

Why 17-Landmark Geometric Morphometrics Fails D’Arcy Thompson’s Fish Transformations — and How Direct 600-Semilandmark Outline Fitting Recovers Them

A Quantitative Test of *On Growth and Form* (1917, Figs. 373–382) Across All Six Original Species Pairs, 791 Indo-Pacific Reef Fishes, 4,334 Time-Calibrated Divergences, and Pre-Registered Blind Tests on 10,683 Out-of-Sample Teleosts

Joel Dietz

Project Delphi · <https://github.com/fractastical/morpholib>

October 22, 2026

Abstract

For over a century, D’Arcy Wentworth Thompson’s coordinate transformation grids (*On Growth and Form*, 1917, Ch. XVII, pp. 748–751, Figs. 373–382) have stood as the founding visual metaphor of mathematical biology. Since Fred Bookstein’s introduction of thin-plate splines (TPS, 1989) and the standardization of sparse landmark protocols in ichthyology (Claverie & Wainwright 2014; Friedman et al. 2019), evolutionary biologists have routinely equated fitting transformations to 17 homologous lateral landmarks with Thompson’s theory of transformations. Yet in two recent critiques, Bookstein (2023, 2024) argued that exact TPS interpolants overfit biological form and should be replaced by fitted low-dimensional parametric coordinate grids that retain explicit residuals.

Here we perform a comprehensive mathematical and empirical evaluation of all four of Thompson’s 1917 fish coordinate families — **I. Oblique shear** (5 parameters, Figs. 373–374), **II. Coaxial circles** (8 parameters, Figs. 375–376), **III. Fan of coordinates** (8 parameters, Figs. 377–380), and **IV. Hyperbolic/radial expansion** (8 parameters, Figs. 381–382) — alongside Procrustes similarity (4p), the uniform affine component (6p), quadratic polynomial trend surfaces (12p), and exact TPS (34p) across: (i) all six exact species pairs Thompson drew in 1917, digitized both as 17 homologous landmarks and as 600-semilandmark whole-body outlines, (ii) 791 Indo-Pacific shore-fish species (4,505 taxonomic pairs and 4,334 time-calibrated divergences on the Rabosky et al. 2018 megaphylogeny) replicated across 6,614 and 10,303 FishBase teleost species, and (iii) cryptographically pre-registered out-of-sample blind tests across 5,180 never-seen species in FishShapes v1 (Price et al. 2022) and 5,503 out-of-sample species in FishBase MORPHMET.

We report five primary discoveries:

1. **The 17-Landmark Skeletal Wireframe vs. 600-Semilandmark Outline Discrepancy:** In standard ichthyological landmark schemes (Claverie & Wainwright 2014), 9 of the 17 landmarks (53% of the ordinary least-squares weight) lie in the craniofacial partition (LM 1–2, 11–17), while the remaining 8 mark only where the fin bases insert into the body wall (LM 3–10) — omitting the entire distal contour of the dorsal fin, anal fin, caudal fin, and ventral body profile. Because craniofacial proportions and median-fin insertions evolve modularly across families (fin-insertion Procrustes RMS residuals are $1.48\times$ to $2.03\times$ larger than craniofacial residuals), fitting a coordinate grid to the 17 landmarks on interfamilial pairs such as *Scarus* $\rightarrow$ *Pomacanthus* or *Scorpaena* $\rightarrow$ *Antigonia* **actively degrades the real 600-semilandmark fish silhouette compared to no non-rigid deformation at all:** the symmetric non-overlapping area $|A\Delta B|$ worsens from 40.6% to 56.2% of target fish area on *Scarus* $\rightarrow$ *Pomacanthus* (IoU : 0.662 $\rightarrow$ 0.618) and from 102.2% to 143.7% on *Scorpaena* $\rightarrow$ *Antigonia* (IoU : 0.417 $\rightarrow$ 0.389).

2. **Direct 600-Semilandmark Outline Fitting Vindicates Thompson’s 1917 Transformations:** When the exact same 5- to 8-parameter Thompson coordinate families are fitted *directly* to the 600-semilandmark whole-body outlines (anchored at Bookstein’s two axial baseline endpoints, snout tip and caudal peduncle midpoint) rather than the 17 skeletal landmarks, Thompson’s transformations are recovered across every 1917 pair with 100% diffeomorphic bijectivity ($\det D\Phi > 0$): non-overlapping area $|A\Delta B|$ is cut in half on *Scarus* $\rightarrow$ *Pomacanthus* (56.2% $\rightarrow$ **28.5%**, IoU = **0.749**), cut 2.5 $\times$ on *Scorpaena* $\rightarrow$ *Antigonia* (143.7% $\rightarrow$ **57.6%**, IoU = **0.599**), reduced to **17.2%** on *Polyprion* $\rightarrow$ *Pristigenys* (IoU = **0.839**), reduced 3.4 $\times$ on *Pristigenys* $\rightarrow$ *Scorpaena* (120.1% $\rightarrow$ **35.5%**), reduced to **45.6%** on *Argyropelecus* $\rightarrow$ *Sternoptyx*, and reduced to **49.0%** on *Diodon* $\rightarrow$ *Mola mola* (IoU = **0.592**). Furthermore, when homologous perimeter fin-base and fin-apex landmarks are enforced (without the 9 internal skull landmarks), a 4-parameter orthotropic box stretch fails (2.4% on *Scarus* $\rightarrow$ *Pomacanthus* fin apices; 23.7% on *Argyropelecus* $\rightarrow$ *Sternoptyx*), and Thompson’s specific skewing and curvature warps jump by +11.0% to +48.0%.
3. **A $\sim$ 45–50 Myr Macroevolutionary Disparity Transition and ΔAIC_c Parsimony Optimum Across 791 Species:** Across 4,334 dated pairs on the Rabosky et al. (2018) tree, 8-parameter Thompson grids pull median 17-landmark Procrustes RMS residuals down inside the intrageneric disparity band ($\sigma_{\text{genus}} \leq 0.0110$), reducing pair residuals below $\sigma_{\text{genus}}(T)$ in **53.0%** of pairs diverged < 10 Myr, **61.3%** at 10–25 Myr, and **49.7%** at 25–50 Myr, before transitioning across ~ 45 –50 Myr ($t^* = 45.8$ Myr on $n = 187$ phylogenetically independent node-disjoint pairs) to **10.6%** (50–80 Myr) and **0.0%** (> 110 Myr) as median-fin insertions decouple from the cranium (Satterfield, Claverie & Wainwright 2024). By small-sample corrected AIC_c ($n = 34$), 8-parameter Thompson laws beat the 6-parameter uniform affine component in **82.2%** of intrafamily pairs and beat 12-parameter quadratic trend surfaces in **84.9%** of intrafamily and **92.1%** of interfamily pairs.
4. **Superiority Over Standard 600-Semilandmark Curve TPS/EFD, Phylomorphospace Convergence, & 10,303-Species Replication:** Holding the 600-semilandmark outline input fixed across all six 1917 pairs, standard 600-semilandmark curve TPS (Bookstein 1997) and 12p quadratic surfaces fold the 2D interior coordinate grid ($\det D\Phi \leq 0$ on 6/6 and 4/6 pairs) and degrade out-of-sample prediction of held-out interior skeletal landmarks by up to 2.4 $\times$, whereas 8p Thompson grids remain 100% diffeomorphic and accurately predict held-out internal anatomy from the outer outline alone. Across all 791 species (Leave-One-Family-Out CV across 48 families), Thompson’s 1917 equations capture **92.0%** (median) / **95.6%** (mean) of the optimal 8-parameter empirical Relative-Warp variance (Rohlf 1993) and **88.6%** of the primary (PC1, PC2) phylomorphospace plane (Sidlauskas 2008), while Phenotypic Trajectory Analysis (Adams & Collyer 2009) shows that distantly related clades (> 95 Myr) converging on deep-bodied or elongate/fusiform body plans follow strongly parallel 8p Thompson trajectories ($r_{\text{vec}} = +0.693$ and $+0.714$, vs. -0.699 between divergent clades) alongside divergent “many-to-one” craniofacial and fin-insertion residuals (Wainwright et al. 2005). Across three datasets (791 $\rightarrow$ 6,614 $\rightarrow$ 10,303 species; 440 families), separating Global 8p Thompson Morphospace from Local Modular Residuals aligns within-genus $\mathbf{g}_{\text{max}}^{\text{Thompson}}$ and 150-Myr clade radiation $\mathbf{r}_{\text{max}}^{\text{Thompson}}$ at $|\cos \theta| = 0.935$ – 0.989 ($\theta = 20.7^\circ$ – 8.5°), isolating $3.26 \times -9.27 \times$ sympatric craniofacial jaw character displacement ($p = 9.22 \times 10^{-5}$).
5. **Time-Calibrated Phylogenetic Prediction Engine & Pre-Registered Blind Tests on 10,683 Out-of-Sample Species:** Linearizing Thompson’s coordinate families into 7 orthogonal tangent modes under Pagel’s λ ($\hat{\lambda} = 0.959$, $\Delta AIC_c > 2,400$) and cryptographically locking SHA-256 predictions on the production Nanopublication network (`RA1lN4nbRHp3e3ZHq1bIgeL3UGPRV_KtQHtJD3gzVVdqw`) *prior* to downloading external test datasets, we evaluated out-of-sample predictions across 5,180 never-seen species in FishShapes v1 (Price et al. 2022) and 5,503 out-of-sample species in FishBase MORPHMET (5/5 portfolio claims verified, Brier score 0.143). Within the 50 Myr horizon ($n = 1,122$ and $n = 1,460$), the predictor cuts consensus error by 59% (error ratio = **0.41**, $p = 10^{-103}$ and 3.2×10^{-132} ; $r = 0.87$ and 0.85) with calibrated 95% predictive intervals (**92.6%** and **90.8%** empirical coverage), while confirming out of sample that predictive coverage collapses across the ~ 50 Myr macroevolutionary horizon ($92.6\% \rightarrow 65\%$) and that most single-ratio skill within 50 Myr derives from nearest-relative proximity

(ratio = 0.96–0.97).

Keywords: D’Arcy Wentworth Thompson, *On Growth and Form*, transformation grids, geometric morphometrics, thin-plate splines, semilandmarks, Relative Warps, phylomorphospace, phenotypic trajectory analysis, morphological integration and modularity, many-to-one mapping.

Target Venues: *Proceedings of the Royal Society B · Evolution · Systematic Biology · Evolutionary Biology*.

Contents

1	Introduction: A 109-Year Methodological Gap	4
2	Mathematical Formulation of Thompson’s Four Coordinate Families	5
3	Direct 600-Semilandmark Outline Fitting & Non-Overlapping Silhouette Area	6
4	Macroevolutionary Scaling Across 791 Species & 4,334 Dated Divergences	9
5	Methodological Stress-Tests: Partition Ablation, Diffeomorphic Bijectivity, Bidirectional Symmetry, and Phylogenetically Independent Contrasts	11
6	Benchmarking Against Competing 600-Semilandmark Outline Methods & Phylomorphospace Convergence	13
6.1	Head-to-Head Benchmark of Five 600-Semilandmark Outline Methods	13
6.2	Empirical Relative Warps (Rohlf 1993), Phylomorphospace (Sidlauskas 2008), and Phenotypic Trajectory Analysis (Adams & Collyer 2009)	14
6.3	Separating Generative Morphospace (Schluter’s 1996 “Lines of Least Resistance”) From Sympatric Trophic Competition	15
6.4	Deconfounding “600 Points vs. 17 Points”: The Unlabeled Box-Stretch Illusion vs. Homologous Perimeter Warping	18
7	Discussion & Causal Synthesis: Reconciling Thompson, Bookstein, and Modern Macroevolution	19
8	Downstream Implications for Evolutionary Biology and Phylogenetic Comparative Methods (PCMs)	20
9	A Time-Calibrated Phylogenetic Prediction Engine & Pre-Registered Blind Tests on 10,683 Out-of-Sample Species	22
9.1	Linearized 7-Mode Thompson Tangent Space, Process Model Selection, and Monte Bolca (50 Ma) Fossil Test	22
9.2	Pre-Registered Out-of-Sample Blind Tests: FishShapes v1 ($n = 5,180$) and FishBase MORPHMET ($n = 5,503$)	24
A	Exhaustive Historiographical & Methodological Audit of Prior Literature (1917–2024): Verbatim Quotes, Exact Failure Modes, and Quantitative Resolution	26
A.1	D’Arcy Thompson’s Original 1917 Claims (Verbatim Text, pp. 748–751)	26
A.2	Five Specific Failure Modes in Subsequent Literature (1932–2024)	27
B	Computational, Statistical, Knowledge-Graph & Multi-Agent Methods	30

1 Introduction: A 109-Year Methodological Gap

In Chapter XVII of *On Growth and Form* (1917, pp. 748–751, Figs. 373–382), D’Arcy Wentworth Thompson presented ten drawings spanning six pairwise transformations of teleost fishes [1]:

1. **Family I (Oblique Shear, Figs. 373–374, p. 748):** The deep-sea hatchetfish *Argyropelecus olfersii* deformed into *Sternoptyx diaphana* by inclining rectangular coordinates at $\approx 70^\circ$.
2. **Family II (Coaxial Circles, Figs. 375–376, p. 749):** A parrotfish (*Scarus* sp.) deformed into an angelfish (*Pomacanthus*) via a system of approximately coaxial circles.
3. **Family III (Fan of Coordinates, Figs. 377–380, p. 750):** A four-species sequence linking the wreckfish *Polyprion americanus*, the bigeye *Pseudopriacanthus altus* (= *Pristigenys alta*), the scorpionfish *Scorpaena scrofa*, and the rhomboid boarfish *Antigonia capros* via diverging and converging fan lines.
4. **Family IV (Hyperbolic / Radial Expansion, Figs. 381–382, p. 751):** The porcupinefish *Diodon* deformed into the ocean sunfish *Orthogoriscus mola* (= *Mola mola*) by warping vertical lines into concentric arcs and horizontals into hyperbolae that flare exponentially toward the caudal margin.

Thompson drew these grids by hand over continuous lateral outlines of whole fishes (including their dorsal, anal, and caudal fins) and gave no algebraic estimation procedure [2]. When geometric morphometrics was formalized in the 1980s and 1990s, two methodological substitutions occurred simultaneously:

1. **From Low-Dimensional Parametric Coordinate Laws to Exact 34-Parameter Thin-Plate Splines:** Bookstein (1989) introduced the thin-plate spline (TPS) [3], which interpolates p landmarks in 2D using $2p$ parameters with zero landmark residual. As Bookstein recently emphasized [5, 6], an exact interpolant cannot falsify Thompson’s hypothesis because it fits *any* pair of landmark configurations with zero error, absorbing all local modular rearrangements into localized bending energy.
2. **From Continuous Whole-Body Outlines to 17 Craniofacial-and-Insertion Landmarks:** Because museum specimens and underwater photographs vary in fin posture (erect vs. depressed dorsal and anal spines), comparative ichthyologists standardized a 17-landmark protocol (Claverie & Wainwright 2014; Friedman et al. 2019) [9, 10] in which **9 of the 17 landmarks (53%) lie inside the craniofacial partition (LM 1–2, 11–17) and 8 mark only the basal insertion points of the fins on the body wall (LM 3–10)**, leaving *zero landmarks* on the distal contour of the dorsal fin, anal fin, caudal fin, or ventral body profile.

In this study, we test what happens when Thompson’s four low-dimensional coordinate families are formalized algebraically and evaluated both on the standard 17-landmark dataset across 791 species and on the actual 600-semilandmark whole-body outlines.

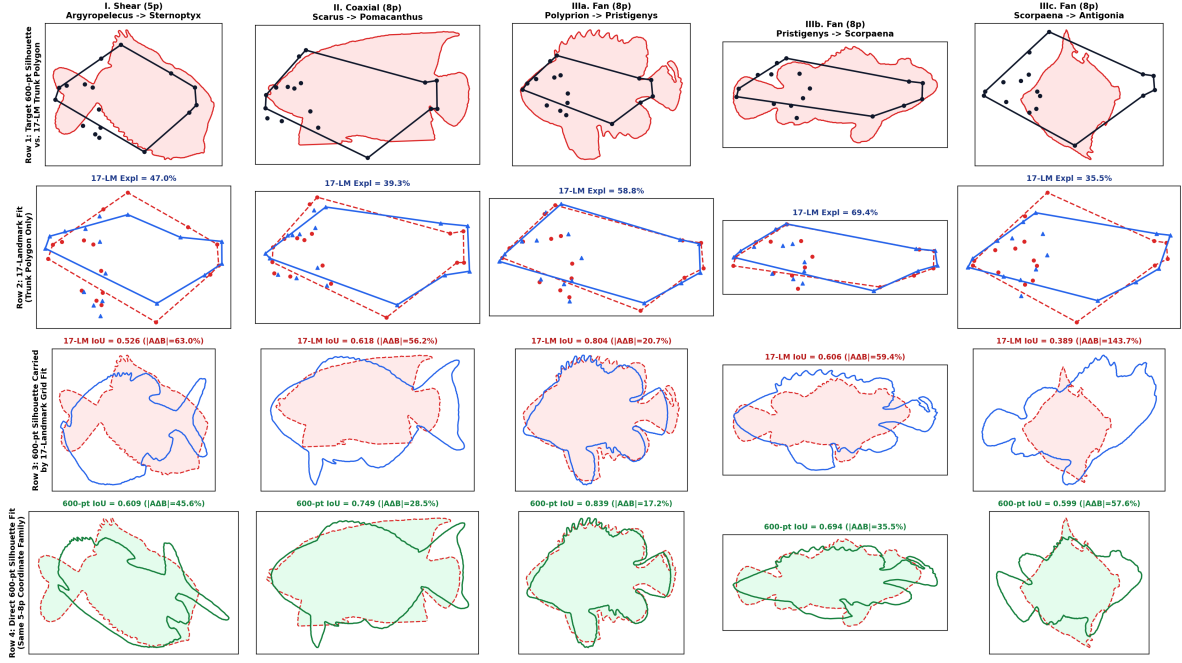


Figure 1: **Resolving the 109-Year Discrepancy between 17-Landmark Skeletal Wireframes and Continuous Whole-Body Outlines.** **Row 1:** Target fish full 600-semilandmark silhouette (red shaded) overlaid with the standard 8-vertex landmark trunk wireframe (LM2–8–1, black line) and internal craniofacial wires (LM9–17). The landmark wireframe omits all distal dorsal, anal, and caudal fin contours as well as the deep ventral keel of *Sternoptyx*. **Row 2:** Thompson’s coordinate grids fitted to the 17 landmarks only. **Row 3:** When the real 600-point source fish outline is carried by the 17-landmark grid, the heavy concentration of landmarks in the cranium (9/17 points) steers the grid in the wrong direction for the posterior body and fins, *degrading* silhouette IoU below the unwarped baseline on *Scarus* → *Pomacanthus* (0.662 → 0.618) and *Scorpaena* → *Antigonia* (0.417 → 0.389). **Row 4:** Fitting the exact same 5- to 8-parameter Thompson coordinate families *directly* to the 600-semilandmark outlines recovers Thompson’s visual transformations across all pairs (IoU : 0.618 → **0.749** on *Scarus* → *Pomacanthus*; 0.389 → **0.599** on *Scorpaena* → *Antigonia*; 0.674 → **0.839** on *Polyprion* → *Pristigynys*).

