Morphometric analysis of *Passiflora* leaves: the relationship between landmarks of the vasculature and elliptical Fourier descriptors of the blade

Daniel H. Chitwood[1, §] and Wagner C. Otoni[2, §]

[1]Donald Danforth Plant Science Center, St. Louis, MO, USA

[2]Departamento de Biologia Vegetal/BIOAGRO, Universidade Federal de Viçosa, Viçosa, MG, Brasil

[§]To whom correspondence should be addressed:

Daniel H. Chitwood
Donald Danforth Plant Science Center
975 North Warson Road
St. Louis, MO 63132
Phone: (314) 587-1609
Fax: (314) 587-1709
e-mail:
dchitwood@danforthcenter.org

Wagner C. Otoni
Departamento de Biologia
Vegetal/BIOAGRO
Universidade Federal de Viçosa
36570-900 Viçosa, MG, Brasil
Phone: +55 (31) 3899 2930
e-mail: wotoni@ufv.br

Abstract

BACKGROUND: Leaf shape among *Passiflora* species is spectacularly diverse. Underlying this diversity in leaf shape are profound changes in the patterning of the primary vasculature and laminar outgrowth. Each of these aspects of leaf morphology—vasculature and blade—provides different insights into leaf patterning.

RESULTS: Here, we morphometrically analyze >3,300 leaves from 40 different *Passiflora* species collected sequentially across the vine. Each leaf is measured in two different ways: using 1) 15 homologous Procrustes-adjusted landmarks of the vasculature, sinuses, and lobes and 2) Elliptical Fourier Descriptors (EFDs), which quantify the outline of the leaf. The ability of landmarks, EFDs, and both datasets together are compared to determine their relative ability to predict species and node position within the vine. Pairwise correlation of x and y landmark coordinates and EFD harmonic coefficients reveals close associations between traits and insights into the relationship between vasculature and blade patterning.

CONCLUSIONS: Landmarks, more reflective of the vasculature, and EFDs, more reflective of the blade contour, describe both similar and distinct features of leaf morphology. Landmarks and EFDs vary in ability to predict species identity and node position in the vine and exhibit a correlational structure (both within landmark or EFD traits and between the two data types) revealing constraints between vascular and blade patterning underlying natural variation in leaf morphology among *Passiflora* species.

Keywords

Passiflora, morphometrics, leaves, leaf shape, leaf morphology, heteroblasty, Procrustes analysis, landmarks, Elliptical Fourier Descriptors

Background

The leaves of *Passiflora* species are remarkably diverse [1-3]. The underlying source of such diversity is ultimately speculative, but diversifying selective pressure from egg-laying *Heliconius* butterflies that use leaf shape as a visual cue has been proposed [4-6]. The leaves not only vary between species, but between successive nodes of a single vine, sometimes dramatically, reflecting both the heteroblastic development of the shoot apical meristem from which they are derived and the ontogeny of individual leaves as they allometrically expand [7-10]. Previous morphometric work using the multiscale Minkowski fractal dimension focused on vein patterning and the contour of the blade to predictively identify *Passiflora* species. Of the 10 species analyzed, some possessed similar leaf morphologies that could be correctly classified using only a small number of leaves per species as a training set [11].

To some degree, the patterning of the vein and blade follow each other, but to what degree they vary independently, or one is the consequence of the other, remains to be determined [12-15]. At a morphometric level, vascular patterning and the contour of the blade are studied separately, as one is a topology and the other a contour. Vasculature and blade can be separated and then analyzed with the same method, and was done using a Fractal-based approach in *Passiflora* previously [11].

Alternatively, traditional morphometric approaches can be applied to vascular patterning and the outline of the blade [16]. Procrustes-adjusted landmarks are coordinate points that correspond between all measured samples, ideally through homology [17]. Homologous landmarks are ideally suited for measuring vein thickness, vascular branch points, and the relative positioning and depth of sinuses and lobes if these features can be found in every sample, as in many palmately-lobed species, such as *Cucurbita*, *Acer*, and *Vitis*. [10, 18-25].

The landmark concept can be applied to contours as well, placing numerous points along a curve and subsequently using a Procrustes superimposition to create a near-continuous analysis of outlines [26]. The pseudo-landmark approach to quantifying contours has been used extensively to study leaf outlines, especially in species like *Antirrhinum* and *Arabidopsis*, where homologous points are lacking [27-29]. Another approach, Elliptical Fourier Descriptors (EFDs), treats an outline as a wave connecting back onto itself and subsequently performs a Fourier transform, decomposing the shape into a harmonic series [30-33]. EFDs have been applied to species from both *Solanum* and *Vitis* [23, 34-38] as well as the study of leaf asymmetry [25, 39, 40], leveraging the ability to separate symmetric from asymmetric sources of variance.

Comparing landmark- and contour-based methods not only provides an integrated perspective on leaf morphology, but can also potentially reveal the extent that patterning of the vasculature and blade are correlated in a quantitative fashion. Understanding the complementary features different morphometric methods detect is relevant to a wide variety of fields that use different approaches to extract information content from leaf shapes, including paleobiology and paleoclimate studies [41], ecology [42], evolution [10, 24, 27, 34, 43], genetics [21, 23, 29, 35, 36, 38], developmental biology [10, 18, 20, 24, 25, 34, 36, 39, 40], and plasticity [19, 20, 24, 37]. *Heliconius* butterflies, too, can even distinguish the shapes of leaves from different *Passiflora* species, presumably using a learning method yet to be determined [6].

Here, we measure landmarks of the vasculature, sinuses, and lobes and EFDs of the blade for >3,300 leaves from 40 *Passiflora* species sampled from successive nodes across vines. Linear Discriminant Analyses (LDAs) are used to determine the capacity of landmarks, EFDs, or both datasets to predict species identity versus node position in the vine. A correlational analysis of landmark and EFD data determines which specific features of leaves change together versus vary independently from each other. Our data reveals the constraints between vascular and blade patterning

underlying natural variation in leaf morphology among *Passiflora* species and provides a critical comparison of complementary morphometric approaches used on the same leaves.

Data Description

The purpose of this manuscript is to compare and contrast landmark and Elliptical Fourier Descriptor (EFD) methods in the genetic and developmental analysis of leaf shape among *Passiflora* species across the sequential nodes of their vines. The dataset released with this manuscript [44] consists of 555 scans, as well as isolated binary outlines of individual leaves, from 40 different species of *Passiflora* in which the order of leaves arising from the vine is recorded, starting with "1" for the youngest leaf scanned from the growing tip of each vine. We importantly note: the numbering of nodes in the raw scans described above, starting at the tip of the shoot, is opposite from the numbering of nodes presented in the manuscript, in which numbering (starting with "1") begins with the oldest leaf at the base of the shoot. The reason for this opposite numbering in the manuscript is that by beginning the counting of nodes with "1" at the shoot base the numbering aligns with the heteroblastic series (which begins with the first emerged leaf at the shoot base). >3,300 leaves are represented in this dataset. The number of vines sampled per a species and the number of nodes sampled for each vine are indicated in the raw data provided with this manuscript [45] and are visually depicted as well (**Fig. S1**). Both landmark data, measuring the vasculature, lobes, and sinuses, and Elliptical Fourier Descriptor (EFD) data, which quantify the leaf outline, can be derived from the provided datasets. EFDs (under PassifloraLeaves/Paper1/Figure1/0.passiflora_nef.txt) and landmarks (under PassifloraLeaves/Paper1/Figure2/0.procrustes_landmarks.txt) are provided with code in a GitHub repository [45]. It is hoped that the release of this data will assist others in developing novel morphometric approaches to better understand the genetic, developmental, and environmental basis of leaf shape.

Analyses

Vascular landmarks and Elliptical Fourier Descriptors (EFDs) of the blade

For the >3,300 leaves measured across the leaf series for 40 different *Passiflora* species, a comparison of homologous landmarks and Elliptical Fourier Descriptors (EFDs) was made (see **Fig. S1** and raw data [45] for the replication associated with each species and the number of nodes for each vine) These two methods globally capture complementary aspects of leaf shape, sensitive to vascular patterning and the shape of the blade, respectively.

15 landmarks were measured for each leaf (**Fig. 1A**). For the proximal veins (near the leaf base) landmarks on each side of the junction of the proximal vein with the petiolar junction (where the major veins meet) were placed (landmarks 1-2 and 5-6), capturing the width of the proximal veins. Landmarks placed at the tip of the proximal vein (landmarks 7 and 15) capture the length and angle of the proximal lobe. On the distal vein (nearer the leaf tip), landmarks were placed only on the distal side of the junction with the midvein (landmarks 3 and 4) as the other side of the base of the distal vein variably intersects the midvein, petiolar junction, and proximal vein (see three examples in **Fig. 1A**, left to right). The landmarks at the tip of the distal veins (landmarks 13 and 9) measure the length and angle of the distal lobe. Additionally, landmarks describe the placement of the leaf tip (landmark 11), distal sinuses (landmarks 10 and 12), and the proximal sinuses (8 and 14).

To determine the extent that landmarks capture qualitative variation in leaf shape among *Passiflora* species, representative leaves were compared to averaged Procrustes-adjusted landmark values (**Fig. 2**). The landmark analysis captures features such as the relative lengths and angular placement of the proximal and distal veins as well as the depth of the sinuses. Visualizing superimposed landmarks for all leaves measured in addition to the averaged landmark values demonstrates substantial sources of shape variance in some species, especially due to changes in

leaf shape across the leaf series, that usually relate to the depth of the sinuses or the number of lobes.

Although landmarks accurately depict information related to vascular patterning and the relative placement of the lobes and sinuses of the blade, they fail to capture more subtle shape variation related to the curvature of the lamina. Elliptical Fourier Descriptors (EFDs) result from a harmonic decomposition of a shape outline. The harmonic contributions to leaf shape can be visualized (**Fig. 1B**), which in *Passiflora* correspond to features reflecting the leaf tip, distal lobes, and proximal lobes (the "trifoliate" features, especially in the lower harmonic ranks) or more local features (the "serrations" represented in the higher harmonic ranks) (**Fig. 1B**). The averaged outlines of leaves capture the curves and lobing of leaves from each species (**Fig. 3**). Species that display leaves with variable numbers of lobes (such as *Passiflora caerulea*, *P. cincinnata*, or *P. suberosa*) have average leaf outlines reflecting this source of shape variance.

The morphospace reflects species and heteroblastic differences in leaf shape

To analyze major sources of shape variance in Procrustes-adjusted landmark values and the harmonic series from the Elliptical Fourier Descriptor (EFD) analysis, a Principal Component Analysis (PCA) was performed to reduce the dimensionality of each dataset. Onto the resulting morphospaces were projected species identity and the node position in the leaf series ("heteroblasty"). Node position is referred to as "heteroblasty" as a shorthand indicating that numbering of nodes begins at the shoot base, with "1" indicating the first emerged leaf at the shoot base. This numbering scheme more closely aligns with the heteroblastic series of leaves compared to the reverse numbering that begins at the growing shoot tip and is more sensitive to the allometric changes in rapidly expanding leaves.

Eigenleaves (theoretical leaf shapes representing the eigenvectors from a Principal Component Analysis) from each PCA reveal the shape features contributing to shape

variance along each Principal Component (PC). The first four landmark PCs (**Fig. 4A**) explain 83.2% of shape variance for the landmark dataset. PC1 reflects shape variance related to long, lance-like leaves versus wider leaves with short midveins and long, extended distal lobes. Both PC2 and PC3 explain shape variance related to leaves with pronounced distal lobes versus more rounder (PC2) or deltoid (PC3) leaves with less lobing. PC4 also explains shape variance related to lobing. A comparison of the landmark eigenleaves (**Fig. 4A**) with the EFD eigenleaves (**Fig. 4B**) shows that the shape variance explained by each respective PC is strikingly similar, especially with respect to lobing and the length-to-width ratio of leaves. This demonstrates a qualitative correspondence between the orthogonal axes of each dataset, including their directionality, which will be subsequently explored in further detail.

Projecting species identity and heteroblastic node onto the landmark and EFD morphospaces reveals that each method separates the shape variance attributable to these variables, but in different ways (**Fig. 5**). Because visualizing 40 distinct species is a challenge, species were assigned to 7 different classes (consistently colored throughout the manuscript) based on a) occupying similar spaces within morphospace and b) qualitative differences in leaf shape (**Fig. 5A**). Species classes show pronounced separation from each other by PC1 and PC2 in both the landmark (**Fig. 5B**) and EFD (**Fig. 5C**) morphospaces. Less separation is observed by species class for PC3 and PC4. When heteroblastic node is projected onto the morphospaces, there is a trend for the leaves originating from high heteroblastic nodes (young leaves towards the growing tip) to occupy the lower PC2 values within each species class. This is especially true for the landmark morphospace (**Fig. 5B**). There is also a trend for leaves originating from high heteroblastic nodes to have low PC3 values, regardless of species class. Both low PC2 and PC3 values correspond to more pronounced distal lobing (**Fig. 4**), a shape feature commonly found in young leaves near the growing tip of the plant, compared to older leaves near the base of the vine that tend to have less lobing. That shape variance attributable to species class and heteroblastic node traverse the morphospace in different ways suggests to some

extent the shape variance for each of these factors is separable, as is discussed in the next section.

Discriminating species vs. node identity

That species class and node identity traverse the morphospace differently (**Fig, 5A-B**) is consistent with previous work demonstrating that shape features can be used to discriminate species independently from node position in grapevine [10, 24]. A Linear Discriminant Analysis (LDA) is used here to determine the extent these two variables can be predicted independently of the other in *Passiflora* using landmark data, Elliptical Fourier Descriptors (EFDs), and both landmark and EFDs together. We stress that the LDA approach taken in this work is fundamentally different from modeling species, node, and interaction effects using linear modeling. Such an approach (which we undertook but the data is not shown here, because it is outside the scope of this manuscript) reveals that for each morphometric trait considered independently, the species and interaction effects are the strongest and the node effect is weak. Rather, an LDA allows explicit questions to be asked regarding all the measured traits together. Can all the traits be used together to discriminate species regardless of node? Using all traits can node be distinguished separately from species? Such a framework is consistent with developmental genetic theory that differences in leaf shape between species versus more conserved heteroblastic changes in leaf shape within individual plants are regulated by distinct genetic pathways [16] that lead to separable morphological effects within single leaves (so called "cryptotypes" [46]). We also note that the LDAs performed use the "leave one out" approach of cross-validation, in which a separate LDA for each leaf, minus the leaf in question, is used to predict the identity of that leaf. Such an approach is designed to compensate for differences in species replication and nodes sampled per a vine in our dataset (see raw data [45] and **Fig. S1**).

An LDA is first performed on species identity, regardless of node position. The resulting discriminants are then used to predict the identity of the species.

Regardless of whether landmarks (**Fig. 6A**), EFDs (**Fig. 6C**), or both landmarks and EFDs are used (**Fig. 6E**) a high proportion of leaves can be correctly reassigned to the correct species. When there is confusion between species, it tends to be within the same species class. This result demonstrates that regardless of the position of a leaf within the heteroblastic series, its species identity can be predicted. For most species classes (all except C and D) the maximum correct prediction is most often achieved with both landmark and EFD data together compared to each data type alone (**Table 1**). For species classes C and D, however, landmark data alone tends to outperform EFD and both data types together. This indicates that for some species, especially those that are highly lobed as in species classes C and D, landmark data is a better indicator of species identity (perhaps because it is more explicitly related to lobing).

