}_1}\}_{k=1}^n$ and $\{l_k^{\text{bdy}_2}\}_{k=1}^n$ be two sets of landmarks on the boundaries ∂S_1 and ∂S_2. Our goal is to find a map $f : S_1 \to S_2$ that satisfies

$$f(l_k^{\text{int}_1}) = l_k^{\text{int}_2}, \quad k = 1, 2, \ldots, m \tag{2.1}$$

and

$$f(l_k^{\text{bdy}_1}) = l_k^{\text{bdy}_2}, \quad k = 1, 2, \ldots, n. \tag{2.2}$$

Moreover, it is desirable that f reflects the difference between S_1 and S_2 under the prescribed landmark correspondences.

(a) Boundary matching based on curvature

To compute a landmark-matching Teichmüller map $f : S_1 \to S_2$ between the two planar shapes, we first need to determine the boundary correspondence $\varphi : \partial S_1 \to \partial S_2$. It is natural to use boundary curvatures to match the boundaries of S_1 and S_2, so that the highly curved parts of ∂S_1 will correspond to the highly curved parts of ∂S_2, while the relatively flat parts of ∂S_1 will correspond to the relatively flat parts of ∂S_2. To represent the curvature variations of ∂S_1 and ∂S_2, we first define the accumulated arc length and the accumulated curvature of a planar curve

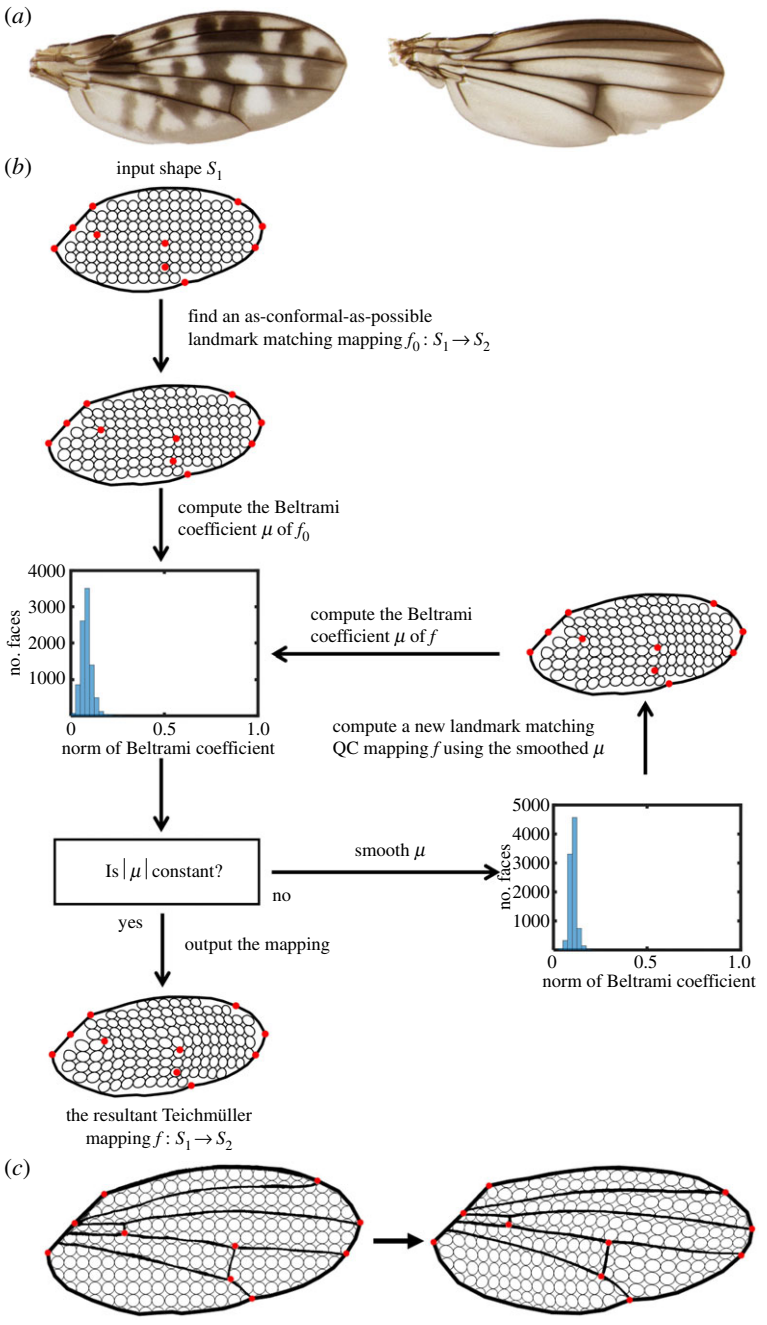

Figure 1. (*a*) Two planar wing shapes that we want to compare. (*b*) A flow chart describing the approach following [23,24] for computing discrete landmark-matching Teichmüller mappings (for details see electronic supplementary material, Algorithm S1). We first discretize the images and use a curvature-guided boundary correspondence to obtain a landmark-matching initial map (see text for details). Then we use an iterative scheme to make it converge to a landmark-matching Teichmüller map. (*c*) A landmark-matching Teichmüller map between two planar wing shapes. Small circles in the left shape are mapped to small ellipses with a uniform aspect ratio in the right shape (see text for details). Red landmark points of vein intersections on the left are also mapped to consistent locations on the right. The wing images are adapted from [28].

C by $x_t = \sum_{i=0}^{t} l(i)$ and $y_t = \sum_{i=0}^{t} \kappa(i)$. Here, $l(i)$ and $\kappa(i)$ are, respectively, the arc length and the curvature approximated at the ith point on C. Then, the function ψ defined by $\psi(x_t) = y_t$ for all t encodes the curvature distribution of C. Using this idea, we obtain two functions ψ_1, ψ_2 that

represent the curvature distributions of $\partial S_1, \partial S_2$, with a reparametrization of the domains of ψ_1, ψ_2 to be $[0,1]$ and consider $\psi_1, \psi_2 : [0,1] \to \mathbb{R}$.