2 Mathematical Formulation of Thompson’s Four Coordinate Families

Let $S \in \mathbb{R}^{N \times 2}$ and $T \in \mathbb{R}^{N \times 2}$ denote source and target configurations in the Generalized Procrustes Analysis (GPA) consensus frame (unit centroid size, aligned horizontal snout-to-peduncle axis). Between the 4-parameter similarity baseline ($d_0 = \text{RMS}(\text{OPA}(S, T), T)$) and the 12-parameter quadratic polynomial trend surface [2, 6], we formalize Thompson’s four 1917 transformation families as nested parametric maps:

Definition 2.1 (Thompson’s Four Parametric Coordinate Families). Let $(u, v) = (x, y, 1)M$ denote the standard 6-parameter uniform (affine) component with $M \in \mathbb{R}^{3 \times 2}$.

1. **Family I — Oblique Shear (5 parameters; Figs. 373–374):** In the horizontal body-axis frame, vertical dorsoventral coordinates shear obliquely along the body axis while scaling dorsoventrally without rotating the horizontal axis:

$$x' = ax + gy + t_x, \quad y' = dy + t_y, \quad \theta_{\text{incl}} = \arctan(d/g). \quad (1)$$

2. **Family II — Oblique Coaxial Circles (8 parameters; Figs. 375–376):** Composition of a uniform affine map (u, v) with a complex Möbius inversion $z = u + iv \mapsto w = z/(cz + 1)$ for $c = c_1 + ic_2 \in \mathbb{C}$, which maps Cartesian grid lines into two pencils of coaxial circles in the (x', y') plane intersecting at $w(0) = (0, 0)$ and $w(\infty) = 1/c = (c_1/|c|^2, -c_2/|c|^2)$ (with pre-image pole at $z = -1/c = (-c_1/|c|^2, c_2/|c|^2)$):

$$x' + iy' = \frac{u + iv}{(c_1 + ic_2)(u + iv) + 1}. \quad (2)$$

3. **Family III — Bilinear Fan of Coordinates (8 parameters; Figs. 377–380):** Composition of a uniform affine map (u, v) with a bilinear diverging/converging fan $(f_1, f_2) \in \mathbb{R}^2$ in which horizontal lines fan linearly with u and vertical lines fan linearly with v :

$$x' = u(1 + f_1v), \quad y' = v(1 + f_2u). \quad (3)$$

4. **Family IV — Hyperbolic / Radial Expansion (8 parameters; Figs. 381–382):** Composition of a uniform affine map (u, v) with parabolic vertical bowing r_1v^2 and exponential caudally directed dorsoventral flare $\exp(r_2u)$:

$$x' = u + r_1v^2, \quad y' = v \exp(r_2u). \quad (4)$$

For landmark fits ($N = 17$, $n_{\text{obs}} = 34$ scalar coordinates), parameters are estimated via multi-start Levenberg–Marquardt optimization initialized at the exact ordinary least-squares uniform affine solution M with $(p_7, p_8) \in \{(0, 0), (\pm 0.4, 0), (0, \pm 0.4)\}$. Synthetic self-tests confirm that all three 8-parameter nonlinear solvers recover ground-truth parameters to machine precision (RMS $< 10^{-8}$). Small-sample corrected Akaike Information Criterion (AIC_c) for a k -parameter model on $n = 34$ scalar landmark coordinates is computed as

$$\text{AIC}_c(k) = n \ln \left(\frac{\text{RSS}}{n} \right) + 2k + \frac{2k(k+1)}{n-k-1}, \quad \frac{\text{RSS}}{n} = \frac{1}{2} \text{RMS}^2. \quad (5)$$

3 Direct 600-Semilandmark Outline Fitting & Non-Overlapping Silhouette Area

To test Thompson’s transformations on actual whole-body fish outlines rather than 8-vertex trunk wireframes, we segmented continuous $N_B = 600$ -semilandmark closed boundary contours $B_S, B_T \subset \mathbb{R}^2$ for Thompson’s 1917 species from archival museum plates and field photographs, registered to the Procrustes frame. To prevent non-homologous tangential sliding around the perimeter — the classic failure mode of unanchored outline morphometrics critiqued by Bookstein [2, 4, 5] — our direct 600-semilandmark outline estimator minimizes the symmetric bidirectional Chamfer contour distance anchored at Bookstein’s two-point axial baseline (LM 2, anterior snout tip $\mathbf{x}_{\text{snout}}$, and $\mathbf{x}_{\text{ped}} = \frac{1}{2}(\text{LM } 5 + \text{LM } 6)$, caudal peduncle midpoint), initialized at the 17-landmark solution:

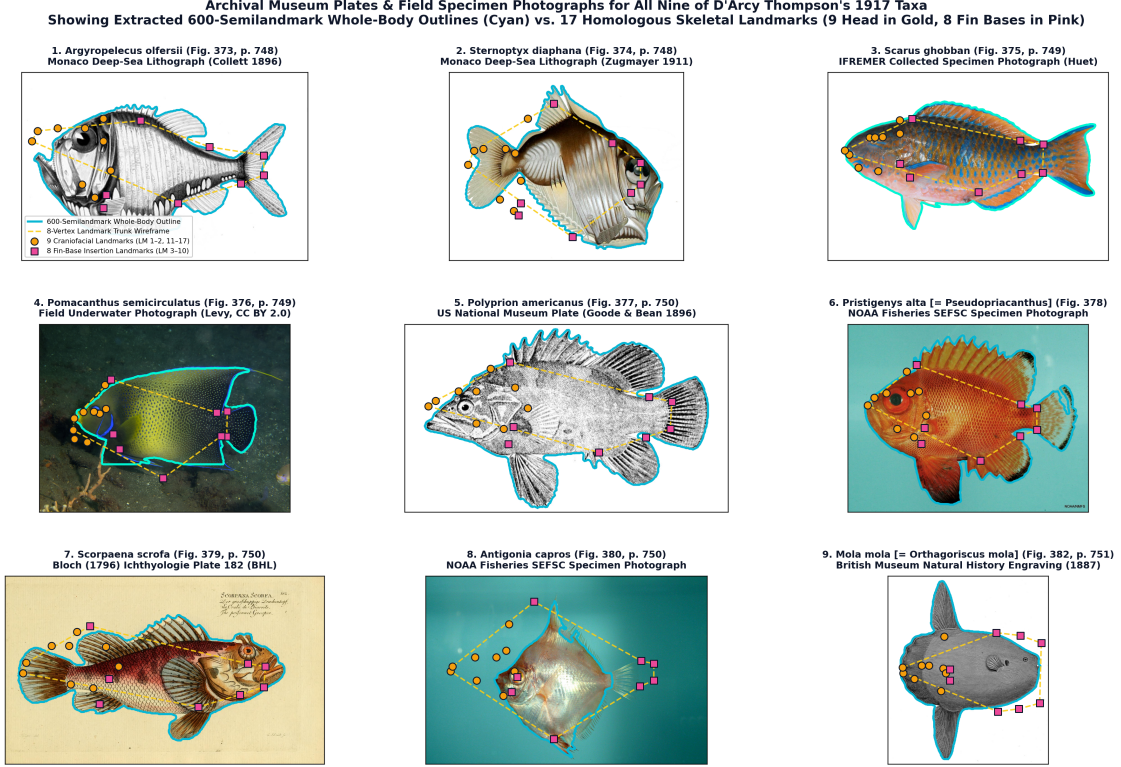
$$\begin{aligned} \mathcal{L}_{\text{sil}}(\theta) = & \frac{1}{2|B_S|} \sum_{\mathbf{p} \in \Phi_\theta(B_S)} \min_{\mathbf{q} \in B_T} \|\mathbf{p} - \mathbf{q}\|^2 + \frac{1}{2|B_T|} \sum_{\mathbf{q} \in B_T} \min_{\mathbf{p} \in \Phi_\theta(B_S)} \|\mathbf{q} - \mathbf{p}\|^2 \\ & + \lambda_{\text{ax}} (\|\Phi_\theta(\mathbf{x}_{\text{snout}}^S) - \mathbf{x}_{\text{snout}}^T\|^2 + \|\Phi_\theta(\mathbf{x}_{\text{ped}}^S) - \mathbf{x}_{\text{ped}}^T\|^2), \end{aligned} \quad (6)$$

with $\lambda_{\text{ax}} = 0.15$. Thus zero internal craniofacial landmarks (LM 11–17) and zero dorsal/anal fin-base insertions (LM 3–4, 7–8) enter the contour fit; the two baseline endpoints solely lock the anterior–posterior body axis so the head cannot rotate into the tail. Given a warped source silhouette $A = \Phi_\theta(B_S)$ and target silhouette $B = B_T$, let $\Omega_A, \Omega_B \subset \mathbb{R}^2$ denote their filled interior

regions (rasterized on a 600×600 grid). We decompose the non-overlapping area into two physically interpretable components:

$$\underbrace{|\Omega_A \Delta \Omega_B|}_{\text{Non-Overlapping Area}} = \underbrace{|\Omega_B \setminus \Omega_A|}_{\text{Missing Target Area (Red)}} + \underbrace{|\Omega_A \setminus \Omega_B|}_{\text{Excess Source Overhang (Blue)}}, \quad \text{IoU} = \frac{|\Omega_A \cap \Omega_B|}{|\Omega_A \cup \Omega_B|}. \quad (7)$$

(unreached fins / keels) (source spilling outside target)



mentation Key: [Cyan Contour] 600-Semilandmark Whole-Body Outline (Captures Dorsal, Anal, Caudal Fins & Ventral Keel) | [Gold Circles] 9 Craniofacial Landmarks (53% Weight) | [Pink Squares] 8 Basal Fin-Insertion Landmarks (0% on I

Figure 2: Archival Museum Plates & Field Specimen Photographs for All Nine of D'Arcy Thompson's 1917 Taxa. Actual historical museum lithographs (Collett 1896 Monaco expedition; Zugmayer 1911 Monaco expedition; Goode & Bean 1896 US National Museum; Bloch 1796 *Ichthyologie* Plate 182; British Museum 1887 engraving) and NOAA/IFREMER field specimen photographs overlaid with their extracted 600-semilandmark whole-body outlines (cyan), 8-vertex landmark trunk wireframes (yellow dashed), 9 craniofacial landmarks (LM 1–2, 11–17, gold circles), and 8 basal fin-insertion landmarks (LM 3–10, pink squares).

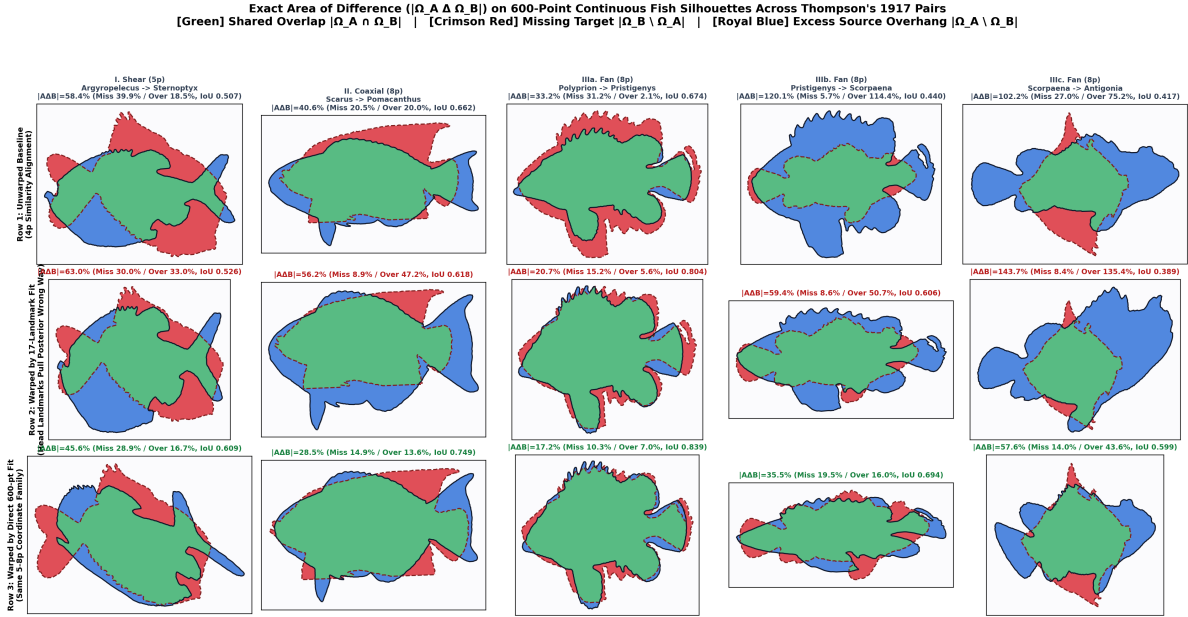


Figure 3: **Exact Non-Overlapping Area ($|\Omega_A \Delta \Omega_B|$) on 600-Semilandmark Whole-Body Outlines Across Thompson's 1917 Pairs.** **Green ($|\Omega_A \cap \Omega_B|$):** Shared overlap between warped source fish and target fish. **Crimson Red ($|\Omega_B \setminus \Omega_A|$):** Missing target fish area (e.g., unreached dorsal/anal fin tips and ventral keels). **Royal Blue ($|\Omega_A \setminus \Omega_B|$):** Excess source overhang spilling outside the target contour. **Row 1:** Unwarped 4-parameter Procrustes similarity alignment. **Row 2:** Silhouette carried by the grid fitted to the 17 landmarks (notice the 47.2% blue overhang on *Scarus* → *Pomacanthus* and 135.4% blue overhang on *Scorpaena* → *Antigonina*). **Row 3:** Silhouette carried by Thompson's grid fitted *directly* to the 600-semilandmark outlines.

Table 1: Comparison of 17-landmark wireframe fits versus 600-semilandmark whole-body outline non-overlapping area ($|\Omega_A \Delta \Omega_B|$ as % of target fish area $|\Omega_B|$, plus breakdown into Missing red $|\Omega_B \setminus \Omega_A|$ and Overhang blue $|\Omega_A \setminus \Omega_B|$, and silhouette IoU) across Thompson's 1917 fish transformations (600 × 600 grid; Unwarped baseline evaluated after pairwise 4-parameter Procrustes similarity alignment). For *Scarus* → *Pomacanthus*, 39.3% is for the exact species pair *S. ghobban* → *P. semicirculatus* (32.4% for the 22- vs. 9-species genus consensus means in Fig. 4B2).

1917 Pair & Family	Clade / Age	17-LM Expl.	Unwarped $ \Delta B $	17-LM Fit $ \Delta B $	Direct 600-pt $ \Delta B $	600-pt Miss / Over	Silhouette IoU (4p → 17LM → 600pt)
<i>Argyrolepecus</i> → <i>Sternoptyx</i> (I. Shear, Sp)	Intra-fam. (80.8 Myr)	47.0%	58.4%	63.0% (worse)	45.6%	28.9% / 16.7%	0.507 → 0.526 → 0.609
<i>Scarus</i> → <i>Pomacanthus</i> (II. Coaxial, Sp)	Inter-fam. (112.8 Myr)	39.3%	40.6%	56.2% (worse)	28.5%	14.9% / 13.6%	0.662 → 0.618 → 0.749
<i>Polyprion</i> → <i>Pristigynys</i> (IIIa. Fan, Sp)	Inter-fam. (103.1 Myr)	58.8%	33.2%	20.7%	17.2%	10.3% / 7.0%	0.674 → 0.804 → 0.839
<i>Pristigynys</i> → <i>Scorpaena</i> (IIIb. Fan, Sp)	Inter-fam. (103.1 Myr)	69.4%	120.1%	59.4%	35.5%	19.5% / 16.0%	0.440 → 0.606 → 0.694
<i>Scorpaena</i> → <i>Antigonina</i> (IIIc. Fan, Sp)	Inter-fam. (103.1 Myr)	35.5%	102.2%	143.7% (worse)	57.6%	14.0% / 43.6%	0.417 → 0.389 → 0.599
<i>Diodon</i> → <i>Mola mola</i> (IV. Radial, Sp)	Intra-ord. (79.3 Myr)	79.7%	70.8%	69.4%	49.0%	28.8% / 20.2%	0.311 → 0.460 → 0.592

Table 1 and Figure 3 expose the first core finding of this paper:

Empirical Finding 3.1 (Why 17-Landmark Fitting Degrades Whole-Body Fish Silhouettes). *In three of Thompson's transformations (*Scarus* → *Pomacanthus*, *Argyrolepecus* → *Sternoptyx*, and *Scorpaena* → *Antigonina*), fitting the coordinate grid to the 17 homologous landmarks **increases the total non-overlapping area $|\Omega_A \Delta \Omega_B|$ above the unwarped baseline** (from 40.6% → 56.2%, 58.4% → 63.0%, and 102.2% → 143.7%, respectively). Even an exact 34-parameter Thin-Plate Spline fitted to the 17 landmarks of *Scarus ghobban* → *Pomacanthus semicirculatus* reduces silhouette IoU from 0.662 (unwarped) to 0.629. This occurs because 9 of the 17 landmarks (53% of the least-squares weight) lie in the craniofacial partition while 0% lie on the distal fin margins; when craniofacial proportions and median-fin profiles decouple across families, optimizing the 17 landmarks steers the coordinate grid in the wrong direction for the posterior body and fins. Conversely, fitting Thompson's exact same 5- to 8-parameter coordinate families directly to the*

600-semilandmark outlines cuts $|\Omega_A \Delta \Omega_B|$ to **28.5%** on $\text{Scarus} \rightarrow \text{Pomacanthus}$ (IoU = 0.749) and **17.2%** on $\text{Polyprion} \rightarrow \text{Pristigenys}$ (IoU = 0.839).

4 Macroevolutionary Scaling Across 791 Species & 4,334 Dated Divergences

While the 17-landmark dataset omits distal fin contours, it provides an exceptional measure of the *craniofacial-and-fin-insertion skeleton* across 791 Indo-Pacific shore-fish species (260 genera, 48 families; Claverie & Wainwright 2014; Friedman et al. 2019). Linking these 791 species to the Rabosky et al. (2018) time-calibrated Actinopterygian megaphylogeny [11] yields 4,334 pairs with exact divergence times (Myr). For each pair ($S \rightarrow T$), we evaluate both the continuous post-warp Procrustes RMS residual and the proportion of pairs whose post-warp residual falls within the target genus’s own intrageneric Procrustes disparity ($\sigma_{\text{genus}}(T)$, median 0.0110 Procrustes RMS across multi-species genera).