Conversely, heteroblastic node position can be predicted independently of species identity, but to a much lesser degree and not equally across the leaf series. The leaves occupying lower node positions (older leaves at the base of the vine) tend to be successfully predicted at a higher rate than the younger leaves of the tip, regardless of whether landmarks (**Fig. 6B**), EFDs (**Fig. 6D**), or both landmarks and EFDS are used (**Fig. 6F**). EFDs, however, overall under-perform landmarks or landmarks and EFDs used together (**Table 2**). This indicates that landmarks are a superior discriminant of node position compared to EFDs. Previous work in grapevine indicates that vein thickness is altered by shoot position [10, 24]. That landmarks measure vein thickness, but not EFDs, may explain the differing abilities of these two shape features to correctly discriminate leaves by heteroblastic node position. That the juvenile leaves at the lower heteroblastic node positions are correctly predicted at higher rates suggests that these leaves are more similar across species (or correspondingly, that leaves at high heteroblastic node positions are more divergent between species).

Correlational matrix between landmarks and Elliptical Fourier Descirptors (EFDs)

Until now, landmarks and Elliptical Fourier Descriptors (EFDs) have either been considered separately or in conjunction together but not compared against each other. The landmarks used in this study tend to represent vascular features of the leaf, the lobes, and the sinuses. The EFDs represent the blade and the continuous contour and curves of lamina. Further, landmark data is represented as (x, y) coordinates, whereas EFD data is a Fourier-based harmonic series. A correlational matrix is used to find strong associations between the components of each dataset and to interpret the features each dataset uniquely quantifies against the other.

The input for the correlation matrix, using Spearman's rho, is each of the fifteen x and y coordinates of the landmark dataset and each of the four harmonic coefficients (A, B, C, D) of the first 20 harmonic ranks from the EFD data, correlated across the >3,300 leaves for all species and heteroblastic node positions used in this study. This correlation matrix was used as a distance matrix to hierarchically cluster these traits and the rho and p values subsequently visualized (**Fig. 7**).

A large set of uncorrelated traits, consisting of the B and C harmonic coefficients and the x11 landmark, end up clustered together (**Fig. 7**). The B and C harmonic coefficients represent asymmetric sources of shape variance [31] and the x11 landmark represents the left-right variance of the leaf tip (**Fig. 1A**), which will mostly be affected by leaf asymmetry. That these shape features are weakly correlated with each other and other traits only implies that they are regulated by an unaccounted source of variance for this particular analysis. In the future, a more in-depth analysis will likely reveal phyllotaxy as modulating leaf asymmetry [39, 40], specifically alternating asymmetry at consecutive nodes, as recently shown in other vines, such as ivy and grapevine [25].

The remaining landmarks and the A and D coefficients of the harmonic series (representing symmetrical shape variation) show various correlational associations with each other (**Figs. 7-8**). Harmonic contributions to leaf shape (**Fig. 1B**) are more abstract and difficult to interpret than the contributions of landmarks to leaf shape,

as the landmarks represent homologous points found in every leaf (**Fig. 1A**). Strong correlations between harmonic coefficients with landmarks can help interpret the context of the harmonic coefficient to leaf shape. Most the harmonic coefficients cluster exclusively together except for landmarks y9 and y13, which represent the proximal-distal displacement of the distal lobes along the leaf length (**Fig. 8**). This suggests that large amounts of the shape variance associated with the contour of the blade are influenced by the relative placement of the distal lobes along the leaf length. The remaining harmonic coefficients that cluster outside most the other coefficients also associate with features of the distal part of the leaf. A1, A3, D2, and D6 associate with the x and y coordinates of the distal sinus (x10, x12, y10, and y12) and D1 and D3 associate with the left-right displacement of the distal lobe (x9 and x13) and the vertical displacement of the leaf tip (y11) (**Fig. 8**). Although difficult to interpret, the correlations of harmonic coefficients suggest that the overall leaf contour is influenced by the placement of the distal lobe and sinus.

The remaining correlations between landmarks reveal interesting constraints governing the shape of *Passiflora* leaves (**Fig. 8**). As mentioned previously, the left-right displacement of the distal lobes (x9 and x13) strongly correlates with the vertical proximal-distal displacement of the leaf tip (y11). The x and y coordinates of the distal lobes (landmarks 10 and 12) are the only features for which the x and y displacement are correlated, suggesting that the distal sinus varies in a diagonal direction. The proximal sinus and lobe (landmarks 7, 8, 14, and 15) and the landmarks at the base of the veins of the petiolar junction (landmarks 1, 2, 3, 4, 5, and 6) form additional groups of associated landmarks, although interestingly the x and y displacement of each of these two groups is distinct in each case (**Fig. 8**)

Discussion

Leaf morphology refers to the totality of leaf architecture, at the cellular, tissue, and organ levels, and distinct attributes of the leaf, both the vasculature and lamina. The topology of the vasculature and contour of the leaf blade are distinct geometric

phenomena that require different morphometric approaches to quantify. Landmarks and Elliptical Fourier Descriptors (EFDs) are ideal methods to analyze the distinct features of leaves contributing to their shape (**Fig. 1**), but rarely are they measured and compared on the same leaves. Our analysis of disparate leaf shapes among *Passiflora* species with landmarks (**Fig. 2**) and EFDs (**Fig. 3**) reveals that both methods capture similar orthogonal axes of shape variation (**Fig. 4**), and separate both species and heteroblastic node identity, but in distinct ways (**Fig. 5**). Landmarks are superior to EFDs in predicting node position compared to species identity, most likely because they describe vascular patterning, which is relatively sensitive to heteroblasty compared to species differences in leaf shape (**Fig. 6; Tables 1-2**). Although most elements of the EFD harmonic series cluster together in a pairwise correlational analysis, a few are closely associated with landmarks (**Fig. 7**). Landmarks exhibit a correlational structure revealing developmental constraints in how leaves vary across *Passiflora* species and the heteroblastic series (**Fig. 8**). Together, our data quantify the relationship between blade and vasculature, revealing that one does not drive the patterning of the other, and although each distinctly varies, many shape features of the leaf change in concert across evolution and development.

Methods

Plant materials and growth conditions

Passiflora germplasm was kindly provided by R. Silva (Viveiros Flora Brasil, Araguari, MG, Brazil), Dr. F.G. Faleiro (EMBRAPA Cerrados, Planaltina, DF, Brazil), Prof. M.M. Souza (Universidade Estadual de Santa Cruz - UESC, Ilhéus, BA, Brazil), M. Peixoto (Mogi das Cruzes, SP, Brazil), Prof. M.L. Silva (Universidade do Estado de Mato Grosso, Tangará da Serra, MT, Brazil), and Prof. C.H. Bruckner (Universidade Federal de Viçosa, Viçosa, MG, Brazil).

The plants were germinated from seed, planted between late October 2015 and early March 2016, in Viçosa, at the Federal University of Viçosa, MG, Brazil. The populations were raised and maintained under polycarbonate-covered greenhouse conditions, equipped with automatic environmental control using exhaust fans and evaporative cooling panels (with expanded clay wettable pads). Seeds for each *Passiflora* species were sown in 128 cell propagation plastic trays (GPlan Comércio de Produtos Agrícola s EIRELI – ME, São Paulo, SP, Brazil) filled with horticultural organic Tropstrato HT Hortaliças substrate (Vida Verde Indústria e Comércio de Insumos Orgânicos Ltda, Mogi Mirim, SP, Brazil). After germination (30-40 days), plantlets were individually transplanted to 5 L capacity plastic pots (EME-A-EME Ind. Com. Ltda., Petrópolis, RJ, Brazil) filled with horticultural substrate. Each pot received 5 g of Osmocote® Plus NPK 15-09-12 3-4 month controlled release fertilizer (Scotts, USA). Plants were irrigated on a daily-basis with tap water, and no phytosanitary control was applied. The germination and growth rates of plants varied widely. The number of replicates for each species and the number of nodes per vine are indicated in the raw data [45] and depicted visually (**Fig. S1**).

For scanning, a multifunction printer (Canon PIXMA MX340 Wireless Office All-in-One Printer, model 4204B019, USA) was used. A 20 cm metallic ruler was positioned at the bottom of each scanned sheet as a size marker. Leaves were carefully detached, from the base to the tip of the shoot, and affixed to an A4 paper sheet, adaxial face down, using 12 mm-double sided tape (Scotch Model 9400, 3M do Brasil, SP, Brazil). The numbers written near each leaf indicate position in the shoot, in a tip-to-base direction, starting with the youngest leaf at the tip of the shoot. It should be noted that the numbering in the scans is opposite from the numbering used in the analysis and figures for this manuscript, in which leaves are numbered with "1" starting at the shoot base. This numbering system more closely aligns with the heteroblastic series than the reverse numbering scheme originally used in the scans.

Morphometric and statistical analyses

All morphometric data and code used for statistical analysis is available on GitHub [45]. All original data is available at GigaDB [44].

Landmarks, as described in the text, were placed on leaves in ImageJ [47]. Procrustes superimposition was performed using the shapes package [48] in R [49] with the procGPA function using reflect=TRUE. Resulting Procrustes-adjusted coordinates and principal component scores (PCs) were written out for subsequent analyses and eigenleaf representations visualized using the shapepca function.

To isolate outlines for Elliptical Fourier Descriptor (EFD) analysis, the "Make Binary" function in ImageJ [47] was found to be sufficient to segment leaves. The wand tool was used to select individual binary leaf outlines, which were pasted into a new canvas, which was subsequently saved as an individual image, which was named by vine and node position from which the leaf was derived. The binary images were batch converted into RGB .bmp files and read into SHAPE, which was used to perform chain-code analysis [31, 32]. The resulting chain-code .chc file was then used to calculate normalized EFDs. The resulting normalized EFD .nef file was then read into Momocs (version 0.2-6) [33] in R. The harmonic contributions to shape were visualized using the hcontrib function. Averaged leaf outlines were calculated using the meanShapes function and Principal Component Analysis (PCA) performed using the pca function and eigenleaves visualized using the PC.contrib function.

Unless otherwise noted, all visualization was performed using ggplot2 in R [50]. Linear Discriminant Analysis (LDA) was performed using the lda function and subsequent prediction of species identity or heteroblastic node position performed using the predict function with MASS [51]. When LDAs were used for prediction, the parameter CV was set to "TRUE", for the "leave one out" cross-validation approach, to help make analyses more robust to differences in replication and node numbers

between species and vines. Hierarchical clustering was performed using the hclust function.

Availability and requirements

Project name: PassifloraLeaves
Project home page: https://github.com/DanChitwood/PassifloraLeaves
Operating system(s): Platform independent
Programming language: R
Other requirements: Not applicable
License: MIT license
Any restriction to use by non-academics: none

Availability of supporting data and materials

The data sets supporting the results of this article are available in the GigaDB repository [44].

Declarations

Funding

Brazilian sponsoring agencies, namely FAPEMIG (Grant no. CBB - APQ-01131-15), CNPq (Grant no. 459.529/2014-5) and CAPES, are acknowledged for financial support.

Authors' contributions

The overall project was conceived by DHC and WCO. WCO grew and scanned all plant material and DHC carried out analysis. DHC and WCO wrote the paper.

Competing interests

The authors declare that they have no competing interests.

References

1. Klucking EP. Leaf venation patterns, vol. 6. 1992; Berlin: J. Cramer. Passifloraceae, 222-262.

2. MacDougal JM. Revision of Passiflora subgenus Decaloba section Pseudodysosmia (Passifloraceae). Syst Bot Monogr. 1994;41:1-146.

3. Ulmer T, Mac Dougal JM. 2004. *Passiflora:* passionflowers of the world. Portland Oregon: Timber Press. 430 p.

4. Gilbert LE. Ecological consequences of a coevolved mutualism between butterflies and plants. Coevolution of animals and plants. 1975;210-240.

5. Gilbert LE. The coevolution of a butterfly and a vine. Sci Amer. 1982;110-121.

6. Dell'aglio DD, Losada ME, Jiggins CD. Butterfly learning and the diversification of plant leaf shape. Frontiers in Ecology and Evolution. 2016;4:81.

7. Allsopp A. Heteroblastic development in vascular plants. Advances in morphogenesis. 1967;6:127-171.

8. Rolland-Lagan AG, Remmler L, Girard-Bock C. Quantifying shape changes and tissue deformation in leaf development. Plant physiology. 2014;165(2):496-505.

9. Gupta MD, Nath U. Divergence in patterns of leaf growth polarity is associated with the expression divergence of miR396. The Plant Cell. 2015;27(10):2785-99.

10. Chitwood DH, Klein LL, O'Hanlon R, Chacko S, Greg M, Kitchen C, Miller AJ, Londo JP. Latent developmental and evolutionary shapes embedded within the grapevine leaf. New Phytol. 2016;210:343-55.

11. Plotze RDO, Falvo M, Pádua JG, Bernacci LC, Vieira MLC, Oliveira GCX, Bruno OM. Leaf shape analysis using the multiscale Minkowski fractal dimension, a new morphometric method: a study with Passiflora (Passifloraceae). Canadian Journal of Botany. 2005;83:287-301.

12. Hagemann W, Gleissberg S. Organogenetic capacity of leaves: the significance of marginal blastozones in angiosperms. Plant Syst Evol. 1996;199:121-152.

13. Nelson T, Dengler N. Leaf vascular pattern formation. Plant Cell. 1997;9:1121-1135.

14. Dengler N, Kang J. Vascular patterning and leaf shape. Curr Opin Plant Biol. 2001;4:50-56.

15. Champagne C, Sinha N. Compound leaves: equal to the sum of their parts? Development. 2004;131:4401-12.

16. Chitwood DH, Sinha NR. Evolutionary and environmental forces sculpting leaf development. Curr Biol. 2016;26:R297-306.

17. Rohlf F, Slice D. Extensions of the Procrustes method for the optimal superimposition of landmarks. Systematic Biology. 1990;39:40-59.

18. Jones CS. Comparative ontogeny of a wild cucurbit and its derived cultivar. Evolution. 1992;46:1827-1847.

19. Jones CS. Does shade prolong juvenile development? A morphological analysis of leaf shape changes in Cucurbita argyrosperma subsp. sororia (Cucurbitaceae). American Journal of Botany. 1995;82:346-359.

20. Young JP, Dickinson TA, Dengler NG. A morphometric analysis of heterophyllous leaf development in Ranunculus flabellaris. Int J Plant Sci. 1995;156:590-602.

21. Jensen RJ, Ciofani KM, Miramontes LC. Lines, outlines, and landmarks: morphometric analyses of leaves of Acer rubrum, Acer saccharinum (Aceraceae) and their hybrid. Taxon. 2002;51:475-492.

22. Klingenberg CP, Duttke S, Whelan S, Kim M. Developmental plasticity, morphological variation and evolability: a multilevel anlaysis of morphometric integration in the shape of compound leaves. J Evol Biol. 2011;25:115-129.

23. Chitwood DH, Ranjan A, Martinez CC, Headland LR, Thiem T, Kumar R, Covington MF, Hatcher T, Naylor DT, Zimmerman S, Downs N, Raymundo N, Buckler ES, Maloof JN, Aradhya M, Prins B, Li L, Myles S, Sinha NR. A modern ampelography: a genetic basis for leaf shape and venation patterning in grape. Plant Physiol. 2014;164:259-72.

