Diffeomorphic Independent Contrasts for Ancestral Reconstruction of Shapes

Michael Lind Severinsen[1,*], Morten Akhøj[2], Rasmus Nielsen [1,3], Stefan Sommer [2], and Christy Anna Hipsley [4]

[1] *GeoGenetics Centre, Globe Institute, University of Copenhagen, Copenhagen, 1350, Denmark*
[2] *Department of Computer Science, University of Copenhagen, Copenhagen, 1350, Denmark*
[3] *Department of Integrative Biology, University of California Berkeley, CA, USA*
[4] *Department of Biology, University of Copenhagen, Copenhagen, 2200, Denmark*

**E-mail: Michael.baand@sund.ku.dk*

Abstract

Ancestral reconstruction is a fundamental challenge in evolutionary biology, requiring methods that can capture complex morphological changes while accounting for phylogenetic relationships. Current approaches are based on linear assumptions that often oversimplify the spatial relationships between anatomical features and fail to account for landmark correlations within shapes. Here, we introduce a novel method that combines the ability of Large Deformation Diffeomorphic Metric Mapping (LDDMM) to model smooth, invertible transformations between shapes while preserving the relationships between landmarks with Felsenstein's Independent Contrasts (IC) to iteratively reconstruct ancestral shapes along the branches of a phylogenetic tree. We call this method Diffeomorphic Independent Contrasts for Ancestral Reconstruction of Shapes (DICAROS). We validate DICAROS against two existing methods: (1) Linear predictors using Ordinary Least Squares and (2) Ancestral character estimation using maximum likelihood under Brownian Motion and apply DICAROS to a dataset of swallowtail butterfly species (Family Papilonidae, Order Lepiodetra) to reconstruct the ancestral shape and visualize evolutionary trajectories in a phylomorphospace from the contrasts. We conclude that

DICAROS outperforms the existing methods in terms of accuracy and provides a more accurate reconstruction of the ancestral shape for non-symmetric phylogenetic trees. With DICAROS we show a transition between un-tailed and tailed papilinodae species while also illustrating how images of modern species would look under the DICAROS ancestral reconstruction

Key words: Ancestral Shape Reconstruction, LDDMM, Phylogenetics, Papilionidae

Evolutionary studies of morphology and omics have experienced rapid advancement in recent years, driven by new analytical methodologies and technological progress. From the increased accessibility of phylogenetic analyses and geometric morphometric (GM) through open-source software packages like Geomorph (Adams (2013, 2024)) and PhyTools (Revell (2024)) in the programming language R. These tools enable researchers to do phylogenetic comparative analysis while accounting for shared inheritance between related species in morphological and phylogenetic analyses, building on foundational work by Felsenstein and Bookstein (Felsenstein (1985); Bookstein (1997)).

Parallel advances in genomic sequencing and assembly (Höhna (2016); Zhang (2025)) have enabled phylogenetic inference across large evolutionary trees (Misof (2014); Kawahara (2023); Stiller (2024)). Which allows for the study of complex shape change in across tree of life, such as the mammalian skull evolution (Goswami (2023)) to bird brain and skull development (Chiappe (2024)) and butterfly wing vein patterns (Chazot (2016); Owens (2020)).

The increasing scale and complexity of these taxonomic datasets have been supported by the digitization of museum collections (Nelson (2019)), with resources like GBIF (GBIF (2024)), Morphosource (Morposource (2024)), and community-driven species mapping platforms (Inaturalist (2024)) providing unprecedented access to information across institutions. With these resources it is possible to combine phylogenetic information with shape data derived from landmarks placed on anatomical correspondences in 2D

images or 3D reconstructions (Gunz (2013); Bardua (2019); Mitterocker (2021)). This enables ancestral shape reconstruction (Schluter (1997); Revell (2024)) and visualization in lower-dimensional phylomorphospaces (Polly (2013); Baken (2021)).

The current paradigm within morphological methods for ancestral reconstruction, including methods such as Ordinary Least Squares (OLS) linear predictors (Adams (2013, 2024)) and Ancestral character estimation using maximum likelihood under Brownian Motion (Revell (2024)), rely on linear assumptions that simplifies the complex spatial relationships between anatomical features. A key limitation is their inability to account for the correlation between landmarks within shapes, neglecting how changes in one landmark's position can influence others during, e.g., reconstruction. This correlation structure is critical for two main reasons. First, biological shapes exhibit inherent modularity through physical proximity and functional relationships between features (Mitterocker (2007); Klingenberg (2008); Adams (2016)). Ignoring these relationships leads to oversimplified models that may miss important biological patterns. Second, understanding landmark correlations provides insights into developmental constraints and evolutionary mechanisms shaping morphological variation (Hallgrímsson (2009); Young (2010)).

To address these limitations, we employ Large Deformation Diffeomorphic Metric Mapping (LDDMM) (Beg (2005); Younes (2010)). In this sense, our work can be seen as a continuation of Tangent Phylogenetic PCA (Akhøj (2023)), which is a method for ancestral state reconstruction based on nodes taking values on any finite-dimensional Riemannian manifold, including the LDDMM landmark manifold. LDDMM provides a mathematically rigorous framework for analyzing landmark configurations by explicitly modeling the correlations between landmarks. Operating as a Hamiltonian framework, LDDMM enables shape registration through smooth, invertible mappings that follow optimal geodesic paths between configurations. This approach preserves the topological relationships between landmarks throughout the transformation process, making it

particularly suitable for biological shape analysis. The framework has been successfully applied in computational anatomy and medical imaging for tasks such as brain mapping (Durrleman (2014); Miller (2015)).

This study leverages LDDMM's geodesic paths to reconstruct ancestral shapes along branches in a phylogenetic tree by combining LDDMM with Felsenstein's *Independent Contrasts* (IC) method (Felsenstein (1985)). We call this novel approach *Diffeomorphic Independent Contrasts for Ancestral Reconstruction of Shapes* (DICAROS). Unlike traditional geometric morphometrics based on Kendall's shape space (Klingenberg (2020)) and Procrustes alignment (Gower (1975)), the LDDMM framework offers several key advantages: it explicitly models landmark correlations, enables stochastic shape development models, and extends beyond landmarks to continuous curves, surfaces, and images. Our primary contribution is reformulating Felsenstein's method using Riemannian manifold operations while maintaining equivalence to the original Euclidean approach when applied to vector data. We validate DICAROS by comparing the root reconstruction against two established methods: (1) Ordinary Least Squares (OLS) linear predictors (Adams (2013, 2024)) and (2) Ancestral character estimation using likelihood under Brownian Motion (Revell (2024)). To demonstrate DICAROS' practical utility, we apply it to reconstruct ancestral wing morphology on swallowtail butterflies (Family Papilionidae, Order Lepidoptera), by showing the phylomorphospace estimated by DICAROS, and the evolutionary shape trajectories from a leaf image to the root shape.

Materials and Methods

Independent contrasts for LDDMM shape observations

The *independent contrasts* (IC) method (Felsenstein (1985)) reconstructs ancestral states and estimates evolutionary covariance matrices. Here, we generalize the IC to handle shape data, where each node value x represents a shape defined by landmarks $x^1, \ldots, x^k \in \mathbb{R}^d$. Rather than treating shapes as vectors in Euclidean space $\mathbb{R}^{d \cdot k}$, we view them as elements of the LDDMM landmark manifold where distances between shapes are measured by geodesics paths. We combine LDDMM with IC by interpreting the phylogenetic tree as a network of geodesic paths on this manifold. Starting from the leaves, we traverse the tree in post-order, computing the diffeomorphism between sister taxa at each internal node. This diffeomorphism defines a geodesic path between the branches of shape x_i and x_j. Given the relative position of their common ancestor y_i along this path, we reconstruct the ancestral shape by deforming x_i toward x_j and stopping at the appropriate position of y_i. Where x_i corresponds to the child of y_i, with the shortest branch lenght. We repeat this process until reaching the root.

The LDDMM framework enables the computation of an evolutionary covariance matrix through the right Lie transport of cotangent vectors, accounting for the nonlinear geometry of shape space. This allows us to perform a generalized version of Phylogenetic PCA (Revell (2009)) that respects the manifold structure of the shape data. The key advantage of our approach is that both the ancestral reconstruction and covariance estimation are performed using operations that are appropriate for shape data, unlike traditional methods that rely on linear approximations. The LDDMM framework introduces a kernel parameter σ which determines the spatial scale of correlations between landmarks through a Gaussian kernel $\kappa(q_i, q_j) := \exp^{-\|q_i - q_j\|^2/2\sigma^2} \in \mathbb{R}$ with parameter $\sigma > 0$ (Miller (2002); Younes (2010); Pennec (2019)). A larger kernel width means that movements of one landmark have a broader influence on surrounding landmarks, while a smaller width allows more localized, independent deformations. This parameter effectively

controls the "stiffness" or "elasticity" of the deformation field but also implies there is no canonical metric on the LDDMM landmark manifold, making σ a hyperparameter.