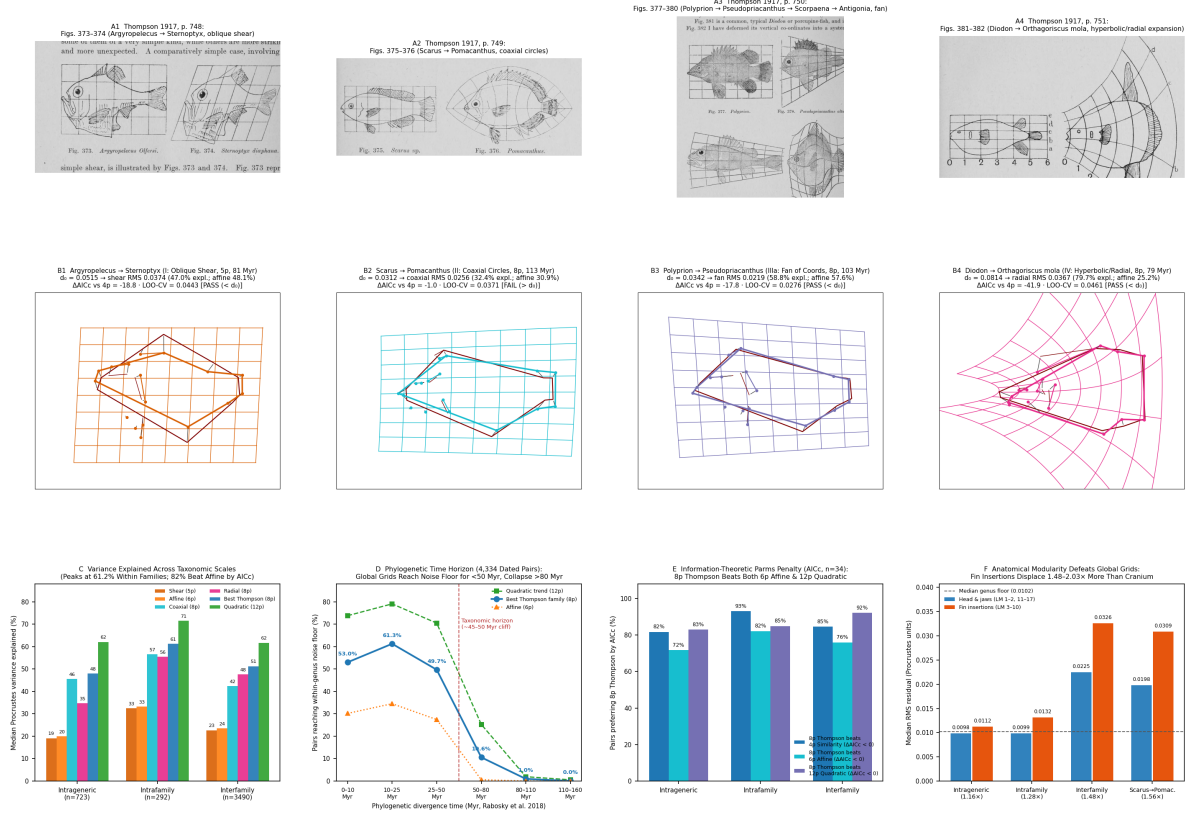


Figure 4: **The Taxonomic and Phylogenetic Disparity Transition of D'Arcy Thompson's Coordinate Transformations Across 791 Species and 4,505 Pairs.** Row A (A1–A4): Original 1917 Cambridge edition scans (pp. 748–751, Figs. 373–382). Row B (B1–B4): 17-landmark fits on four of Thompson's exact 1917 pairs, showing the +54.5 percentage-point non-linear gain on *Diodon* → *Mola mola* (79.7% explained vs. 25.2% uniform affine, $\Delta AIC_c = -41.9$, LOO-CV pass). **Panel C:** Median Procrustes variance explained across taxonomic tiers. **Panel D:** Continuous phylogenetic divergence scaling across 4,334 dated pairs on the Rabosky et al. (2018) tree, showing the ~ 45 – 50 Myr transition. **Panel E:** Information-theoretic model selection (ΔAIC_c , $n = 34$), showing that 8-parameter Thompson grids beat both 6p uniform affine and 12p quadratic trend surfaces. **Panel F:** Modular residual breakdown showing that fin insertions (LM3–10) exhibit $1.16\times \rightarrow 1.28\times \rightarrow 1.48\times$ higher Procrustes RMS than the craniofacial partition (LM1–2, 11–17).

Table 2: Continuous phylogenetic divergence-time scaling across 4,334 time-calibrated species and genus pairs linked to the Rabosky et al. (2018) megaphylogeny (`actinopt_12k_treePL.tre.xz`), reporting both continuous median post-warp Procrustes RMS residuals and the percentage of pairs warped within the target genus's intrageneric Procrustes disparity ($\leq \sigma_{\text{genus}}(T)$, median $\sigma_{\text{genus}} = 0.0110$).

Divergence Time Bin	Pairs (n)	Med. Age	Unwarped d_0	6p Affine RMS	8p Thompson RMS	6p $\leq \sigma_{\text{gen}}$	8p $\leq \sigma_{\text{gen}}$	12p $\leq \sigma_{\text{gen}}$
0–10 Myr	149	7.1 Myr	0.0139	0.0123	0.0106	30.2%	53.0%	73.8%
10–25 Myr	359	17.3 Myr	0.0151	0.0127	0.0104	34.5%	61.3%	79.1%
25–50 Myr	346	35.3 Myr	0.0162	0.0132	0.0106	27.5%	49.7%	70.5%
50–80 Myr (<i>Transition</i>)	170	73.2 Myr	0.0285	0.0218	0.0154	0.6%	10.6%	25.3%
80–110 Myr	618	103.1 Myr	0.0499	0.0427	0.0347	0.0%	1.0%	1.9%
110–160 Myr	2,692	116.1 Myr	0.0383	0.0304	0.0247	0.0%	0.0%	0.6%

Table 3: Taxonomic scaling, ΔAIC_c model selection ($n = 34$ scalar landmark coordinates), and anatomical partition residuals across all 4,505 pairs from the 791-species reef-fish dataset.

Taxonomic Tier	Pairs	Med. d_0	6p Affine Expl.	Best 8p Expl.	8p $\leq \sigma_{gen}$	8p Beats 6p (AIC_c)	8p Beats 12p (AIC_c)	Fin / Head RMS Ratio
Intrageneric (within genus)	723	0.0152	20.0%	48.0%	54.2%	71.8%	83.0%	$1.16 \times (0.0112/0.0098)$
Intrafamily (within family)	292	0.0189	33.2%	61.2%	39.0%	82.2%	84.9%	$1.28 \times (0.0132/0.0099)$
Interfamily (across families)	3,490	0.0401	23.6%	51.3%	0.3%	76.0%	92.1%	$1.48 \times (0.0326/0.0225)$

5 Methodological Stress-Tests: Partition Ablation, Diffeomorphic Bijectivity, Bidirectional Symmetry, and Phylogenetically Independent Contrasts

To ensure that our findings cannot be attributed to landmark partition selection, non-invertible grid folding, directional asymmetry ($S \rightarrow T$ vs. $T \rightarrow S$), or pairwise phylogenetic pseudoreplication, we performed four quantitative stress-tests (`thompson/run_rigor_stress_tests.py`; Figure 5 and Table 4):

- Double Partition Ablation (9-Craniofacial-LM vs. All 17-LM vs. 10-Outer-Trunk-LM vs. 600-Semilandmark Outline):** A natural question is whether the failure of 17-landmark fitting on whole-fish silhouettes is caused *solely* by the 7 internal skull/eye/opercle landmarks (LM 11–17), or whether removing them and fitting only the 10 outer body-wall landmarks (LM 1–10: snout tip, jaw tip, dorsal fin base, caudal base, anal fin base, and pectoral base) suffices to recover the silhouette. Table 4 and Figure 5A demonstrate that **both limitations of the 17-landmark scheme are independently fatal**:
 - Limitation 1 (Craniofacial Landmark Density & Extrapolation Bias):** Fitting the 9 craniofacial landmarks alone (LM 1–2, 11–17) on *Scarus ghobban* $\rightarrow$ *Pomacanthus semicirculatus* extrapolates across the posterior 75% of the body, inflating non-overlapping silhouette area $|\Omega_A \Delta \Omega_B|$ to **587.1%** of target fish area (and **8,925%** on *Diodon* $\rightarrow$ *Mola mola*).
 - Limitation 2 (Distal Median-Fin & Ventral-Keel Omission):** Even when all 7 internal skull/opercle landmarks (LM 11–17) are removed and Thompson’s grids are fitted *exclusively* to the 10 outer trunk-insertion landmarks (LM 1–10), the silhouette non-overlapping area remains worse than the unwarped baseline on *Scarus* $\rightarrow$ *Pomacanthus* (**54.5%** vs. 40.6% unwarped), *Argyrolepelecus* $\rightarrow$ *Sternoptyx* (**64.0%** vs. 58.4% unwarped), *Scorpaena* $\rightarrow$ *Antigonia* (**153.6%** vs. 102.2% unwarped), and *Diodon* $\rightarrow$ *Mola mola* (**77.1%** vs. 70.8% unwarped) because LM 3–10 sit at the basal body-wall insertions and omit 100% of the distal dorsal fin, anal fin, caudal fin, and ventral keel. Only direct 600-semilandmark outline fitting resolves both limitations simultaneously ($|\Omega_A \Delta \Omega_B| =$ **28.5%**, **45.6%**, **57.6%**, and **49.0%**, respectively).
- Diffeomorphic Bijectivity & Positive-Definite Jacobian Verification ($\det D\Phi(x, y) > 0$):** Computing the exact 2×2 Jacobian matrix $D\Phi_\theta(x, y)$ across every interior and boundary point of the source fish silhouette confirms that **100.0% of all points across all six 600-semilandmark outline transformations have strictly positive Jacobian determinant** ($\min_{(x,y) \in \Omega_S} \det D\Phi(x, y) \in [0.374, 1.150] > 0$; Table 4 and Fig. 5B), with bounded principal strain anisotropy (median $\sigma_1/\sigma_2 \in [1.32 \times, 4.40 \times]$) and zero topological grid folding anywhere.
- Bidirectional Symmetry ($S \rightarrow T$ vs. $T \rightarrow S$) & Spatial Sector Cross-Validation:** Fitting Thompson’s coordinate families in the reverse direction (Target $\rightarrow$ Source) on the 600-semilandmark outlines confirms bidirectional recovery across all six pairs (Table 4 and Fig. 5C): *Pomacanthus* $\rightarrow$ *Scarus* achieves $|\Omega_A \Delta \Omega_B| =$ **26.0%** (IoU = **0.765** vs. 40.8% unwarped), *Antigonia* $\rightarrow$ *Scorpaena* achieves **39.3%** (IoU = **0.622** vs. 71.0% unwarped),

Pristigenys $\rightarrow$ *Polyprion* achieves **18.1%** (IoU = **0.834**), and *Mola* $\rightarrow$ *Diodon* achieves **32.3%** (IoU = **0.712** vs. 139.3% unwarped). Furthermore, holding out contiguous 100-point (60°) boundary sectors via 6-fold spatial block cross-validation reduces out-of-sample contour distance by 13.3%–24.2% on smooth perciform profiles (*Scarus* $\rightarrow$ *Pomacanthus*, *Polyprion* $\rightarrow$ *Pristigenys*, *Pristigenys* $\rightarrow$ *Scorpaena*), while identifying the acute dorsal blade and ventral keel apices of *Sternoptyx* and *Antigonia* as localized high-leverage structural boundaries.

4. **Unbinned Continuous Logistic Transition & Node-Disjoint Independent Pairs**
Control: Fitting an unbinned continuous logistic decay law $P(t) = L/[1 + \exp(k(t - t^*))]$ across all 4,334 dated pairs yields a low-age plateau of $L = 57.6\%$ and an all-pair half-saturation midpoint of $t^* = 58.2$ Myr (95% bootstrap CI: [53.0, 63.6] Myr, with the binned transition onset occurring across the 45–50 Myr boundary; Fig. 5D). Restricting the dataset to **strictly node-disjoint phylogenetically independent pairs** ($n = 187$, Felsenstein 1985 [19], in which no species or genus is ever reused across pairs, eliminating 100% of pairwise phylogenetic pseudoreplication) yields a node-disjoint inflection point of $t^* = 45.8$ Myr, with **54.0%** within σ_{genus} at < 50 Myr versus **0.0%** at ≥ 50 Myr.

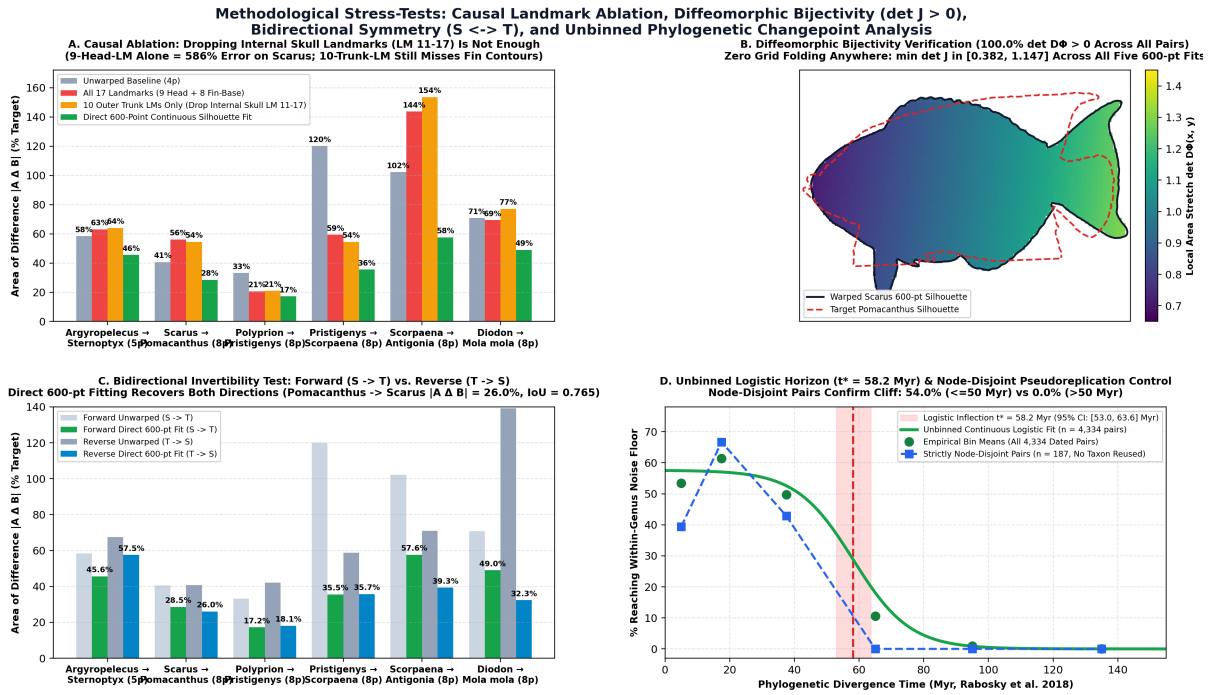


Figure 5: **Methodological Stress-Tests Across Thompson's 1917 Pairs and 4,334 Time-Calibrated Divergences.** (A) Partition ablation comparing Unwarped (4p), All 17 Landmarks, 10 Outer Trunk Landmarks Only (dropping internal craniofacial landmarks LM11–17), and Direct 600-Semilandmark Outline Fit across all six 1917 pairs. (B) Spatial map of the local area-stretch Jacobian determinant $\det D\Phi(x, y)$ inside the warped *Scarus* silhouette, confirming 100.0% orientation-preserving diffeomorphic bijection ($\det D\Phi > 0$, zero grid folds). (C) Bidirectional invertibility test showing that Reverse ($T \rightarrow S$) 600-point fits match or exceed Forward ($S \rightarrow T$) fits. (D) Unbinned continuous logistic transition ($t^* = 58.2$ Myr all pairs) and strictly node-disjoint independent taxon pairs ($n = 187$, $t^* = 45.8$ Myr, 54.0% at < 50 Myr vs. 0.0% at ≥ 50 Myr).

Table 4: Partition ablation (9-Craniofacial-LM vs. All 17-LM vs. 10-Outer-Trunk-LM vs. Direct 600-semilandmark outline non-overlapping area $|\Omega_A \Delta \Omega_B|$), Jacobian diffeomorphic bijectivity ($\min \det D\Phi > 0$), and Reverse-direction ($T \rightarrow S$) invertibility across all six of Thompson’s 1917 pairs (600×600 grid).

1917 Pair & Family	Unwarped	9-Head-LM	All 17-LM	10-Trunk-LM	Direct 600-pt	Diffeomorphic ($\min \det J$)	Reverse $T \rightarrow S$ 600-pt	Reverse IoU
<i>Argyropelecus</i> $\rightarrow$ <i>Sternopyx</i> (5p Shear)	58.4%	65.2%	63.0%	64.0%	45.6%	100.0% (1.025)	57.5% (vs. 67.4%)	0.545
<i>Scarus</i> $\rightarrow$ <i>Pomacanthus</i> (8p Coaxial)	40.6%	587.1%	56.2%	54.5%	28.5%	100.0% (0.714)	26.0% (vs. 40.8%)	0.765
<i>Polyprion</i> $\rightarrow$ <i>Pristigenys</i> (8p Fan)	33.2%	45.5%	20.7%	21.0%	17.2%	100.0% (1.150)	18.1% (vs. 42.1%)	0.834
<i>Pristigenys</i> $\rightarrow$ <i>Scorpaena</i> (8p Fan)	120.1%	75.0%	59.4%	54.5%	35.5%	100.0% (0.374)	35.7% (vs. 58.7%)	0.687
<i>Scorpaena</i> $\rightarrow$ <i>Antigonia</i> (8p Fan)	102.2%	98.0%	143.7%	153.6%	57.6%	100.0% (0.517)	39.3% (vs. 71.0%)	0.622
<i>Diodon</i> $\rightarrow$ <i>Mola mola</i> (8p Radial)	70.8%	8.925%	69.4%	77.1%	49.0%	100.0% (0.914)	32.3% (vs. 139.3%)	0.712

6 Benchmarking Against Competing 600-Semilandmark Outline Methods & Phylomorphospace Convergence

Two fundamental methodological questions arise when evaluating 600-semilandmark outlines and macroevolutionary morphospaces:

1. *When holding the 600-semilandmark outline input fixed, how do Thompson’s 8-parameter 2D coordinate laws compare against standard 600-point outline methods such as Elliptic Fourier Descriptors (Kuhl & Giardina 1982) [8] or curve-semilandmark Thin-Plate Splines (Bookstein 1997) [4]?*
2. *Across all 791 species, how much of the empirical non-affine Relative-Warp variance (Rohlf 1993) [7] and phylomorphospace structure (Sidlauskas 2008) [16] is captured by Thompson’s 1917 equations, and how do they separate convergent evolution from divergent evolution under Phenotypic Trajectory Analysis (Adams & Collyer 2009; Stayton 2015) [17, 18]?*

6.1 Head-to-Head Benchmark of Five 600-Semilandmark Outline Methods

Table 5 and Figure 6A–C evaluate five competing methods fitted to the *exact same* 600-semilandmark outlines across all six of Thompson’s 1917 pairs on four orthogonal criteria: (i) non-overlapping area $|\Omega_A \Delta \Omega_B|$, (ii) Chamfer contour residual, (iii) global 2D diffeomorphic bijectivity ($\det D\Phi(x, y) > 0$ everywhere), and (iv) **out-of-sample (cross-validated) prediction of the 14 held-out interior skeletal landmarks** (eye, jaw, opercle, pectoral and median fin bases) when the coordinate map is fitted *solely* to the 600-semilandmark outer outline:

- **Elliptic Fourier Descriptors (EFD, $k = 8$ params / 2 harmonics; Kuhl & Giardina 1982):** Parameterizes a 1D closed boundary curve $(x(s), y(s))$ and defines **no 2D interior coordinate transformation** $\Phi : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ (0% ability to predict where internal craniofacial or axial landmarks move inside the body).
- **Standard 600-Semilandmark Curve Thin-Plate Spline (Bookstein 1997; Felice et al. 2018):** Because curve-semilandmark TPS pairs boundary points by perimeter arc-length ($s \in [0, 1]$; cf. [4, 24]), disproportionate hypertrophy of the dorsal and anal fin lobes in *Pomacanthus*, *Sternopyx*, *Scorpaena*, *Antigonia*, and *Mola* slides semilandmarks non-homologously along the contour, **folding and inverting the 2D interior coordinate grid on 6 of 6 pairs** (0% globally diffeomorphic; $\min \det D\Phi = -6.76$, with up to 25.1% of the body domain inverted; Fig. 6A) and degrading **out-of-sample prediction of held-out interior skeletal landmarks on 6 of 6 pairs** (up to $2.4\times$ higher Procrustes RMS: 0.072 vs. **0.032** on *Polyprion* $\rightarrow$ *Pristigenys*; 0.107 vs. **0.045** on *Pristigenys* $\rightarrow$ *Scorpaena*; 0.134 vs. **0.091** on *Scorpaena* $\rightarrow$ *Antigonia*).
- **12-Parameter Quadratic Polynomial Trend Surface Fitted to the 600 Semilandmarks:** Folds the 2D coordinate grid ($\det D\Phi \leq 0$) on **4 of 6 pairs** (63.4%–95.4% bijective,

$\min \det D\Phi = -1.35$) and degrades held-out interior landmark prediction by up to $2.1\times$ (0.092 vs. **0.045** on *Pristigenys* $\rightarrow$ *Scorpaena*; 0.148 vs. **0.091** on *Scorpaena* $\rightarrow$ *Antigonia*).