Our next step is then to match the curvature distributions by aligning ψ_1 and ψ_2. To do this in a reparametrization independent way, we use the square root velocity function (SRVF) dynamic warping method [19,20] that considers a bijection from ψ_1, ψ_2 to the SRVFs $q_1, q_2 : [0,1] \to \mathbb{R}$ defined by

$$q_1 = \mathrm{sgn}(\dot{\psi_1})\sqrt{|\dot{\psi_1}|} \quad \text{and} \quad q_2 = \mathrm{sgn}(\dot{\psi_2})\sqrt{|\dot{\psi_2}|}. \tag{2.3}$$

Instead of aligning ψ_1, ψ_2 directly, the SRVF dynamic warping method aims to find the optimal alignment of ψ_1, ψ_2 by aligning the SRVFs q_1, q_2 using a warping function

$$\gamma^* \in \Gamma := \{\gamma : [0,1] \to [0,1] \,|\, \gamma(0) = 0, \gamma(1) = 1, \gamma \text{ is a diffeomorphism}\}. \tag{2.4}$$

An optimal γ^* can be found by solving the following minimization problem

$$\gamma^* = \mathrm{argmin}_{\gamma \in \Gamma} \|q_1 - (q_2 \circ \gamma)\sqrt{\dot{\gamma}}\|_2, \tag{2.5}$$

where $\|\cdot\|_2$ denotes the Euclidean 2-norm. Using dynamic programming to solve this problem, we obtain a curvature-guided correspondence

$$\psi_1(t) \leftrightarrow \psi_2(\gamma(t)), t \in [0,1]. \tag{2.6}$$

This gives us the desired curvature-guided map $\varphi : \partial S_1 \to \partial S_2$.

In case there are boundary landmark constraints as specified in equation (2.2), we partition ∂S_1 and ∂S_2 according to the boundary landmarks $\{l_k^{\mathrm{bdy}_1}\}_{k=1}^n$ and $\{l_k^{\mathrm{bdy}_2}\}_{k=1}^n$. This gives us n pairs of corresponding boundary segments. For each pair of segments, we deploy the above procedures and match them based on their curvature distributions. This results in a curvature-guided boundary mapping $\varphi : \partial S_1 \to \partial S_2$ that satisfies

$$\varphi\left(l_k^{\mathrm{bdy}_1}\right) = l_k^{\mathrm{bdy}_2}, \quad k = 1, 2, \ldots, n. \tag{2.7}$$

(b) Quasi-conformal theory and Teichmüller maps

Conformal maps preserve angles and thus infinitesimal circles are mapped onto infinitesimal circles. Relaxing this condition allows us to define quasi-conformal maps as follows—they are homeomorphisms $f : D \subset \mathbb{C} \to \mathbb{C}$ satisfying the Beltrami equation $\partial f/\partial \bar{z} = \mu_f(z)(\partial f/\partial z)$ for some complex-valued function $\mu_f(z)$ with $\|\mu_f(z)\|_\infty < 1$. Here μ_f is called the Beltrami coefficient of f. Intuitively, this allows quasi-conformal maps to map infinitesimal circles to infinitesimal ellipses with bounded eccentricity. To see this, let z_0 be a point in D. The first-order approximation of f around z_0 is given by

$$\begin{aligned} f(z) &\approx f(z_0) + f_z(z_0)(z - z_0) + f_{\bar{z}}(z_0)\overline{z - z_0} \\ &= f(z_0) + f_z(z_0)(z - z_0 + \mu_f(z_0)\overline{z - z_0}), \end{aligned} \tag{2.8}$$

i.e. infinitesimal circles are mapped to infinitesimal ellipses with the maximum magnification $|f_z|(1 + |\mu_f|)$ and the maximum shrinkage $|f_z|(1 - |\mu_f|)$, so that the aspect ratio of the ellipses, the dilatation, is $(1 + |\mu_f|)/(1 - |\mu_f|)$. Thus, the maximal dilatation of the quasi-conformal map f is defined by $K_f = (1 + \|\mu_f\|_\infty)/(1 - \|\mu_f\|_\infty)$. A classical result is that among all quasi-conformal maps, Teichmüller maps achieve a constant $|\mu_f(z)|$ over the entire domain D, so that every infinitesimal circle on D is mapped to an infinitesimal ellipse with a constant aspect ratio [21].

Teichmüller map-based morphometrics have a number of natural advantages. First, given a boundary correspondence $\varphi : \partial S_1 \to \partial S_2$ and landmark constraints $\{l_i^1\} \subset S_1 \leftrightarrow \{l_i^2\} \subset S_2$, there exists a unique landmark-matching Teichmüller map f that achieves the minimum maximal dilatation over the space of all landmark-matching quasi-conformal maps [22]. In other words,

$f = \text{argmin}_{h:h|_{\partial S_1}=\varphi,\ h(l_i^1)=l_i^2} K_h$. Second, the bijectivity of Teichmüller map is guaranteed [21], giving a 1-1 correspondence between every part of two shapes. Third, the Beltrami coefficient μ_f with constant norm over the whole domain gives us a measure of how similar two shapes are. Since $|\mu_f|$ always lies within $[0, 1)$, and equals 0 if and only if f is conformal, if $1 - |\mu_f| = 1$, then the two shapes are identical up to conformal maps. If $1 - |\mu_f| \ll 1$, there is a large local quasi-conformal dissimilarity between the two shapes. Besides, if we denote $w = f(z)$, by the definition of the Beltrami coefficient we have

$$0 = \mu_{f^{-1}\circ f} = \frac{(f^{-1}\circ f)_{\bar{z}}}{(f^{-1}\circ f)_z} = \frac{((f^{-1})_w \circ f) f_{\bar{z}} + ((f^{-1})_{\bar{w}} \circ f)\bar{f}_{\bar{z}}}{((f^{-1})_w \circ f) f_z + ((f^{-1})_{\bar{w}} \circ f)\bar{f}_z} = \frac{\mu_f + (\mu_{f^{-1}} \circ f)\frac{\bar{f}_{\bar{z}}}{f_z}}{1 + (\mu_{f^{-1}} \circ f)\frac{\bar{f}_{\bar{z}}}{f_z}}, \tag{2.9}$$

which implies that $|\mu_f(z)| = |\mu_{f^{-1}}(f(z))|$ for all z. Therefore, $1 - |\mu_f|$ provides an inverse consistent measurement of similarity between the two shapes.