The LDDMM landmark manifold is a Riemannian manifold, where standard vector space operations like addition and multiplication are not available. Instead, the *Riemannian exponential* map $\mathrm{Exp}_x(v)$ and *Riemannian logarithmic* map is used $\mathrm{Log}_x(y)$ as generalizations of addition and subtraction. The logarithm $\mathrm{Log}_{x_i}(x_j)$ between two shapes x_i and x_j yields a vector v in the *tangent space* $T_{x_i}\mathbb{R}^{d\times k}$ at x_i, providing a local linear approximation of the curved manifold (Michor (2020)). The exponential map $\mathrm{Exp}_{x_i}(x_j)$ takes this tangent vector and maps it back to a point on the manifold.

When computing these maps for independent contrasts, the kernel width σ directly influences how shape differences are interpreted through the cometric $g_m^\star(x)$, which is a block matrix of Gaussian kernels, here we choose σ as the mean distance between each landmark and its nearest neighbor. Since the contrasts are computed as cotangent (momentum) vectors in different cotangent spaces $T_x^\star\mathbb{R}^{d\times k}$, we need to map them to a common space to compute the phylogenetic covariance matrix. For the LDDMM landmark manifold, this mapping between cotangent spaces can be achieved by composing linear transformations based on the Jacobians of the optimal diffeomorphisms along the geodesic path between shapes (see Supplementary Section 1 for mathematical details).

To adapt the IC method for this manifold structure, we modify the computation of covariance matrices by replacing standard vector outer products with Riemannian metric-weighted outer products, as shown in Equation (1). Our algorithm for IC on the LDDMM manifold follows the same structure as the Euclidean version, differing only in three key steps that employ Riemannian operations: step *3* (raw contrast computation between nodes), step *5* (weighted mean computation for inner nodes), and step *9* (evolutionary covariance matrix computation using standardized contrasts) as detailed in Algorithm 1. With the evolutionary covariance matrix we project all the evolutionary shape trajectories with Principal Component Analysis (PCA) into a so-called Phylomorphospace.

Algorithm 1 Independent contrasts on the LDDMM landmark manifold

Require: A phylogenetic tree with $|V|$ nodes, $|E|$ branches and leaf node observations $x_1 \dots x_N \in \mathbb{R}^{d \cdot k}$.

Ensure: Estimates of the root value and evolutionary covariance matrix.

1: **while** $|V| > 1$ **do**

2: Choose two leaf nodes x_i and x_j with common parent node.

3: Compute the raw contrast $c_{ij} = \mathrm{Log}_{x_i}(x_j)$, where Log is the Riemannian logarithm.

4: Compute the standardized contrast

$$s_{ij} = \frac{c_{ij}}{e_i + e_j},$$

where e_i, e_j are the lengths of the branches leading to their common parent node.

5: Remove nodes i, j and branches e_i, e_j from the tree, so that $|V|$ is lowered by 2. The parent node k now becomes a leaf. Assign to it the node value

$$x_k \leftarrow \mathrm{Exp}_{x_i}\left(\frac{1/e_j}{1/e_i + 1/e_j} \cdot c_{ij}\right) = \mathrm{Exp}_{x_j}\left(\frac{1/e_i}{1/e_i + 1/e_j} \cdot c_{ji}\right),$$

where Exp is the Riemannian exponential.

6: Increase the length of the edge leading from node k to its parent as follows,

$$e_k \leftarrow e_k + \frac{e_i \cdot e_j}{e_i + e_j}.$$

7: **end while**

8: Let $\hat{x}_r \in \mathbb{R}^{d \cdot k}$ be the assigned node value of the single remaining node in the tree. This is the root estimate.

9: Let $s_1, \dots, s_{|E|/2}$ be the standardized contrasts computed above. Map each contrast to the tangent space at $\hat{x}_r$ and denote the resulting vectors by $\tilde{s}_1, \dots, \tilde{s}_{|E|/2}$ (see Supplementary Section 2 for details on this map). The estimated evolutionary covariance matrix is

$$\hat{\Sigma}^2 = \frac{\sum_{i,j} \tilde{s}_{ij} \tilde{s}_{ij}^T g(\hat{x}_r)}{|E|/2}, \tag{1}$$

where $g(\hat{x}_r)$ is the Riemannian metric matrix at $\hat{x}_r$.

return $\hat{x}_r$ and $\hat{\Sigma}^2$

Given two landmark shapes x_i and x_j, the geodesic contrast $c_{ij} = \mathrm{Log}_{x_i}(x_j)$ defines a geodesic path of shapes from x_i to x_j. This geodesic path can be extended beyond the landmarks to include all points in the surrounding domain, such as a rectangle encompassing both shapes. The entire domain is transformed by a diffeomorphism ϕ that maps the k landmarks of shape x_i to their corresponding landmarks in shape x_j, i.e., $\phi(x_i^r) = x_j^r$ for $r = 1, \ldots, k$. The movement of surrounding points is determined smoothly through an interpolation kernel (see Supplementary Section 2). In practice, the Gaussian kernel width parameter $\sigma_{DICAROS}$.

For image reconstruction along the phylogenetic tree, consider x_i as the estimated parent node of a child node x_j. If x_j has an associated image of G pixels $\{p_i\}_{i=1,\ldots,G}$ arranged on a rectangular grid around its landmarks, we construct the parent node image as follows: For each pixel p'_i in the parent image, we compute its color as a weighted average of the colors of nearby pixels in the child image after applying the transformation $\phi(p'_i)$. This process propagates iteratively along the branches of the tree, ultimately yielding the reconstructed image at the estimated root shape and visualizing the evolutionary trajectory through the tree.

In terms of implementations, DICAROS builds upon the numerical differential geometry algorithms contained in JaxGeometry (Kühnel (2017, 2019)). These algorithms are implemented using JAX (Bradbury (2018)), a library that enables highly optimized and efficient numerical computations. The phylogenetic components of our method, including the crucial "Upwards Pass" algorithm, which transverse the tree from leaves to root, are implemented in Hyperiax `https://github.com/ComputationalEvolutionaryMorphometry/hyperiax`. Hyperiax leverages JAX to efficiently process independent branches in the phylogenetic tree in parallel, significantly boosting computational performance.

Shape simulation and comparison of method

To evaluate DICAROS, we simulate shape evolution along phylogenetic trees using the Kunita Flow(Stroustrup (2025)). The simulated leaf shapes are then used to reconstruct the root shape, allowing us to compare DICAROS reconstruction ability, and furhter compare against established reconstruction methods. The Kunita flow provides key advantages over a Brownian Motion (BM) model, which would likely produce biologically unrealistic shapes(Diaz-Uriarte (1996)). Instead, the Kunita flow models a shape as a set of k landmarks $x_1, \ldots, x_k$ in a domain $D \subset \mathbb{R}^d$ (for $d = 2$ or 3), where landmark movements are spatially correlated based on their proximity. This correlation ensures that nearby landmarks move coherently and that shapes maintain biological plausibility as they evolve (Arnaudon (2019, 2021); Sommer (2021); Stroustrup (2025)). The spatial correlation is modeled using a scaled Gaussian kernel $k_{p_j}(x) = \alpha \cdot \exp^{-\frac{\|p_j - x\|^2}{2\sigma^2}}$, where σ controls the spatial extent of correlation (kernel width) and α determines the size of the process.

We benchmark DICAROS against two standard approaches in the field: (1) Ordinary least squares estimation using *gm.prcomp* and *shape.predictor* from Geomorph v4.0.6 (Adams (2013)) and (2) Maximum likelihood estimation of ancestral states under Brownian motion using *anc.ML* from Phytools v2.3-0 (Revell (2024); Schluter (1997)). For fair comparison, all reconstructed shapes are aligned to the original root via general Procrustes alignment (Rohlf (1990)) using *gpagen* from Geomorph (Adams (2013, 2024)).To evaluate reconstruction accuracy, we compare the reconstructed root shapes to the original root using normalized Procrustes distance. Specifically, we calculate the mean least squared error (MLSE) between corresponding landmarks using Equation 2:

$$\epsilon = \frac{\sum_{i=1}^{n} \sqrt{(\mathbf{X}_{i_{original}} - \mathbf{X}_{i_{reconstructed}})^2}}{n} \tag{2}$$

where ϵ is the error, $\mathbf{X}_i$ is the coordinate vector of landmark i, and n is the total number of landmarks. The relative performance between methods is quantified as $\varepsilon = \epsilon_{\textbf{establish method}} - \epsilon_{\textbf{DICAROS}}$.

Biological data

As a biological application of DICAROS, we analyzed Lepidoptera wings from the Papilionidae family. The phylogenetic tree was derived from (Kawahara (2023)) and matched with image data through the Global Biodiversity Information Facility (GBIF.org) (GBIF (2024)). To ensure data quality, we restricted our selection to museum collections of male specimens to account for sexual dimorphism. Images showing damaged wings or improper orientation were manually excluded.