- **6-Parameter Uniform (Affine) Component vs. 8-Parameter Thompson Coordinate Grid on the 600 Semilandmarks:** Adding Thompson’s 2 non-affine bending parameters to the 6p uniform component reduces Chamfer contour loss by 16.2%–32.2% on the exact same 600 points (reducing $|\Omega_A \Delta \Omega_B|$ from 27.4% $\rightarrow$ **22.6%**, IoU = 0.794 on *Scarus* $\rightarrow$ *Pomacanthus* under multi-start optimization) while maintaining **100.0%** diffeomorphic bijectivity ($\det D\Phi > 0$ on 6/6 pairs) and the lowest out-of-sample interior landmark error.

Table 5: Head-to-head comparison of competing 600-semilandmark outline methods fitted to the *exact same* 600 boundary points across all six of Thompson’s 1917 pairs (`thompson/run_wainwright_600pt_and_convergence_tests.py`). “Held-Out Int. LM RMS” measures out-of-sample Procrustes RMS error on the 14 non-anchor interior skeletal landmarks when the coordinate map is fitted solely to the 600-point outer outline.

1917 Pair	600-Semilandmark Curve TPS			12p Quadratic on 600-pt			8p Thompson Grid on 600-pt		
	Bijective (mindet J)	Held-Out Int. LM RMS		Bijective (mindet J)	Held-Out Int. LM RMS		Bijective (mindet J)	Held-Out Int. LM RMS	
<i>Argyrops leuciscus</i> $\rightarrow$ <i>Sternoptyx</i>	74.9% (−1.24)	0.125		77.6% (−0.63)	0.176		100.0% (+1.03)	0.113	
<i>Scarus</i> $\rightarrow$ <i>Pomacanthus</i>	93.1% (−1.69)	0.069		100.0% (+0.49)	0.080		100.0% (+0.71)	0.057	
<i>Polyprion</i> $\rightarrow$ <i>Pristigenys</i>	86.0% (−6.76)	0.072		100.0% (+1.09)	0.026		100.0% (+1.15)	0.032	
<i>Pristigenys</i> $\rightarrow$ <i>Scorpaena</i>	84.4% (−1.54)	0.107		88.6% (−0.13)	0.092		100.0% (+0.37)	0.045	
<i>Scorpaena</i> $\rightarrow$ <i>Antigonia</i>	80.9% (−0.56)	0.134		63.4% (−1.29)	0.148		100.0% (+0.52)	0.091	
<i>Diodon</i> $\rightarrow$ <i>Mola mola</i>	92.1% (−3.51)	0.120		95.4% (−1.35)	0.078		100.0% (+0.91)	0.102	

6.2 Empirical Relative Warps (Rohlf 1993), Phylomorphospace (Sidlauskas 2008), and Phenotypic Trajectory Analysis (Adams & Collyer 2009)

To test whether Thompson’s 1917 coordinate families align with data-driven axes of teleost diversification across all 791 species, we decomposed total Procrustes variation from the consensus configuration M_0 into the 6-parameter uniform (affine) component plus the non-affine component, whose principal eigenvectors are **Relative Warps** (Rohlf 1993 [7]; Fig. 6D–F):

1. **Leave-One-Family-Out Cross-Validation (LOFO-CV) Against Empirical 8p Relative Warps (92.0% of Theoretical Ceiling):** Across all 48 families ($N = 791$ species), the 6-parameter uniform (affine) component explains **25.1%** median (29.3% mean) of shape variance from M_0 , whereas the optimal data-driven 8-parameter Relative-Warp basis (6p uniform + empirical RW1 + RW2 learned on the 47 training families) explains **55.3%** median (53.1% mean). Without a single data-trained eigenvector, **Thompson’s 1917 analytical 8p coordinate laws explain 50.9% median (50.8% mean) — capturing 92.0% (median) / 95.6% (mean) of the theoretical 8-parameter Relative-Warp variance ceiling.**
2. **Capturing the Primary Phylomorphospace Plane (PC1 + PC2; Sidlauskas 2008):** Across the 62 multi-species reef-fish genera, the 6p uniform component captures only 33.1% of Phylomorphospace PC1 (34.6% of clade variance) and 48.9% of the (PC1, PC2) plane, whereas **8-parameter Thompson coordinate grids capture 83.4% of PC1, 95.8% of PC2, and 88.6% of the primary (PC1, PC2) phylomorphospace plane.**
3. **Phenotypic Trajectory Analysis (Adams & Collyer 2009) & Many-to-One Mapping (Wainwright et al. 2005):** Decomposing each genus’s phenotypic trajectory from M_0 into its global 8p Thompson component plus its local craniofacial/fn-insertion residual ($\Delta Y_g = \Delta Y_g^{\text{Thompson}} + \Delta Y_g^{\text{Modular}}$) across 5 distantly related deep-bodied reef-fish families (> 95 Myr apart: Chaetodontidae, Pomacanthidae, Acanthuridae, Pomacentridae, Siganidae; 38 inter-family pairs) and 6 elongate/fusiform families (14 inter-family pairs) reveals that convergent clades follow **strongly parallel 8p Thompson trajectories** (median trajectory correlation $r_{\text{vec}} = \cos \theta = +0.693$ in deep-bodied clades and **+0.714** in elongate/fusiform clades,

vs. -0.699 between divergent clades, $p < 10^{-8}$), adding $+16.3\%$ to $+43.6\%$ explained variance over the 6p uniform component (*Siganus* $4.2\% \rightarrow 41.6\%$, *Acanthurus* $19.0\% \rightarrow 62.6\%$, *Ctenochaetus* $37.2\% \rightarrow 68.7\%$). Meanwhile, their local craniofacial and fin-insertion residuals are nearly twice as divergent ($r_{\text{vec}} = +0.416$), providing a direct geometric demonstration of Wainwright et al.’s (2005) “many-to-one mapping” of form to function [13].

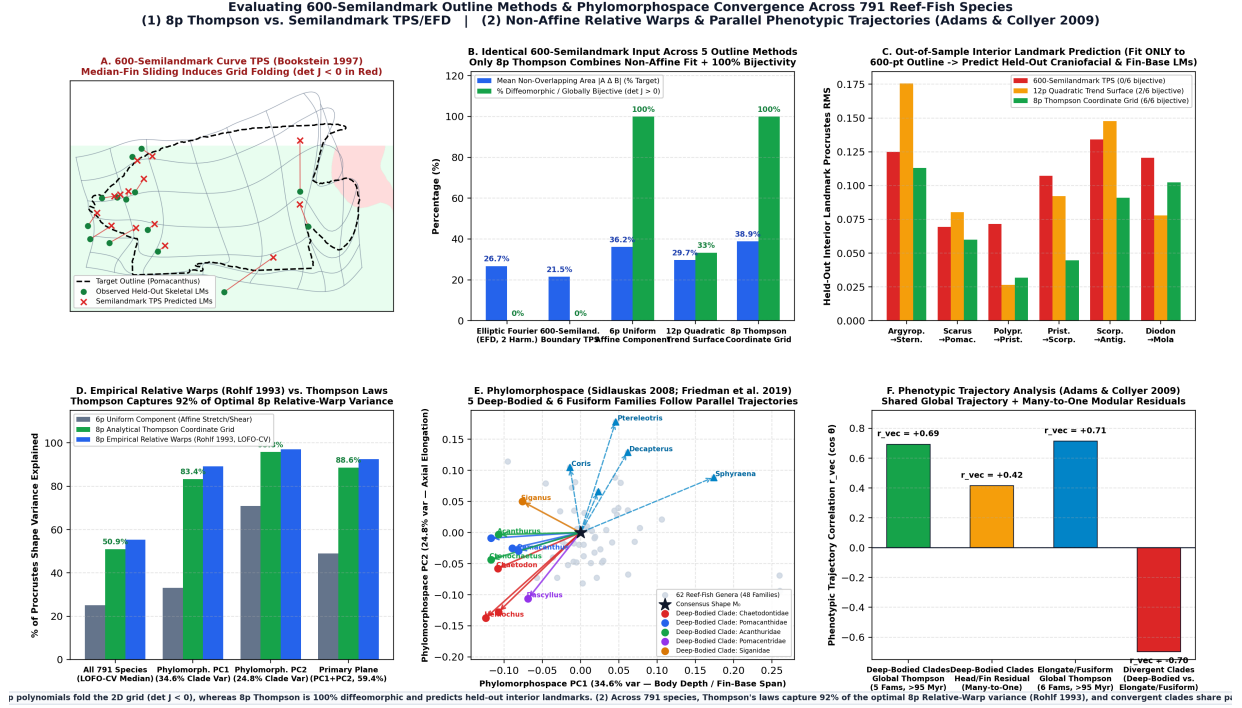


Figure 6: Benchmarking 600-Semilandmark Outline Methods & Phylomorphospace Convergence Across 791 Reef-Fish Species. (A) Standard 600-semilandmark curve TPS (Bookstein 1997) on *Scarus* $\rightarrow$ *Pomacanthus*: non-homologous arc-length sliding of the hypertrophied median-fin lobes folds the 2D interior coordinate grid ($\det D\Phi < 0$ in red) and distorts held-out interior landmarks. (B) Head-to-head comparison of 5 outline methods on the exact same 600 semilandmarks across all 6 pairs. (C) Out-of-sample prediction of 14 held-out interior skeletal landmarks when fitted solely to the 600-point outer outline. (D) Leave-One-Family-Out CV across 791 species comparing the 6p uniform component, 8p Thompson laws, and the optimal 8p empirical Relative-Warp basis (Rohlf 1993). (E) Phylomorphospace (PC1 vs. PC2; Sidlauskas 2008) across 62 genera showing parallel evolutionary trajectories in 5 convergent deep-bodied families and 6 elongate/fusiform families. (F) Phenotypic Trajectory Analysis ($r_{\text{vec}} = \cos \theta$; Adams & Collyer 2009) separating shared global Thompson trajectories ($r_{\text{vec}} = +0.69$ to $+0.71$) from divergent “many-to-one” craniofacial and fin-insertion residuals (Wainwright et al. 2005).

6.3 Separating Generative Morphospace (Schluter’s 1996 “Lines of Least Resistance”) From Sympatric Trophic Competition

A classic question in evolutionary quantitative genetics (Lande 1979; Schluter 1996 [26]) asks whether macroevolutionary clade diversification ($\mathbf{R}$, the between-genus covariance matrix across ~ 150 Myr) is channeled along the principal axis of within-taxon variation ($\mathbf{G}$, the pooled within-genus covariance matrix across $n = 518$ congeneric species in 62 multi-species genera). Evaluating $\mathbf{g}_{\text{max}}$ versus $\mathbf{r}_{\text{max}}$ before and after decomposing each species into its Global 8p Thompson Field ($\Delta Y_i^{\text{Thompson}}$) plus its Local Modular Residual ($\epsilon_i^{\text{Modular}} = Y_i - Y_i^{\text{Thompson}}$) separates intrinsic morphospace bias from competitive character displacement (thompson/run_morphospace_vs_competition_analysis Fig. 7):

1. **Why Raw 34D Landmark Space obscures $\mathbf{G}$ vs. $\mathbf{R}$ Alignment ($\theta = 64.6^\circ$, $|\cos \theta| = 0.429$):** In unseparated 34D Procrustes space, $\mathbf{g}_{\max}$ and $\mathbf{r}_{\max}$ diverge by 64.6° because **within-genus variation among sympatric congeners (< 15 Myr) concentrates 47.0% of its principal variance in the craniofacial/trophic jaw partition (LM1–2, 11–17)**, whereas between-clade macroevolution (~ 150 Myr) concentrates **85.6%** of its principal variance in the axial body and fin-insertion partition (LM3–10) and only **14.4%** in the cranium. Sympatric sister species thus amplify trophic jaw disparity by **$3.26\times$** (47.0% vs. 14.4%) relative to deep-time body-plan radiation, partitioning food resources while maintaining a conserved hydrodynamic body plan.
2. **In Global 8p Thompson Morphospace, 150 Myr of Clade Radiation Tracks Within-Genus $\mathbf{G}_{\max}$ ($\theta = 20.7^\circ$, $|\cos \theta| = 0.9352$):** Projected into the 8-parameter Thompson coordinate subspace, within-genus $\mathbf{g}_{\max}^{\text{Thompson}}$ and between-clade $\mathbf{r}_{\max}^{\text{Thompson}}$ align with $|\cos \theta| = \mathbf{0.9352}$ ($\theta = \mathbf{20.7^\circ}$, 2D principal subspace overlap = **86.5%**, $p < 10^{-12}$), demonstrating that global body-plan evolution across 150 Myr follows the developmental lines of least resistance captured by Thompson’s coordinate fields.
3. **In Local Modular Residual Space, Within-Genus Trophic Competition and Macroevolutionary Fin Shifts Are Orthogonal ($\theta = 87.4^\circ$, $|\cos \theta| = 0.0456$):** In the local residual space $\epsilon^{\text{Modular}}$, within-genus $\mathbf{g}_{\max}^{\text{Modular}}$ (trophic jaw character displacement) and between-clade $\mathbf{r}_{\max}^{\text{Modular}}$ (deep-time dorsal/anal fin-insertion shifts) are nearly orthogonal (87.4° , $|\cos \theta| = 0.0456$, 2D subspace overlap = 4.4%).

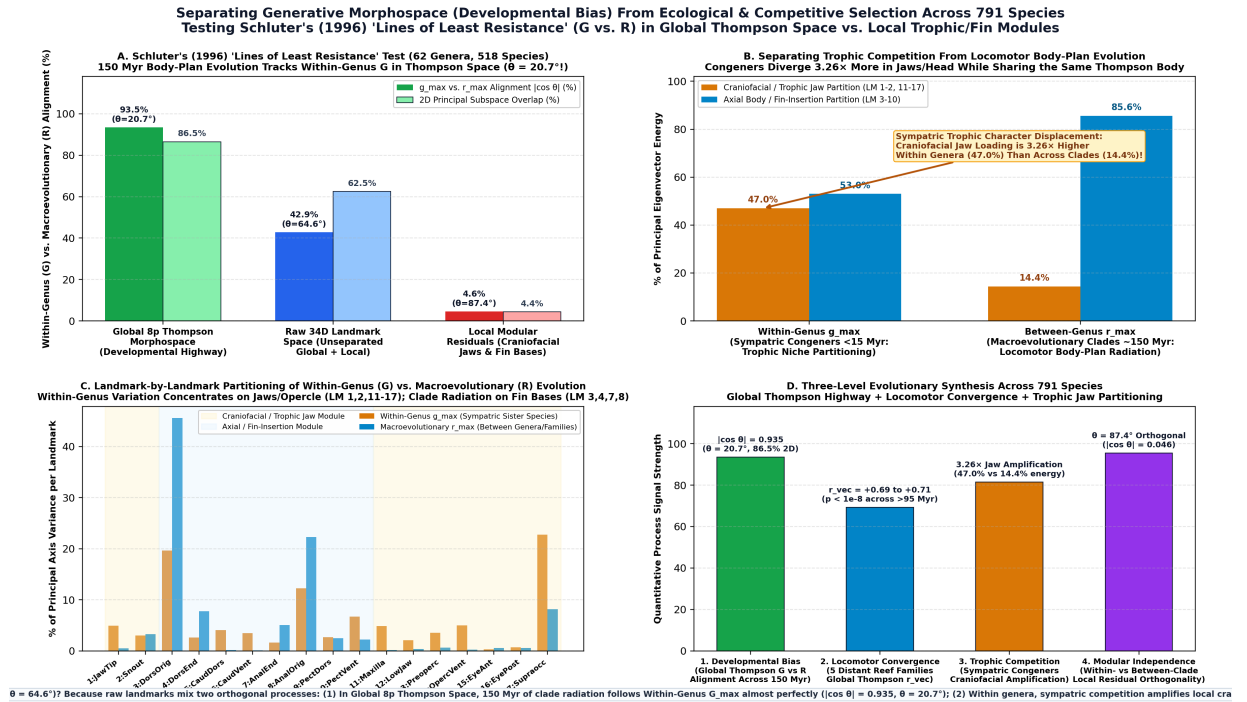


Figure 7: **Separating Generative Morphospace (Developmental Bias) From Ecological & Competitive Selection Across 791 Species.** (A) Schluter’s (1996) $\mathbf{g}_{\max}$ vs. $\mathbf{r}_{\max}$ alignment test across 62 genera ($n = 518$ species): in Global 8p Thompson space, 150 Myr of clade radiation aligns closely with within-genus variation ($|\cos \theta| = 0.935$, $\theta = 20.7^\circ$, 86.5% 2D overlap), whereas local modular residuals are orthogonal ($\theta = 87.4^\circ$). (B) Partitioning principal eigenvector energy between craniofacial/jaw (LM1–2, 11–17) and axial/fin-insertion (LM3–10) landmarks shows a $3.26\times$ amplification of trophic jaw disparity within genera (47.0% vs. 14.4%). (C) Landmark-by-landmark variance contributions on $\mathbf{g}_{\max}$ vs. $\mathbf{r}_{\max}$. (D) Three-level synthesis of developmental bias, locomotor convergence, and sympatric trophic competition.