Using the boundary correspondence $f|_{\partial S_1} = \varphi$ established in the last subsection, along with efficient iterative and provably convergent algorithms [23–25], we can obtain a landmark-matching, curvature-guided Teichmüller map $f : S_1 \to S_2$ with the interior landmark constraints in equation (2.1). Since f is Teichmüller, the norm of the associated Beltrami coefficient $|\mu_f|$ is a constant, and the quantity $1 - |\mu_f|$ is a measure of quasi-conformal similarity between S_1 and S_2. The procedure is summarized in electronic supplementary material, Algorithm S1, and depicted pictorially in figure 1, along with an example showing the map linking two planar shapes corresponding to insect wings.

3. Statistical analysis of the quasi-conformal similarity matrix

Given a set of planar shapes $\{S_i\}_{i=1}^p$, we can construct a $p \times p$ similarity matrix M, where the (i, j)th entry of M represents a measure of similarity between S_i and S_j defined in terms of the magnitude of the Beltrami coefficients connecting these shapes. The values of all entries of M are within the range $[0, 1]$, where a larger value indicates a higher level of similarity. This sets the stage for performing a statistical analysis on M and ultimately clustering the set of shapes.

(a) Adaptive thresholding

Graphically, the $p \times p$ similarity matrix M can be represented by a weighted directed graph with p vertices, where every vertex represents a shape in the set $\{S_i\}_{i=1}^p$, and every pair of vertices are connected by two directed weighted edges. The weight of the directed edge $[i, j]$ is given by M_{ij}. Intuitively, a larger weight M_{ij} represents a higher level of similarity between S_i and S_j from the perspective of S_i. Note that the similarity matrix M is dense since the graph is complete with $p^2 - p$ edges. To highlight the important information in it, we use an adaptive thresholding algorithm that iteratively modifies and sparsifies M.

More explicitly, given a thresholding parameter λ, to judge the importance of the edge $[i, j]$ from vertex i, we consider the quantity $v_i = \bar{M}_i + \lambda\sigma_i$, where $\bar{M}_i$ and σ_i are, respectively, the mean and the standard deviation of $\{M_{ik}\}_{k=1}^p$. If $M_{ij} > v_i$ we set $M_{ij} = 1$. If not, we neglect the edge by setting $M_{ij} = 0$. For a pair of directed edges $[i, j]$ and $[j, i]$, there are exactly three possibilities: (i) $M_{ij} = M_{ji} = 1$, i.e. S_i and S_j are similar to each other, (ii) $M_{ij} = M_{ji} = 0$, i.e. S_i, S_j are dissimilar and (iii) $M_{ij} = 1, M_{ji} = 0$ or $M_{ij} = 0, M_{ji} = 1$, i.e. it is not clear that whether S_i, S_j are sufficiently similar to be grouped in one community. To better represent this, we symmetrize M by taking $M \longleftarrow (M + M^T)/2$, so that $M_{ij} \in \{0, \frac{1}{2}, 1\}$ indicates the relationship of S_i, S_j for all $i, j = 1, \ldots, p$. Algorithmically, we repeat the thresholding and symmetrizing steps on M until the result converges. The final matrix M serves as a weighted adjacency matrix characterizing the relationship of the shapes, and an algorithm for its construction is summarized in electronic supplementary material, Algorithm S2, along with convergence proofs and the discussion on the thresholding parameter λ.

(b) Clustering and community detection

Finally, to cluster all shapes into several communities based on M, we apply a recent community detection method [26] that accounts for the non-locality and asymmetry of the connections between edges [27], noting that the importance of an edge to the two nodes it connects may be different. Denoting the set of shapes $\mathcal{S} = \{S_i\}_{i=1}^{p}$, the similarity between two communities $C_I, C_J \subset \mathcal{S}$ is given by $\mathrm{sim}_{IJ} = \mathrm{sim}^{\mathrm{in}} - \mathrm{sim}^{\mathrm{out}}$, where $\mathrm{sim}^{\mathrm{in}} = (1/\|C_I \cup C_J\|^2) \sum_{i,j \in C_I \cup C_J} g_{ij}$ represents the average similarity score inside the communities, and $\mathrm{sim}^{\mathrm{out}} = (1/(\|C_I \cup C_J\| \cdot \|\mathcal{S} \backslash C_I \cup C_J\|)) \sum_{i \in \mathcal{S} \backslash C_I \cup C_J} \sum_{j \in C_I \cup C_J} g_{ij}$ represents the average similarity score outside the communities. In each iteration, the communities with high sim_{IJ} are combined until the grouping result stabilizes. The final communities formed represent the clustering result based on our Teichmüller morphometric method. We remark that the clustering result is robust to the choice of the community detection method. A comparison of the clustering results obtained using different community detection methods is provided in the electronic supplementary material.

4. Quantifying wing shape in evolution and development

(a) Phenotypic variation of Hawaiian *Drosophila* wings

As an application of our Teichmüller-map morphometric framework, we study the *Drosophila* wings in the Hawaiian Drosophila Wing Database [28], which consists of data drawn from four 'picture wing' phylogenetic groups, including the *adiastola* group, the *planitibia* group, the *glabriapex* group and the *grimshawi* group. We discretize every wing image from the different species using a triangular mesh with approximately 9000 triangle elements. For each wing, seven points are manually chosen as boundary landmarks, including the intersections between the longitudinal veins L2, L3, L4, L5 and the wing boundary. Three intersections between the veins are manually chosen as interior landmarks, including the intersections between L4 and the anterior cross-vein (ACV) and the posterior cross-vein (PCV), and the intersection between L5 and PCV. The 10 landmark locations are shown in electronic supplementary material, figure S1. In figure 2, we show our landmark-matching Teichmüller map comparing a pair of wings from the *D. punalua* and the *D. silvestris* species and observe that our approach is able to match boundary and interior landmarks accurately. A comparison between our method and other existing morphometric methods is provided in electronic supplementary material, Section S4 and figure S6. A mathematical approach for assessing the difference between the veins of two wings using the landmark-matching Teichmüller map is also discussed in electronic supplementary material, Section S8.