The phylogenetic tree was pruned using the *drop.tip* function from Ape (Paradis (2019)) to remove taxa without shape data. We used *Segment Anything* (SAM) (Kirillov (2023)) with *Grounding Dino* (Liu (2023)) to automatically extract the butterfly from its background, using the prompt "Butterfly". Images were cropped to include only the butterfly. Cases where SAM failed to identify a butterfly or identified multiple specimens were excluded.

To differentiate between morphological features, hind-/fore-wings and thorax, six anatomical landmarks were annotated on the outline for each image in a similar setting as seen in (Chan (2022)). The annotation was done with the software *Supervisely*, by external consultants hired through *fiverr.com* The butterfly contour was extracted using *cv2.findContours* from the Python package *opencv* (Bradski (2008)). The contour was matched with the six anatomical landmarks, and for the regions covering the wing sections, 30 equidistance landmarks were equidistance resampled with interpolation on the wing contour The methodology is detailed in Supplementary Section 3.

Each specimen was manually categorized based on tail morphology into three categories: untailed, tailed, or long-tailed. Using DICAROS, a full ancestral reconstruction was executed for the entire phylogeny, and the morphological changes from the leaves to the root. Further, we use the covariance to visualize the entire evolutionary trajectory by projecting the covariances and shape into the phylomorphospace.

Results

The evaluation of DICAROS is based on the simulated leave shapes from applying Kunita Flow on different phylogenies with different parameters, root shapes, and sizes. Figure 1 shows the steps of this comparison. First, a root shape (here, a butterfly) is simulated on a phylogeny with four leaves. The leaf shapes are then used to reconstruct the root shape and finally compare the reconsturcted root the original root shape.

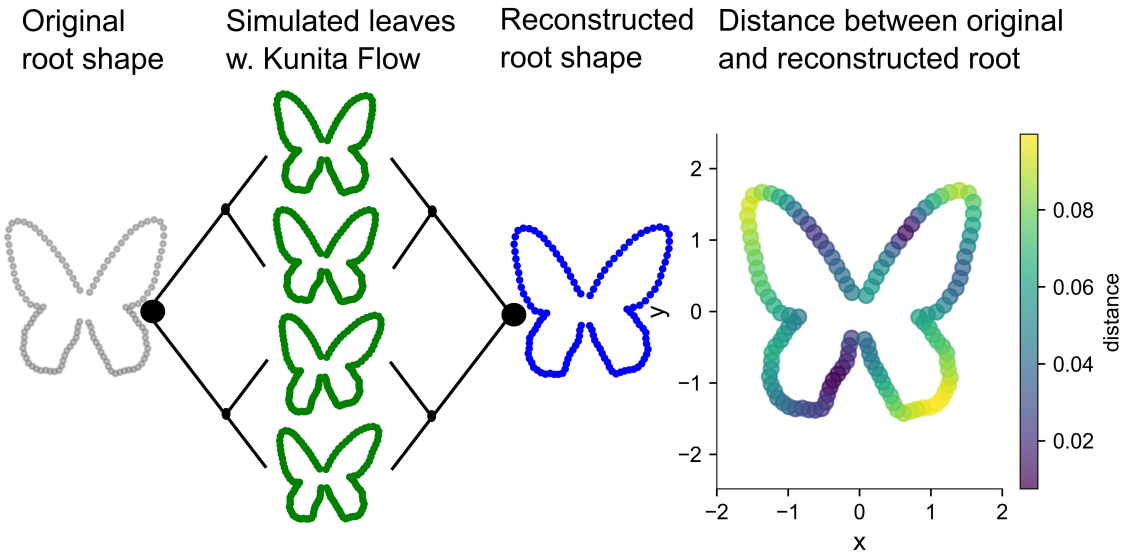

Fig. 1: The grey butterfly is used as the root shape for the Kunita Flow on a symmetric ultrametric phylogeny with four leaves. The root shape is simulated with Kunita Flow on the phylogeny to obtain the green leaf shapes. These leaf shapes are used for the reconstruction of the root shape using DICAROS on the same phylogeny. The reconstructed root shape in blue is then compared to the original root shape using the Euclidean distance between each landmark in the reconstructed and original root shape

To repeat this experiment, four different root shapes are used. Two calculated shapes (a circle and a sphere) and two real shapes (a butterfly and a bird beak), described in Table 1.

The robustness of DICAROS is tested by varying the $\sigma_{DICAROS}$ by a factor of 2 and 10, with $\sigma_{DICAROS}$ being the mean squared distance to the closest neighboring landmark. The test is executed on both symmetric and asymmetric ultrametric trees with a fixed total length. When the tree size is increased, the individual branch length decreases to

Table 1: The four shapes and their dimension (d) and number of landmarks (m), and the origin of the root shape

Shape	(d,m)	origin
Circle	(2,30)	Calculated
Sphere	(3,50)	Calculated
Butterfly	(2,118)	*Atrophaneura dixoni* (own data)
Birdbeak	(3,79)	*Fratercula arctica* (Cooney (2017))

represent a greater species diversification within the same period of time. The tree size vary by$2^k, (k = 2, .., 8)$ leaves and is repeated for all four shapes. For each configuration, the simulation is repeated 100 times. (Fig. 2)

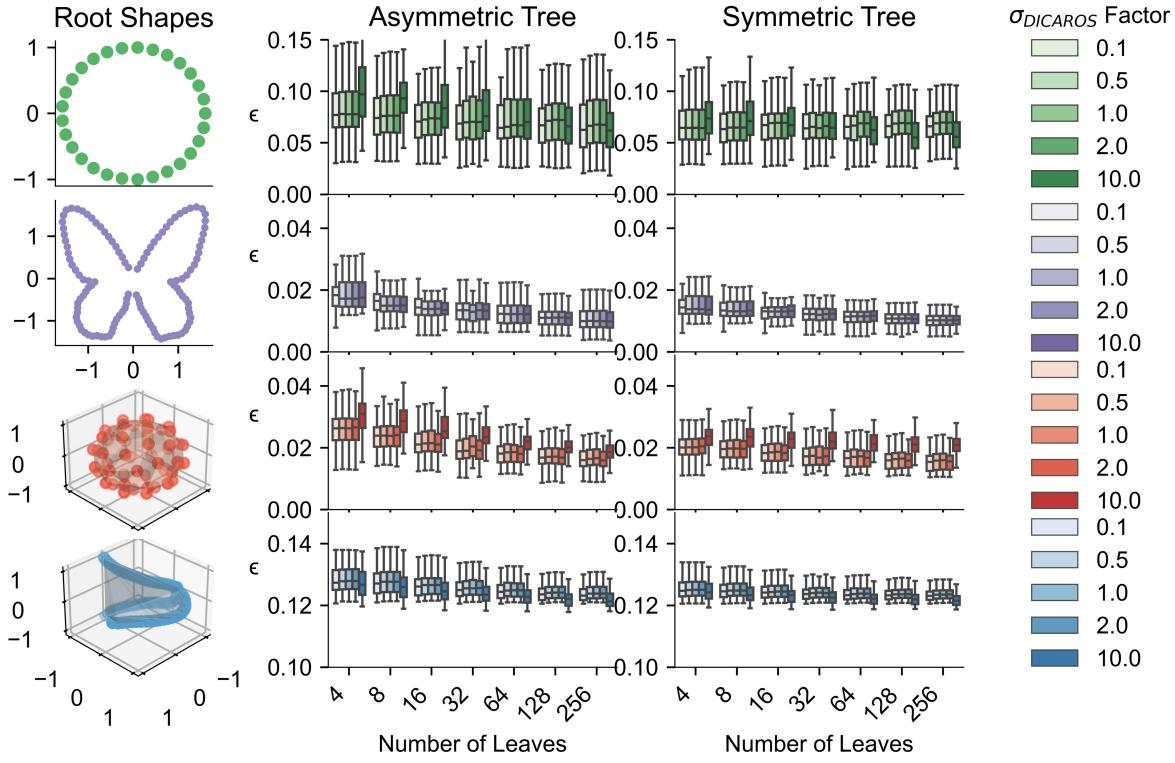

Fig. 2: The landmark configuration of each root shape is shown to the left, with the MSLE between the original and reconstructed root for both types of phylogenies, tree sizes, and scaled $\sigma_{DICAROS}$. The Y-axis is adjusted for each root shape, but the tendencies show a higher variance for the MSLE under an asymmetric phylogeny compared to the symmetric phylogeny

Figure 2 shows the robustness of the method for different $\sigma_{DICAROS}$, and the MSLE remains consistent between symmetric and asymmetric phylogeny across all 100

repetitions. We denote $\sigma_{DICAROS}$ as not to confuse it with the kernel width in the Kunita Flow, noted as σ_{KF}. Figure 2 demonstrates the robustness of the method regardless of tree topology. Based on these results, we fixed sigma at its base value without applying scaling factors for subsequent analyses.