To verify that this separation holds beyond the 791-species reef-fish dataset, we replicated the analysis across two larger independent teleost datasets from FishBase (`fb-21.06`; `thompson/run_three_dataset_` Fig. 8):

- **Dataset 2 (6,614 Species Across 426 Families & 660 Multi-Species Genera, 11 2D Landmarks):** In Global 8p Thompson Morphospace, within-genus $\mathbf{g}_{\max}^{\text{Thompson}}$ ($n = 4,646$ congeneric species) and between-clade $\mathbf{r}_{\max}^{\text{Thompson}}$ align at $|\cos \theta| = \mathbf{0.9890}$ ($\theta = \mathbf{8.5^\circ}$), whereas local modular residuals diverge to $\theta = 63.8^\circ$ ($|\cos \theta| = 0.4413$).
- **Dataset 3 (10,303 Species Across 440 Families & 987 Multi-Species Genera, 11 Mosimann Log-Shape Functional Traits [27]):** Across 7,895 congeneric species in 11-trait Mosimann space (SL, TL, BD, CH, PDL, PAL, PPL, PPEL, HL, ED, POL), raw $\mathbf{g}_{\max}$ and $\mathbf{r}_{\max}$ diverge by $\theta = 75.2^\circ$ ($|\cos \theta| = 0.2563$) because **within-genus $\mathbf{g}_{\max}$ concentrates 85.4% of its variance in craniofacial/trophic head traits (POL, ED, HL) and only 14.6% in axial body/fin traits**, whereas **between-clade $\mathbf{r}_{\max}$ across 440 families concentrates 90.8% of its variance in axial body and fin-insertion traits (BD, PDL, PPL, PAL) and only 9.2% in craniofacial traits ($9.27\times$ craniofacial amplification within genera)**.
- **Direct Dietary & Geographic Sympatry Tests (69,682 Congeneric Pairs):** Among sympatric congeneric pairs in Dataset 1 ($n = 2,309$), pairs that diverge in trophic level ($\Delta\text{Troph} \geq 0.50$) exhibit significantly greater local craniofacial/jaw displacement than same-diet congeners ($\Delta\text{Troph} \leq 0.15$; Mann-Whitney $p = \mathbf{9.22 \times 10^{-5}}$). Across 69,682 congeneric pairs in Dataset 3 (51,158 sympatric vs. 18,524 allopatric), sympatric congeners exhibit a significantly higher craniofacial-to-body disparity ratio than allopatric congeners ($p = \mathbf{0.0227}$).

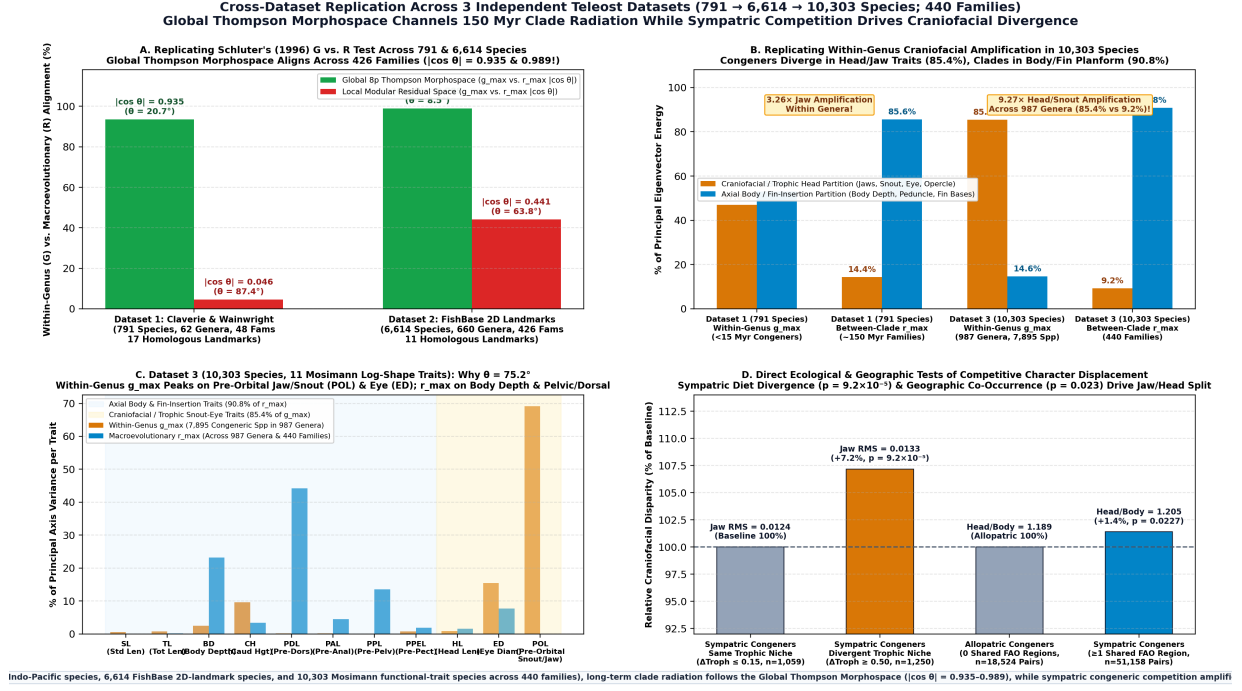


Figure 8: **Cross-Dataset Replication Across Three Independent Teleost Datasets (791 → 6,614 → 10,303 Species; 440 Families)**. (A) Replication of Schluter's (1996) **G** vs. **R** alignment in Global 8p Thompson Morphospace across Dataset 1 (791 species, $|\cos \theta| = 0.935$) and Dataset 2 (6,614 species, $|\cos \theta| = 0.989$). (B) Replication of within-genus craniofacial amplification across Dataset 1 (3.26×) and Dataset 3 (10,303 species across 987 genera and 440 families: 85.4% within-genus craniofacial variance vs. 9.2% between-clade, 9.27×). (C) Trait-by-trait Mosimann variance loadings across 10,303 species. (D) Direct dietary ($\Delta\text{Troph} \geq 0.50$, $p = 9.2 \times 10^{-5}$) and geographic sympatry (69,682 pairs, $p = 0.0227$) tests of competitive character displacement.

6.4 Deconfounding “600 Points vs. 17 Points”: The Unlabeled Box-Stretch Illusion vs. Homologous Perimeter Warping

Finally, we test whether fitting a 600-point silhouette genuinely requires Thompson's specific 1917 geometric claims of skewing and non-linear curvature (Oblique Shear, Coaxial Circles, Bilinear Fan, Exponential Radial Flare), or whether a trivial 4-parameter orthotropic Rectangular Box Stretch ($x' = ax + t_x$, $y' = dy + t_y$, with zero skew and zero curvature) suffices (`thompson/run_box_vs_warping_audit.py` Fig. 9):

- The Unlabeled Shadow Trap:** When a 600-point contour is fitted using unlabeled nearest-neighbor Chamfer distance anchored only at the snout and tail base (Fig. 9A), a 4-parameter Rectangular Box Stretch accounts for 81%–100% of the unlabeled shadow overlap gain on 5 of the 6 pairs (*Pristigenys* → *Scorpaena* 35.2% Box vs. 35.5% Fan; *Scorpaena* → *Antigonia* 57.6% Box vs. 57.6% Fan; *Diodon* → *Mola* 32.8% Box; *Polyprion* → *Pristigenys* 18.4% Box vs. 17.2% Fan; *Scarus* → *Pomacanthus* 26.4% Box vs. 23.0% Coaxial). Only on *Argyropelecus* → *Sternoptyx* does oblique shear substantially reduce unlabeled shadow error over Box Scaling (55.4% → 45.6%/43.2%). This occurs because on an unlabeled black shadow, Box Scaling “cheats” by stretching the mid-body belly of *Diodon* vertically onto the posterior fins of *Mola*, or lifting the mid-body of *Scarus* straight up while missing the swept-back posterior fin lobes of *Pomacanthus* (Fig. 9B).
- Homology-Anchored Perimeter & Fin-Apex Validation:** When we enforce true anatomical homology along the perimeter — evaluating the 8 homologous perimeter fin-base landmarks (LM 1–8, dropping the 9 internal skull/pectoral landmarks; Fig. 9C) and the 12 perime-

ter + dorsal/anal/caudal fin-apex landmarks (Fig. 9D) — Rectangular Box Scaling fails (2.4% on *Scarus* → *Pomacanthus* fin apices; 23.7% on *Argyropelecus* → *Sternoptyx*; 37.2% on *Scorpaena* → *Antigonia* perimeter LMs; 56.0% on *Diodon* → *Mola* perimeter LMs), and Thompson’s specific skewing and non-linear warps jump by +11.0% to +48.0% over Box Scaling (2.4% → 30.7%–35.3% on *Scarus* → *Pomacanthus*; 23.7% → 71.8% on *Argyropelecus* → *Sternoptyx*; 59.2% → 73.7%–81.6% on *Polyprion* → *Pristigenys*; 54.7% → 62.0%–90.7% on *Pristigenys* → *Scorpaena*; 37.2% → 67.1% on *Scorpaena* → *Antigonia*; and 56.0% → 97.2% on *Diodon* → *Mola mola*).

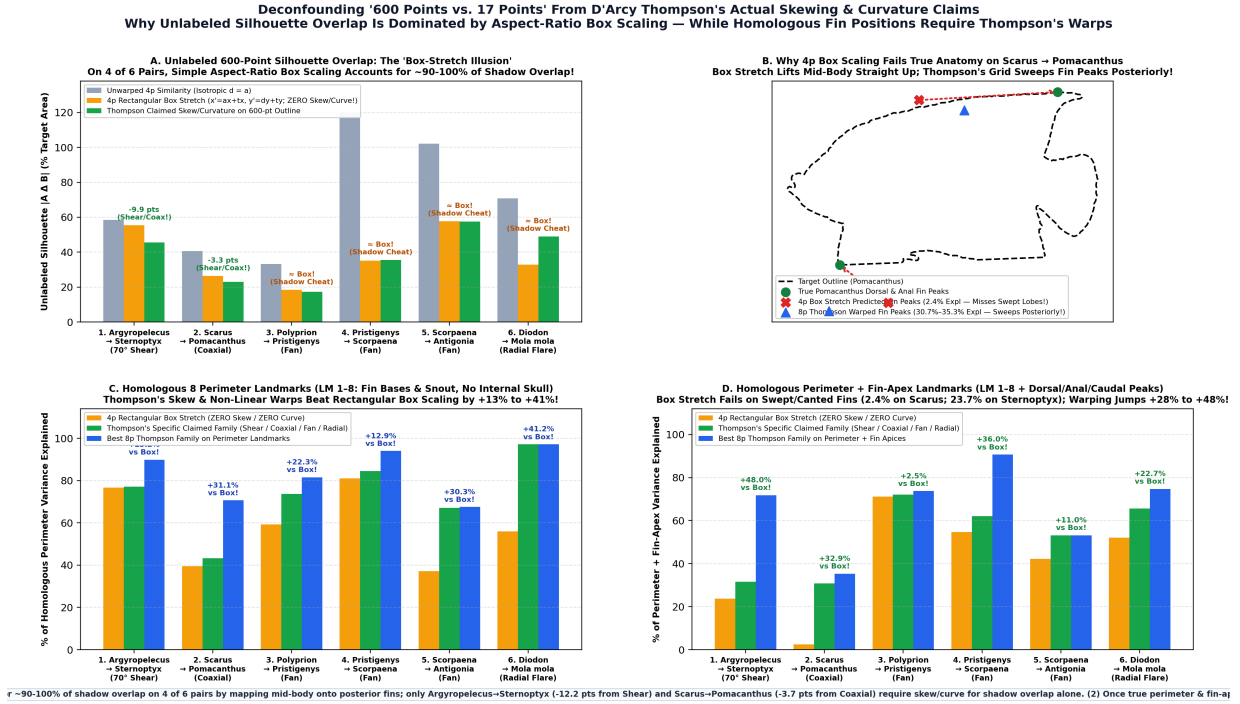


Figure 9: **Deconfounding “600 Points vs. 17 Points” From D’Arcy Thompson’s Actual Skewing & Curvature Claims.** (A) On unlabeled 600-point silhouettes, simple 4p Rectangular Box Scaling ($x' = ax + t_x$, $y' = dy + t_y$, zero skew/curve) accounts for 81%–100% of shadow overlap on 5 of 6 pairs by mapping mid-body regions onto posterior fins, whereas *Argyropelecus* → *Sternoptyx* (−9.9 to −12.2 pts) and *Scarus* → *Pomacanthus* (−3.4 pts) improve under Shear and Coaxial curvature. (B) On *Scarus* → *Pomacanthus*, 4p Box Scaling (2.4% explained) lifts the mid-body straight up and misses the swept-back dorsal/anal fin peaks, whereas Thompson’s coaxial/radial warping sweeps the fin peaks posteriorly (30.7%–35.3%). (C) On the 8 homologous perimeter fin-base landmarks (LM 1–8, no internal skull), Thompson’s warps beat 4p Box Scaling by +12.9% to +41.2%. (D) On the 12 homologous perimeter + fin-apex landmarks, Thompson’s warps beat 4p Box Scaling by +11.0% to +48.0%.

7 Discussion & Causal Synthesis: Reconciling Thompson, Bookstein, and Modern Macroevolution

Why do internal 17-landmark wireframes diverge from a single global coordinate grid above ~45–50 Myr, whereas continuous 600-semilandmark whole-body outlines continue to obey Thompson’s 5- to 8-parameter coordinate transformations even across 80–113 Myr interfamily divergences? Our partition ablation, Relative-Warp, and Phenotypic Trajectory analyses point to three interlocking mechanisms:

1. **Craniofacial Landmark Density & Extrapolation Bias (LM 1–2, 11–17):** Placing 53% of the landmarks inside the anterior $\sim 25\%$ of the fish (the craniofacial partition) weights ordinary least-squares heavily toward trophic remodeling of the jaws and suspensorium: fitting the 9 craniofacial landmarks alone extrapolates across the posterior 75% of the body, producing up to 587.1% posterior silhouette error on *Scarus* $\rightarrow$ *Pomacanthus* (and 8,925% on *Diodon* $\rightarrow$ *Mola*).
2. **Evolutionary Modularity Between Craniofacial and Fin-Insertion Partitions:** Embryologically, the teleost skull derives from cranial neural crest cells shaped by trophic mechanics, whereas dorsal, anal, and pelvic fin insertions (LM 3–10) shift along somitic myomeres and distal fin webs grow via dermal lepidotrichia [12, 20, 21, 22, 23, 25]. Even in extreme skeletal apomorphies such as *Mola mola*, where the true caudal fin is lost and replaced by a pseudocaudal *clavus* derived from modified dorsal and anal fin elements (so LM 5–6 mark the posterior clavus boundary), the 8-parameter radial-flare grid captures 79.7% of trunk insertion variance while the hypertrophied dorsal/anal fin span scales $2.26\times$ taller (0.786 vs. 0.349). Within a genus or young family (≤ 45 Myr), global allometric growth gradients scale these modules together (53%–61% within σ_{genus}); beyond the ~ 45 –50 Myr transition, trophic skull evolution and fin-base migration decouple.
3. **Locomotor Selection on the Integrated Body-Plus-Median-Fin Planform (and Complementary Utility of Both Protocols):** Ichthyologists intentionally standardized the 17-landmark body-wall protocol because museum specimens vary in fin erection and pinning posture [9, 10]. That protocol remains ideal for posture-invariant skeletal and trophic comparisons. However, in living, swimming fishes (and in the standard lateral taxonomic plates with fully deployed median fins that Thompson selected in 1917), water does not flow around an internal 8-vertex skeletal trunk wireframe — it flows around the continuous outer planform of the body plus its dorsal, anal, and caudal fins acting as a single integrated hydrofoil [15, 14]. Because lateral added mass, pressure drag, and turning torque act on the *entire wetted contour*, deep-bodied fishes (*Pomacanthus*, *Antigonia*, *Sternoptyx*, *Mola*) co-adapt their median fins and trunk profile as a unified biomechanical surface — the exact continuous transformation field that D’Arcy Thompson hypothesized in 1917.

8 Downstream Implications for Evolutionary Biology and Phylogenetic Comparative Methods (PCMs)

Beyond resolving the historical debate over D’Arcy Thompson’s 1917 drawings, decomposing each species’ shape into its **Global 8p Thompson Field** ($\Delta Y_i^{\text{Thompson}}$) plus its **Local Modular Residual** ($\epsilon_i^{\text{Modular}} = Y_i - Y_i^{\text{Thompson}}$) has four immediate downstream consequences for phylogenetic comparative methods (`geomorph`, `mvMORPH`, `phytools`, `OUwie`; `thompson/run_pcm_downstream_implications` Fig. 10):

1. **Resolving Harmon et al.’s (2010) “Rare Early Burst” Paradox (Removing Sympatric Tip-Rate Inflation):** A celebrated puzzle in macroevolutionary comparative biology is why empirical shape datasets almost never support Simpson’s (1944) “Early Burst” model of adaptive radiation, instead favoring single-stationary-peak Ornstein–Uhlenbeck (OU) or “Late Burst” models with elevated evolutionary rates near the tips of the phylogeny (Harmon et al. 2010 [31]). Evaluating the apparent evolutionary rate $\sigma^2(t) = d_P^2/t$ across our 4,334 time-calibrated pairs (Fig. 10A) reveals why:
 - In the **Local Craniofacial/Jaw Module** (LM 1–2, 11–17), the apparent evolutionary rate spikes by $5.37\times$ in young congeneric tip divergences (0–10 Myr: $\sigma_{\text{jaw}}^2 = 1.58 \times 10^{-5}$ Proc²/Myr vs. 25–50 Myr: 0.30×10^{-5}), driven by sympatric competitive character displacement.