In figure 3, we depict the 128×128 similarity matrix whose elements $M_{ij} = 1 - |\mu_f(i,j)|$ are calculated as discussed above (electronic supplementary material, Algorithm S1). This serves as the initial step to apply our proposed adaptive thresholding algorithm (electronic supplementary material, Algorithm S2) with the thresholding parameter $\lambda = 1$. Finally, we apply the community detection method [26] on the thresholded similarity matrix. To visualize the result, we make use of a multidimensional scaling (MDS) coordinate plane which projects the information of a dissimilarity matrix onto the Euclidean plane, with the similarity information preserved as distances between nodes. Every specimen in our dataset is visualized as a node on the plane. In figure 4, we show our results and observe that the nodes are clustered into three groups, with the species *D. glabriapex* (denoted by Community 1), *D. planitibia* (denoted by Community 2) and *D. grimshawi* (denoted by Community 3) being a representative in each of them.

Of the 22 specimens in the *glabriapex* phylogenetic group, 20 (91%) of them are classified into Community 1. Of the 33 specimens in the *planitibia* phylogenetic group, 28 (85%) of them are classified into Community 2. Of the 41 specimens in the *grimshawi* phylogenetic group, 21 (51%) of them are classified into Community 3 and 18 (44%) are classified into Community 1. Of the 32 specimens in the *adiastola* phylogenetic group, 18 (56%) of them are classified into Community 3 and 12 (38%) are classified into Community 1. From the above, we observe that Community 2

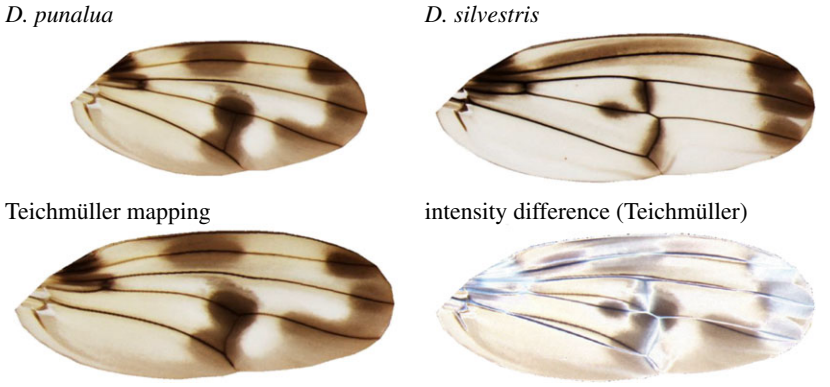

Figure 2. The landmark-matching, curvature-guided Teichmüller mapping between the species *D. punalua* and the species *D. silvestris*. The top row shows the *D. punalua* and *D. silvestris* wings. The bottom row shows the Teichmüller map of *D. punalua* onto *D. silvestris*, and the intensity difference between the Teichmüller mapping result and the *D. silvestris* wing. The *D. punalua* and *D. silvestris* wing images are adapted from [28].

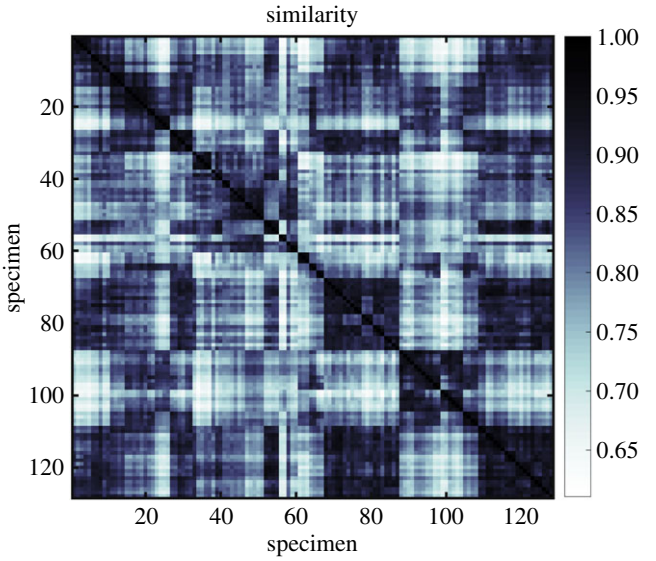

Figure 3. The quasi-conformal similarity matrix obtained by our landmark-matching, curvature-guided Teichmüller mapping method (electronic supplementary material, Algorithm S1) for the comparison between the wings in the Hawaiian Drosophila Wing Database [28].

primarily consists of wings from the *planitibia* phylogenetic group but not the other three groups. We deduce that the wings in the *planitibia* phylogenetic group share highly similar phenotypic features and are very different from the wings in all other phylogenetic groups. Because of the high percentage of specimens in the *glabriapex* phylogenetic group classified into Community 1, there is also a high level of similarity among the wings in the *glabriapex* phylogenetic group. By contrast, the *adiastola* phylogenetic group and the *grimshawi* phylogenetic group demonstrate a higher level of shape diversity as both of the two groups are primarily classified into two communities.

To understand the community detection results, we further analyse the phenotypic features of the three communities photographically. The images of the wings in Community 1, Community 2 and Community 3 are shown in the electronic supplementary material. It can be observed that

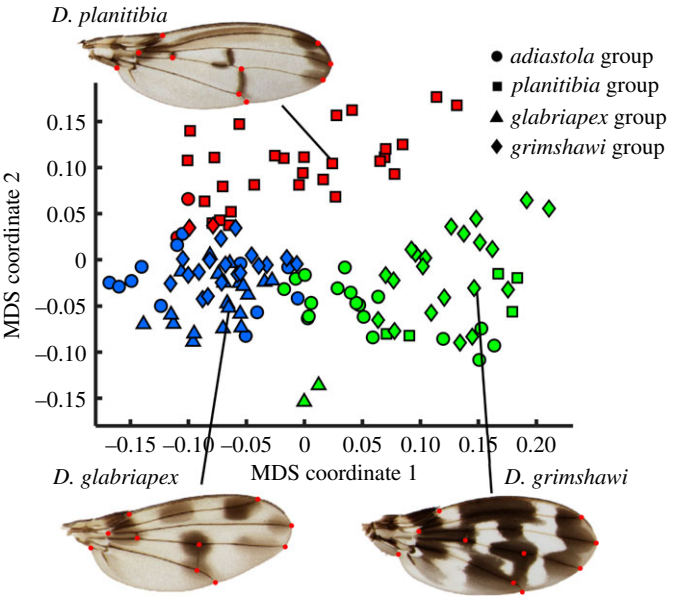

Figure 4. Visualizing the community detection result obtained by our Teichmüller morphometric framework on the MDS coordinate plane. Every specimen is represented as a node on the plane constructed by multidimensional scaling. The nodes are coloured based on their communities. Blue: Community 1. Red: Community 2. Green: Community 3. The shapes of the nodes represent their phylogenetic groupings. Circle: *adiastola* group. Square: *planitibia* group. Triangle: *glabriapex* group. Diamond: *grimshawi* group. The wing images shown are adapted from [28].