24. Chitwood DH, Rundell SM, Li DY, Woodford QL, Yu TT, Lopez JR, Greenblatt D, Kang J, Londo JP. Climate and developmental plasticity: interannual variability in grapevine leaf morphology. Plant Physiol. 2016;170:1480-91.

25. Martinez CC, Chitwood DH, Smith RS, Sinha NR. Left-right leaf asymmetry in decussate and distichous phyllotactic systems. bioRxiv. 2016; dx.doi.org/10.1101/043869

26. Bookstein FL. Landmark methods for forms without landmarks: morphometrics of group differences in outline shape. Medical image analysis. 1997;1(3):225-43.

27. Langlade NB, Feng X, Dransfield T, Copsey L, Hanna AI, Thebaud C, Bangham A, Hudson A, Coen E. Evolution through genetically controlled allometry space. Proc Natl Acad Sci USA. 2005; 102:10221-10226.

28. Weight C, Parnham D, Waites R. LeafAnalyser: a computational method for rapid and large-scale analyses of leaf shape variation. Plant Journal. 2007;53:578-586.

29. Bensmihen S, Hanna AI, Langlade NB, Micol JL, Bangham A, Coen ES. Mutational spaces for leaf shape and size. Mutational spaces for leaf shape and size. HFSP Journal. 2008;2:110-120.

30. Kuhl FP, Giardina CR. Elliptic Fourier features of a closed contour. Computer graphics and image processing. 1982;18:236-258.

31. Iwata H, Niikura S, Matsuura S, Takano Y, Ukai Y. Evaluation of variation of root shape of Japanese radish (Raphanus sativus L.) based on image analysis using elliptic Fourier descriptors. Euphytica. 1998;102:143-9.

32. Iwata H, Ukai Y. SHAPE: a computer program package for quantitative evaluation of biological shapes based on elliptic Fourier descriptors. Journal of Heredity. 2002;93:384-385.

33. Bonhomme V, Picq S, Gaucherel C, Claude J. Momocs: outline analysis using R. Journal of Statistical Software. 2014;56:1-24.

34. Chitwood DH, Headland LR, Kumar R, Peng J, Maloof JN, Sinha NR. The developmental trajectory of leaflet morphology in wild tomato species. Plant Physiol. 2012;158:1230-40.

35. Chitwood DH, Kumar R, Headland LR, Ranjan A, Covington MF, Ichihashi Y, Fulop D, Jimenez-Gomez JM, Peng J, Maloof JN, Sinha NR. A quantitative genetic basis for leaf morphology in a set of precisely defined tomato introgression lines. Plant Cell. 2013;25:2465-81.

36. Chitwood DH, Ranjan A, Kumar R, Ichihashi Y, Zumstein K, Headland LR, Ostria-Gallardo E, Aguilar-Martinez JA, Bush S, Carriedo L, Fulop D, Martinez CC, Peng J, Maloof JN, Sinha NR. Resolving distinct genetic regulators of tomato leaf shape within a heteroblastic and ontogenetic context. Plant Cell. 2014;26:3616-29.

37. Chitwood DH, Kumar R, Ranjan A, Pelletier JM, Townsley BT, Ichihashi Y, Martinez CC, Zumstein K, Harada JJ, Maloof JN, Sinha NR. Light-Induced

Indeterminacy Alters Shade-Avoiding Tomato Leaf Morphology. Plant physiology. 2015;169(3):2030-47.

38. Fulop D, Ranjan A, Ofner I, Covington MF, Chitwood DH, West D, Ichihashi Y, Headland L, Zamir D, Maloof JN, Sinha NR. A new advanced backcross tomato population enables high resolution leaf QTL mapping and gene identification. G3 (Bethesda). 2016;6(10):3169-3184.

39. Chitwood DH, Headland LR, Ranjan A, Martinez CC, Braybrook SA, Koenig DP, Kuhlemeier C, Smith RS, Sinha NR. Leaf asymmetry as a developmental constraint imposed by auxin-dependent phyllotactic patterning. Plant Cell. 2012;24:2318-27.

40. Chitwood DH, Naylor DT, Thammapichai P, Weeger AC, Headland LR, Sinha NR. Conflict between intrinsic leaf asymmetry and phyllotaxis in the resupinate leaves of Alstroemeria psittacina. Front Plant Sci. 2012;3:182.

41. Bailey IW, Sinnott EW. A botanical index of Cretaceous and Tertiary climates. Science. 1915;831-4.

42. Peppe DJ, Royer DL, Cariglino B, Oliver SY, Newman S, Leight E, Enikolopov G, Fernandez-Burgos M, Herrera F, Adams JM, Correa E. Sensitivity of leaf size and shape to climate: global patterns and paleoclimatic applications. New Phytologist. 2011;190(3):724-39.

43. Schmerler SB, Clement WL, Beaulieu JM, Chatelet DS, Sack L, Donoghue MJ, Edwards EJ. Evolution of leaf form correlates with tropical–temperate transitions in Viburnum (Adoxaceae). Proceedings of the Royal Society of London B: Biological Sciences. 2012;279(1744):3905-13.

44. Chitwood DH, Otoni WC. Supporting data for "Morphometric analysis of Passiflora leaves I: the relationship between landmarks of the vasculature and elliptical Fourier descriptors of the blade" GigaScience Database. 2016. http://dx.doi.org/10.5524/100251.

45. Chitwood DH. PassifloraLeaves. GitHub. 2016. https://github.com/DanChitwood/PassifloraLeaves

46. Chitwood DH, Topp CN. Revealing plant cryptotypes: defining meaningful phenotypes among infinite traits. Current opinion in plant biology. 2015;24:54-60.

47. Abràmoff MD, Magalhães PJ, Ram SJ. Image processing with ImageJ. Biophotonics international. 2004;11(7):36-42.

48. Dryden IL. shapes: Statistical Shape Analysis. R package version 1.1-11. 2015;https://CRAN.R-project.org/package=shapes

49. R Development Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2016;http://www.R-project.org.

50. Wickham H. ggplot2: Elegant Graphics for Data Analysis. 2009;Springer-Verlag, New York.

51. Venables WN, Ripley BD. Modern Applied Statistics with S. Fourth Edition. 2002;Springer, New York.

52. Chitwood DH, Otoni WC. Divergent heteroblastic trajectories underlie disparate leaf shapes among Passiflora species. bioRxiv. 2016; http://dx.doi.org/10.1101/067520

Figure Legends

Figure 1: Landmarks and harmonic contributions to shape. A) The 15 landmarks used for analysis. Left to right, landmark placement when the distal and proximal veins l) pinnately emerge from the midvein, m) both originate from the petiolar junction, or r) the proximal vein branches from the distal. **B)** Harmonic contributions to shape resulting from Elliptical Fourier Descriptor (EFD) analysis. The harmonic rank is arranged horizontally and the amplification factor (which multiplies the harmonic contributions to shape by the indicated amount) vertically. Note: for convenience to the reader, these panels are recapitulated in the companion manuscript [52].

Figure 2: The shapes of *Passiflora* leaves measured using landmarks. For the 40 species analyzed in this study, both a representative leaf and landmark data are shown. For the landmark data, the mean leaf for the species in shown in black, whereas all data for the species is depicted in semi-transparent blue.

Figure 3: The shapes of *Passiflora* leaves measured using Elliptical Fourier Descriptors (EFDs). Mean leaves calculated for each of 40 species analyzed in this study from the harmonic series resulting from an Elliptical Fourier Descriptor (EFD) analysis of the leaf contours. **A-G)** Classes of species are indicated by their

respective panels. Species classes were determined by neighboring position in the Principal Component Analysis (PCA) morphospace, described in **Figs. 4-5**. Color indicates class: class A, teal; class B, orange; class C, lavender; class D, magenta; class E, green; class F, yellow; class G, brown.

Figure 4: Principal Components (PCs) and eigenleaves. A) Principal components (PCs) representing shape variance in landmark data. Eigenleaf representations (theoretical leaf shapes representing the eigenvectors from a Principal Component Analysis) at +/- 1.5 standard deviations (s.d.) are shown for the first four PCs. Percent variance explained by each PC indicated. **B)** PCs representing shape variance in Elliptical Fourier Descriptor (EFD) data. Eigenleaf representations at +/- 1 s.d. are shown for the first four PCs. Percent variance explained by each PC indicated.

Figure 5: Morphospace by species and heteroblastic node. A) Key, showing species classes and averaged leaf contours for each species. Color indicates class, which is used in other panels. **B)** Principal Component Analysis (PCA) of landmark data. Graphs for PC2 vs. PC1 and PC4 vs. PC3 are colored by species class and by heteroblastic node. Percent variance explained by each PC indicated. **C)** PCA of Elliptical Fourier Descriptor (EFD) data. Graphs for PC2 vs. PC1 and PC4 vs. PC3 are colored by species class and by heteroblastic node. Percent variance explained by each PC indicated. Heteroblastic node position is numbered "1" starting from the shoot base. Class color scheme: class A, teal; class B, orange; class C, lavender; class D, magenta; class E, green; class F, yellow; class G, brown. Heteroblastic node color scheme: shoot base, black; middle shoot, blue; shoot tip, yellow.

Figure 6: Linear Discriminant Analysis (LDA). Linear Discriminant Analysis (LDA) using **A-B)** landmark data, **C-D)** Elliptical Fourier Descriptor (EFD) data, and **E-F)** both datasets. For each set of LDAs, analysis was performed to discriminate species (ignoring heteroblastic node information) or to discriminate heteroblastic node (ignoring species information). Subsequent prediction of species or

heteroblastic node identity is then visualized using confusion matrices, where actual identity is oriented vertically, predicted identity horizontally, and the proportion assigned indicated as fill. Species LDAs are broken up by species class. For heteroblastic node LDAs, Spearman's rho and associated p values calculated from correlating actual and predicted node identities are provided. Predictions carried out using LDA use the "leave one out" approach cross-validation approach. Heteroblastic node position is numbered "1" starting from the shoot base. Color scheme: low assigned proportion, white; high assigned proportion, black.

Figure 7: Correlational matrix of landmark and Elliptical Fourier Descriptor (EFD) traits. Spearman's correlation matrix for morphometric features analyzed in this study. Upper half indicates $-\log_{10}$ p value and lower half Spearman's rho between indicated traits. Morphometric traits, both landmark and the harmonic series, are indicated along the sides, arranged using hierarchical clustering, the topology of which is depicted as a dendrogram. Key groupings of landmarks indicating correlational associations with each other or EFD harmonics are indicated. Spearman's rho: low values, green; middle values, white; high values, magenta. $-\log_{10}$ p values: low values, purple; high values, yellow; $p < 0.05$, no color.

Figure 8: Correlational relationships between vascular landmarks and leaf contours. Correlational relationships between x and y components of landmarks and Elliptical Fourier Descriptor (EFD) harmonics are indicated by dendrogram (left) and landmarks qualitatively on a representation of a leaf (right). x and y landmark components are independently depicted by arrows and colored as indicated to show major correlational sources of shape variance within *Passiflora* leaves.

Supplemental Information

Figure S1: Species replication and number of nodes sampled. A) Dotplot showing the number of vines sampled for each species. **B)** Boxplot showing the

number of nodes sampled for vines from each species. The largest red line is the median nodes sampled (14 nodes), the medium sized redlines the 25th and 75th quantiles (12 and 16 nodes, respectively), and the thin red lines the minimum and maximum (7 and 28 nodes, respectively).

Table 1: Predictive power of different morphometric methods to discriminate *Passiflora* species.

Species	Class	Landmark	EFD	Both	Max
P. coriacea	A	83.2%	81.7%	88.1%	Both
P. misera	A	77.0%	71.6%	76.9%	Landmark
P. biflora	B	84.1%	75.4%	92.1%	Both
P. capsularis	B	77.2%	72.0%	77.3%	Both
P. micropetala	B	69.1%	81.8%	92.4%	Both
P. organensis	B	89.4%	70.7%	96.6%	Both
P. pohlii	B	54.5%	77.8%	77.8%	EFD
P. rubra	B	60.3%	59.7%	71.6%	Both
P. tricuspis	B	49.0%	67.8%	69.2%	Both
P. caerulea	C	0.0%	15.1%	9.4%	EFD
P. cincinnata	C	71.4%	59.3%	59.3%	Landmark
P. edmundoi	C	72.8%	78.8%	83.8%	Both
P. gibertii	C	84.0%	72.2%	81.9%	Landmark
P. hatschbachii	C	72.0%	65.4%	67.9%	Landmark
P. kermesina	C	71.0%	43.6%	70.9%	Landmark
P. mollissima	C	53.6%	35.5%	67.7%	Both
P. setacea	C	81.9%	64.4%	77.8%	Landmark
P. suberosa	C	52.3%	63.4%	66.9%	Both
P. tenuifila	C	68.8%	65.1%	79.4%	Both
P. amethystina	D	69.2%	53.8%	66.7%	Landmark
P. foetida	D	88.6%	71.2%	90.1%	Both
P. gracilis	D	67.6%	88.9%	86.1%	EFD
P. morifolia	D	92.6%	77.8%	81.5%	Landmark
P. actinia	E	81.1%	44.1%	86.0%	Both
P. miersii	E	59.4%	77.5%	79.8%	Both
P. sidifolia	E	68.1%	69.7%	77.1%	Both
P. triloba	E	34.1%	70.3%	59.5%	EFD
P. alata	F	58.5%	73.3%	80.0%	Both
P. edulis	F	72.7%	15.9%	75.0%	Both
P. ligularis	F	84.0%	62.9%	85.7%	Both
P. nitida	F	60.0%	27.5%	67.5%	Both
P. racemosa	F	40.6%	60.6%	55.1%	EFD

P. villosa	F	82.8%	59.6%	84.2%	Both
P. coccinea	G	46.2%	51.1%	56.5%	Both
P. cristalina	G	75.0%	65.4%	79.8%	Both
P. galbana	G	17.4%	65.1%	33.9%	EFD
P. malacophylla	G	70.1%	67.4%	83.7%	Both
P. maliformis	G	36.0%	36.0%	60.0%	Both
P. miniata	G	71.0%	22.0%	72.5%	Both
P. mucronata	G	88.6%	41.4%	90.8%	Both

For each species its class and percent correct prediction using the indicated morphometric features (landmarks, EFDs, or both) with linear discriminants is provided. "Max" indicates the set of morphometric features providing the maximum discrimination of species identity.

Table 2: Predictive power of different morphometric methods to discriminate heteroblastic node.

Heteroblasty	Landmark	EFD	Both	Max
1	49.1%	33.2%	47.9%	Landmark
2	22.9%	19.5%	27.0%	Both
3	12.3%	16.7%	15.3%	EFD
4	13.3%	8.2%	12.7%	Landmark
5	6.2%	12.2%	9.0%	EFD
6	7.0%	9.0%	9.9%	Both
7	15.9%	10.2%	10.7%	Landmark
8	5.9%	13.6%	14.2%	Both
9	10.9%	7.8%	12.0%	Both
10	12.8%	11.1%	11.6%	Landmark

For each heteroblastic node its percent correct prediction using the indicated morphometric features (landmarks, EFDs, or both) with linear discriminants is provided. "Max" indicates the set of morphometric features providing the maximum discrimination of heteroblastic node identity.