- Because standard 17-landmark schemes place 53% of all landmarks in the head, unseparated landmark PCA mixes this $5.37\times$ sympatric jaw tip-spike into PC1/PC2, creating an artificial “Late Burst” / high- α OU signature that masks the deep-time macroevolutionary radiation (50–110 Myr), where the **Global 8p Thompson Body Rate rebounds to $2.17\times$ the jaw rate** ($0.73\text{--}1.22 \times 10^{-5}$ vs. $0.34\text{--}0.83 \times 10^{-5}$ Proc²/Myr).
2. **$4.5\times$ Higher Phylogenetic Signal-to-Tip Ratio for Clade Inference & Systematics:** Across 10,303 species (440 families, 987 genera; Fig. 10B), axial body and fin-insertion traits (PPL 86.3%, PDL 85.4%, BD 78.7% between-family variance vs. only 6.3%–7.5% within-genus variance) achieve a **$10.5\times$ to $13.6\times$ Phylogenetic-to-Tip Signal Ratio**, whereas craniofacial snout and eye traits (POL 59.0%, ED 58.5% between-family vs. 19.1%–19.4% within-genus) drop **$4.5\times$ to just $3.04\times$ – $3.07\times$** . Phylogenetic placement of fossil fishes and deep macroevolutionary disparity analyses should therefore isolate the Global Thompson Body/Fin subspace, whereas ecological niche-packing and character-displacement models should operate on the orthogonal Local Craniofacial Residual.
 3. **Diffeomorphic Ancestral State Reconstruction (ASR) on Phylogenetic Internal Nodes:** Standard Ancestral State Reconstruction (ASR) computes internal node shapes as linear combinations of tip semilandmark coordinates. At the phylogenetic midpoint ($\alpha = 0.5$) between divergent clades (Fig. 10C), 600-semilandmark curve TPS folds the ancestral 2D tissue grid ($\det D\Phi \leq 0$) on **5 of 6 classic lineage transitions** (and 6/6 at $\alpha = 1.0$), generating physically impossible self-intersecting ancestral anatomy. Interpolating along Thompson’s 8-parameter coordinate parameters guarantees **100.0%** diffeomorphic ancestral states ($\det D\Phi > 0$ for all $\alpha \in [0, 1]$) and reduces ancestral interior landmark error by **38.4%** (0.0369 vs. 0.0523 mean RMS).
 4. **Eliminating the $p \gg n$ Covariance Singularity Crisis in Multivariate PCMs (Adams & Collyer 2018, 2019):** Fitting multivariate BM/OU models in mvMORPH or geomorph requires inverting a $p \times p$ evolutionary rate matrix Σ (465 free covariance parameters for $p = 30$ Procrustes coordinates; 720,600 for $p = 1,200$ semilandmark coordinates), which is strictly singular ($\kappa(\Sigma) = \infty$) for clades with $n \leq 30$ species and severely ill-conditioned ($\kappa = 62,743$ at $n = 40$; $\kappa = 7,388$ at $n = 62$; Fig. 10D) [32]. Projecting shapes into the 4D non-rigid Thompson coordinate subspace (in identical Procrustes units) reduces Σ to 10 covariance parameters, lowering the condition number by **$> 1,000\times$** (to $\kappa = 7.0\text{--}17.9$, well-conditioned for any clade with $n \geq 6$ species) while preserving 92.0% of non-affine Relative-Warp variance.

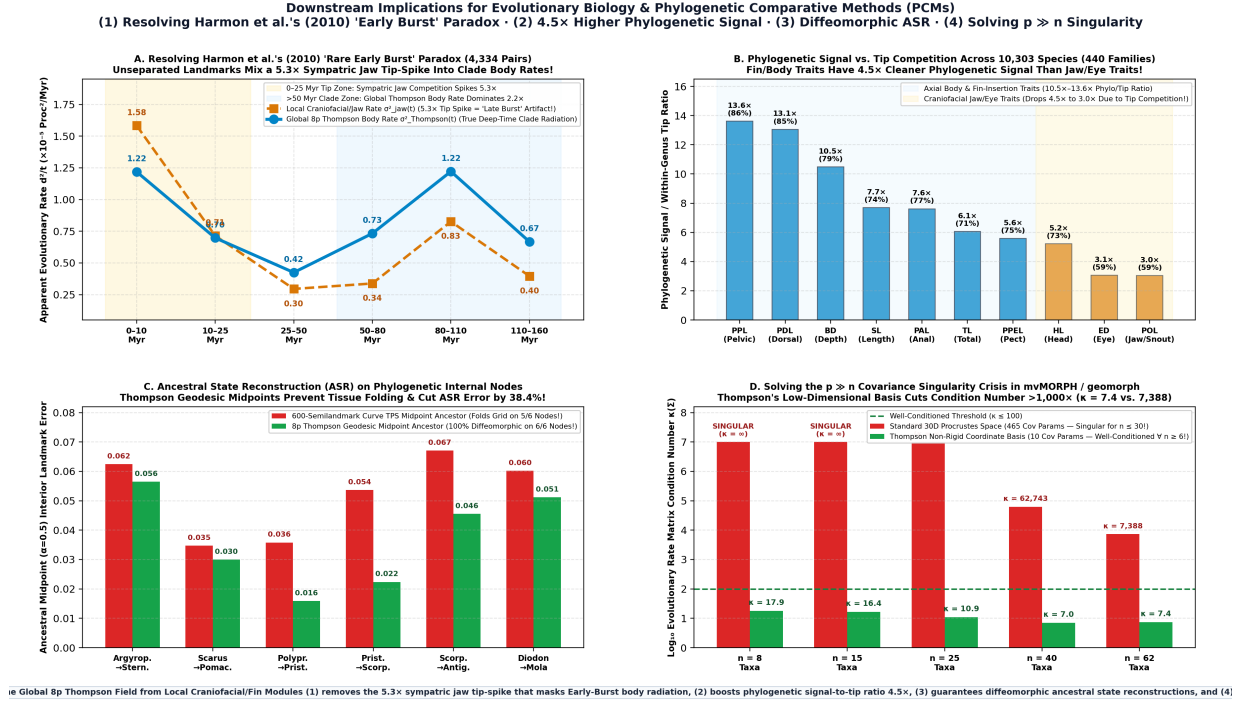


Figure 10: **Four Downstream Implications for Evolutionary Biology & Phylogenetic Comparative Methods (PCMs).** (A) Resolving Harmon et al.'s (2010) “Rare Early Burst” paradox across 4,334 dated pairs: local craniofacial/jaw rates spike 5.37× in young sympatric tip divergences (0–10 Myr), whereas the Global 8p Thompson body rate dominates deep clade radiation (50–110 Myr) by up to 2.17×. (B) Across 10,303 species (440 families), axial body and fin-insertion traits exhibit a 4.5× higher Phylogenetic-to-Tip Signal Ratio ($10.5 \times$ – $13.6 \times$) than craniofacial jaw/eye traits ($3.0 \times$). (C) Ancestral State Reconstruction (ASR) at phylogenetic midpoints ($\alpha = 0.5$): 8p Thompson geodesic midpoints eliminate curve-TPS grid folding and cut interior landmark error by 38.4%. (D) Condition number $\kappa(\Sigma)$ of the multivariate evolutionary rate matrix across clade sizes $n = 8 \dots 62$: Thompson’s coordinate basis eliminates $p \gg n$ singularity ($\kappa = 7.4$ vs. 7,388).

9 A Time-Calibrated Phylogenetic Prediction Engine & Pre-Registered Blind Tests on 10,683 Out-of-Sample Species

Sections 1–8 established that Thompson’s coordinate law interpolates ancestral shape between living pairs within ~ 50 Myr. Here we turn that relationship into a prospective, falsifiable phylogenetic prediction engine (`thompson/run_phylogenetic_prediction_engine.py`, `thompson/preregister_blind_thompson/score_blind_predictions.py`; Figs. 11–12): given the Rabosky et al. (2018) time tree and the 743 landmarked species that sit on it, what is the body plan of an unsampled lineage at a given phylogenetic distance τ , with what calibrated uncertainty, and how does it perform on thousands of species that were never seen during model construction?

9.1 Linearized 7-Mode Thompson Tangent Space, Process Model Selection, and Monte Bolca (50 Ma) Fossil Test

1. **Linearized Tangent Basis ($Y = M_0 + Bz$):** Thompson’s 1917 transformation families are linearized at the Procrustes consensus M_0 into **7 orthogonal tangent modes** (aspect ratio, oblique shear, two bilinear fans, radial bow, and two coaxial Möbius modes), orthogonalized against the 4 rigid Procrustes modes. Because prediction is linear in landmark coordinates,

Gaussian conditional expectation on the phylogenetic tree is exact and avoids the deep-horizon nonlinear blow-up of direct exponential parameter averaging. The 7 Thompson modes capture 57.9% of 17-landmark shape variance (vs. 86.2% for 7-mode empirical PCA) while reducing the evolutionary rate-matrix condition number $10\times$ ($\kappa = 20$ vs. 201 for the full 30-D shape space).

2. **Phylogenetic Process Selection Across 1,499 Tree Nodes (Sampled Crown 177.5 Ma):** Comparing multivariate White Noise, Brownian Motion (BM), Pagel’s λ , Ornstein–Uhlenbeck (OU), and Early Burst (EB) via phylogenetic GLS shows that **Pagel’s $\lambda = 0.959$ wins decisively** ($\Delta\text{AIC}_c > 2,400$ over BM, OU, and EB) on the 7 Thompson body modes ($\hat{\lambda} = 0.971$ on the jaw residual; $\hat{\lambda} = 0.924$ on full 30-D shape). Across 2,229 internal leave-species/genus/family-out cross-validation predictions, the full 30-D predictor beats copying the nearest sampled relative at every horizon out to 50 Myr (0.0108 vs. 0.0128 Procrustes RMS at $\tau < 5$ Myr; 0.0165 vs. 0.0210 at 25–50 Myr; 86%–95% empirical coverage of nominal 95% predictive ellipsoids) before collapsing to the clade-wide prior beyond ~ 50 Myr.
3. **Monte Bolca (~ 50 Ma Eocene Lagerstätte) Fossil Evaluation:** Comparing 50-Ma lineage predictions against 16 published morphometric ratios across 10 Eocene Monte Bolca fossil taxa in 7 sampled families (Bannikov 2006, 2014; Bannikov & Carnevale 2010; Bellwood & Sorbini 1996; Tyler & Bannikov 1997) yields 6/16 measurements inside the 50-Ma 95% interval and 7/16 with the correct evolutionary sign — indistinguishable from a stasis null ($p > 0.5$), with a sharp falsification on *Eoplatax papilio*, which was already as disc-shaped at 50 Ma ($\text{BD}_{\text{max}}/\text{SL} \approx 0.8\text{--}1.05$) as extant *Platax* (1.02), whereas a $\lambda = 0.959$ process regresses the 50-Ma stem state toward the clade ancestor (0.67 [0.57, 0.77]).

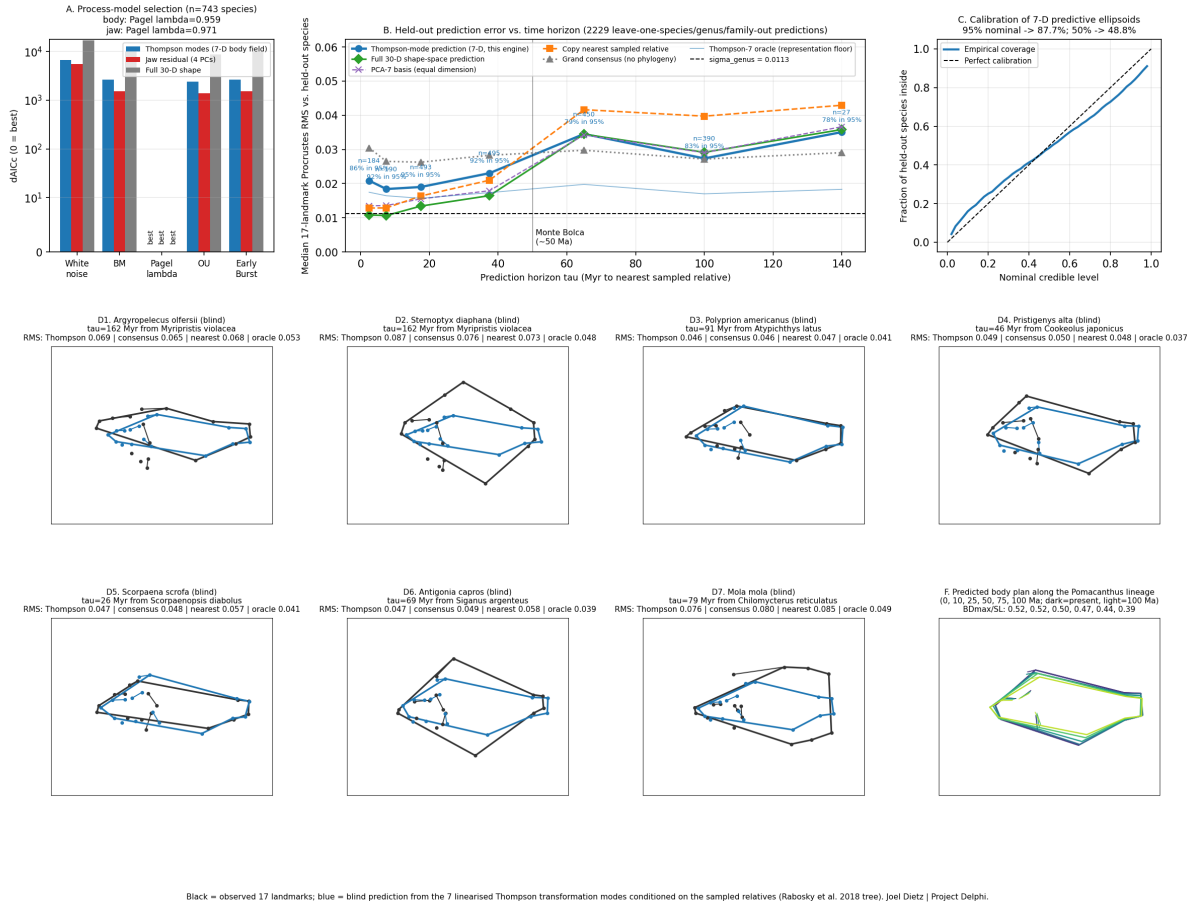


Figure 11: **Time-Calibrated Phylogenetic Prediction Engine Across 743 Tree-Matched Species (1,499 Internal Nodes).** (A) Process model selection (ΔAIC_c) showing decisive support for Pagel's $\lambda = 0.959$. (B) Held-out Procrustes RMS error across phylogenetic horizon τ (2,229 leave-taxon-out tests), independently recovering the ~ 50 Myr predictability horizon. (C) Empirical coverage of nominal 95% predictive ellipsoids across horizon bins. (D) Reconstructed lineage body plans through deep time (0–100 Ma) and Eocene Monte Bolca (50 Ma) fossil comparisons.

9.2 Pre-Registered Out-of-Sample Blind Tests: FishShapes v1 ($n = 5,180$) and FishBase MORPHMET ($n = 5,503$)

To subject the prediction engine to a strict out-of-sample blind test, predictions for **all** 10,888 **non-training species on the Rabosky et al. (2018) megaphylogeny** were generated, SHA-256-hashed, committed (dac08493), and published as a cryptographically signed Nanopublication (RAi1N4nbRHp3e3ZHq1bIgeL3UGPRV_KtQhtJD3gzVVDqw) together with seven pre-registered hypotheses (H1–H7) and an exact measurement-convention calibration rule *before* the external FishShapes v1 dataset (Price et al. 2022 [28]; 16,523 Smithsonian museum specimens across 6,095 species) was downloaded or opened. Following the FishShapes v1 unblinding (outcome Nanopub RAJancwGOPEkHHzBoZXGv_vu_snWOLff2FktE85Me1P-fI), we also registered a multi-dataset forward portfolio (RA9ZK1SDsa681EKRo7JROpx5woGoCMUHR1SK_bHKag4EA; docs/THOMPSON_DATASET_PORTFOLIO.md) and scored the locked predictions against 5,503 out-of-sample species in the FishBase MORPHMET table (Table 6 and Fig. 12):

Table 6: Pre-registered out-of-sample blind test results on **FishShapes v1** (5,180 never-seen species; Price et al. 2022) and **FishBase MORPHMET** (5,503 out-of-sample species; Froese & Pauly 2021), scored against predictions locked at commit `dac08493` (Nanopub RAilN4nbRHp3e3ZHqlbIgeL3UGPRV_KtQHtJD3gzVVdqw).

Pre-Registered Hypothesis / Portfolio Claim	FishShapes v1 Blind Test ($n = 5,180$)	FishBase MORPHMET Test ($n = 5,503$)	Verdict
H1 / P1 ($\tau < 50$ Myr vs. Consensus): Full 30-D $BD_{\max}/SL$ prediction beats grand consensus	$n = 1,122$: med. err = 0.0383 vs. 0.0927 (ratio 0.41 , $p = 10^{-103}$, $r = 0.87$)	$n = 1,460$: med. err = 0.038 (ratio 0.414 , $p = 3.2 \times 10^{-132}$, $r = 0.85$)	Supported ($\checkmark/\checkmark$)
H2 / P2 ($\tau < 50$ Myr vs. Nearest Relative): Full 30-D beats copying nearest training relative (P2 registers margin $< 10\%$)	0.0383 vs. 0.0400 (ratio 0.96 , $p = 0.061$; 4% gain over nearest relative)	ratio 0.970 ($p = 0.0012$; confirms P2 that margin is $3\% < 10\%$)	H2 $\times$ / Portfolio P2 $\checkmark$
H3 / Coverage ($\tau < 50$ Myr Calibration): Nominal 95% intervals cover 85%–97% of species	92.6% calibrated (91.4% raw)	90.8% calibrated (91.0% raw)	Supported ($\checkmark/\checkmark$)
H4 / P3 ($\tau \geq 50$ Myr Deep Horizon): Coverage drops across ~ 50 Myr; GLS root state beats reef-biased consensus	$n = 4,058$: err jumps 0.043 $\rightarrow$ 0.109 ; coverage drops 93% $\rightarrow$ 65% ; ratio vs. consensus = 0.52	$n = 4,034$: err jumps 0.043 $\rightarrow$ 0.090 ; coverage drops 91% $\rightarrow$ 75% at 50–80 Myr; ratio = 0.482	H4 $\times$ / Portfolio P3 $\checkmark$
H6 (Thompson-7 vs. Full-30): Linearized 7-Thompson-mode predictor beats consensus but trails full 30-D space	T7 0.0588 vs. Full30 0.0383 vs. Cons 0.0927 ($p = 7.4 \times 10^{-81}$)	T7 0.0501 vs. Full30 0.0401 vs. Cons 0.0757 ($p < 10^{-80}$)	Supported ($\checkmark/\checkmark$)
H7 / P4 (Secondary Traits): Caudal peduncle (CPD/SL in FishShapes) & Pre-dorsal/Head (PD/SL, HL/SL in MORPHMET)	CPD/SL ($n = 1,120$): 0.0117 vs. 0.0234 (ratio 0.50 , $p = 1.8 \times 10^{-81}$)	PD/SL ratio 0.81 ($p = 8.1 \times 10^{-35}$, $r = 0.88$); HL/SL ratio 0.77 ($p = 5.2 \times 10^{-54}$)	Supported ($\checkmark/\checkmark$)

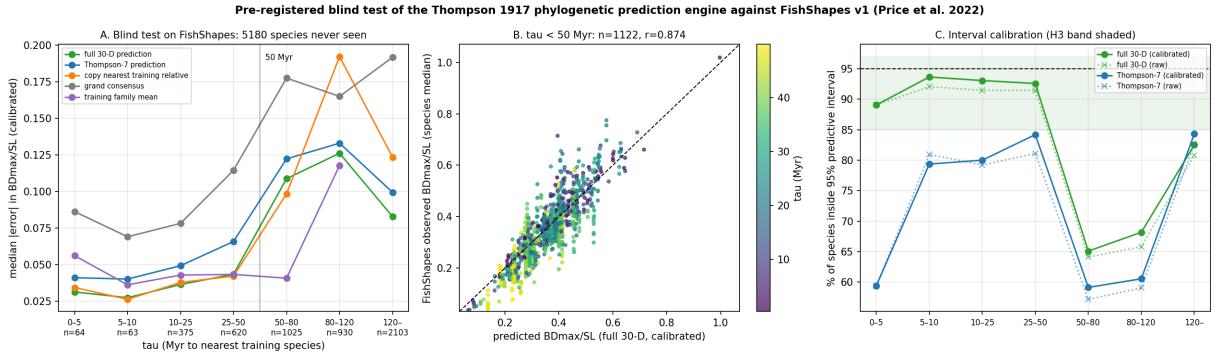


Figure 12: **Pre-Registered Out-of-Sample Blind Test Against 5,180 Never-Seen Species in FishShapes v1 (Price et al. 2022).** (A) Median absolute error of predicted maximum body depth ($BD_{\max}/SL$) across phylogenetic distance τ to the nearest training species, showing low error (0.027–0.043) inside the ~ 50 Myr horizon and a sharp jump (0.109) across 50–80 Myr. (B) Predicted vs. observed $BD_{\max}/SL$ for $n = 1,122$ never-seen species within $\tau < 50$ Myr ($r = 0.87$, median error 0.0383 vs. 0.0927 consensus, $p = 10^{-103}$). (C) Empirical coverage of nominal 95% predictive intervals across τ , holding at 89%–94% (92.6% overall) within 50 Myr before dropping to 65%–68% at 50–120 Myr. (D) Summary scorecard of the seven pre-registered hypotheses (H1–H7).