the wing geometries in the three communities are different in terms of the wing shapes and the relative locations of the landmarks. The wings in Community 1 are typically with a relatively round shape at the bottom of the wing boundary, and a relatively sharp wing tip. The intersection between L5 and PCV is relatively far away from the wing boundary. The wings in Community 2 are typically with an elongated shape, and the intersection between L5 and PCV is relatively close to the wing boundary. The wings in Community 3 are typically with a round silhouette. Also, the intersection between L4 and ACV is relatively distal when compared with that in the other two communities.

For the two species *D. hemipeza* and *D. picticornis*, which belong to the *planitibia* phylogenetic group but are classified into Community 3, we note that their wings are different from the other wings in the *planitibia* phylogenetic group. More specifically, the *D. picticornis* wing has a relatively round shape and is more pigmented than all other wings in the *planitibia* phylogenetic group. This can be explained by the early division in the phylogeny of the *planitibia* phylogenetic group [29] that separates *D. picticornis* from the other species. For *D. hemipeza*, we observe that the intersection between L4 and ACV is distal relative to the same landmark for the other wings in the *planitibia* phylogenetic group. The above observations might explain the exceptions in the community detection result for the *planitibia* phylogenetic group.

Our focus so far has been on landmarks based on signatures associated with venation network motifs. Since our methods are agnostic to the type of pattern motif used as landmarks, we also deployed it to see if it might shed light on wing pigmentation patterns, a trait that is controlled by just a few genes [30], and is likely to be more labile. In the wings of the three communities we considered, there are essentially three different pigment patterns. Community 1 consists primarily of most specimens from the *glabriapex* phylogenetic group, around half of the specimens from the *grimshawi* phylogenetic group, and around half of the specimens from the *adiastola* phylogenetic group. The majority in this community possesses a moderate number of pigment spots, which occur near the top part of the wing, the intersection between L4 and PCV, and the intersection between L5 and PCV. Community 2 which primarily consists of most specimens from the *planitibia*

phylogenetic group has wings that possess a small number of pigment spots, which occur near the distal tips of L2, L3 and L4. Community 3 primarily consists of half of the specimens from the *grimshawi* phylogenetic group, and half of the specimens from the *adiastola* phylogenetic group. The majority in this community possesses a high level of pigmentation. It is perhaps surprising that our Teichmüller-based classification which is solely based on wing boundary shape and venation landmark positions also shows a clustering effect with respect to pigmentation patterns. A plausible explanation for this is that there is crosstalk between the genes encoding for both phenotypes, and may be worth exploring further using phylogenetic approaches in the future.

Going beyond venation network and wing pigmentation patterns, our Teichmüller map-based methods can also be used for detecting subtle shape dissimilarities such as microevolutionary patterns and bilateral asymmetry; an example of analysing the bilateral asymmetry of Hawaiian *Drosophila* wings is provided in electronic supplementary material, Section S7, as an illustration that might perhaps be expanded in the future.

A natural question that arises at this point is the connection between community detection using Teichmüller morphometrics and other classification approaches using macroevolutionary cladistics. While they are not the same, for comparison, we decided to start with our similarity matrix and apply the leave-one-out cross-validation (LOOCV) [31], which uses one wing shape as the validation data and the remaining wing shapes as the training data to classify it. We obtain a classification accuracy of around 70% with respect to the phylogenetic groupings. For example of 'misclassification' in our community detection result, wings of the species *D. adiastola* are split into two communities. To further investigate this, we compare our clustering result with the result obtained by the traditional geometric morphometrics approach [32] (see electronic supplementary material, Section S9) and find that the traditional approach also splits the *D. adiastola* wings. This suggests that the 'misclassification' of *D. adiastola* wings in our clustering result is not an error, and instead that *D. adiastola* indeed exhibits multimodal wing patterns.

(b) Temporal development of Lepidoptera wings

Our proposed Teichmüller morphometric method (electronic supplementary material, Algorithm S1) can also be used for analysing the temporal development of Lepidoptera wings. We deploy our method on the forewings of the species *Manduca sexta* and *Junonia coenia* in [33] at the larval, prepupal, pupal and adult stages. A Teichmüller map is computed between every two successive stages. The constant norm of the Beltrami coefficient of the Teichmüller map between the two stages, denoted by $\Delta_{|\mu|}$, represents the local shear and the quasi-conformal dissimilarity between the two shapes. While the local shear is a constant over the entire domain, the local area change and orientation change may vary. We analyse the changes by constructing a circle packing on the wing at the earlier temporal stage. Then we map the younger wing onto the older wing using the Teichmüller map, and visualize it via the deformation of the circle packing on the older shape. In general, the circles are deformed to ellipses with different size and orientation. The change in size relative to the original circles reflects the local area change, and the change in orientation relative to the original circles characterizes the local rotation. Here we quantify the change in size by the quantity $\Delta_A =$ (Area of ellipse)/(Area of circle) for each small circle, and the change in orientation by $\Delta_p =< 2\cos^2\theta - 1 >$ which averages over the 1-ring neighbourhood of every circle, where θ is the orientation change of the circle under the mapping. Proximal-distal orientation of the major axis is denoted by $\Delta_p = 1$ while an anterior–posterior orientation of the major axis yields $\Delta_p = -1$.