Figure 1

Click here to download Figure Figure1.tif

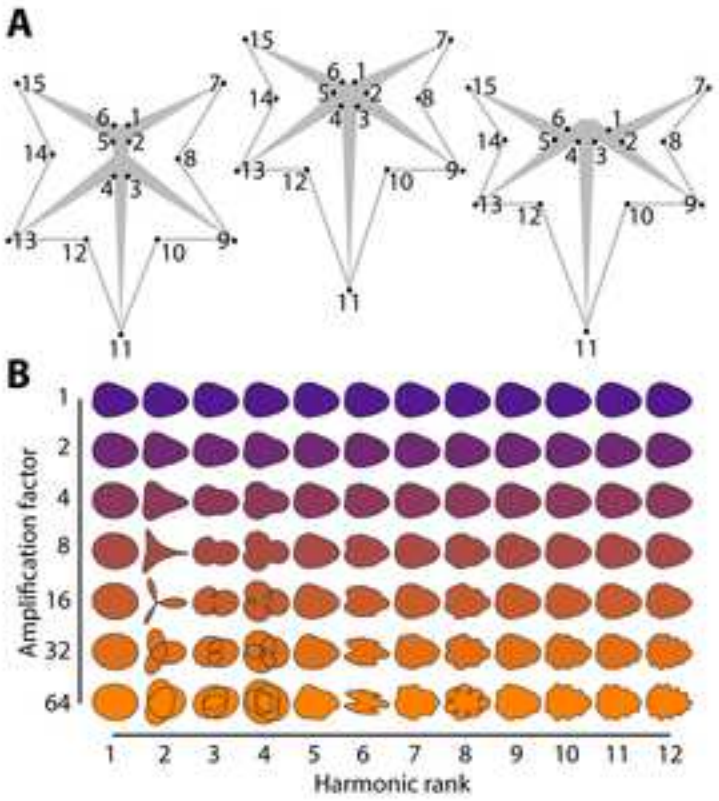


Figure 2

Click here to download Figure Figure2.tif

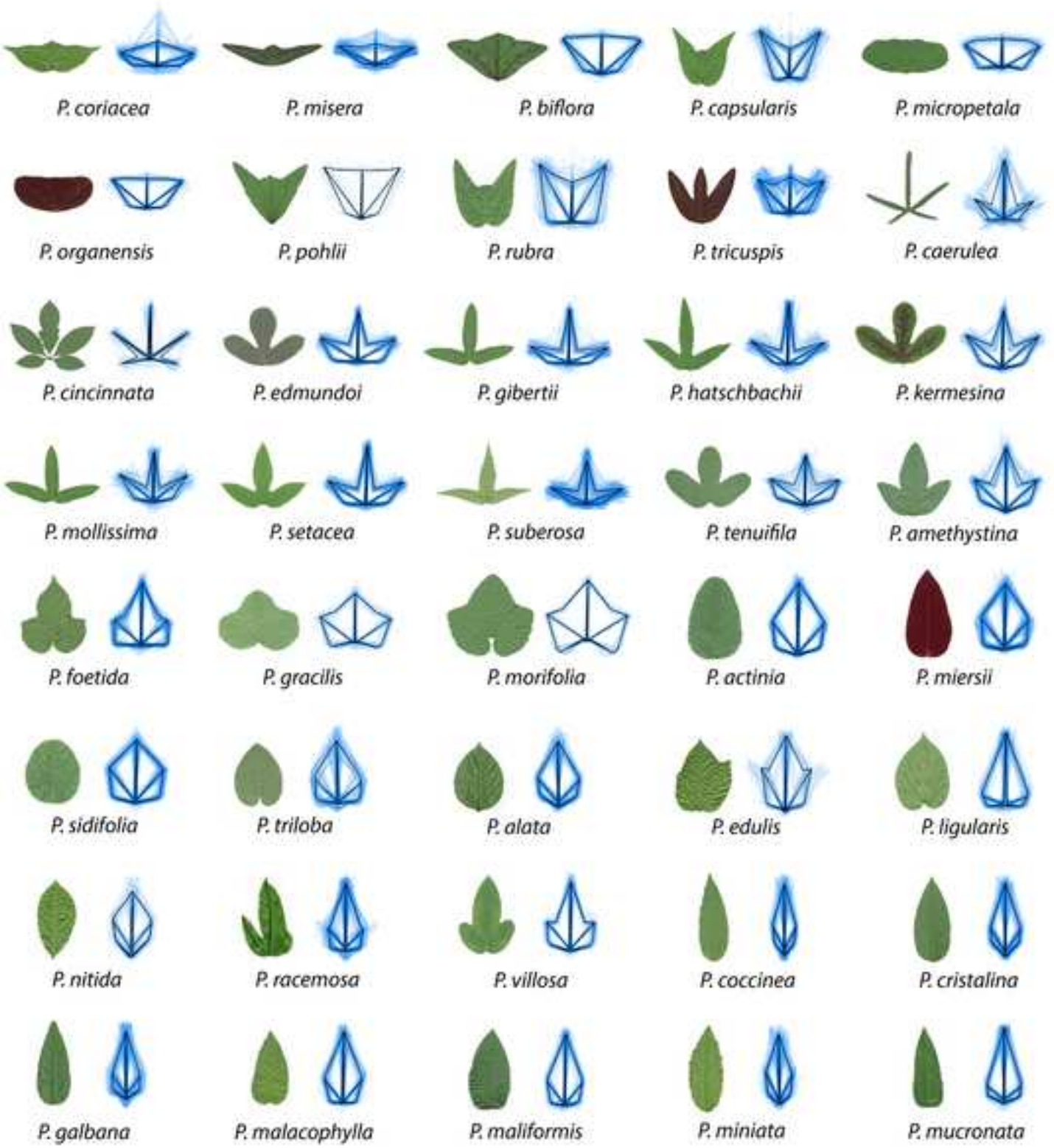


Figure 3

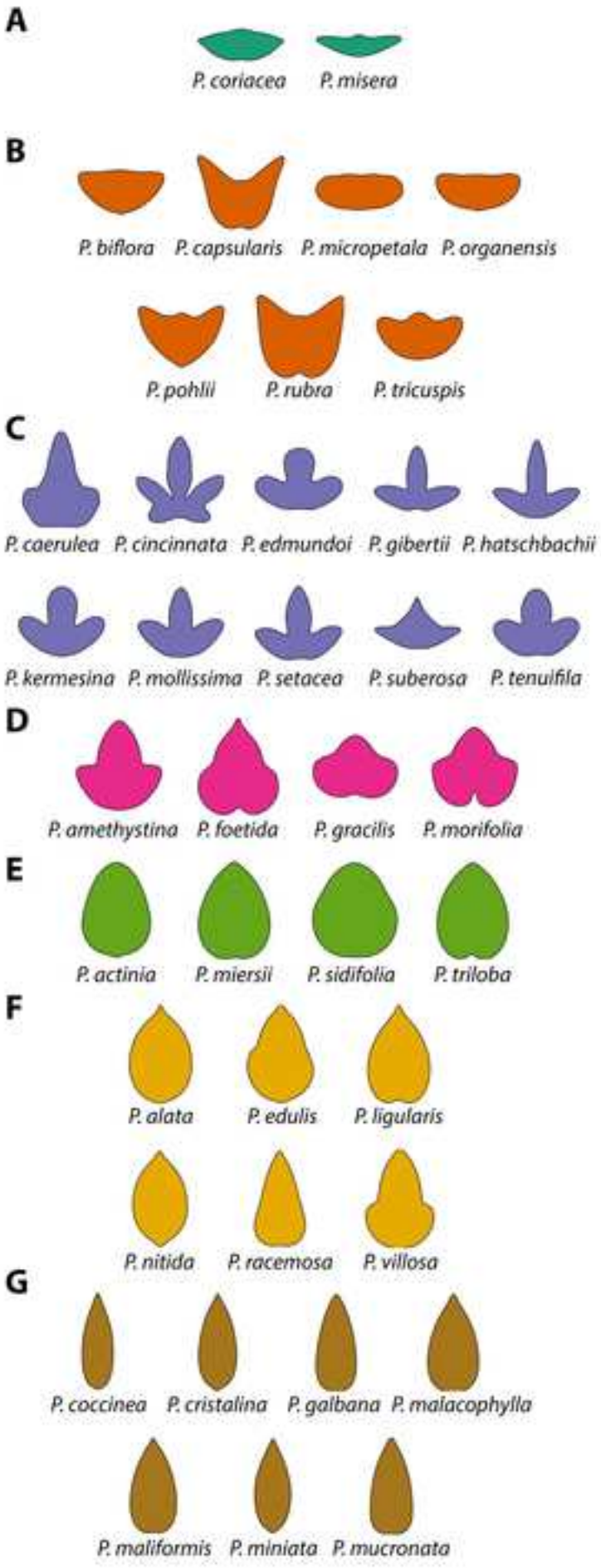


Figure 4

Click here to download Figure Figure4.tif

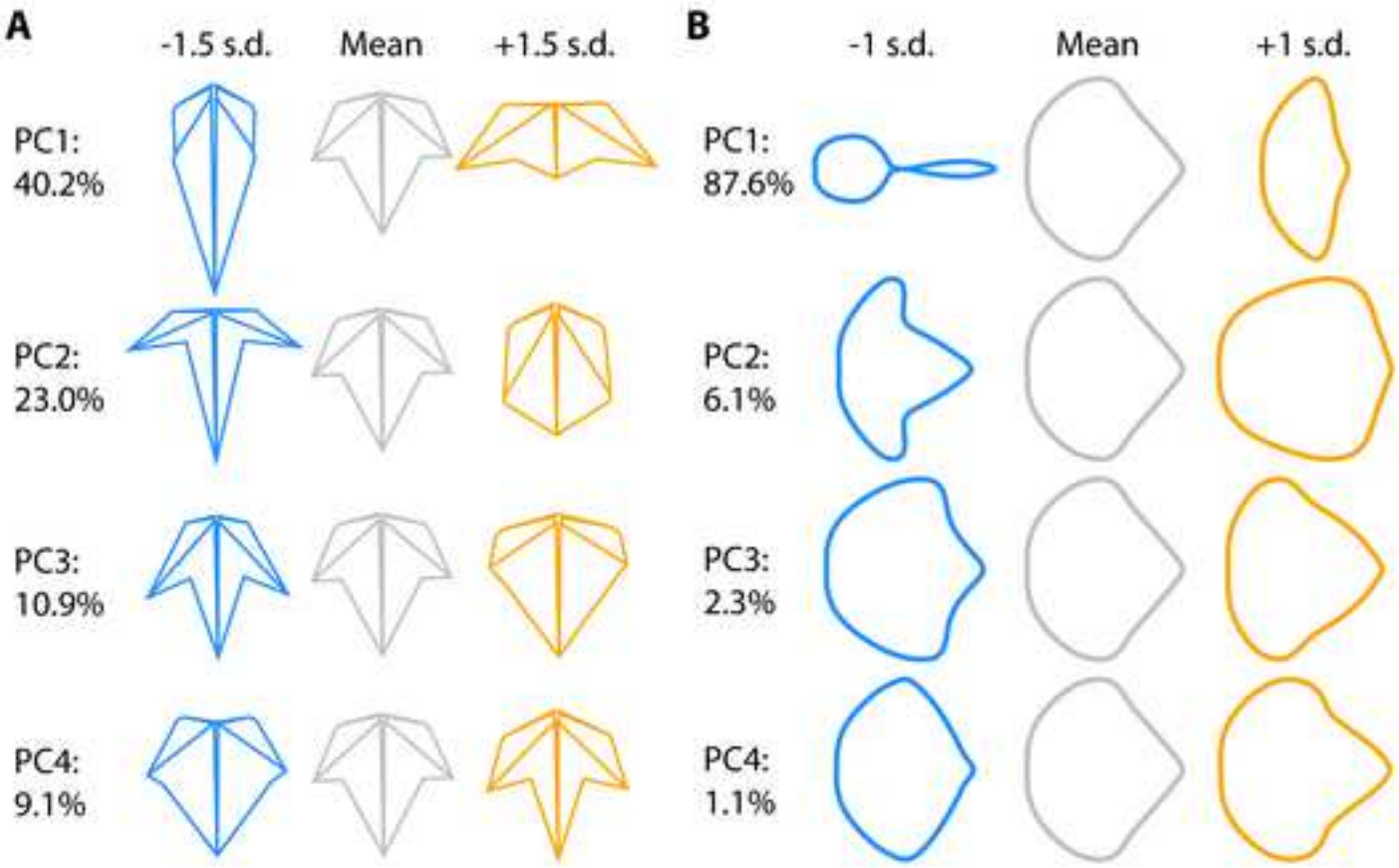


Figure 5

Click here to download Figure Figure5.tif

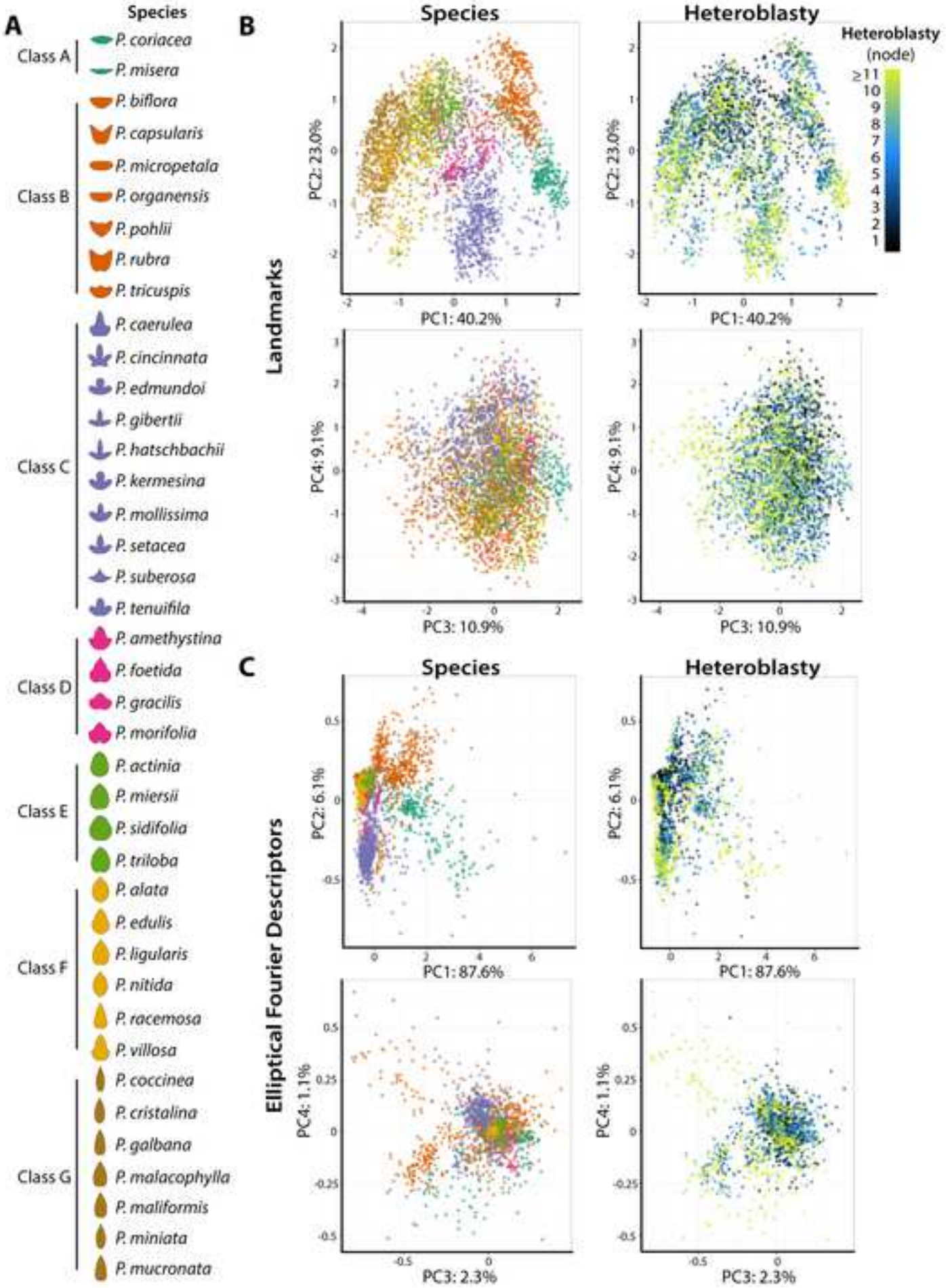


Figure 6

Click here to download Figure Figure6.tif

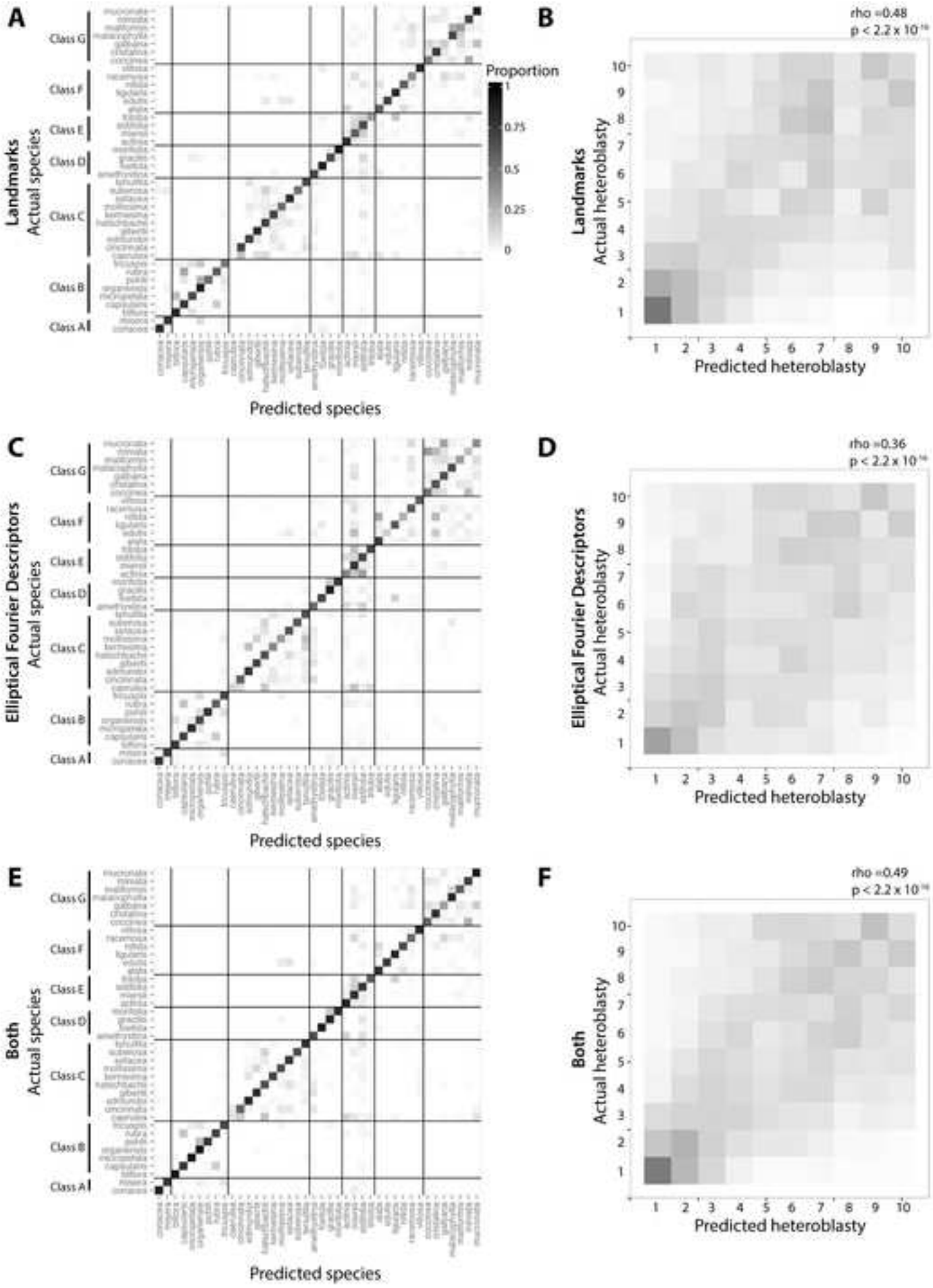


Figure 7

Click here to download Figure Figure7.tif

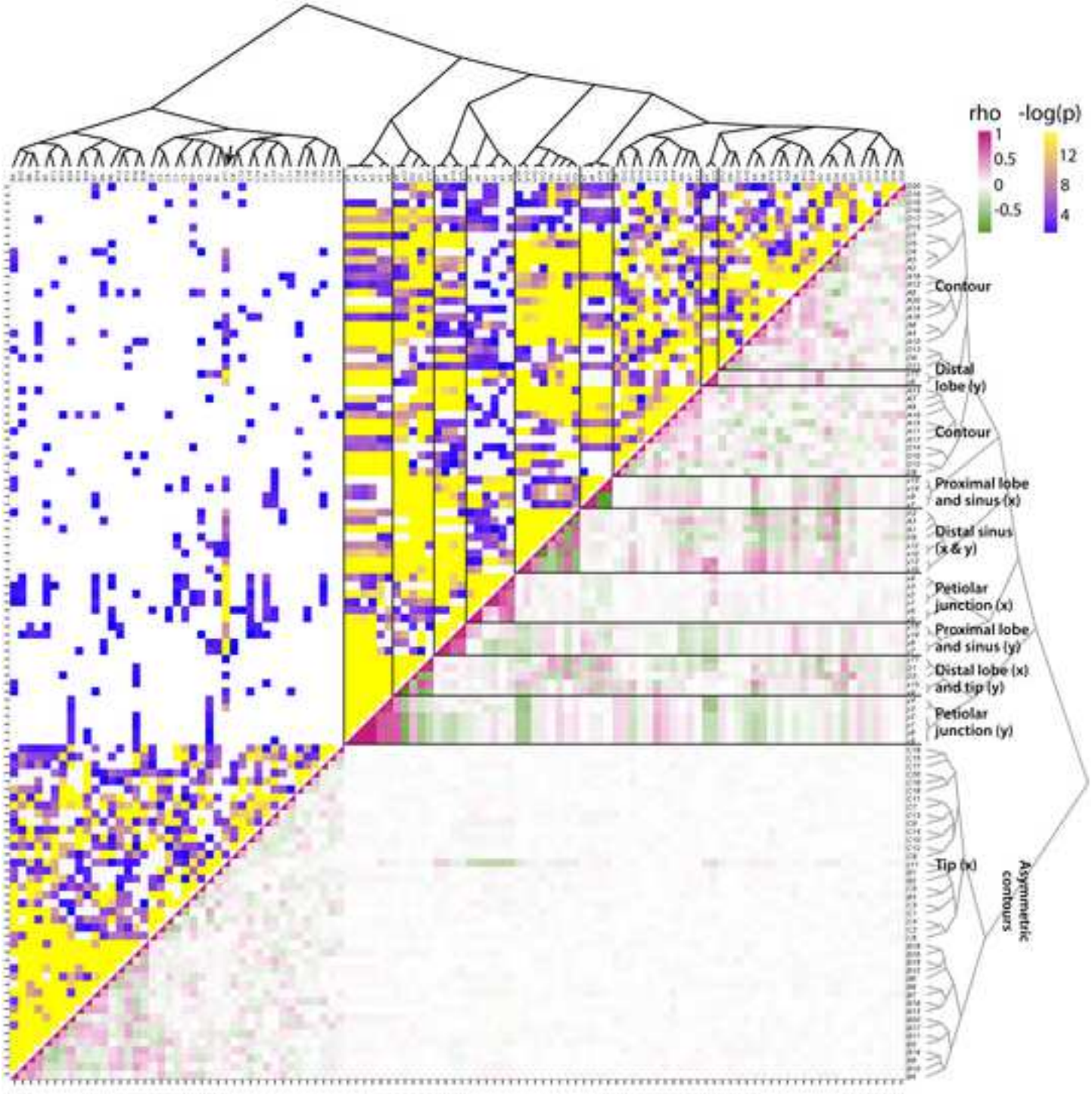
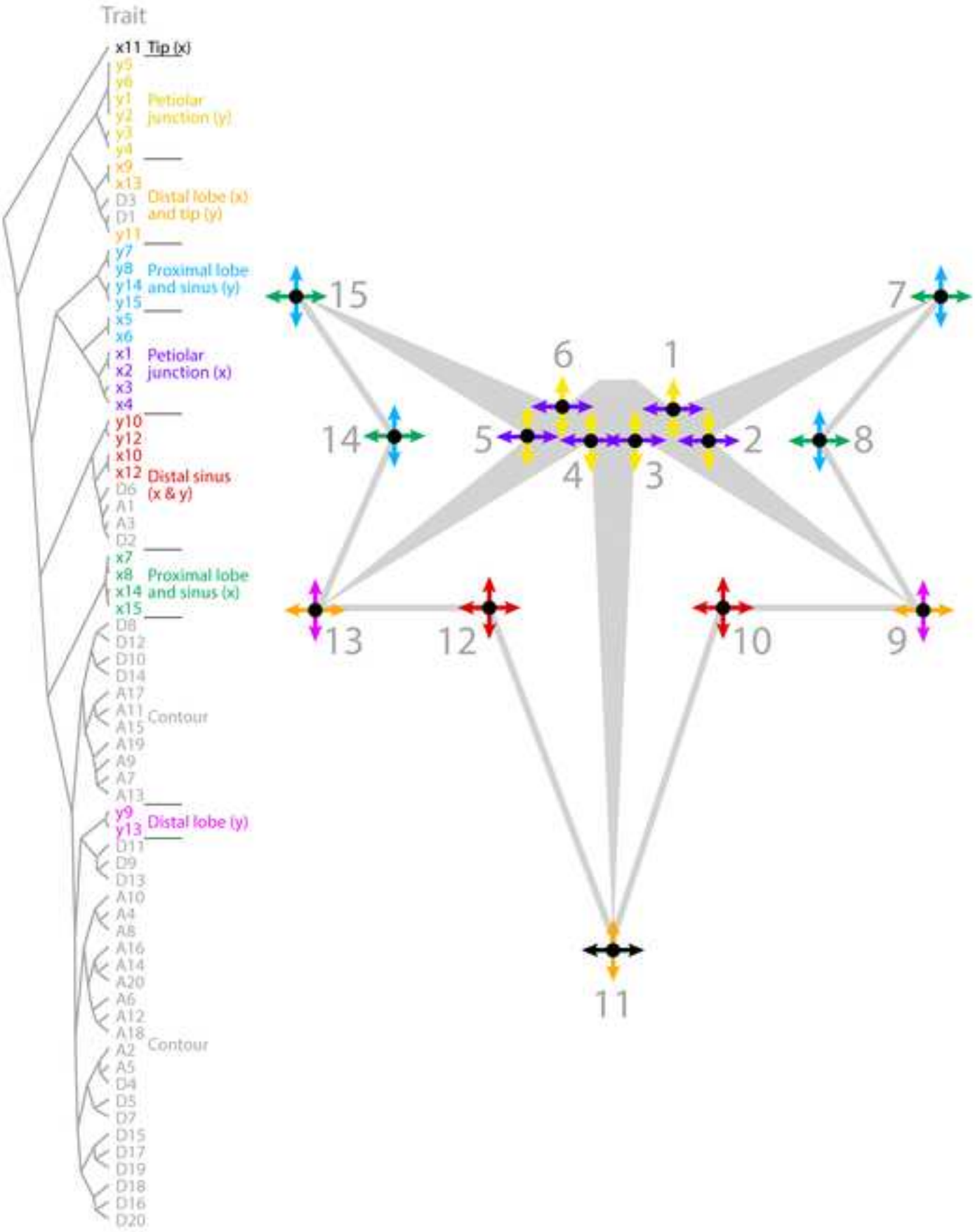


Figure 8

Click here to download Figure Figure8.tif



Morphometric analysis of *Passiflora* leaves~~I~~: the relationship between landmarks of the vasculature and elliptical Fourier descriptors of the blade

Daniel H. Chitwood[1, §] and Wagner C. Otoni[2, §]

[1]Donald Danforth Plant Science Center, St. Louis, MO, USA

[2]Departamento de Biologia Vegetal/BIOAGRO, Universidade Federal de Viçosa, Viçosa, MG, Brasil

[§]To whom correspondence should be addressed:

Daniel H. Chitwood
Donald Danforth Plant Science Center
975 North Warson Road
St. Louis, MO 63132
Phone: (314) 587-1609
Fax: (314) 587-1709
e-mail:
dchitwood@danforthcenter.org

Wagner C. Otoni
Departamento de Biologia
Vegetal/BIOAGRO
Universidade Federal de Viçosa
36570-900 Viçosa, MG, Brasil
Phone: +55 (31) 3899 2930
e-mail: wotoni@ufv.br

Abstract