Continuing, with a fixed $\sigma_{DICAROS}$ at the factor equal to one, the experiment is repeated for different parameters of the Kunita Flow. Here, the variance and noise parameter (σ_{KF}, α) is changed to illustrate less and larger shape changes across the phylogeny. For each simulation, the root is now reconstructed using DICAROS, OLS, and BM. For each evaluation, the relative procrustes distance $\epsilon = \epsilon_{\textbf{establish method}} - \epsilon_{\textbf{DICAROS}}$, is shown. Where a positive value on the y-axis, indicates a better reconstruction of the root shape by DICAROS, compared to the other method. The relative error between OLS and DICAROS is shown in Figure 3, where DICAROS are more accurate than OLS for asymmetric chronograms, but they are similar for the symmetric chronogram. Except a few outliers for the sphere where the OLS is not able to reconstruct the root as good as for the other simulation.

The relative error between methods shown in Figures 3 and 4 corresponds to $\approx 10\%$ of the total error magnitude, when comparing the base case ($\sigma_{DICAROS}$ factor = 1) in Figure 2, which shows absolute MSLE values for the circle is within [0.05;0.15], to the relative differences in Figures 3 and 4 within [-0.005;0.015]. We observe this consistent 10% relationship for all root shapes. This indicates that while DICAROS shows improvement over existing methods, the magnitude of improvement is moderate compared to the overall reconstruction error.

Diffeomorphic Independent Contrasts for wing shape analysis

The butterfly phylogeny with example images and shapes is shown in Figure 5. The entire dataset encompasses 992 individual butterflies distributed across the 49 Papilionidae species. The phylogenetic analysis contains three main morphological categories based on

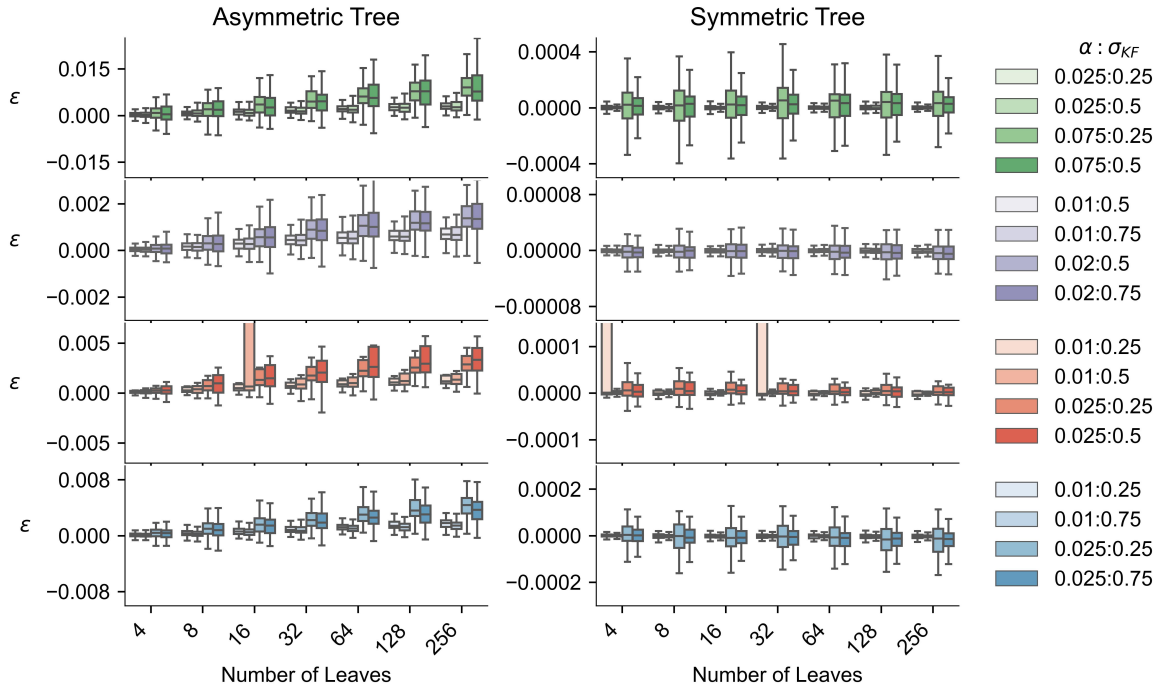

Fig. 3: The relative distance $\varepsilon = \epsilon_{\mathbf{OLS}} - \epsilon_{\mathbf{DICAROS}}$ is shown for four different root shapes (circle, butterfly, sphere, and bird beak) under varying Kunita Flow parameters. Positive values indicate that DICAROS outperforms OLS, which is seen for the asymmetric trees and increasing the tree sizes. For symmetric phylogenies (right column), both methods perform similarly with error distributions centered around zero. The boxplots show the distribution of relative errors across 100 simulations for each configuration

hindwing tail length: untailed, tailed, and long-tailed species. All mean shapes can be found in Supplementary Section 5 and origins of all images can be found in Supplementary Section 6. The entire dataset is shared in DataDryad;

`https://doi.org/10.5061/dryad.41ns1rnr7`

With the mean shape of each species, we can visualize the phylomorphospace of the butterfly wing shape. The phylomorphospace is shown in Figure 6, where it illustrates the transition from the untailed to the long-tailed species. The color of the leaves matches the color of the species in Figure 5, and the root is placed as a star. For each branch in the phylogeny, a grey line is drawn and selected leaf mean shapes is shown across the path of least resistance.

DICAROS reveals distinct evolutionary patterns in butterfly wing morphology

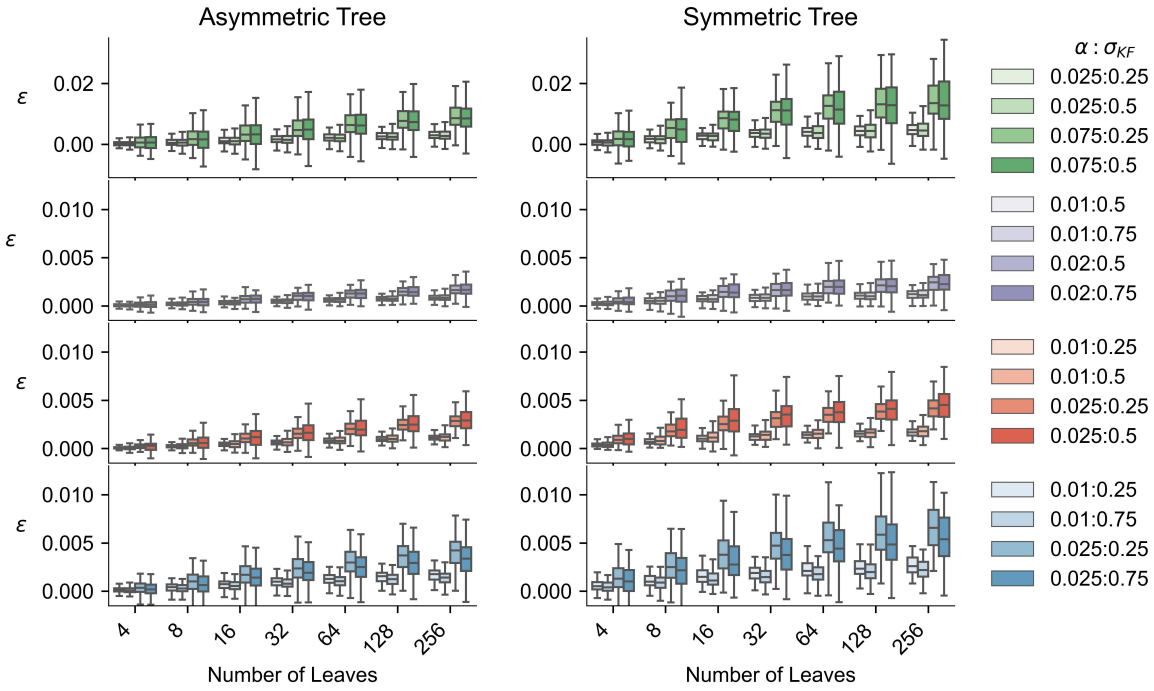

Fig. 4: Comparison of reconstruction accuracy between the Ancestral character estimation using maximum likelihood under Brownian Motion and DICAROS using relative MSLE ($\varepsilon = \epsilon_{\mathbf{BM}} - \epsilon_{\mathbf{DICAROS}}$). The consistently positive values across both symmetric and asymmetric phylogeny indicate that DICAROS achieves better reconstruction accuracy compared to this method. This pattern holds true across all tested shapes and parameter configurations, demonstrating the robust performance advantage of DICAROS

demonstrating a clear separation between long-tailed and untailed species. To visualize the evolutionary transitions, we applied the estimated deformation fields from DICAROS to transform modern specimen images along their phylogenetic branches back to the root. Figure 7 shows this process for eight individuals across different species. The full animation, going from the leaf shape and continuous to the root shape, is visible at the GitHub supplementary repository `https://github.com/MichaelSev/DICAROS`.