Figures 5 and 6 show the Teichmüller maps between successive stages of *Manduca sexta* and *Junonia coenia* wings, respectively. We first analyse the overall change between every two successive stages quantified by the quasi-conformal dissimilarity $\Delta_{|\mu|}$, the average local area change mean(Δ_A) and the average local orientation change mean(Δ_p). For *Manduca sexta*, the largest $\Delta_{|\mu|}$ occurs between prepupa and pupa. By contrast, for *Junonia coenia*, $\Delta_{|\mu|}$ increases throughout the development. The average local area changes of the two species are also different. For *Manduca sexta*, mean(Δ_A) decreases throughout the development, while for *Junonia coenia* the greatest mean(Δ_A) occurs between prepupa and pupa. It is also noteworthy that the magnitude

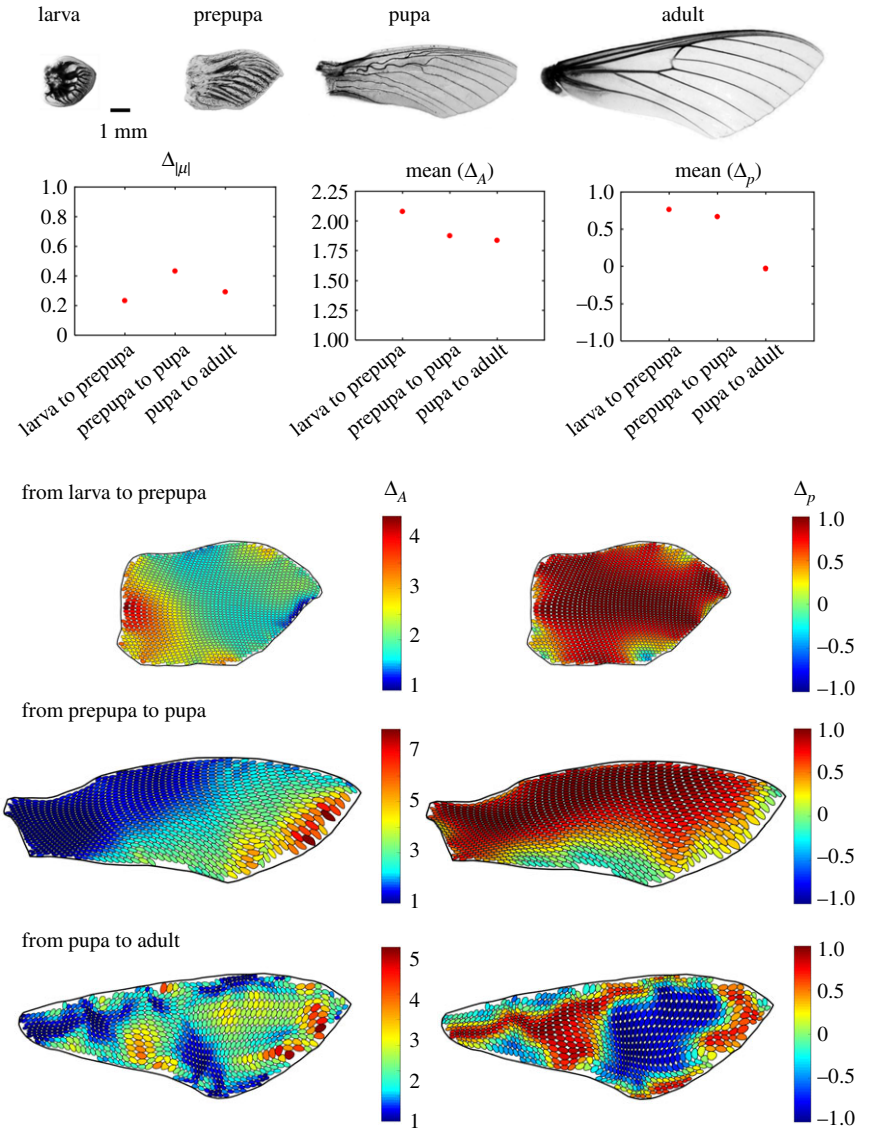

Figure 5. Analysing the temporal development of *Manduca sexta* wings via Teichmüller map. The top row shows the wing images adapted from [33] at different developmental stages (displayed to scale). A Teichmüller map is computed between every two successive stages. The second row shows the quasi-conformal dissimilarity $\Delta_{|\mu|}$, the average of the local size change Δ_A and the average of the local orientation change Δ_p between every two successive stages. The last three rows show the deformation of the circle packing under the Teichmüller map between every two successive stages, visualized on the latter stage. The left column shows the resulting ellipses colour-coded by the local size change Δ_A. The right column shows the resulting ellipses colour-coded by the local orientation change Δ_p. For visualization purpose, the shapes in the last three rows are rescaled to have the same height, with their aspect ratio kept unchanged.

of mean(Δ_A) for *Junonia coenia* is greater than that for *Manduca sexta* between every two successive stages. For the local orientation change, both *Manduca sexta* and *Junonia coenia* wings undergo an overall proximal-distal orientation change at the earlier stages of the development, as mean(Δ_p) > 0. However, the two species have a notable difference in the overall orientation change between pupa and adult. For *Manduca sexta* we have mean(Δ_p) ≈ 0, which indicates that the overall orientation change is small, while for *Junonia coenia* there is a notable overall proximal-distal orientation change.