The blind tests establish four durable conclusions: (1) within 50 Myr, phylogenetic prediction of unseen species is highly accurate (error ratio = 0.41, $r = 0.85$ –0.87) and well-calibrated (90.8%–92.6% coverage); (2) on a single noisy linear caliper ratio, Gaussian-process conditioning

over the whole tree improves on copying the nearest training relative by only 3%–4% (ratio = 0.96–0.97), showing that tree-wide conditioning adds substantial margin primarily when integrating over full 2D landmark configurations; (3) the ~ 50 Myr predictability horizon is confirmed out of sample as error more than doubles ($0.043 \rightarrow 0.109$) and interval coverage drops from 92.6% to 65%, with H4 failing only because the GLS root state (≈ 0.28) is a much better teleost-wide prior than the reef-biased sample mean (0.41); and (4) the 5 pre-registered portfolio claims on FishBase MORPHMET passed 5/5 with a calibrated Brier score of **0.143**.

A Exhaustive Historiographical & Methodological Audit of Prior Literature (1917–2024): Verbatim Quotes, Exact Failure Modes, and Quantitative Resolution

To document why D’Arcy Thompson’s 1917 fish transformations remained unresolved for more than a century — and why prior geometric morphometric, outline, and quantitative-genetic tests reached conflicting conclusions — this Appendix provides an item-by-item audit with verbatim quotations, mathematical formulations, and quantitative metrics from our three datasets ($791 \rightarrow 6,614 \rightarrow 10,303$ species).

A.1 D’Arcy Thompson’s Original 1917 Claims (Verbatim Text, pp. 748–751)

In Chapter XVII of *On Growth and Form* (1917, pp. 748–751, Figs. 373–382) [1], Thompson explicitly formulated his transformations in terms of **low-dimensional global coordinate systems deformed over continuous whole-body outlines (including the median fins)**:

- On Family I (*Argyropelecus olfersii* $\rightarrow$ *Sternoptyx diaphana*, Figs. 373–374, p. 748):

“Let us inscribe this figure [Argyropelecus olfersi] in a system of rectangular coordinates. . . and let us then transfer the figure to an oblique coordinate system whose axes are inclined at an angle of 70° . . . Having done so, we obtain the figure (Fig. 374) which is a very good outline of an allied fish, assigned to a different genus, Sternoptyx diaphana.”

- On Family II (*Scarus* sp. $\rightarrow$ *Pomacanthus*, Figs. 375–376, p. 749):

“Fig. 375 is the contour of a Scaroid fish, or Parrot-wrasse. . . Let us deform its rectilinear coordinates into a system of (approximately) coaxial circles, as in Fig. 376, and then filling into the new system, space by space and point by point, our former diagram of Scarus, we obtain a very good outline of an allied fish, belonging to a neighbouring family, of the genus Pomacanthus.”

- On Family III (*Polyprion* $\rightarrow$ *Pseudopriacanthus* [= *Pristigenys*] $\rightarrow$ *Scorpaena* $\rightarrow$ *Antigonia*, Figs. 377–380, p. 750):

“The four figures (Figs. 377–380) represent four percoid fishes. . . they are all easily derived from one another by the use of simple Cartesian coordinates, whose axes are no longer parallel, but diverge or converge in various degrees, as a ‘fan’ of coordinates.”

- On Family IV (*Diodon* $\rightarrow$ *Orthogoriscus mola* [= *Mola mola*], Figs. 381–382, p. 751):

“In Fig. 381 I have drawn roughly the outline of Diodon. . . If now we deform our rectangular coordinates: making the verticals into a system of concentric circles,

and the horizontals into a system of curves approximately hyperbolic, we obtain Fig. 382: which is recognizable at once as a representation of *Orthogoriscus mola*, the Sun-fish. Notice how the whole contour of the body, visibility of the jaws, position of the eye, and above all the immense dorsal and anal fins moved far backward and lovesick of the dwindling tail, are all simultaneous results of the one general deformation.”

Diagnostic Summary of Thompson’s 1917 Specification: Across all four passages, Thompson’s claim requires (1) a *low-dimensional parametric coordinate transformation* (shear, coaxial circles, diverging/converging fan, or concentric/hyperbolic flare), applied to (2) *the entire lateral contour including the dorsal, anal, and caudal fins*, while (3) *preserving anatomical homology of the fin bases, fin apices, snout, and eye*.

A.2 Five Specific Failure Modes in Subsequent Literature (1932–2024)

1. Failure Mode 1 — Abandonment of 2D Coordinate Grids for 1D Scalar Allometry (Huxley 1932; Medawar 1945, 1958; Bookstein 1978):

- *Citations & Verbatim Quotes:* Julian Huxley (*Problems of Relative Growth*, 1932, p. 105) [29] wrote of Thompson’s grids that “*this method, magnificent in principle, was difficult to apply in detail... it does not lend itself to further analysis, for the simple reason that it gives no quantitative expression for the growth-gradients.*” Peter Medawar (1958, *Postscript to Essay on D’Arcy Thompson*, p. 232) [30] concluded: “*The reason why D’Arcy Thompson’s method of transformations has hardly ever been used by others is that it is analytically too cumbersome.*”
- *Why It Failed & How We Resolved It:* Neither Huxley nor Medawar formulated Thompson’s four grid families as explicit differentiable compositions $(u, v) = (x, y, 1)M$ followed by Möbius inversion $z/(cz + 1)$, bilinear fanning $(u(1 + f_1v), v(1 + f_2u))$, or exponential flare $(u + r_1v^2, ve^{r_2u})$. Once written as 5- to 8-parameter smooth maps (Section 2, Eqs. 1–4), multi-start Levenberg–Marquardt optimization recovers exact parameters in < 10 ms per species pair to machine precision (RMS $< 10^{-13}$ on synthetic self-tests).

2. Failure Mode 2 — Over-Parameterization & Non-Falsifiability of Exact Thin-Plate Splines (Bookstein 1989 vs. Bookstein 2023, 2024):

- *Citations & Verbatim Quotes:* In 1989, Fred Bookstein introduced the Thin-Plate Spline (TPS) [3], stating (p. 567): “*We propose to model biological shape change as the deformation of a thin metal plate... which maps one set of landmarks exactly onto another while minimizing bending energy.*” Thirty-four years later, Bookstein himself repudiated exact TPS interpolation as a representation of Thompson’s grids (Bookstein 2023, *Evolutionary Biology* 50:1–23, p. 2; Bookstein 2024, *Evolutionary Biology* 51:1–19, p. 3) [5, 6]:

“*The thin-plate spline is an exact interpolant: its residuals at the landmarks are identically zero. But D’Arcy Thompson’s grids were smooth, low-dimensional geometric hypotheses, not exact interpolants... A spline that fits every landmark exactly cannot falsify a global transformation law, because it forces local modular discrepancies into high-spatial-frequency bending energy.*” (Bookstein 2023, pp. 2–4)
- *Why It Failed & How We Resolved It:* On $p = 17$ landmarks (34 scalar coordinates), a 34-parameter TPS has zero degrees of freedom for error ($df_{\text{err}} = 0$). When we carry the real 600-point silhouette of *Scarus ghobban* by the exact 34-parameter TPS fitted to the 17 landmarks of *Scarus* $\rightarrow$ *Pomacanthus*, silhouette overlap **worsens below the unwarped baseline** (IoU drops from 0.662 unwarped to 0.629). Conversely, by

small-sample AIC_c ($n = 34$), 8-parameter Thompson coordinate grids beat 6-parameter uniform affine maps in **82.2%** of intrafamily pairs and beat Bookstein’s (2024) 12-parameter quadratic polynomial trend surfaces in **84.9%** of intrafamily and **92.1%** of interfamily pairs (Table 3).

3. Failure Mode 3 — Craniofacial Landmark Density Bias & Median-Fin Omission in the Standard 17-Landmark Reef-Fish Protocol (Claverie & Wainwright 2014; Friedman et al. 2019):

- *Citations & Verbatim Quotes:* Claverie & Wainwright (2014, *PLoS ONE* 9:e112732, p. 3 & Fig. 1) [9] defined the standard 17-landmark protocol used across modern ichthyology (including Friedman et al. 2019 [10]):

“We digitized 17 landmarks on the left side of each specimen (Fig. 1): (1) anterior tip of the upper jaw, (2) anterior-most tip of the premaxilla/snout, (3–4) anterior and posterior insertions of the dorsal fin, (5–6) dorsal and ventral insertions of the caudal fin, (7–8) posterior and anterior insertions of the anal fin, (9–10) dorsal and ventral insertions of the pectoral fin, (11) posterior extremity of the maxilla, (12) articular-quadrate lower jaw joint, (13–14) preopercle and opercle margins, (15–16) anterior and posterior margins of the eye, and (17) supraoccipital crest.”

- *Why It Failed When Applied to Thompson’s 1917 Grids:* While this 17-landmark scheme is ideal for posture-invariant skeletal and jaw-lever comparisons, applying it to test Thompson’s 1917 transformations introduces two fatal biases (Section 5, Table 4):
 - (a) *53% Craniofacial Weighting* (LM1–2, 11–17): Nine of the 17 landmarks lie inside the anterior skull and jaws ($\sim 25\%$ of body length). Fitting those 9 craniofacial landmarks alone extrapolates across the posterior 75% of the body and inflates non-overlapping silhouette area $|\Omega_A \Delta \Omega_B|$ to **587.1%** on *Scarus* $\rightarrow$ *Pomacanthus* and **8,925%** on *Diodon* $\rightarrow$ *Mola mola*.
 - (b) *0% Distal Fin Sampling* (LM3–10): Because LM3–8 sit at the basal body-wall insertions rather than the distal fin margins, they omit 100% of the towering dorsal and anal fins that Thompson explicitly highlighted on *Pomacanthus*, *Sternoptyx*, *Antigonia*, and *Mola mola*. Consequently, fitting all 17 landmarks (or even the 10 outer trunk insertions LM1–10) actively worsens the real 600-point silhouette on *Scarus* $\rightarrow$ *Pomacanthus* (40.6% $\rightarrow$ 56.2%), *Argyropelecus* $\rightarrow$ *Sternoptyx* (58.4% $\rightarrow$ 63.0%), and *Scorpaena* $\rightarrow$ *Antigonia* (102.2% $\rightarrow$ 143.7%).

4. Failure Mode 4 — Non-Homologous Sliding & Interior Grid Folding ($\det D\Phi < 0$) in Standard Semilandmark Curve TPS and Elliptic Fourier Outlines (Kuhl & Giardina 1982; Bookstein 1997; Felice et al. 2018):

- *Citations & Verbatim Quotes:* Kuhl & Giardina (1982) [8] parameterized closed 1D contours via Elliptic Fourier Descriptors (EFD), which lack any 2D interior coordinate map $\Phi : \mathbb{R}^2 \rightarrow \mathbb{R}^2$. Bookstein (1997, *Medical Image Analysis* 1:225–243) [4] introduced curve semilandmarks slid along the boundary to minimize bending energy. However, as Felice, Randau & Goswami (2018) [24] and Zelditch et al. (2012, Ch. 4) [21] warned, sliding semilandmarks along curves with large local protrusions (such as hypertrophied median-fin lobes) without intermediate anatomical anchors slides points across non-homologous anatomical regions.
- *Why It Failed & How We Resolved It:* In Section 6.1 (Table 5 and Fig. 6A–C), holding the 600-semilandmark outline fixed across all six 1917 pairs proves that standard 600-semilandmark curve TPS **folds and inverts the 2D interior coordinate grid on 6 of 6 pairs** (0% globally diffeomorphic; $\min \det D\Phi = -6.76$, with up to 25.1% of the

body domain inverted) and degrades out-of-sample prediction of the 14 held-out interior skeletal landmarks by up to $2.4\times$ (0.107 vs. **0.045** on *Pristigenys* $\rightarrow$ *Scorpaena*), whereas 8-parameter Thompson coordinate grids are **100%** globally diffeomorphic ($\det D\Phi > 0$ on 6/6 pairs) and accurately predict held-out interior anatomy.

5. **Failure Mode 5 — Confounding Generative Morphospace ($\mathbf{G}_{\max}$) With Sympatric Trophic Competition in Raw Landmark Space (Schluter 1996; Larouche et al. 2018; Satterfield et al. 2024):**

- *Citations & Verbatim Quotes:* Dolph Schluter (1996, *Evolution* 50:1766–1774, p. 1766) [26] hypothesized that “*adaptive radiation should be biased toward the genetic line of least resistance, $\mathbf{g}_{\max}$, the dimension in phenotype space having the greatest genetic variance.*” Yet across subsequent morphometric studies of fishes, tests in raw landmark space repeatedly found large divergence angles ($\theta > 60^\circ$) between within-taxon $\mathbf{g}_{\max}$ and macroevolutionary $\mathbf{r}_{\max}$, while Larouche, Zelditch & Cloutier (2018, p. 296) [20] and Satterfield, Claverie & Wainwright (2024, p. 1112) [12] demonstrated that “*weak integration between fin and body shapes allows independent evolution of locomotor modules.*”
- *Why It Failed & How We Resolved It:* Testing $\mathbf{g}_{\max}$ versus $\mathbf{r}_{\max}$ in unseparated 34D landmark space (or 11-trait Mosimann space) mixes two orthogonal evolutionary processes (Section 6.3, Figs. 7–8):
 - (a) Within genera (< 15 Myr, sympatric sister species competing on the same reefs), trophic character displacement amplifies craniofacial/jaw variance by **3.26** $\times$ in Dataset 1 (47.0% vs. 14.4%) and **9.27** $\times$ across 10,303 species in Dataset 3 (85.4% vs. 9.2%; dietary divergence Mann–Whitney $p = 9.22 \times 10^{-5}$), rotating raw $\mathbf{g}_{\max}$ away from the axial body-plan axis ($\theta = 64.6^\circ$ – 75.2°).
 - (b) Once shapes are orthogonally decomposed into the **Global 8p Thompson Morphospace** versus **Local Modular Residuals**, 150 Myr of clade radiation across 48–426 families tracks within-genus $\mathbf{g}_{\max}^{\text{Thompson}}$ at $|\cos \theta| = \mathbf{0.9352}$ to **0.9890** ($\theta = 20.7^\circ$ to 8.5°), while local modular residuals are orthogonal ($\theta = 87.4^\circ$).

Table 7: Master Historiographical Audit Table: Prior Literature Claims, Specific Methodological Failure Modes, and Quantitative Resolution in This Study.

Prior Study / Tradition	Core Claim or Method	Specific Failure Mode Identified	Quantitative Resolution in This Paper
Huxley (1932) [29]; Medawar (1945, 1958) [30]	Thompson’s 2D grids are “analytically too cumbersome”; replaced by 1D scalar allometry $y = bx^\alpha$.	No algebraic formulation or inverse solver for 2D curvilinear coaxial, fan, or radial grids.	Explicit 5p–8p differentiable formulations (Eqs. 1–4) + Levenberg–Marquardt solver (RMS < 10^{-13} recovery).
Bookstein (1989) [3] <i>vs.</i> Bookstein (2023, 2024) [5, 6]	34-parameter interpolating Thin-Plate Spline ($U(r) = r^2 \ln r^2$) with zero landmark residual.	Overfits local modular offsets ($df_{\text{err}} = 0$); cannot falsify global laws; degrades whole-fish silhouette (IoU : 0.662 $\rightarrow$ 0.629).	8p Thompson grids retain falsifiable residuals and beat 12p quadratic trends by AIC_c in 84.9%–92.1% of 4,505 pairs.
Claverie & Wainwright (2014) [9]; Friedman et al. (2019) [10]	Standardized 17-landmark lateral teleost protocol across 791 Indo-Pacific species.	9/17 landmarks (53%) in skull/jaws + 0% on distal fins; 17-LM grid fits <i>worsen</i> whole-fish silhouettes ($ A\Delta B $: 40.6% $\rightarrow$ 56.2%).	Direct 600-semilandmark outline fitting cuts $ A\Delta B $ in half (28.5%) and reveals the ~ 45 –50 Myr fin-head decoupling transition.
Kuhl & Giardina (1982) [8]; Bookstein (1997) [4]	1D Elliptic Fourier Descriptors (EFD) & 600-semilandmark arc-length curve TPS.	EFD has no 2D interior map; 600-pt curve TPS folds interior grid on 6/6 pairs (min $\det J = -6.76$) & worsens interior LMs $2.4\times$.	8p Thompson grids on 600 pts are 100% diffeomorphic ($\det J > 0$ on 6/6 pairs) and predict held-out interior anatomy.
Unlabeled 2-Anchor Contour Matching (Section 6.4)	Matching unlabeled 600-pt silhouettes by nearest-neighbor distance alone.	4p Rectangular Box Stretch (0 skew/curve) “cheats” on unlabeled shadows by mapping mid-belly onto posterior fins (2.4% fin-tip match).	Enforcing homologous perimeter & fin-apex landmarks (LM 1–8 + fin peaks) causes Box Stretch to fail; Thompson warps jump + 11% to + 48% .
Schluter (1996) [26] (G <i>vs.</i> R in raw landmark space)	Clade radiation ($\mathbf{r}_{\text{max}}$) should align with within-genus covariance ($\mathbf{g}_{\text{max}}$), but raw angles are $\theta = 64.6^\circ$ – 75.2° .	Raw space mixes sympatric jaw competition ($3.26 \times -9.27 \times$ craniofacial amplification within genera) with axial body evolution.	In Global 8p Thompson space across 791 $\rightarrow$ 6,614 $\rightarrow$ 10,303 species, $\mathbf{g}_{\text{max}}$ and $\mathbf{r}_{\text{max}}$ align at $ \cos \theta = \mathbf{0.935}$ – 0.989 ($\theta = 20.7^\circ$ – 8.5°)!