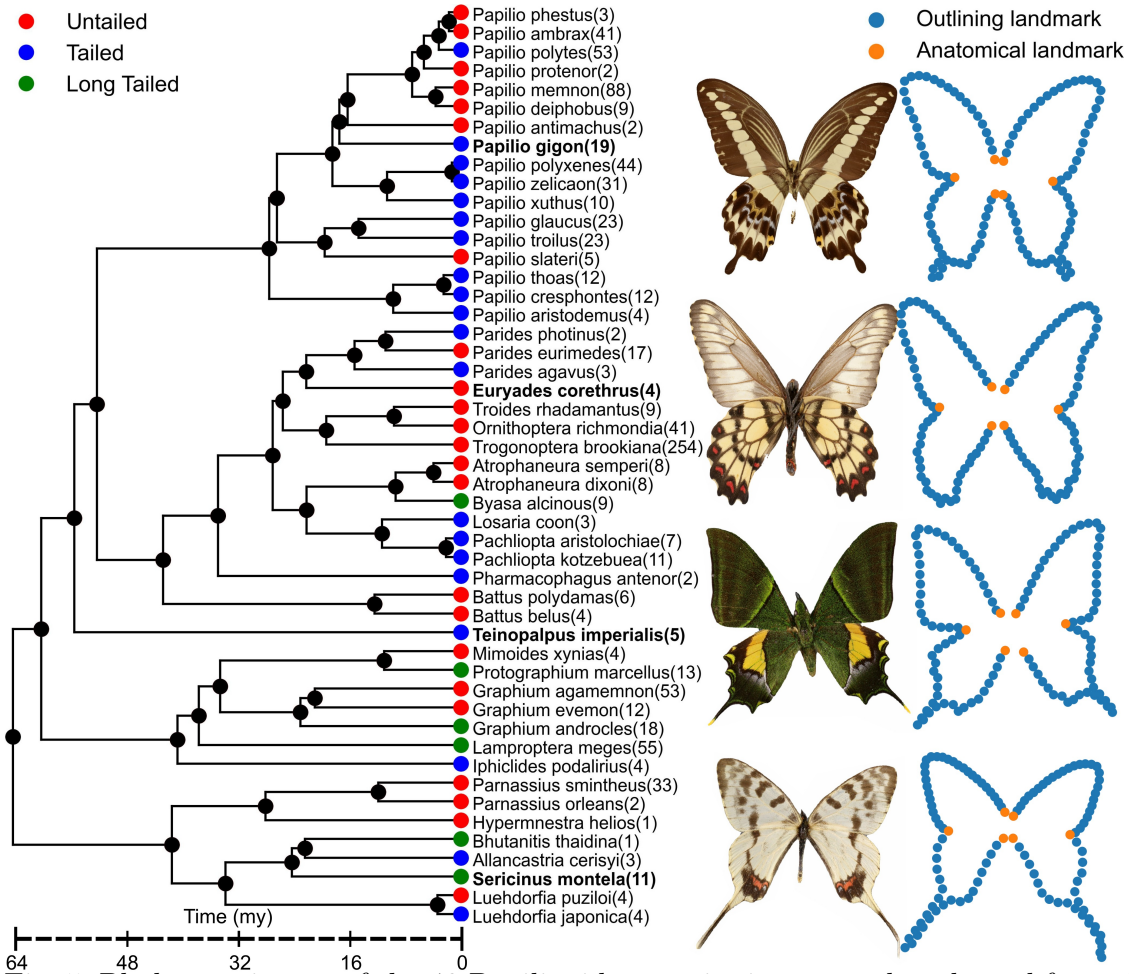

Fig. 5: Phylogenetic tree of the 49 Papilionidae species in our study, adapted from the larger Lepidoptera phylogeny from (Kawahara (2023)), with the number of specimens pr species in parenthesis. Species are categorized based on hindwing tail length: untailed, tailed, and long-tailed, with examples of four species specimens marked in bold *Papilio gigon*, *Euryades corethrus*, *Teinopalpus imperialis* and *Sericinus montela*.

Discussion

DICAROS represents an advancement in the field of ancestral shape reconstruction by demonstrating the novel use of LDDMM in a phylogenetic setting, utilizing tree topology as a map to reconstruct ancestral shapes. A key feature of DICAROS is its ability to incorporate correlations between landmarks, allowing for reconstruction to be applied to entire images through image manipulation (Fig. 7). Additionally, if the landmark data are only aligned using translation and rotation, the ancestral size can also be reconstructed.

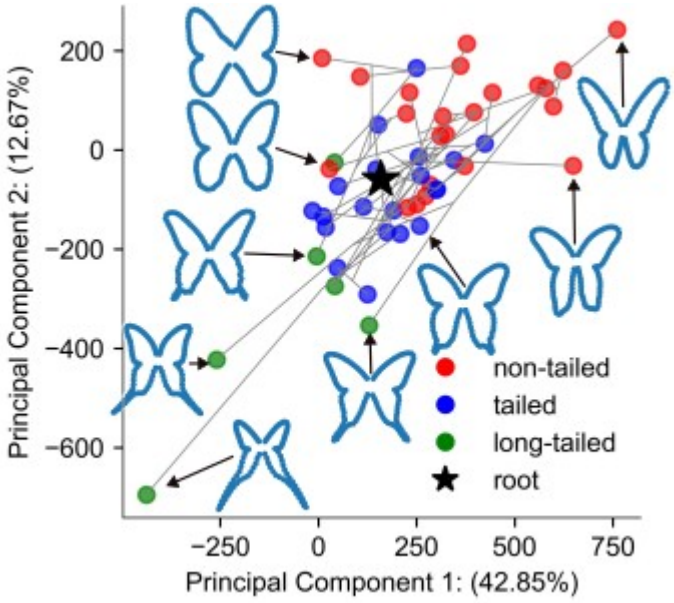

Fig. 6: The DICAROS Phylomorphospace of butterfly shapes across the Papilionidae phylogeny. The plot shows the distribution of species in the morphological space, with the ancestral root state marked as a star. Grey lines represent phylogenetic branches connecting species, and colors correspond to tail morphology categories as defined in Figure 5. The visualization reveals the transition of shape between long-tailed and untailed species

Using the methodology from Independent Contrasts(Felsenstein (1985)), it estimates the evolutionary covariance matrices to project shape data of both leaves and reconstructed inner node/root into phylomorphospace. The DICAROS implementation utilizes modern frameworks including *JAX* (Bradbury (2018)), *Jaxgeometry* (Kühnel (2017, 2019)), and *Hyperiax*, providing good performance and stability across different dataset sizes and shape complexities. The application of DICAROS to butterfly wings, combined with semi-automated segmentation and landmarking tools (Kirillov (2023)), not only provided insight into the tail evolution/devolution of the swallowtail but also demonstrates the potential for creating larger biological geometric morphometric datasets with greater time efficacy compared to traditional human labor approaches.

DICAROS naturally operates under a design constraint that can only be used with a rooted phylogeny. In contrast, other methods can illustrate the morphospace with and without phylogeny(Baken (2021)). Furthermore, the implementation of DICAROS only