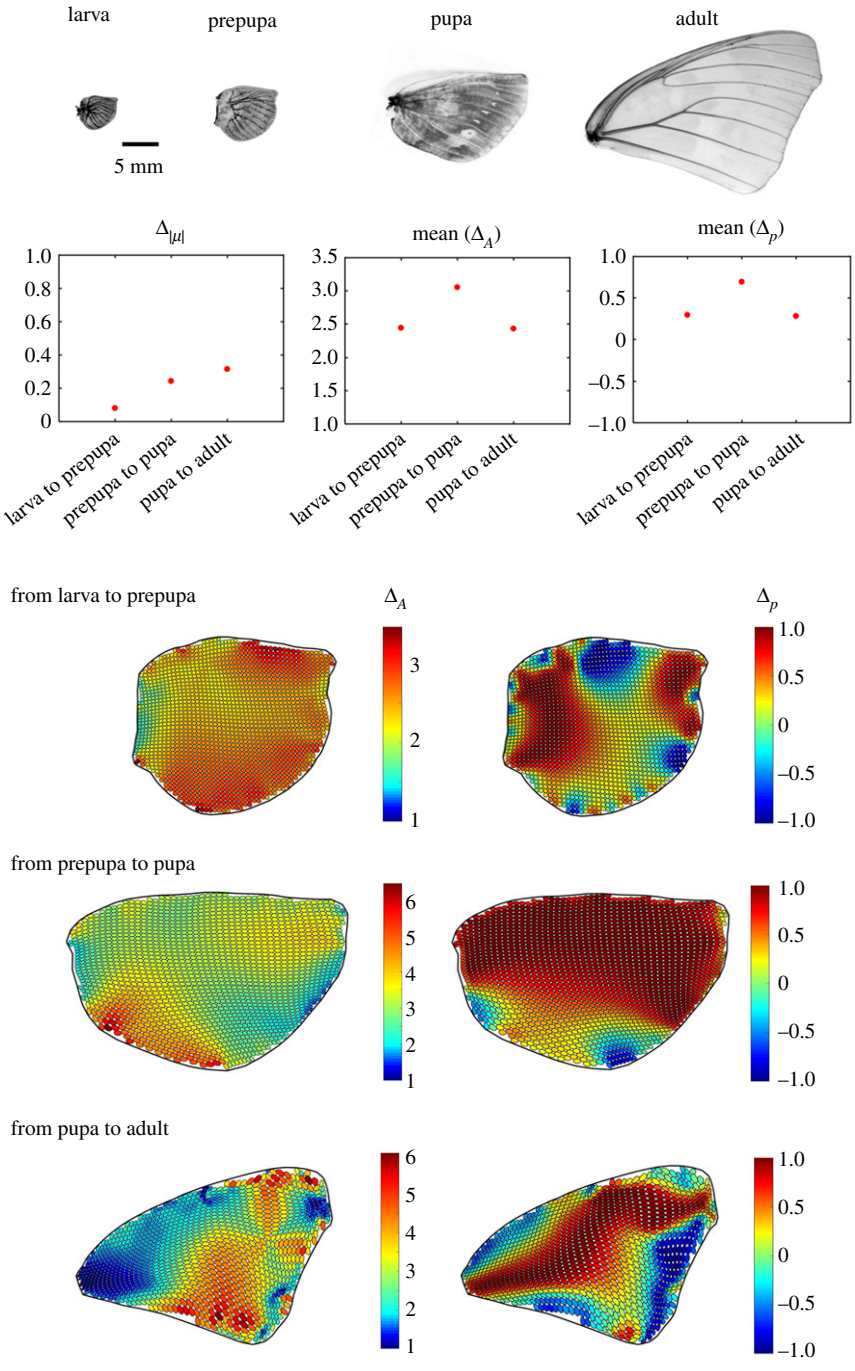

Figure 6. Analysing the temporal development of *Junonia coenia* wings via Teichmüller map. We use the same notation as in figure 5.

We then analyse the variation of the local area and orientation changes across each deformed circle packing. For both *Manduca sexta* and *Junonia coenia*, it can be observed from the heat maps of Δ_A that the most significant local area change between pupa and adult occurs at the distal half of the wings. However, it is noteworthy that the regions with the greatest Δ_A at the earlier stages are different for the two species. From the heat maps of Δ_p, the patterns of the orientation change of the two species are also different. The *Manduca sexta* wing primarily undergoes a proximal-distal

orientation change from larva to pupa, followed by a significant anterior–posterior orientation change at the central region from pupa to adult. By contrast, the *Junonia coenia* wing undergoes a more diversified local orientation change from larva to pupa, followed by a notable anterior–posterior change at the distal region from pupa to adult.

The local change reflected by Teichmüller maps throughout the development of *Manduca sexta* appears to be correlated with the local mitotic density measured in [33]. Specifically, the Teichmüller map between prepupa and pupa given in figure 5 shows a small Δ_A at the top part of the wing and a large Δ_A at the distal part. This distribution agrees with the mitotic density distribution at that period. Also, both Δ_A and the mitotic density for the period from pupa to adult achieve the greatest value around the tip of the wing and the smallest value at the proximal part. By contrast, the correlation between the local change under Teichmüller maps and the local mitotic density for *Junonia coenia* is less significant.

As illustrated in the above study, the Teichmüller-based method provides a way to assess, identify and remove allometry. Under Teichmüller maps, all changes are captured by the three quantities $\Delta_{|\mu|}$, Δ_A and Δ_p, where $\Delta_{|\mu|}$ and Δ_p together describe the shape change, and Δ_A describes the size change. Hence, shape and size can be analysed separately.

5. Discussion

To study the variation of complex data such as that embodied in shape is an old and vexing problem. Our geometric morphometric approach combines complex analysis, computations and statistics and provides a framework for the quantification, comparison and classification of planar biological shape. In particular, it improves on previous methods by allowing for arbitrary boundary and bulk landmark-matching with provably convergent, fast algorithms. By using our method for the analysis of wing shape across species, we can begin to link phenotypes to genotypes, and furthermore, we can describe the process of wing shaping developmentally via variations in mitotic density variations. More generally, our method for quantifying planar shapes might be useful beyond morphometrics, in any instance of image recognition and classification in physical, engineering and biological settings.

Data accessibility. This work does not contain any additional data.

Authors' contributions. G.P.T.C. and L.M. conceived the mathematical models, interpreted the computational results and wrote the paper. G.P.T.C. implemented and performed the simulations in consultation with L.M. Both authors gave their final approval for publication.

Competing interests. We declare we have no competing interests.

Funding. We thank the Croucher Foundation, the MacArthur Foundation, and the Harvard NSF-Simons Center for Mathematical and Statistical Analysis of Biology (DMS 1764269) for partial financial support.