BACKGROUND: Leaf shape among *Passiflora* species is spectacularly diverse. Underlying this diversity in leaf shape are profound changes in the patterning of the primary vasculature and laminar outgrowth. Each of these aspects of leaf morphology—vasculature and blade—provides different insights into leaf patterning.

RESULTS: Here, we morphometrically analyze >3,300 leaves from 40 different *Passiflora* species collected sequentially across the vine. Each leaf is measured in two different ways: using 1) 15 homologous Procrustes-adjusted landmarks of the vasculature, sinuses, and lobes and 2) Elliptical Fourier Descriptors (EFDs), which quantify the outline of the leaf. The ability of landmarks, EFDs, and both datasets together are compared to determine their relative ability to predict species and node position within the vine. Pairwise correlation of x and y landmark coordinates and EFD harmonic coefficients reveals close associations between traits and insights into the relationship between vasculature and blade patterning.

CONCLUSIONS: Landmarks, more reflective of the vasculature, and EFDs, more reflective of the blade contour, describe both similar and distinct features of leaf morphology. Landmarks and EFDs vary in ability to predict species identity and node position in the vine and exhibit a correlational structure (both within landmark or EFD traits and between the two data types) revealing constraints between vascular and blade patterning underlying natural variation in leaf morphology among *Passiflora* species.