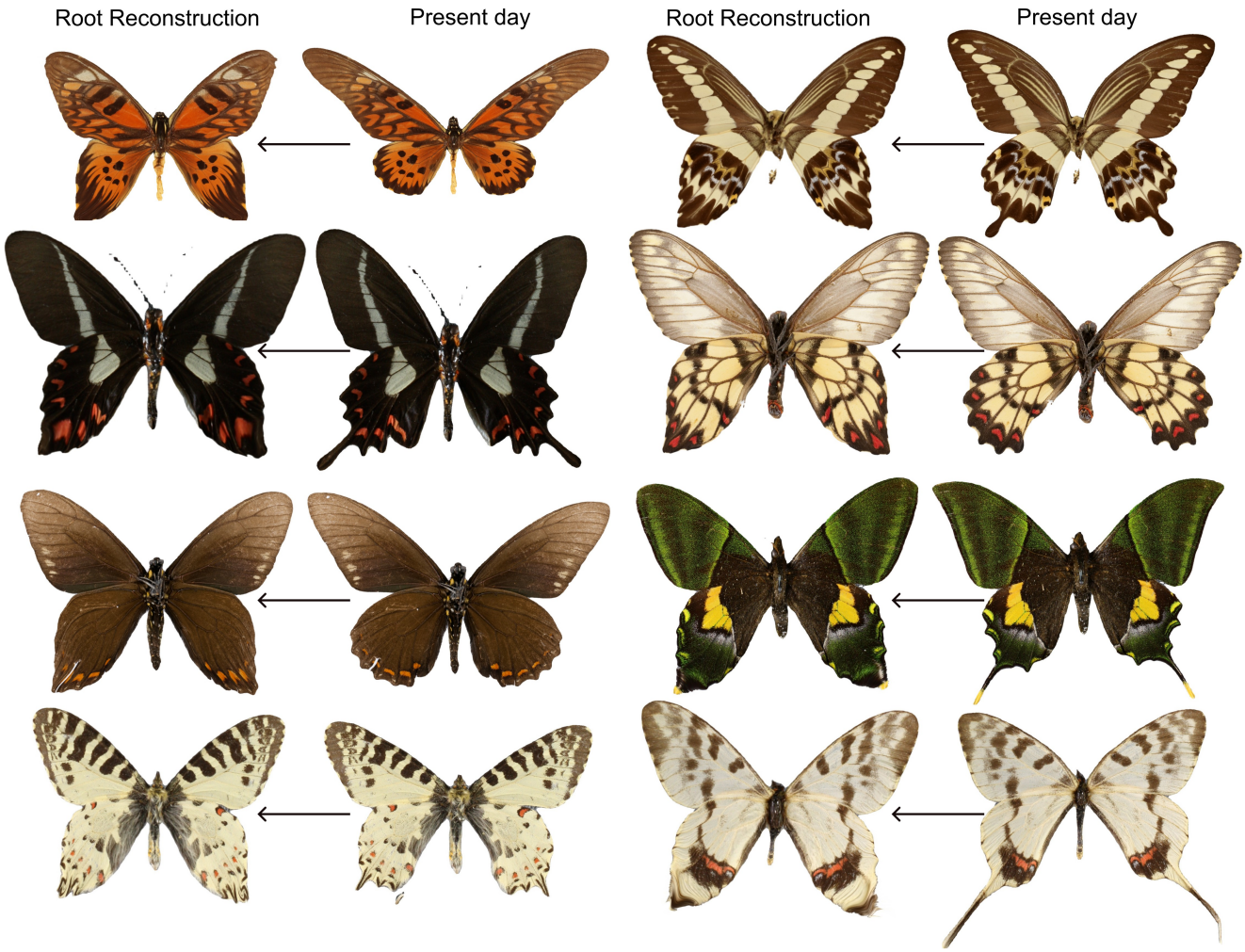

Fig. 7: The application of DICAROS on the image for the eight individuals across different species. With the right column corresponding to the individuals from Figure 5, *Papilio gigon*, *Euryades corethrus*, *Teinopalpus imperialis* and *Sericinus montela* and the right column corresponding to these species in bold; *Papilio antimachus*, *Parides agavus*, *Battus belus* and *Allancastria cerisyi*. The origin of the species is detailed in supplementary materials. A full transformation for the leaf species to the root species can be found at `https://github.com/MichaelSev/DICAROS`

uses bifurcating nodes, but the method can be generalized to include multifurcating nodes. For the bifurcation nodes, the leaves can also be represented from a fossil or prehistoric species, these tips can be inserted in the already existing inner nodes to make it the socalled total evidence phylogeny. This will likely make the inner nodes more exact in their reconstruction. Using the iteration of local computations that form the IC algorithm it is a

computational advantage instead of simply evaluating the closed-form MLE estimates. When dealing with large phylogenies, DICAROS avoids the inversion of large matrices, as seen for the MLE estimates. In the case of manifold valued nodes, as assumed by DICAROS, the iteration of local computations is an advantage since the local computation of inner nodes (step *5* in Algorithm 1) can be solved by a single evaluation of the exponential map without the need of larger optimization schemes that would otherwise be necessary for estimating the root directly from the leaves. The latter is done in Tangent Phylogenetic PCA Akhøj (2023)), which formulates the root estimate as a weighted Frechét mean (i.e. a mean on a Riemannian manifold), the computation which demands the the same number of exponential map evaluations as the number of leaf nodes at every iteration until convergence. This number of exponantial map evaluations well exceed that of DICAROS.

To evaluate the performance of DICAROS relative to existing methods, we conducted four simulation experiments (Fig. 1. First, we simulated shapes evolving along phylogenetic trees using the Kunita flow. Secondly, these simulated leaf shapes were used to reconstruct the root shape using three methods: DICAROS, Ordinary least squares(Adams (2013, 2024)), and Brownian Motion Maximum Likelihood(Revell (2024)). Third, we compared the reconstructed root shapes to the true ancestral shape to calculate reconstruction error (Eq. 2), and finally repeated this process under different topologies, tree sizes, and root shapes to assess method performance across varying conditions (Fig 3,4). We show that DICAROS perform slightly better in ancestral shape reconstruction under an asymmetric chronogram and for increasing tree sizes (Fig. 3,4). The results were close to identical for a symmetric chronogram (Fig. 3 relative to the OLS from Geomorph and DICAROS are more accurate in shape compared to the BM(Fig 4). This may be related to the reconstruction under a symmetric tree roughly corresponds to the equal weighted average among species. Whereas for the asymmetric tree, the morphological evolution and tree topology is crucial, because its determine how the evidence from

different branches are weighted in the reconstruction process.

While our simulations provide insights into comparing DICAROS with existing methods by the ancestral root shape, it is not possible to directly evaluate evolutionary covariance estimation, as the Kunita flow used does not have a parameter directly corresponding to the estimated evolutionary covariance matrix (Equation 1). Furthermore The Kunita Flow arrives from a stochastic process tailored for shapes rather than a BM model. When we evaluate our method based on simulations, we note that neither the standard IC nor the DICAROS estimates correspond to maximum likelihood estimates under this stochastic model, and therefore, is it fair to compare DICAROS and the other methods with these simulated shapes.

To evaluate the sensitivity of DICAROS to the kernel size, we tested different values of σ_{DICAR} by varying the default value (average distance of neighboring landmarks) by factors of 2 and 10. Note that this σ_{DICAR} does not correspond to the kernel size used in the simulations (σ_{KF}) . The method showed robust performance across a range of kernel sizes (Fig. 2). With very small kernel sizes approximating Euclidean space, performance remained similar to the default value, demonstrating that the combination of LDDMM with phylogenetic shape estimation provides better ancestral reconstruction compared to OLS and BM models. However, performance degraded with excessively large kernel sizes, as these caused all landmarks to move uniformly rather than capturing local shape differences. This occurs because large kernels encompass most or all landmarks simultaneously, preventing the LDDMM method from modeling independent landmark movements needed for accurate shape reconstruction.

Butterflies are often used as model organisms in fields of biology, such as conservation biology (Collins (1985); Sánchez-Bayo (2019)), phylogenetic inference (Caterino (1999); Kawahara (2023)) and studies of independent evolution between fore- and hind-wings (Chai (1990); Chotard (2022)) and how to do high throughput imaging of museums specimens (Chan (2022)). Here we focus on the swallow-tailed family

Papilonidae, which has long fascinated researchers due to their many speciation events and mimicry abilities (Mallet (1998); Carvalho (2024)), morphs (Le Roy (2019)), and sexual dimorphism (Condamine (2012)). The tails appear independently throughout the phylogeny and are not restrained to a monophyletic group. The tail can most often be on the fourth vein, but this also differs in placement, count, shape, and occurrence between species within and across genera and sexes (Nakae (2021)). Here, we aim to reconstruct the trajectory of shape evolution for the swallowtail family using DICAROS and illuminate that change using modern species images.

The biological dataset comprises wing shape data from 49 species of Papilionidae butterflies (Fig. 5), representing $\approx 20\%$ of the recognized species in this family (Condamine (2023)), and $\approx 68\%$ of the Papilionidae family from Kawahara et.al. 2023 (Kawahara (2023)). After sorting the obtained images through GBIF, we obtained a total of 992 specimens on to base our mean shape. While our sample size for some species is limited, it provides a foundation for exploring shape evolution within this group while showcasing the power of the DICAROS method. The application of semi-automated landmark placement on butterfly wing outlines utilized state-of-the-art segmentation tools (Kirillov (2023))(Fig. 5), together with using the contour instead of vein points with none/less homology across the family allowed us to showcase a more comprehensive representation of Papilionidae with a high number of individuals. For visualization purposes in the phylomorphospace, we categorized hindwing types into three groups: untailed, tailed, and long-tailed. While this classification simplifies the presentation in Figure 6, it is important to note that tail length exists on a continuous spectrum rather than as discrete categories. For this purpose the "tailed" group represents an intermediate state that shows overlap with both untailed and long-tailed morphologies. This categorization was used solely for visualization and was not incorporated into any quantitative analyses.