References

1. Zelditch ML, Swiderski DL, Sheets HD. 2012 *Geometric morphometrics for biologists: a primer*. New York, NY: Academic Press.
2. Ramsay J, Silverman BW. 2005 *Functional data analysis*. New York, NY: Springer.
3. Srivastava A, Klassen EP. 2016 *Functional and shape data analysis*. Heidelberg, Germany: Springer.
4. Thompson DW. 1917 *On growth and form*. Cambridge, UK: Cambridge University Press.
5. Theodoridis S. 2015 *Machine learning: a Bayesian and optimization perspective*. New York, NY: Academic Press.
6. Gower JC. 1975 Generalized Procrustes analysis. *Psychometrika* **40**, 33–51. (doi:10.1007/BF02291478)
7. Hoffmann AA, Shirriffs J. 2002 Geographic variation for wing shape in *Drosophilia serrata*. *Evolution* **56**, 1068–1073. (doi:10.1111/j.0014-3820.2002.tb01418.x)
8. Debat V, Debelle A, Dworkin I. 2009 Plasticity, canalization, and developmental stability of the Drosophila wing: joint effects of mutations and developmental temperature. *Evolution* **63**, 2864–2876. (doi:10.1111/j.1558-5646.2009.00774.x)

9. Bookstein FL. 1989 Principal warps: thin-plate splines and the decomposition of deformations. *IEEE Trans. Pattern Anal. Mach. Intell.* **11**, 567–585. (doi:10.1109/34.24792)
10. Marquez EJ, Cabeen R, Woods RP, Houle D. 2012 The measurement of local variation in shape. *Evol. Biol.* **39**, 419–439. (doi:10.1007/s11692-012-9159-6)
11. Christensen GE, Rabbitt RD, Miller MI. 1996 Deformable templates using large deformation kinematics. *IEEE Trans. Image Process.* **5**, 1435–1447. (doi:10.1109/83.536892)
12. Joshi SC, Miller MI. 2000 Landmark matching via large deformation diffeomorphisms. *IEEE Trans. Image Process.* **9**, 1357–1370. (doi:10.1109/83.855431)
13. Beg MF, Miller MI, Trouvé A, Younes L. 2005 Computing large deformation metric mappings via geodesic flows of diffeomorphisms. *Int. J. Comput. Vis.* **61**, 139–157. (doi:10.1023/B:VISI.0000043755.93987.aa)
14. Rolland-Lagan AG, Remmler L, Girard-Bock C. 2014 Quantifying shape changes and tissue deformation in leaf development. *Plant Physiol.* **165**, 496–505. (doi:10.1104/pp.113.231258)
15. Alim K, Armon S, Shraiman BI, Boudaoud A. 2016 Leaf growth is conformal. *Phys. Biol.* **13**, 05LT01. (doi:10.1088/1478-3975/13/5/05LT01)
16. Mitchison G. 2016 Conformal growth of Arabidopsis leaves. *J. Theor. Biol.* **408**, 155–166. (doi:10.1016/j.jtbi.2016.08.023)
17. Lui LM, Wong TW, Thompson P, Chan T, Gu X, Yau ST. 2010 Shape-based diffeomorphic registration on hippocampal surfaces using Beltrami holomorphic flow. In *Medical image computing and computer-assisted intervention – MICCAI 2010*. Lecture Notes in Computer Sciences, vol. 6362, pp. 323–330. Springer.
18. Jones GW, Mahadevan L. 2013 Planar morphometry, shear and optimal quasi-conformal mappings. *Proc. R. Soc. A* **469**, 20120653. (doi:10.1098/rspa.2012.0653)
19. Srivastava A, Wu W, Kurtek S, Klassen E, Marron JS. 2011 Registration of functional data using Fisher-Rao metric. Preprint (http://arxiv.org/abs/1103.3817v2).
20. Kurtek SA, Srivastava A, Wu W. 2011 Signal estimation under random time-warpings and nonlinear signal alignment. In *Advances in Neural Information Processing Systems 24: 25th Annual Conference on Neural Information Processing Systems 2011, Granada, Spain, December 2011*, pp. 675–683.
21. Gardiner FP, Lakic N. 2000 *Quasiconformal Teichmüller theory*. Mathematical Surveys and Monographs, vol. 76. Providence, RI: American Mathematical Soc.
22. Strebel K. 1978 On quasiconformal mappings of open Riemann surfaces. *Comment. Math. Helv.* **53**, 301–321. (doi:10.1007/BF02566081)
23. Lui LM, Lam KC, Gu X, Yau ST. 2014 Teichmüller mapping (T-map) and its applications to landmark matching registration. *SIAM J. Imaging Sci.* **7**, 391–426. (doi:10.1137/120900186)
24. Meng TW, Choi GPT, Lui LM. 2016 TEMPO: feature-endowed Teichmüller extremal mappings of point clouds. *SIAM J. Imaging Sci.* **9**, 1922–1962. (doi:10.1137/15M1049117)
25. Lui LM, Gu X, Yau ST. 2015 Convergence of an iterative algorithm for Teichmüller maps via harmonic energy optimization. *Math. Comp.* **84**, 2823–2842. (doi:10.1090/S0025-5718-2015-02962-7)
26. Morrison G, Mahadevan L. 2012 Discovering communities through friendship. *PLoS ONE* **7**, e38704. (doi:10.1371/journal.pone.0038704)
27. Morrison G, Mahadevan L. 2011 Asymmetric network connectivity using weighted harmonic averages. *EPL* **93**, 40002. (doi:10.1209/0295-5075/93/40002)
28. Edwards KA, Doescher LT, Kaneshiro KY, Yamamoto D. 2007 A database of wing diversity in the Hawaiian *Drosophila*. *PLoS ONE* **2**, e487. (doi:10.1371/journal.pone.0000487)
29. Bonacum J, O'Grady PM, Kambysellis M, Desalle R. 2005 Phylogeny and age of diversification of the *planitibia* species group of the Hawaiian *Drosophilia*. *Mol. Phylogenet. Evol.* **37**, 73–82. (doi:10.1016/j.ympev.2005.03.008)
30. Parchem RP, Perry MW, Patel NH. 2007 Patterns on the insect wing. *Curr. Opin. Genet. Dev.* **17**, 300–308. (doi:10.1016/j.gde.2007.05.006)
31. Stone M. 1974 Cross-validatory choice and assessment of statistical predictions. *J. R. Stat. Soc. B Stat. Methodol.* **36**, 111–147.
32. Blackith RE, Reyment RA. 1971 *Multivariate morphometrics*. New York, NY: Academic Press.
33. Nijhout HF, Cinderella M, Grunert LW. 2014 The development of wing shape in Lepidoptera: mitotic density, not orientation, is the primary determinant of shape. *Evol. Dev.* **16**, 68–77. (doi:10.1111/ede.12065)