~~This manuscript is the first of two companion pieces, the second describing divergent heteroblastic trajectories underlying the disparate leaf shapes among *Passiflora* species.~~

Keywords

Passiflora, morphometrics, leaves, leaf shape, leaf morphology, heteroblasty, Procrustes analysis, landmarks, Elliptical Fourier Descriptors

Background

The leaves of *Passiflora* species are remarkably diverse [1-3]. The underlying source of such diversity is ultimately speculative, but diversifying selective pressure from egg-laying *Heliconius* butterflies that use leaf shape as a visual cue has been proposed [4-6, 5]. The leaves not only vary between species, but between successive nodes of a single vine, sometimes dramatically, reflecting both the heteroblastic development of the shoot apical meristem from which they are derived and the ontogeny of individual leaves as they allometrically expand [67-10]. Previous morphometric work using the multiscale Minkowski fractal dimension focused on vein patterning and the contour of the blade to predictively identify *Passiflora* species. Of the 10 species analyzed, some possessed similar leaf morphologies that could be correctly classified using only a small number of leaves per species as a training set [117].

To some degree, the patterning of the vein and blade follow each other, but to what degree they vary independently, or one is the consequence of the other, remains to be determined [8 11 12-15]. At a morphometric level, vascular patterning and the contour of the blade are studied separately, as one is a topology and the other a contour. Vasculature and blade can be separated and then analyzed with the same method, and was done using a Fractal-based approach in *Passiflora* previously [711].

Alternatively, traditional morphometric approaches can be applied to vascular patterning and the outline of the blade [162]. Procrustes-adjusted landmarks are coordinate points that correspond between all measured samples, ideally through homology [173]. Homologous landmarks are ideally suited for measuring vein

thickness, vascular branch points, and the relative positioning and depth of sinuses and lobes if these features can be found in every sample, as in many palmately-lobed species, such as *Cucurbita*, *Acer*, and *Vitis*. [~~14-22~~10, 18-25].

The landmark concept can be applied to contours as well, placing numerous points along a curve and subsequently using a Procrustes superimposition to create a near-continuous analysis of outlines [26~~23~~]. The pseudo-landmark approach to quantifying contours has been used extensively to study leaf outlines, especially in species like *Antirrhinum* and *Arabidopsis*, where homologous points are lacking [~~24-26~~27-29]. Another approach, Elliptical Fourier Descriptors (EFDs), treats an outline as a wave connecting back onto itself and subsequently performs a Fourier transform, decomposing the shape into a harmonic series [~~27-30~~30-33]. EFDs have been applied to species from both *Solanum* and *Vitis* [~~19, 31-33~~23, 34-38] as well as the study of leaf asymmetry [~~22, 34, 35~~25, 39, 40], leveraging the ability to separate symmetric from asymmetric sources of variance.

Comparing landmark- and contour-based methods not only provides an integrated perspective on leaf morphology, but ~~also can~~can also potentially reveal the extent that patterning of the vasculature and blade are correlated in a quantitative fashion. Understanding the complementary features different morphometric methods detect is relevant to a wide variety of fields that use different approaches to extract information content from leaf shapes, including paleobiology and paleoclimate studies [41], ecology [42], evolution [10, 24, 27, 34, 43], genetics [21, 23, 29, 35, 36, 38], developmental biology [10, 18, 20, 24, 25, 34, 36, 39, 40], and plasticity [19, 20, 24, 37]. *Heliconius* butterflies, too, can even distinguish the shapes of leaves from different *Passiflora* species, presumably using a learning method yet to be determined [6].

Formatted: Font: Italic

Formatted: Font: Italic

-Here, we measure landmarks of the vasculature, sinuses, and lobes and EFDs of the blade for >3,300 leaves from 40 *Passiflora* species sampled from successive nodes across vines. Linear Discriminant Analyses (LDAs) are used to determine the

capacity of landmarks, EFDs, or both datasets to predict species identity versus node position in the vine. A correlational analysis of landmark and EFD data determines which specific features of leaves change together versus vary independently from each other. Our data reveals the constraints between vascular and blade patterning underlying natural variation in leaf morphology among *Passiflora* species and provides a critical comparison of complementary morphometric approaches used on the same leaves.

Data Description

The purpose of this manuscript is to compare and contrast landmark and Elliptical Fourier Descriptor (EFD) methods in the genetic and developmental analysis of leaf shape among *Passiflora* species across the sequential nodes of their vines. The dataset released with this manuscript [44~~36~~] consists of 555 scans, as well as isolated binary outlines of individual leaves, ~~of leaves~~ from 40 different species of *Passiflora* in which the order of leaves arising from the vine is recorded, ~~(~~starting with "1" for the youngest leaf scanned from the growing tip of each vine~~)~~. We importantly note: the numbering of nodes in the raw scans described above, starting at the tip of the shoot, is opposite from the numbering of nodes presented in the manuscript, in which numbering (starting with "1") begins with the oldest leaf at the base of the shoot. The reason for this opposite numbering in the manuscript is that by beginning the counting of nodes with "1" at the shoot base the numbering aligns with the heteroblastic series (which begins with the first emerged leaf at the shoot base). >3,300 leaves are represented in this dataset. The number of vines sampled per a species and the number of nodes sampled for each vine are indicated in the raw data provided with this manuscript [45] and are visually depicted as well (**Fig. S1**). Both landmark data, measuring the vasculature, lobes, and sinuses, and Elliptical Fourier Descriptor (EFD) data, which quantify the leaf outline, can be derived from the provided~~is~~ datasets. EFDs (under PassifloraLeaves/Paper1/Figure1/0.passiflora_nef.txt) and landmarks (under PassifloraLeaves/Paper1/Figure2/0.procrustes_landmarks.txt) ~~Isolated outlines of~~

~~leaves used to calculate EFD data are also released with this dataset. Code used in the statistical analysis of data is also provided~~ are provided with code in a GitHub repository [45~~37~~]. It is hoped that the release of this data will assist others in developing novel morphometric approaches to better understand the genetic, developmental, and environmental basis of leaf shape. ~~This dataset is also used in the companion piece to this manuscript [38].~~

Analyses

Vascular landmarks and Elliptical Fourier Descriptors (EFDs) of the blade

For the >3,300 leaves measured across the leaf series for 40 different *Passiflora* species, a comparison of homologous landmarks and Elliptical Fourier Descriptors (EFDs) was made (see **Fig. S1** and raw data [45] for the replication associated with each species and the number of nodes for each vine)~~.~~ These two methods globally capture complementary aspects of leaf shape, sensitive to vascular patterning and the shape of the blade, respectively.

Formatted: Font: Bold

15 landmarks were measured for each leaf (**Fig. 1A**). For the proximal veins (near the leaf base) landmarks on each side of the junction of the proximal vein with the petiolar junction (where the major veins meet) were placed (landmarks 1-2 and 5-6), capturing the width of the proximal veins. Landmarks placed at the tip of the proximal vein (landmarks 7 and 15) capture the length and angle of the proximal lobe. On the distal vein (nearer the leaf tip), landmarks were placed only on the distal side of the junction with the midvein (landmarks 3 and 4) as the other side of the base of the distal vein variably intersects the midvein, petiolar junction, and proximal vein (see three examples in **Fig. 1A**, left to right). The landmarks at the tip of the distal veins (landmarks 13 and 9) measure the length and angle of the distal lobe. Additionally, landmarks describe the placement of the leaf tip (landmark 11), distal sinuses (landmarks 10 and 12), and the proximal sinuses (8 and 14).

To determine the extent that landmarks capture qualitative variation in leaf shape among *Passiflora* species, representative leaves were compared to averaged Procrustes-adjusted landmark values (**Fig. 2**). The landmark analysis captures features such as the relative lengths and angular placement of the proximal and distal veins as well as the depth of the sinuses. Visualizing superimposed landmarks for all leaves measured in addition to the averaged landmark values demonstrates substantial sources of shape variance in some species, especially due to changes in leaf shape across the leaf series, that usually relate to the depth of the sinuses or the number of lobes.

Although landmarks accurately depict information related to vascular patterning and the relative placement of the lobes and sinuses of the blade, they fail to capture more subtle shape variation related to the curvature of the lamina. Elliptical Fourier Descriptors (EFDs) result from a harmonic decomposition of a shape outline. The harmonic contributions to leaf shape can be visualized (**Fig. 1B**), which in *Passiflora* correspond to features reflecting the leaf tip, distal lobes, and proximal lobes (the "trifoliate" features, especially in the lower harmonic ranks) or more local features (the "serrations" represented in the higher harmonic ranks) (**Fig. 1B**). The averaged outlines of leaves capture the curves and lobing of leaves from each species (**Fig. 3**). Species that display leaves with variable numbers of lobes (such as *Passiflora caerulea*, *P. cincinnata*, or *P. suberosa*) have average leaf outlines reflecting this source of shape variance.

The morphospace reflects species and heteroblastic differences in leaf shape

To analyze major sources of shape variance in Procrustes-adjusted landmark values and the harmonic series from the Elliptical Fourier Descriptor (EFD) analysis, a Principal Component Analysis (PCA) was performed to reduce the dimensionality of each dataset. Onto the resulting morphospaces were projected species identity and the node position in the leaf series ("heteroblast~~t~~sty"). Node position is referred to as "heteroblasty" as a shorthand indicating that numbering of nodes begins at the

shoot base, with "1" indicating the first emerged leaf at the shoot base. This numbering scheme more closely aligns with the heteroblastic series of leaves compared to the reverse numbering that begins at the growing shoot tip and is more sensitive to the allometric changes in rapidly expanding leaves.

Eigenleaves (theoretical leaf shapes representing the eigenvectors from a Principal Component Analysis) from each PCA reveal the shape features contributing to shape variance along each Principal Component (PC). The first four landmark PCs (**Fig. 4A**) explain 83.2% of shape variance for the landmark dataset. PC1 reflects shape variance related to long, lance-like leaves versus wider leaves with short midveins and long, extended distal lobes. Both PC2 and PC3 explain shape variance related to leaves with pronounced distal lobes versus more rounder (PC2) or deltoid (PC3) leaves with less lobing. PC4 also explains shape variance related to lobing. A comparison of the landmark eigenleaves (**Fig. 4A**) with the EFD eigenleaves (**Fig. 4B**) shows that the shape variance explained by each respective PC is strikingly similar, especially with respect to lobing and the length-to-width ratio of leaves. This demonstrates a qualitative correspondence between the orthogonal axes of each dataset, including their directionality, which will be subsequently explored in further detail.

Projecting species identity and heteroblastic node onto the landmark and EFD morphospaces reveals that each method separates the shape variance attributable to these variables, but in different ways (**Fig. 5**). Because visualizing 40 distinct species is a challenge, species were assigned to 7 different classes (consistently colored throughout the manuscript) based on a) occupying similar spaces within morphospace and b) qualitative differences in leaf shape (**Fig. 5A**). Species classes show pronounced separation from each other by PC1 and PC2 in both the landmark (**Fig. 5B**) and EFD (**Fig. 5C**) morphospaces. Less separation is observed by species class for PC3 and PC4. When heteroblastic node is projected onto the morphospaces, there is a trend for the leaves originating from high heteroblastic nodes (young leaves towards the growing tip) to occupy the lower PC2 values within each species

class. This is especially true for the landmark morphospace (**Fig. 5B**). There is also a trend for leaves originating from high heteroblastic nodes to have low PC3 values, regardless of species class. Both low PC2 and PC3 values correspond to more pronounced distal lobing (**Fig. 4**), a shape feature commonly found in young leaves near the growing tip of the plant, compared to older leaves near the base of the vine that tend to have less lobing. That shape variance attributable to species class and heteroblastic node traverse the morphospace in different ways suggests to some extent the shape variance for each of these factors is separable, as is discussed in the next section.

Discriminating species vs. ~~heteroblastic~~ node identity

That species class and ~~heteroblastic~~ node identity traverse the morphospace differently (**Fig, 5A-B**) is consistent with previous work demonstrating that shape features can be used to discriminate species independently from node position in grapevine [~~20, 21~~10, 24]. A Linear Discriminant Analysis (LDA) is used here to determine the extent these two variables can be predicted independently of the other in *Passiflora* using landmark data, Elliptical Fourier Descriptors (EFDs), and both landmark and EFDs together. We stress that the LDA approach taken in this work is fundamentally different from modeling species, node, and interaction effects using linear modeling. Such an approach (which we undertook but the data is not shown here, because it is outside the scope of this manuscript) reveals that for each morphometric trait considered independently, the species and interaction effects are the strongest and the node effect is weak. Rather, an LDA allows explicit questions to be asked regarding all the measured traits together. Can all the traits be used together to discriminate species regardless of node? Using all traits can node be distinguished separately from species? Such a framework is consistent with developmental genetic theory that differences in leaf shape between species versus more conserved heteroblastic changes in leaf shape within individual plants are regulated by distinct genetic pathways [16] that lead to separable morphological effects within single leaves (so called "cryptotypes" [46]). We also note that the

LDAs performed use the "leave one out" approach of cross-validation, in which a separate LDA for each leaf, minus the leaf in question, is used to predict the identity of that leaf. Such an approach is designed to compensate for differences in species replication and nodes sampled per a vine in our dataset (see raw data [45] and **Fig. S1**).

An LDA is first performed on species identity, regardless of node position. The resulting discriminants are then used to predict the identity of the species. Regardless of whether landmarks (**Fig. 6A**), EFDs (**Fig. 6C**), or both landmarks and EFDs are used (**Fig. 6E**) a high proportion of leaves can be correctly reassigned to the correct species. When there is confusion between species, it tends to be within the same species class. This result demonstrates that regardless of the position of a leaf within the heteroblastic series, its species identity can be predicted. For most species classes (all except C and D) the maximum correct prediction is most often achieved with both landmark and EFD data together compared to each data type alone (**Table 1**). For species classes C and D, however, landmark data alone tends to outperform EFD and both data types together. This indicates that for some species, especially those that are highly lobed as in species classes C and D, landmark data is a better indicator of species identity (perhaps because it is more explicitly related to lobing).

Conversely, heteroblastic node position can be predicted independently of species identity, but to a much lesser degree and not equally across the leaf series. The leaves occupying lower node positions (older leaves at the base of the vine) tend to be successfully predicted at a higher rate than the younger leaves of the tip, regardless of whether landmarks (**Fig. 6B**), EFDs (**Fig. 6D**), or both landmarks and EFDS are used (**Fig. 6F**). EFDs, however, overall under-perform landmarks or landmarks and EFDs used together (**Table 2**). This indicates that landmarks are a superior discriminant of node position compared to EFDs. Previous work in grapevine indicates that vein thickness is altered by shoot position [10, 24~~20~~]. That landmarks measure vein thickness, but not EFDs, may explain the differing abilities

of these two shape features to correctly discriminate leaves by heteroblastic node position. That the juvenile leaves at the lower heteroblastic node positions are correctly predicted at higher rates suggests that these leaves are more similar across species (or correspondingly, that leaves at high heteroblastic node positions are more divergent between species)~~, a hypothesis that we explore more fully in the companion manuscript [38].~~

Correlational matrix between landmarks and Elliptical Fourier Descirptors (EFDs)

Until now, landmarks and Elliptical Fourier Descriptors (EFDs) have either been considered separately or in conjunction together but not compared against each other. The landmarks used in this study tend to represent vascular features of the leaf, the lobes, and the sinuses. The EFDs represent the blade and the continuous contour and curves of lamina. Further, landmark data is represented as (x, y) coordinates, whereas EFD data is a Fourier-based harmonic series. A correlational matrix is used to find strong associations between the components of each dataset and to interpret the features each dataset uniquely quantifies against the other.