The phylomorphospace analysis (Fig. 6) revealed distinct patterns in the distribution of butterfly wing shapes across Papilionidae. It has a clear separation between

long-tailed and untailed species, with a gradual transition of tail length in between. This indicates that the presence of tails is a factor in shape variation with small tail length representing intermediate evolutionary stages towards the long tail. With the root shape represented as tailed and is within the tailed group. This conforms with (Chotard (2022)), who employed a binary classification of tailed/untailed species with landmarks placed manually on each vein. The image reconstruction (Fig. 7) provides a visualization of shape evolution across evolutionary time. While this visual reconstruction effectively captures changes in wing morphology, it is important to note its limitations. The reconstruction is based solely on landmark shape data and does not account for other aspects of wing morphology, such as pigmentation patterns or wing-scale structures. The color scheme in the deformations is arbitrary and does not reflect actual evolutionary changes in wing coloration, but only from the leaf image. Despite these constraints, this approach represents a significant advancement in the quantitative analysis and visualization of morphological evolution in Lepidoptera. The reconstruction supports the hypothesis that the ancestral state of Papilionidae was likely tailed.

Conclusion

In conclusion, this study introduces DICAROS as a novel method using LDDMM for ancestral shape reconstruction in phylogenetic contexts. The method offers significant advantages, including the incorporation of correlations between landmarks. While computationally more intensive than traditional approaches, DICAROS leverages modern tools like JAX, Jaxgeometry, and Hyperiax to enhance performance, mainly through GPU acceleration. As such, it represents an advancement in the field of geometric morphometrics evolution studies.

Our validation through simulations demonstrated that DICAROS obtained similar results and outperformed established methods in the case of asymmetric trees(Fig 3,4). The method's accuracy improved with increasing numbers of leaf nodes, highlighting its

potential for large-scale analyses. These simulations provide a solid foundation for the method's reliability in reconstructing ancestral shapes while preserving both form and shape.

Our application of DICAROS to Papilionidae butterfly wing shapes supported insights into their evolution. We hypothesize that the evolution of modern non-tailed species originates from a tailed ancestor and that similar structures in phylomorphospace (Fig. 6) are observed for independence of hindwing and forewing shape evolution. The novel implementation of image reconstruction by DICAROS offers a new tool for studying any shape evolution on a phylogeny, which will contribute to our understanding of shape evolution across time.

Acknowledgments

We express our gratitude to GBIF.org for its extensive work in assembling information across diverse organizations and data structures. We also acknowledge the valuable contributions of organizations, curators, taxonomists, and authors across museums who classified and photographed the butterflies. Our work has been enhanced by the tools developed by Google (JAX) and Meta (Segment Anything), which greatly improved our computational capabilities. We are also grateful to the authors of Geomorph for developing their R package and for their openness and fast responses. We also express our gratitude for software engineers and staff at AI companies, who build and designed; auto completions, spelling checkers and text generators for speeding up both coding and writing.

Funding

This work was supported by a research grant (VIL40582) from VILLUM FONDEN and Danmarks Frie Forskningsfond (Independent Research Fund Denmark) grant ID 10.46540/3103-00296B.

Data Availability

Code for figures and both simulated and biological dataset are shared in the data dryid DOI; `https://doi.org/10.5061/dryad.41ns1rnr7`. Further illustration and biological examples is available from Github repository; `https://github.com/MichaelSev/DICAROS`. For Hyperiax, we advertise for `https://github.com/ComputationalEvolutionaryMorphometry/hyperiax`, to follow the most recent updates.

References

Adams, D. C., & Otárola-Castillo, E. (2013). geomorph: an R package for the collection and analysis of geometric morphometric shape data. Methods in ecology and evolution, 4(4), 393-399.

Adams, D.C., 2016. Evaluating modularity in morphometric data: challenges with the RV coefficient and a new test measure. Methods in Ecology and Evolution, 7(5), pp.565-572.

Adams, D.C., Collyer, M., Kaliontzopoulou, A. and Sherratt, E., 2016. Geomorph: Software for geometric morphometric analyses.

Akhøj, M., Pennec, X., and Sommer, S. 2023. Tangent phylogenetic PCA. In Scandinavian Conference on Image Analysis (pp. 77–90).

Arnaudon, A., Holm, D.D. and Sommer, S., 2019. A geometric framework for stochastic shape analysis. Foundations of Computational Mathematics, 19, pp.653-701.

Arnaudon,A., Holm D.D. and Sommer, S., 2021. "Stochastic flows and shape bridges." In Statistics of Stochastic Differential Equations on Manifolds and Stratified Spaces - number 48 in Oberwolfach Reports 48, pp. 18–21

E. K. Baken, M. L. Collyer, A. Kaliontzopoulou, and D. C. Adams 2021. geomorph v4.0 and gmShiny: enhanced analytics and a new graphical interface for a comprehensive morphometric experience. Methods in Ecology and Evolution, 12, p.2355-2363.

Bardua, C., Felice, R., Watanabe, A., Fabre, A.C., and Goswami, A. 2019. A practical guide to sliding and surface semilandmarks in morphometric analyses. Integrative Organismal Biology, 1(1), p.obz016.

Beg, M., Miller, M., Trouvé, A., and Younes, L. 2005. Computing large deformation metric mappings via geodesic flows of diffeomorphisms. International journal of computer vision, 61, p.139–157.

467 bookstein, F. 1997. . Morphometric tools for landmark data

468 Bradbury, J., Frostig, R., Hawkins, P., Johnson, M.J., Leary, C., Maclaurin, D., Necula,
469 G., Paszke, A., VanderPlas, J., Wanderman-Milne, S. and Zhang, Q., 2018. JAX:
470 composable transformations of Python+ NumPy programs.

471 Bradski, G. and Kaehler, A., 2008. Learning OpenCV: Computer vision with the OpenCV
472 library. " O'Reilly Media, Inc.".

473 Carvalho, A., Owens, H., St Laurent, R., Earl, C., Dexter, K., Messcher, R., Willmott, K.,
474 Aduse-Poku, K., Collins, S., Homziak, N., and others 2024. Comprehensive phylogeny of
475 Pieridae butterflies reveals strong correlation between diversification and temperature.
476 Iscience, 27(4).

477 Caterino, M., and Sperling, F. 1999. PapilioPhylogeny based on mitochondrial cytochrome
478 oxidase I and II genes. Molecular Phylogenetics and Evolution, 11(1), p.122–137.

479 Chai, P., and Srygley, R. 1990. Predation and the flight, morphology, and temperature of
480 neotropical rain-forest butterflies. The American Naturalist, 135(6), p.748–765.

481 Chan, W.P., Rabideau Childers, R., Ashe, S., Tsai, C.C., Elson, C., Keleher, K.J., Sipe,
482 R.L.H., Maier, C.A., Sourakov, A., Gall, L.F. and Bernard, G.D., 2022. A
483 high-throughput multispectral imaging system for museum specimens. Communications
484 Biology, 5(1), p.1318.

485 Chazot, N., Panara, S., Zilbermann, N., Blandin, P., Le Poul, Y., Cornette, R., Elias, M.,
486 and Debat, V. 2016. Morpho morphometrics: shared ancestry and selection drive the
487 evolution of wing size and shape in Morpho butterflies. Evolution, 70(1), p.181–194.

488 Chiappe, L.M., Navalón, G., Martinelli, A.G., Carvalho, I.D.S., Miloni Santucci, R., Wu,
489 Y.H. and Field, D.J., 2024. Cretaceous bird from Brazil informs the evolution of the
490 avian skull and brain. Nature, 635(8038), pp.376-381.

491 Chotard, A., Ledamoisel, J., Decamps, T., Herrel, A., Chaine, A., Llaurens, V., and Debat,
492 V. 2022. Evidence of attack deflection suggests adaptive evolution of wing tails in
493 butterflies. Proceedings of the Royal Society B, 289(1975), p.20220562.

494 Collins, N., and Morris, M. 1985. . Threatened swallowtail butterflies of the world: the
495 IUCN Red Data Book. Iucn.

496 Condamine, F., Sperling, F., Wahlberg, N., Rasplus, J.Y., and Kergoat, G. 2012. What
497 causes latitudinal gradients in species diversity? Evolutionary processes and ecological
498 constraints on swallowtail biodiversity. Ecology letters, 15(3), p.267–277.

499 Condamine, F., Allio, R., Reboud, E., Dupuis, J., Toussaint, E., Mazet, N., Hu, S.J., Lewis,
500 D., Kunte, K., Cotton, A., and others 2023. A comprehensive phylogeny and revised
501 taxonomy illuminate the origin and diversification of the global radiation of Papilio
502 (Lepidoptera: Papilionidae). Molecular Phylogenetics and Evolution, 183, p.107758.

503 Cooney, C.R., Bright, J.A., Capp, E.J., Chira, A.M., Hughes, E.C., Moody, C.J., Nouri,
504 L.O., Varley, Z.K. and Thomas, G.H., 2017. Mega-evolutionary dynamics of the adaptive
505 radiation of birds. Nature, 542(7641), pp.344-347.

506 Diaz-Uriarte, R., and Garland Jr, T. (1996). Testing hypotheses of correlated evolution
507 using phylogenetically independent contrasts: sensitivity to deviations from Brownian
508 motion. Systematic Biology, 45(1), 27-47.

509 Durrleman, S., Prastawa, M., Charon, N., Korenberg, J.R., Joshi, S., Gerig, G. and
510 Trouvé, A., 2014. Morphometry of anatomical shape complexes with dense deformations
511 and sparse parameters. NeuroImage, 101, pp.35-49.

512 Felsenstein, J. 1985. Phylogenies and the comparative method. The American Naturalist,
513 125(1), p.1–15.

514 GBIF.org (03 October 2024) GBIF Occurrence Download
515 https://doi.org/10.15468/dl.tseaw8

Garland Jr, T., Midford, P.E. and Ives, A.R., 1999. An introduction to phylogenetically based statistical methods, with a new method for confidence intervals on ancestral values. American Zoologist, 39(2), pp.374-388.

Goswami, A., Noirault, E., Coombs, E., Clavel, J., Fabre, A.C., Halliday, T., Churchill, M., Curtis, A., Watanabe, A., Simmons, N., and others 2023. Developmental origin underlies evolutionary rate variation across the placental skull. Philosophical Transactions of the Royal Society B, 378(1880), p.20220083.

Gower, J.C., 1975. Generalized procrustes analysis. Psychometrika, 40, pp.33-51.

Gunz, P., and Mitteroecker, P. (2013). Semilandmarks: a method for quantifying curves and surfaces.

Hallgrímsson, B., Jamniczky, H., Young, N.M., Rolian, C., Parsons, T.E., Boughner, J.C. and Marcucio, R.S., 2009. Deciphering the palimpsest: studying the relationship between morphological integration and phenotypic covariation. Evolutionary biology, 36, pp.355-376.

Höhna, S., Landis, M., Heath, T., Boussau, B., Lartillot, N., Moore, B., Huelsenbeck, J., and Ronquist, F. 2016. RevBayes: Bayesian phylogenetic inference using graphical models and an interactive model-specification language. Systematic biology, 65(4), p.726–736.

Inaturalist.org. (2024). Inaturalist Home Page. Available from: https://www.Inaturalist.org.

Kawahara, A., Storer, C., Carvalho, A., Plotkin, D., Condamine, F., Braga, M., Ellis, E., St Laurent, R., Li, X., Barve, V., and others 2023. A global phylogeny of butterflies reveals their evolutionary history, ancestral hosts and biogeographic origins. Nature Ecology & Evolution, p.1–11.

Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., Xiao, T., Whitehead, S., Berg, A., Lo, W.Y., and others 2023. Segment anything. arXiv preprint arXiv:2304.02643.

Klingenberg, C.P., 2008. Morphological integration and developmental modularity. Annual review of ecology, evolution, and systematics, 39(1), pp.115-132.

Klingenberg, C.P., 2020. Walking on Kendall's shape space: understanding shape spaces and their coordinate systems. Evolutionary Biology, 47(4), pp.334-352.

Kühnel, L., and Sommer, S. 2017. Computational anatomy in Theano. In Graphs in Biomedical Image Analysis, Computational Anatomy and Imaging Genetics: First International Workshop, GRAIL 2017, 6th International Workshop, MFCA 2017, and Third International Workshop, MICGen 2017, Held in Conjunction with MICCAI 2017, Québec City, QC, Canada, September 10–14, 2017, 1 (pp. 164–176).

Kühnel, L., Sommer, S., and Arnaudon, A. 2019. Differential geometry and stochastic dynamics with deep learning numerics. Applied Mathematics and Computation, 356, p.411–437.

Le Roy, C., Debat, V., and Llaurens, V. 2019. Adaptive evolution of butterfly wing shape: from morphology to behaviour. Biological Reviews, 94(4), p.1261–1281.

Liu, S., Zeng, Z., Ren, T., Li, F., Zhang, H., Yang, J., Li, C., Yang, J., Su, H., Zhu, J., and others 2023. Grounding dino: Marrying dino with grounded pre-training for open-set object detection. arXiv preprint arXiv:2303.05499.

Joron, M., and Mallet, J. 1998. Diversity in mimicry: paradox or paradigm?. Trends in ecology & evolution, 13(11), p.461–466.

Michor, P.W., 2020. Manifolds of mappings for continuum mechanics. Geometric Continuum Mechanics, pp.3-75.

564 Miller, M.I., Trouvé, A. and Younes, L., 2015. Hamiltonian systems and optimal control in
565 computational anatomy: 100 years since D'Arcy Thompson. Annual review of biomedical
566 engineering, 17(1), pp.447-509. Miller, M.I., Trouvé, A. and Younes, L., 2002. On the
567 metrics and Euler-Lagrange equations of computational anatomy. Annual review of
568 biomedical engineering, 4(1), pp.375-405.

569

570 Misof, B., Liu, S., Meusemann, K., Peters, R., Donath, A., Mayer, C., Frandsen, P., Ware,
571 J., Flouri, T., Beutel, R., and others 2014. Phylogenomics resolves the timing and
572 pattern of insect evolution. Science, 346(6210), p.763–767.

573 Mitteroecker, P. and Bookstein, F., 2007. The conceptual and statistical relationship
574 between modularity and morphological integration. Systematic biology, 56(5),
575 pp.818-836.

576 Mitterocker, P. 2021. Morphometrics in evolutionary developmental biology. Journal is
577 required!, p.941–951.

578 morposource.org. (2024). morposource Home Page.

579 Nakae, M., 2021. Illustrated guide to swallowtail butterflies around the world: Papilionidae
580 of the world. Roppon-Ashi Entomological Books.

581 Nelson, G., and Ellis, S. 2019. The history and impact of digitization and digital data
582 mobilization on biodiversity research. Philosophical Transactions of the Royal Society B,
583 374(1763), p.20170391.

584 Owens, H., Lewis, D., Condamine, F., Kawahara, A., and Guralnick, R. 2020. Comparative
585 phylogenetics of Papilio butterfly wing shape and size demonstrates independent
586 hindwing and forewing evolution. Systematic biology, 69(5), p.813–819.

587 Paradis, E., and Schliep, K. 2019. ape 5.0: an environment for modern phylogenetics and
588 evolutionary analyses in R. Bioinformatics, 35(3), p.526–528.

589 Pennec, X., Sommer, S. and Fletcher, T. eds., 2019. Riemannian geometric statistics in
590 medical image analysis. Academic Press.

591 Polly, P., Lawing, A., Fabre, A.C., and Goswami, A. 2013. Phylogenetic principal
592 components analysis and geometric morphometrics. Hystrix, 24(1), p.33.

593 Revell, L. 2009. Size-correction and principal components for interspecific comparative
594 studies. Evolution, 63(12), p.3258–3268.

595 Revell, L. 2024. phytools 2.0: an updated R ecosystem for phylogenetic comparative
596 methods (and other things). PeerJ, 12, p.e16505.

597 Rohlf, F.J. and Slice, D., 1990. Extensions of the Procrustes method for the optimal
598 superimposition of landmarks. Systematic zoology, 39(1), pp.40-59.

599 Schluter, D., Price, T., Mooers, A., and Ludwig, D. 1997. Likelihood of ancestor states in
600 adaptive radiation. Evolution, 51(6), p.1699–1711.

601 Sommer, S. Schauer, M., and van der Meulen, Frank. 2021. Stochastic flows and shape
602 bridges. Statistics of Stochastic Differential Equations on Manifolds and Stratified
603 Spaces (hybrid meeting). Mathematisches Forschungsinstitut Oberwolfach.
604 (48)Oberwolfach Reports, p.18–21.

605 Stiller, J., Feng, S., Chowdhury, A.A., Rivas-González, I., Duchene, D., Fang, Q., Deng,
606 Y., Kozlov, A., Stamatakis, A., Claramunt, S., and others 2024. Complexity of avian
607 evolution revealed by family-level genomes. Nature, 629(8013), p.851–860.

608 Stroustrup, Sofia, Akhøj Pedersen, Morten, van der Meulen, Frank, Sommer, Stefan and
609 Nielsen, Rasmus. Stochastic phylogenetic models of shape, In prep

610 Sánchez-Bayo, F., and Wyckhuys, K. 2019. Worldwide decline of the entomofauna: A
611 review of its drivers. Biological conservation, 232, p.8–27.

612 Younes, L., 2010. Shapes and diffeomorphisms (Vol. 171, pp. xviii+-434). Berlin: Springer.

Young, N.M., Chong, H.J., Hu, D., Hallgrímsson, B. and Marcucio, R.S., 2010. Quantitative analyses link modulation of sonic hedgehog signaling to continuous variation in facial growth and shape. Development, 137(20), pp.3405-3409.

Zhang, Chao, Rasmus Nielsen, and Siavash Mirarab. "CASTER: Direct species tree inference from whole-genome alignments." Science (2025): eadk9688.

2024. Exciting times for evolutionary biology. Nature Ecology and Evolution,593-594,8.