The input for the correlation matrix, using Spearman's rho, is each of the fifteen x and y coordinates of the landmark dataset and each of the four harmonic coefficients (A, B, C, D) of the first 20 harmonic ranks from the EFD data, correlated across the >3,300 leaves for all species and heteroblastic node positions used in this study. This correlation matrix was used as a distance matrix to hierarchically cluster these traits and the rho and p values subsequently visualized (**Fig. 7**).

A large set of uncorrelated traits, consisting of the B and C harmonic coefficients and the x11 landmark, end up clustered together (**Fig. 7**). The B and C harmonic coefficients represent asymmetric sources of shape variance [31~~28~~] and the x11 landmark represents the left-right variance of the leaf tip (**Fig. 1A**), which will mostly be affected by leaf asymmetry. That these shape features are weakly correlated with each other and other traits only implies that they are regulated by

an unaccounted source of variance for this particular analysis. In the future, a more in-depth analysis will likely reveal phyllotaxy as modulating leaf asymmetry [~~34, 35~~39, 40], specifically alternating asymmetry at consecutive nodes, as recently shown in other vines, such as ivy and grapevine [2~~52~~].

The remaining landmarks and the A and D coefficients of the harmonic series (representing symmetrical shape variation) show various correlational associations with each other (**Figs. 7-8**). Harmonic contributions to leaf shape (**Fig. 1B**) are more abstract and difficult to interpret than the contributions of landmarks to leaf shape, as the landmarks represent homologous points found in every leaf (**Fig. 1A**). Strong correlations between harmonic coefficients with landmarks can help interpret the context of the harmonic coefficient to leaf shape. Most the harmonic coefficients cluster exclusively together except for landmarks y9 and y13, which represent the proximal-distal displacement of the distal lobes along the leaf length (**Fig. 8**). This suggests that large amounts of the shape variance associated with the contour of the blade are influenced by the relative placement of the distal lobes along the leaf length. The remaining harmonic coefficients that cluster outside most the other coefficients also associate with features of the distal part of the leaf. A1, A3, D2, and D6 associate with the x and y coordinates of the distal sinus (x10, x12, y10, and y12) and D1 and D3 associate with the left-right displacement of the distal lobe (x9 and x13) and the vertical displacement of the leaf tip (y11) (**Fig. 8**). Although difficult to interpret, the correlations of harmonic coefficients suggest that the overall leaf contour is influenced by the placement of the distal lobe and sinus.

The remaining correlations between landmarks reveal interesting constraints governing the shape of *Passiflora* leaves (**Fig. 8**). As mentioned previously, the left-right displacement of the distal lobes (x9 and x13) strongly correlates with the vertical proximal-distal displacement of the leaf tip (y11). The x and y coordinates of the distal lobes (landmarks 10 and 12) are the only features for which the x and y displacement are correlated, suggesting that the distal sinus varies in a diagonal direction. The proximal sinus and lobe (landmarks 7, 8, 14, and 15) and the

landmarks at the base of the veins of the petiolar junction (landmarks 1, 2, 3, 4, 5, and 6) form additional groups of associated landmarks, although interestingly the x and y displacement of each of these two groups is distinct in each case (**Fig. 8**)

Discussion

Leaf morphology refers to the totality of leaf architecture, at the cellular, tissue, and organ levels, and distinct attributes of the leaf, both the vasculature and lamina. The topology of the vasculature and contour of the leaf blade are distinct geometric phenomena that require different morphometric approaches to quantify. Landmarks and Elliptical Fourier Descriptors (EFDs) are ideal methods to analyze the distinct features of leaves contributing to their shape (**Fig. 1**), but rarely are they measured and compared on the same leaves. Our analysis of disparate leaf shapes among *Passiflora* species with landmarks (**Fig. 2**) and EFDs (**Fig. 3**) reveals that both methods capture similar orthogonal axes of shape variation (**Fig. 4**), and separate both species and heteroblastic node identity, but in distinct ways (**Fig. 5**). Landmarks are superior to EFDs in predicting node position compared to species identity, most likely because they describe vascular patterning, which is relatively sensitive to heteroblast~~yic~~ compared to species differences in leaf shape (**Fig. 6; Tables 1-2**). Although most elements of the EFD harmonic series cluster together in a pairwise correlational analysis, a few are closely associated with landmarks (**Fig. 7**). Landmarks exhibit a correlational structure revealing developmental constraints in how leaves vary across *Passiflora* species and the heteroblastic series (**Fig. 8**). Together, our data quantify the relationship between blade and vasculature, revealing that one does not drive the patterning of the other, and although each distinctly varies, many shape features of the leaf change in concert across evolution and development.

Methods

Plant materials and growth conditions

Passiflora germplasm was kindly provided by R. Silva (Viveiros Flora Brasil, Araguari, MG, Brazil), Dr. F.G. Faleiro (EMBRAPA Cerrados, Planaltina, DF, Brazil), Prof. M.M. Souza (Universidade Estadual de Santa Cruz - UESC, Ilhéus, BA, Brazil), M. Peixoto (Mogi das Cruzes, SP, Brazil), Prof. M.L. Silva (Universidade do Estado de Mato Grosso, Tangará da Serra, MT, Brazil), and Prof. C.H. Bruckner (Universidade Federal de Viçosa, Viçosa, MG, Brazil).

The plants were germinated from seed, planted between late October 2015 and early March 2016, in Viçosa, at the Federal University of Viçosa, MG, Brazil. The populations were raised and maintained under polycarbonate-covered greenhouse conditions, equipped with automatic environmental control using exhaust fans and evaporative cooling panels (with expanded clay wettable pads). Seeds for each *Passiflora* species were sown in 128 cell propagation plastic trays (GPlan Comércio de Produtos Agrícola s EIRELI – ME, São Paulo, SP, Brazil) filled with horticultural organic Tropstrato HT Hortaliças substrate (Vida Verde Indústria e Comércio de Insumos Orgânicos Ltda, Mogi Mirim, SP, Brazil). After germination (30-40 days), plantlets were individually transplanted to 5 L capacity plastic pots (EME-A-EME Ind. Com. Ltda., Petrópolis, RJ, Brazil) filled with horticultural substrate. Each pot received 5 g of Osmocote® Plus NPK 15-09-12 3-4 month controlled release fertilizer (Scotts, USA). Plants were irrigated on a daily-basis with tap water, and no phytosanitary control was applied. The germination and growth rates of plants varied widely. The number of replicates for each species and the number of nodes per vine are indicated in the raw data [45] and depicted visually (**Fig. S1**).

Formatted: Font: Bold

Formatted: Font: Bold

For scanning, a multifunction printer (Canon PIXMA MX340 Wireless Office All-in-One Printer, model 4204B019, USA) was used. A 20 cm metallic ruler was positioned at the bottom of each scanned sheet as a size marker. Leaves were carefully detached, from the base to the tip of the shoot, and affixed to an A4 paper sheet, adaxial face down, using 12 mm-double sided tape (Scotch Model 9400, 3M do Brasil, SP, Brazil). The numbers written near each leaf indicate position in the

shoot, in a tip-to-base direction, starting with the youngest leaf at the tip of the shoot. It should be noted that the numbering in the scans is opposite from the numbering used in the analysis and figures for this manuscript, in which leaves are numbered with "1" starting at the shoot base. This numbering system more closely aligns with the heteroblastic series than the reverse numbering scheme originally used in the scans.

Morphometric and statistical analyses

All morphometric data and code used for statistical analysis is available on GitHub [45~~37~~]. All original data is available at GigaDB [44~~36~~].

Landmarks, as described in the text, were placed on leaves in ImageJ [47~~39~~]. Procrustes superimposition was performed using the shapes package [48~~0~~] in R [49~~1~~] with the procGPA function using reflect=TRUE. Resulting Procrustes-adjusted coordinates and principal component scores (PCs) were written out for subsequent analyses and eigenleaf representations visualized using the shapepca function.

To isolate outlines for Elliptical Fourier Descriptor (EFD) analysis, the "Make Binary" function in ImageJ [47~~39~~] was found to be sufficient to segment leaves. The wand tool was used to select individual binary leaf outlines, which were pasted into a new canvas, which was subsequently saved as an individual image, which was named by vine and node position from which the leaf was derived. The binary images were batch converted into RGB .bmp files and read into SHAPE, which was used to perform chain-code analysis [~~28, 29~~31, 32]. The resulting chain-code .chc file was then used to calculate normalized EFDs. The resulting normalized EFD .nef file was then read into Momocs (version 0.2-6) [33~~0~~] in R. The harmonic contributions to shape were visualized using the hcontrib function. Averaged leaf outlines were calculated using the meanShapes function and Principal Component Analysis (PCA) performed using the pca function and eigenleaves visualized using the PC.contrib function.

Unless otherwise noted, all visualization was performed using ggplot2 in R [5042]. Linear Discriminant Analysis (LDA) was performed using the lda function and subsequent prediction of species identity or heteroblastic node position performed using the predict function with MASS [5143]. When LDAs were used for prediction, the parameter CV was set to "TRUE", for the "leave one out" cross-validation approach, to help make analyses more robust to differences in replication and node numbers between species and vines. Hierarchical clustering was performed using the hclust function.

Availability and requirements

Project name: PassifloraLeaves
Project home page: https://github.com/DanChitwood/PassifloraLeaves
Operating system(s): Platform independent
Programming language: R
Other requirements: Not applicable
License: MIT license
Any restriction to use by non-academics: none

Availability of supporting data and materials

The data sets supporting the results of this article are available in the GigaDB repository [4436].

Declarations

Funding

Brazilian sponsoring agencies, namely FAPEMIG (Grant no. CBB - APQ-01131-15), CNPq (Grant no. 459.529/2014-5) and CAPES, are acknowledged for financial support.

Authors' contributions

The overall project was conceived by DHC and WCO. WCO grew and scanned all plant material and DHC carried out analysis. DHC and WCO wrote the paper.

Competing interests

The authors declare that they have no competing interests.

References

1. Klucking EP. Leaf venation patterns, vol. 6. 1992; Berlin: J. Cramer. Passifloraceae, 222-262.

2. MacDougal JM. Revision of Passiflora subgenus Decaloba section Pseudodysosmia (Passifloraceae). Syst Bot Monogr. 1994;41:1-146.

3. Ulmer T, Mac Dougal JM. 2004. *Passiflora:* passionflowers of the world. Portland Oregon: Timber Press. 430 p.

4. Gilbert LE. Ecological consequences of a coevolved mutualism between butterflies and plants. Coevolution of animals and plants. 1975;210-240.

5. Gilbert LE. The coevolution of a butterfly and a vine. Sci Amer. 1982;110-121.

6. Dell'aglio DD, Losada ME, Jiggins CD. Butterfly learning and the diversification of plant leaf shape. Frontiers in Ecology and Evolution. 2016;4:81.

76. Allsopp A. Heteroblastic development in vascular plants. Advances in morphogenesis. 1967;6:127-171.

8. Rolland-Lagan AG, Remmler L, Girard-Bock C. Quantifying shape changes and tissue deformation in leaf development. Plant physiology. 2014;165(2):496-505.

Formatted: Font: Cambria, 12 pt

Formatted: Font: Cambria

Formatted: Font: Cambria, 12 pt

9. Gupta MD, Nath U. Divergence in patterns of leaf growth polarity is associated with the expression divergence of miR396. The Plant Cell. 2015;27(10):2785-99.

10. Chitwood DH, Klein LL, O'Hanlon R, Chacko S, Greg M, Kitchen C, Miller AJ, Londo JP. Latent developmental and evolutionary shapes embedded within the grapevine leaf. New Phytol. 2016;210:343-55.

~~7~~11. Plotze RDO, Falvo M, Pádua JG, Bernacci LC, Vieira MLC, Oliveira GCX, Bruno OM. Leaf shape analysis using the multiscale Minkowski fractal dimension, a new morphometric method: a study with Passiflora (Passifloraceae). Canadian Journal of Botany. 2005;83:287-301.

~~8~~12. Hagemann W, Gleissberg S. Organogenetic capacity of leaves: the significance of marginal blastozones in angiosperms. Plant Syst Evol. 1996;199:121-152.

~~9~~13. Nelson T, Dengler N. Leaf vascular pattern formation. Plant Cell. 1997;9:1121-1135.

~~10~~14. Dengler N, Kang J. Vascular patterning and leaf shape. Curr Opin Plant Biol. 2001;4:50-56.

15~~1~~. Champagne C, Sinha N. Compound leaves: equal to the sum of their parts? Development. 2004;131:4401-12.

16~~2~~. Chitwood DH, Sinha NR. Evolutionary and environmental forces sculpting leaf development. Curr Biol. 2016;26:R297-306.

17~~3~~. Rohlf F, Slice D. Extensions of the Procrustes method for the optimal superimposition of landmarks. Systematic Biology. 1990;39:40-59.

18~~4~~. Jones CS. Comparative ontogeny of a wild cucurbit and its derived cultivar. Evolution. 1992;46:1827-1847.

19~~5~~. Jones CS. Does shade prolong juvenile development? A morphological analysis of leaf shape changes in Cucurbita argyrosperma subsp. sororia (Cucurbitaceae). American Journal of Botany. 1995;82:346-359.

~~16~~20. Young JP, Dickinson TA, Dengler NG. A morphometric analysis of heterophyllous leaf development in Ranunculus flabellaris. Int J Plant Sci. 1995;156:590-602.

~~17~~21. Jensen RJ, Ciofani KM, Miramontes LC. Lines, outlines, and landmarks: morphometric analyses of leaves of Acer rubrum, Acer saccharinum (Aceraceae) and their hybrid. Taxon. 2002;51:475-492.

Formatted: Font: Cambria, 12 pt

Formatted: Font: Cambria

2218. Klingenberg CP, Duttke S, Whelan S, Kim M. Developmental plasticity, morphological variation and evolability: a multilevel anlaysis of morphometric integration in the shape of compound leaves. J Evol Biol. 2011;25:115-129.

2319. Chitwood DH, Ranjan A, Martinez CC, Headland LR, Thiem T, Kumar R, Covington MF, Hatcher T, Naylor DT, Zimmerman S, Downs N, Raymundo N, Buckler ES, Maloof JN, Aradhya M, Prins B, Li L, Myles S, Sinha NR. A modern ampelography: a genetic basis for leaf shape and venation patterning in grape. Plant Physiol. 2014;164:259-72.

~~20. Chitwood DH, Klein LL, O'Hanlon R, Chacko S, Greg M, Kitchen C, Miller AJ, Londo JP. Latent developmental and evolutionary shapes embedded within the grapevine leaf. New Phytol. 2016;210:343-55.~~

241. Chitwood DH, Rundell SM, Li DY, Woodford QL, Yu TT, Lopez JR, Greenblatt D, Kang J, Londo JP. Climate and developmental plasticity: interannual variability in grapevine leaf morphology. Plant Physiol. 2016;170:1480-91.

252. Martinez CC, Chitwood DH, Smith RS, Sinha NR. Left-right leaf asymmetry in decussate and distichous phyllotactic systems. bioRxiv. 2016; dx.doi.org/10.1101/043869

263. Bookstein FL. Landmark methods for forms without landmarks: morphometrics of group differences in outline shape. Medical image analysis. 1997;1(3):225-43.

274. Langlade NB, Feng X, Dransfield T, Copsey L, Hanna AI, Thebaud C, Bangham A, Hudson A, Coen E. Evolution through genetically controlled allometry space. Proc Natl Acad Sci USA. 2005; 102:10221-10226.

285. Weight C, Parnham D, Waites R. LeafAnalyser: a computational method for rapid and large-scale analyses of leaf shape variation. Plant Journal. 2007;53:578-586.

296. Bensmihen S, Hanna AI, Langlade NB, Micol JL, Bangham A, Coen ES. Mutational spaces for leaf shape and size. Mutational spaces for leaf shape and size. HFSP Journal. 2008;2:110-120.

3027. Kuhl FP, Giardina CR. Elliptic Fourier features of a closed contour. Computer graphics and image processing. 1982;18:236-258.

3128. Iwata H, Niikura S, Matsuura S, Takano Y, Ukai Y. Evaluation of variation of root shape of Japanese radish (Raphanus sativus L.) based on image analysis using elliptic Fourier descriptors. Euphytica. 1998;102:143-9.

32~~29~~. Iwata H, Ukai Y. SHAPE: a computer program package for quantitative evaluation of biological shapes based on elliptic Fourier descriptors. Journal of Heredity. 2002;93:384-385.

33~~0~~. Bonhomme V, Picq S, Gaucherel C, Claude J. Momocs: outline analysis using R. Journal of Statistical Software. 2014;56:1-24.

34~~1~~. Chitwood DH, Headland LR, Kumar R, Peng J, Maloof JN, Sinha NR. The developmental trajectory of leaflet morphology in wild tomato species. Plant Physiol. 2012;158:1230-40.

35~~2~~. Chitwood DH, Kumar R, Headland LR, Ranjan A, Covington MF, Ichihashi Y, Fulop D, Jimenez-Gomez JM, Peng J, Maloof JN, Sinha NR. A quantitative genetic basis for leaf morphology in a set of precisely defined tomato introgression lines. Plant Cell. 2013;25:2465-81.

36~~3~~. Chitwood DH, Ranjan A, Kumar R, Ichihashi Y, Zumstein K, Headland LR, Ostria-Gallardo E, Aguilar-Martinez JA, Bush S, Carriedo L, Fulop D, Martinez CC, Peng J, Maloof JN, Sinha NR. Resolving distinct genetic regulators of tomato leaf shape within a heteroblastic and ontogenetic context. Plant Cell. 2014;26:3616-29.

37. Chitwood DH, Kumar R, Ranjan A, Pelletier JM, Townsley BT, Ichihashi Y, Martinez CC, Zumstein K, Harada JJ, Maloof JN, Sinha NR. Light-Induced Indeterminacy Alters Shade-Avoiding Tomato Leaf Morphology. Plant physiology. 2015;169(3):2030-47.

38. Fulop D, Ranjan A, Ofner I, Covington MF, Chitwood DH, West D, Ichihashi Y, Headland L, Zamir D, Maloof JN, Sinha NR. A new advanced backcross tomato population enables high resolution leaf QTL mapping and gene identification. G3 (Bethesda). 2016;6(10):3169-3184.

39~~4~~. Chitwood DH, Headland LR, Ranjan A, Martinez CC, Braybrook SA, Koenig DP, Kuhlemeier C, Smith RS, Sinha NR. Leaf asymmetry as a developmental constraint imposed by auxin-dependent phyllotactic patterning. Plant Cell. 2012;24:2318-27.

40~~35~~. Chitwood DH, Naylor DT, Thammapichai P, Weeger AC, Headland LR, Sinha NR. Conflict between intrinsic leaf asymmetry and phyllotaxis in the resupinate leaves of Alstroemeria psittacina. Front Plant Sci. 2012;3:182.

41. Bailey IW, Sinnott EW. A botanical index of Cretaceous and Tertiary climates. Science. 1915;831-4.

42. Peppe DJ, Royer DL, Cariglino B, Oliver SY, Newman S, Leight E, Enikolopov G, Fernandez-Burgos M, Herrera F, Adams JM, Correa E. Sensitivity of leaf size and

Formatted: Font: Cambria, 12 pt

Formatted: Font: Cambria, 12 pt

Formatted: Font color: Auto

Formatted: Font: Cambria

Formatted: Font: Cambria

Formatted: Font: Cambria, 12 pt

shape to climate: global patterns and paleoclimatic applications. New Phytologist. 2011;190(3):724-39.

43. Schmerler SB, Clement WL, Beaulieu JM, Chatelet DS, Sack L, Donoghue MJ, Edwards EJ. Evolution of leaf form correlates with tropical–temperate transitions in Viburnum (Adoxaceae). Proceedings of the Royal Society of London B: Biological Sciences. 2012;279(1744):3905-13.

4436. GigaDB reference

4537. Chitwood DH. PassifloraLeaves. GitHub. 2016. https://github.com/DanChitwood/PassifloraLeaves

46. Chitwood DH, Topp CN. Revealing plant cryptotypes: defining meaningful phenotypes among infinite traits. Current opinion in plant biology. 2015;24:54-60.

38. Chitwood DH, Otoni WC. Morphometric analysis of Passiflora leaves II: divergent heteroblastic trajectories underlie the disparate leaf shape of Passiflora species. bioRxiv. 2016;

4379. Abràmoff MD, Magalhães PJ, Ram SJ. Image processing with ImageJ. Biophotonics international. 2004;11(7):36-42.

480. Dryden IL. shapes: Statistical Shape Analysis. R package version 1.1-11. 2015;https://CRAN.R-project.org/package=shapes

491. R Development Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2016;http://www.R-project.org.

5042. Wickham H. ggplot2: Elegant Graphics for Data Analysis. 2009;Springer-Verlag, New York.

5143. Venables WN, Ripley BD. Modern Applied Statistics with S. Fourth Edition. 2002;Springer, New York.

52. Chitwood DH, Otoni WC. Divergent heteroblastic trajectories underlie disparate leaf shapes among Passiflora species. bioRxiv. 2016; http://dx.doi.org/10.1101/067520

Figure Legends

Formatted: Font: Cambria, 12 pt

Formatted: Font: Cambria, 12 pt

Figure 1: Landmarks and harmonic contributions to shape. A) The 15 landmarks used for analysis. Left to right, landmark placement when the distal and proximal veins l) pinnately emerge from the midvein, m) both originate from the petiolar junction, or r) the proximal vein branches from the distal. **B)** Harmonic contributions to shape resulting from Elliptical Fourier Descriptor (EFD) analysis. The harmonic rank is arranged horizontally and the amplification factor (which multiplies the harmonic contributions to shape by the indicated amount) vertically. Note: for convenience to the reader, these panels are recapitulated in the companion manuscript [52].

Figure 2: The shapes of *Passiflora* leaves measured using landmarks. For the 40 species analyzed in this study, both a representative leaf and landmark data are shown. For the landmark data, the mean leaf for the species in shown in black, whereas all data for the species is depicted in semi-transparent blue.

Figure 3: The shapes of *Passiflora* leaves measured using Elliptical Fourier Descriptors (EFDs). Mean leaves calculated for each of 40 species analyzed in this study from the harmonic series resulting from an Elliptical Fourier Descriptor (EFD) analysis of the leaf contours. **A-G)** Classes of species are indicated by their respective panels. Species classes were determined by neighboring position in the Principal Component Analysis (PCA) morphospace, described in **Figs. 4-5**. Color indicates class: class A, teal; class B, orange; class C, lavender; class D, magenta; class E, green; class F, yellow; class G, brown.

Figure 4: Principal Components (PCs) and eigenleaves. A) Principal components (PCs) representing shape variance in landmark data. Eigenleaf representations (theoretical leaf shapes representing the eigenvectors from a Principal Component Analysis) at +/- 1.5 standard deviations (s.d.) are shown for the first four PCs. Percent variance explained by each PC indicated. **B)** PCs representing shape variance in Elliptical Fourier Descriptor (EFD) data. Eigenleaf representations at +/-

1 s.d. are shown for the first four PCs. Percent variance explained by each PC indicated.

Figure 5: Morphospace by species and heteroblastic node. A) Key, showing species classes and averaged leaf contours for each species. Color indicates class, which is used in other panels. **B)** Principal Component Analysis (PCA) of landmark data. Graphs for PC2 vs. PC1 and PC4 vs. PC3 are colored by species class and by heteroblastic node. Percent variance explained by each PC indicated. **C)** PCA of Elliptical Fourier Descriptor (EFD) data. Graphs for PC2 vs. PC1 and PC4 vs. PC3 are colored by species class and by heteroblastic node. Percent variance explained by each PC indicated. Heteroblastic node position is numbered "1" starting from the shoot base. Class color scheme: class A, teal; class B, orange; class C, lavender; class D, magenta; class E, green; class F, yellow; class G, brown. Heteroblastic node color scheme: shoot base, black; middle shoot, blue; shoot tip, yellow.

Figure 6: Linear Discriminant Analysis (LDA). Linear Discriminant Analysis (LDA) using **A-B)** landmark data, **C-D)** Elliptical Fourier Descriptor (EFD) data, and **E-F)** both datasets. For each set of LDAs, analysis was performed to discriminate species (ignoring heteroblastic node information) or to discriminate heteroblastic node (ignoring species information). Subsequent prediction of species or heteroblastic node identity is then visualized using confusion matrices, where actual identity is oriented vertically, predicted identity horizontally, and the proportion assigned indicated as fill. Species LDAs are broken up by species class. For heteroblastic node LDAs, Spearman's rho and associated p values calculated from correlating actual and predicted node identities are provided. Predictions carried out using LDA use the "leave one out" approach cross-validation approach. Heteroblastic node position is numbered "1" starting from the shoot base. Color scheme: low assigned proportion, white; high assigned proportion, black.

Figure 7: Correlational matrix of landmark and Elliptical Fourier Descriptor (EFD) traits. Spearman's correlation matrix for morphometric features analyzed in

this study. Upper half indicates $-\log_{10}$ p value and lower half Spearman's rho between indicated traits. Morphometric traits, both landmark and the harmonic series, are indicated along the sides, arranged using hierarchical clustering, the topology of which is depicted as a dendrogram. Key groupings of landmarks indicating correlational associations with each other or EFD harmonics are indicated. Spearman's rho: low values, green; middle values, white; high values, magenta. $-\log_{10}$ p values: low values, purple; high values, yellow; $p < 0.05$, no color.

Figure 8: Correlational relationships between vascular landmarks and leaf contours. Correlational relationships between x and y components of landmarks and Elliptical Fourier Descriptor (EFD) harmonics are indicated by dendrogram (left) and landmarks qualitatively on a representation of a leaf (right). x and y landmark components are independently depicted by arrows and colored as indicated to show major correlational sources of shape variance within *Passiflora* leaves.

Supplemental Information

Figure S1: Species replication and number of nodes sampled. A) Dotplot showing the number of vines sampled for each species. **B)** Boxplot showing the number of nodes sampled for vines from each species. The largest red line is the median nodes sampled (14 nodes), the medium sized redlines the 25th and 75th quantiles (12 and 16 nodes, respectively), and the thin red lines the minimum and maximum (7 and 28 nodes, respectively).

Table 1: Predictive power of different morphometric methods to discriminate *Passiflora* species.

Species	Class	Landmark	EFD	Both	Max
P. coriacea	A	83.2%	81.7%	88.1%	Both
P. misera	A	77.0%	71.6%	76.9%	Landmark
P. biflora	B	84.1%	75.4%	92.1%	Both
P. capsularis	B	77.2%	72.0%	77.3%	Both

Formatted: Font: Bold

Formatted: Font: Bold

Formatted: Superscript

Formatted: Superscript

P. micropetala	B	69.1%	81.8%	92.4%	Both
P. organensis	B	89.4%	70.7%	96.6%	Both
P. pohlii	B	54.5%	77.8%	77.8%	EFD
P. rubra	B	60.3%	59.7%	71.6%	Both
P. tricuspis	B	49.0%	67.8%	69.2%	Both
P. caerulea	C	0.0%	15.1%	9.4%	EFD
P. cincinnata	C	71.4%	59.3%	59.3%	Landmark
P. edmundoi	C	72.8%	78.8%	83.8%	Both
P. gibertii	C	84.0%	72.2%	81.9%	Landmark
P. hatschbachii	C	72.0%	65.4%	67.9%	Landmark
P. kermesina	C	71.0%	43.6%	70.9%	Landmark
P. mollissima	C	53.6%	35.5%	67.7%	Both
P. setacea	C	81.9%	64.4%	77.8%	Landmark
P. suberosa	C	52.3%	63.4%	66.9%	Both
P. tenuifila	C	68.8%	65.1%	79.4%	Both
P. amethystina	D	69.2%	53.8%	66.7%	Landmark
P. foetida	D	88.6%	71.2%	90.1%	Both
P. gracilis	D	67.6%	88.9%	86.1%	EFD
P. morifolia	D	92.6%	77.8%	81.5%	Landmark
P. actinia	E	81.1%	44.1%	86.0%	Both
P. miersii	E	59.4%	77.5%	79.8%	Both
P. sidifolia	E	68.1%	69.7%	77.1%	Both
P. triloba	E	34.1%	70.3%	59.5%	EFD
P. alata	F	58.5%	73.3%	80.0%	Both
P. edulis	F	72.7%	15.9%	75.0%	Both
P. ligularis	F	84.0%	62.9%	85.7%	Both
P. nitida	F	60.0%	27.5%	67.5%	Both
P. racemosa	F	40.6%	60.6%	55.1%	EFD
P. villosa	F	82.8%	59.6%	84.2%	Both
P. coccinea	G	46.2%	51.1%	56.5%	Both
P. cristalina	G	75.0%	65.4%	79.8%	Both
P. galbana	G	17.4%	65.1%	33.9%	EFD
P. malacophylla	G	70.1%	67.4%	83.7%	Both
P. maliformis	G	36.0%	36.0%	60.0%	Both
P. miniata	G	71.0%	22.0%	72.5%	Both
P. mucronata	G	88.6%	41.4%	90.8%	Both

For each species its class and percent correct prediction using the indicated morphometric features (landmarks, EFDs, or both) with linear discriminants is provided. “Max” indicates the set of morphometric features providing the maximum discrimination of species identity.

Table 2: Predictive power of different morphometric methods to discriminate heteroblastic node.

Heteroblasty	Landmark	EFD	Both	Max
1	49.1%	33.2%	47.9%	Landmark
2	22.9%	19.5%	27.0%	Both
3	12.3%	16.7%	15.3%	EFD
4	13.3%	8.2%	12.7%	Landmark
5	6.2%	12.2%	9.0%	EFD
6	7.0%	9.0%	9.9%	Both
7	15.9%	10.2%	10.7%	Landmark
8	5.9%	13.6%	14.2%	Both
9	10.9%	7.8%	12.0%	Both
10	12.8%	11.1%	11.6%	Landmark

For each heteroblastic node its percent correct prediction using the indicated morphometric features (landmarks, EFDs, or both) with linear discriminants is provided. “Max” indicates the set of morphometric features providing the maximum discrimination of heteroblastic node identity.