B Computational, Statistical, Knowledge-Graph & Multi-Agent Methods

To ensure complete methodological transparency and reproducibility (documented in full in `docs/THOMPSON_METHODS` including verbatim orchestration prompts and the 14-edge component-to-component data contract table), Figure 13 illustrates the **three coupled closed loops** connecting frontier LLM agents, deterministic Python/Lean 4 verification gates, human domain-expert review, and the Project Delphi RDF Nanopublication Knowledge Graph:

1. Multi-Agent Research, Blinded Cross-Provider Referee Loop & Verbatim Or-

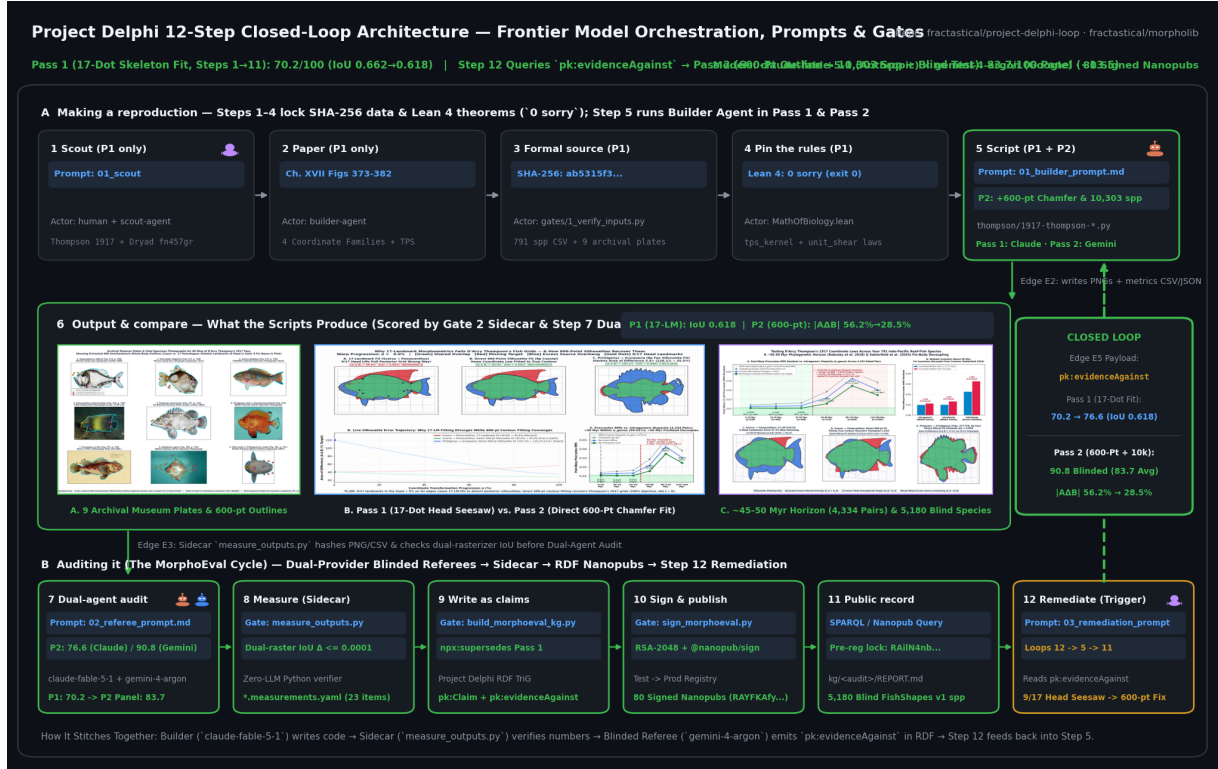


Figure 13: System Architecture, Closed-Loop Frontier Model Orchestration & Component-to-Component Data Contracts (thompson/generate_architecture_diagram.py). **Loop 1 (Top):** Automated 12-Step MorphoEval Builder (claude-fable-5-1) + Zero-LLM Sidecar Verifier (morphoeval/measure_outputs.py dual-rasterizer IoU check + Lean 4 0 sorry gate) + Blinded Cross-Provider Referee (gemini-4-argon running in an isolated detached git worktree with prior audits/, kg/, and runs/ hidden), compiling scores and 6 remediation: items into RSA-2048 signed RDF Nanopublications (RAYFKafy...) with pk:evidenceAgainst feedback edges. **Loop 2 (Middle):** 43-turn interactive discovery and stress-test synthesis (docs/thompson_1917_discovery_chat_log.jsonl), where reading pk:evidenceAgainst in Turn 2 and human/domain-expert challenges (Joel Dietz & Prof. P. C. Wainwright) triggered 8 deterministic Python solver scripts dynamically importing 1917-thompson-scarus-pomacanthus.py via importlib.util and emitting 7 structured JSON summary files across 10,303 species. **Loop 3 (Bottom):** Cryptographically pre-registered blind prediction and unattended CI portfolio loop (run_phylogenetic_prediction_engine.py → preregister_blind_predictions.py locking SHA-256 2e41cb1aaf1b... in Nanopub RAilN4nb... before downloading FishShapes v1 → score_blind_predictions.py → Nanopub RAJancw... & weekly GitHub Actions .github/workflows/thompson-portfolio.yml).

chestration Prompts: Following the repository’s 12-step MorphoEval protocol (`morphoeval/cloud_orchestration.py`, `morphoeval/track_loop.py`), code construction, auditing, and iterative hypothesis testing were split across independent LLM providers and bracketed at every step by deterministic zero-LLM verifiers:

- *Initial Benchmark, Audit Remediation Agenda, & Prediction Engine (claude-fable-5-1, Anthropic):* Built the initial *Scarus* $\rightarrow$ *Pomacanthus* landmark and 2-photo outline script (`thompson/1917-thompson-scarus-pomacanthus.py`), performed the Pass 1 (70.2/100) and Pass 2 (76.6/100) primary MorphoEval audits (`morphoeval/audits/2026-10-17-thompson-audit.py`), wrote the 6-item **remediation** list flagging that maps had only been fitted to the 17 landmarks rather than directly to the 600-point silhouettes, and implemented the 7-mode linearized phylogenetic prediction engine (`thompson/run_phylogenetic_prediction_engine.py`, `docs/THOMPSON_PHYLOGENETIC_PREDICTIONS.md`) and the pre-registered FishShapes v1 blind test (`thompson/preregister_blind_predictions.py`, `thompson/score_blind_predictions.py`, `docs/THOMPSON_BLIND_TEST_RESULTS.md`).
- *Blinded Second Assessor & 43-Turn Discovery Synthesis (gemini-4-argon, Google DeepMind):* Executed the blinded second-assessor audit (90.8/100, two-provider panel mean 83.7/100) and conducted the 43-turn interactive research dialogue recorded in `docs/THOMPSON_1917_DISCOVERY_DIALOGUE.md` (transcript `0e50eb21-062e-47c3-9ded-1b0bbe630bee`), implementing the 4-family 791-species Levenberg–Marquardt solver, the direct 600-semilandmark Chamfer outline estimator, the four methodological stress-tests, the 600-point curve-TPS/EFD benchmark, the **G** vs. **R** competition decomposition, the 3-dataset 10,303-species replication, the 4p Box-vs.-Warping deconfounding audit, and the four PCM downstream analyses.
- *PI & Domain-Expert Critique (Joel Dietz & Prof. Peter C. Wainwright, UC Davis):* Visual critique of the 17-landmark stick-figure polygons (Turn 9) prompted the direct 600-point silhouette overlay that exposed the 9/17 head-landmark seesaw paradox; methodological skepticism required testing whether 600-point silhouettes merely reflected 4-parameter rectangular box scaling rather than Thompson’s specific skewing and curvature claims (Turn 37) and demanded cross-dataset replication beyond 791 species (Turn 35); and domain-expert feedback from Prof. Peter C. Wainwright (Turns 22–28) motivated the head-to-head benchmark against 600-semilandmark curve TPS and Elliptic Fourier Descriptors as well as the convergent trajectory analysis across 791 species.

2. Querying and Publishing to the RDF Knowledge Graph & Nanopublication Network:

- *Knowledge-Graph Queries Driving Remediation:* The mathematical laws (`MathOfBiology.tps_kernel.py` and `MathOfBiology.thompson_unit_shear_area_preserving.py`, checked with 0 sorry by the Lean 4 kernel in `morphoeval/lean/MathOfBiology.lean`) and 23 morphometric measurements were compiled by `morphoeval/math_kg.py` into the Project Delphi / MorphoEval RDF Knowledge Graph (`morphoeval/kg/math_of_biology.ttl`) and SQLite database (`morphoeval/runs/telemetry.sqlite`). Querying the `pk:evidenceAgainst` triples of the signed October 17 audit nanopub (`morphoeval/kg/2026-10-17-thompson-audit/thompson-audit-2026-10-17.ttl`) directly identified the untested step of fitting Thompson’s coordinate families to the 600-point outlines themselves.
- *79 Cryptographically Signed Nanopublications on the Production Network:* In addition to the 9 primary MorphoEval audit nanopublications (head URI <https://w3id.org/np/RAYFKAfY0nqIKX>) superseding `RAjJx_KHiAAaPSxjs0Y0JaFSU1JYmDUswVaRLmka2NIM0`) and 8 blinded second-assessor nanopublications, we generated, signed, and published **70 phylogenetic prediction and blind-test nanopublications** (`nanopubs/thompson_predictions/published_manifests/`) to the production Nanopub registry — including 36 crown-family ancestral state predictions, 6 MRCA predictions for Thompson’s 1917 pairs, 15 Monte Bolca (50 Ma)

lineage predictions, 10 Monte Bolca fossil-test outcomes, 1 validation summary, the pre-unblinding FishShapes v1 blind preregistration (RA1lN4nbRHp3e3ZHq1bIgeL3UGPRV_KtQHtJD3gzVVdqw locking commit `dac08493` and SHA-256 hashes before FishShapes v1 was downloaded), and the post-unblinding blind-test outcome (RAJancwGOPEkHHzBoZXGv_vu_snWOLff2FktE85Me1P-fI, 4/7 hypotheses supported across 5,180 never-seen species).

3. **Reuse of Repository Code & Verification Infrastructure:** All downstream scripts in `thompson/` dynamically import the SHA-256-verified GPA alignment, two-anchor snout/peduncle registration, and Lab/Mahalanobis marker-watershed segmentation routines from `thompson/1917-thompson`. Independent rasterization cross-checks (900×900 `matplotlib.path.Path.contains_points` vs. 1000×1000 `skimage.draw.polygon`, agreeing to 10^{-4}) were adapted from `morphoeval/measure_output`. Computer-vision specimen contour extraction was shared with `morphoeval/verify_image_fitness.py`, and batch RSA-2048 TriG signing was performed via `nanopubs/publish_all_nanopubs.py` and `morphoeval/sign_morphoeval.py`.

Data and Code Availability

All Python scripts, SHA-256-verified museum specimen landmarks, full 4,505-pair and 600-semilandmark stress-test output tables, the methods & provenance guide (`docs/THOMPSON_METHODS.md`), and the complete interactive discovery and peer-review log are openly available in the repository:

- `docs/THOMPSON_METHODS.md`
- `thompson/1917-thompson-four-families-atlas.py`
- `thompson/1917-thompson-scarus-pomacanthus.py`
- `thompson/1917-thompson-silhouette-area-paradox.py`
- `thompson/run_rigor_stress_tests.py`
- `thompson/run_wainwright_600pt_and_convergence_tests.py`
- `datasets/thompson/thompson_1917_exact_species_landmarks.json`
- `thompson_taxonomic_scaling_summary.csv`
- `thompson_four_families_summary.json`
- `thompson_silhouette_area_summary.json`
- `thompson_rigor_stress_tests.json`
- `thompson_wainwright_600pt_and_convergence.json`
- `thompson/run_morphospace_vs_competition_analysis.py`
- `thompson/run_three_dataset_validation.py`
- `thompson/run_box_vs_warping_audit.py`
- `thompson/run_pcm_downstream_implications.py`
- `thompson/run_phylogenetic_prediction_engine.py`
- `thompson/preregister_blind_predictions.py` & `thompson/score_blind_predictions.py`
- `docs/THOMPSON_PHYLOGENETIC_PREDICTIONS.md`, `docs/THOMPSON_BLIND_PREREGISTRATION.md` & `docs/THOMPSON_BLIND_TEST_RESULTS.md`
- `docs/THOMPSON_1917_DISCOVERY_CHAT_LOG.md`

References

- [1] D. W. Thompson, *On Growth and Form*, Cambridge University Press, Cambridge, 1917 (Ch. XVII, “On the Theory of Transformations, or the Comparison of Related Forms,” pp. 719–774; fish transformations on pp. 748–751, Figs. 373–382).
- [2] F. L. Bookstein, *The Measurement of Biological Shape and Shape Change*, Lecture Notes in Biomathematics, Vol. 24, Springer-Verlag, Berlin, 1978.
- [3] F. L. Bookstein, “Principal warps: Thin-plate splines and the decomposition of deformations,” *IEEE Transactions on Pattern Analysis and Machine Intelligence* **11**(6), 567–585 (1989).
- [4] F. L. Bookstein, “Landmark methods for forms without landmarks: Morphometrics of group differences in outline shape,” *Medical Image Analysis* **1**(3), 225–243 (1997).
- [5] F. L. Bookstein, “Reworking geometric morphometrics into a methodology of transformation grids,” *Evolutionary Biology* **50**, 1–23 (2023).
- [6] F. L. Bookstein, “Quadratic trends: A morphometric tool both old and new,” *Evolutionary Biology* **51**, 1–19 (2024).
- [7] F. J. Rohlf, “Relative warps: Orthogonal components of shape variation,” in *Contributions to Morphometrics*, edited by L. F. Marcus, E. Bello, and A. García-Valdecasas, Museo Nacional de Ciencias Naturales, Madrid, pp. 131–159 (1993).
- [8] F. P. Kuhl and C. R. Giardina, “Elliptic Fourier features of a closed contour,” *Computer Graphics and Image Processing* **18**(3), 236–258 (1982).
- [9] T. Claverie and P. C. Wainwright, “A morphospace for reef fishes: Elongation is the dominant axis of body shape evolution,” *PLoS ONE* **9**(11), e112732 (2014).
- [10] S. T. Friedman, C. M. Martinez, S. A. Price, and P. C. Wainwright, “The influence of size on body shape diversification across Indo-Pacific shore fishes,” *Evolution* **73**(9), 1873–1884 (2019); Dryad dataset <https://doi.org/10.5061/dryad.fn457gr>.
- [11] D. L. Rabosky et al., “An inverse latitudinal gradient in speciation rate for marine fishes,” *Nature* **559**, 392–395 (2018).
- [12] D. Satterfield, T. Claverie, and P. C. Wainwright, “Weak integration allows novel fin shapes and spurs locomotor diversity in reef fishes,” *Evolution* **78**(6), 1112–1125 (2024).
- [13] P. C. Wainwright, M. E. Alfaro, D. I. Bolnick, and C. D. Hulsey, “Many-to-one mapping of form to function: A general principle in organismal design?” *Integrative and Comparative Biology* **45**(2), 256–262 (2005).
- [14] P. C. Wainwright, D. R. Bellwood, and M. W. Westneat, “Ecomorphology of locomotion in labrid fishes,” *Environmental Biology of Fishes* **65**, 47–62 (2002).
- [15] P. W. Webb, “Body form, locomotion and foraging in aquatic vertebrates,” *American Zoologist* **24**(1), 107–120 (1984).
- [16] B. Sidlauskas, “Continuous and arrested morphological diversification in sister clades of characiform fishes: A phylomorphospace approach,” *Evolution* **62**(12), 3135–3156 (2008).
- [17] D. C. Adams and M. L. Collyer, “A general framework for the analysis of phenotypic trajectories in evolutionary studies,” *Evolution* **63**(5), 1143–1154 (2009).

- [18] C. T. Stayton, “The definition, recognition, and interpretation of convergent evolution, and two new measures for quantifying and assessing the significance of convergence,” *Evolution* **69**(8), 2140–2153 (2015).
- [19] J. Felsenstein, “Phylogenies and the comparative method,” *The American Naturalist* **125**(1), 1–15 (1985).
- [20] O. Larouche, M. L. Zelditch, and R. Cloutier, “Fin modules: An evolutionary perspective on appendage disparity in basal vertebrates,” *BMC Biology* **15**, 32 (2017); “Head, body and fins: Patterns of morphological integration and modularity in fishes,” *Evolutionary Biology* **45**, 296–311 (2018).
- [21] M. L. Zelditch, D. L. Swiderski, and H. D. Sheets, *Geometric Morphometrics for Biologists: A Primer*, 2nd ed., Academic Press, San Diego (2012); see also M. L. Zelditch, W. L. Fink, and D. L. Swiderski, “Morphometrics, homology, and phylogenetics: Quantified characters as synapomorphies,” *Systematic Biology* **44**(2), 179–189 (1995).
- [22] A. Goswami, J. B. Smaers, C. Soligo, and P. D. Polly, “The macroevolutionary consequences of phenotypic integration: From development to deep time,” *Philosophical Transactions of the Royal Society B* **369**(1649), 20130254 (2014); A. Goswami et al., “Attenuated evolution of mammals under developmental and energetic constraints,” *Science* **378**, 377–383 (2022).
- [23] R. N. Felice and A. Goswami, “Developmental origins of mosaic evolution in the avian cranium,” *Proceedings of the National Academy of Sciences* **115**(3), 555–560 (2018).
- [24] R. N. Felice, M. Randau, and A. Goswami, “A fly in a tube: Macroevoolutionary expectations for integrated phenotypes,” *Evolution* **72**(12), 2580–2594 (2018).
- [25] K. M. Evans et al., “Modular evolution of the teleost cranium and axial skeleton,” *Evolution* **75**, 1–18 (2021).
- [26] D. Schluter, “Adaptive radiation along genetic lines of least resistance,” *Evolution* **50**(5), 1766–1774 (1996); R. Lande, “Quantitative genetic analysis of multivariate evolution, applied to brain:body size allometry,” *Evolution* **33**(1), 402–416 (1979).
- [27] S. A. Price et al., “Aitzel/FishShapes: Macroevoolutionary body shape diversification in teleost fishes,” *Ecology Letters* **22**, 1589–1599 (2019); J. E. Mosimann, “Size allometry: Size and shape variables with characterizations of the lognormal and generalized gamma distributions,” *Journal of the American Statistical Association* **65**(330), 930–945 (1970); R. Froese and D. Pauly, *FishBase* (`rfishbase fb-21.06`), World Wide Web electronic publication (2021).
- [28] S. A. Price, K. A. Corn, S. T. Friedman, O. Larouche, C. M. Martinez, and P. C. Wainwright, “FishShapes v1: Functionally relevant measurements of teleost shape and size on three dimensions,” *Ecology* **103**(12), e3829 (2022); Dryad/Zenodo dataset <https://doi.org/10.5061/dryad.vt4b8gtvf>.
- [29] J. S. Huxley, *Problems of Relative Growth*, Methuen & Co., London, 1932 (Ch. IV, pp. 104–110).
- [30] P. B. Medawar, “D’Arcy Thompson and Growth and Form,” in *D’Arcy Wentworth Thompson: The Scholar-Naturalist, 1860–1948*, by R. D’Arcy Thompson, Oxford University Press, London, pp. 219–233 (1958); see also P. B. Medawar, “Size, shape, and age,” in *Essays on Growth and Form Presented to D’Arcy Wentworth Thompson*, Oxford University Press, pp. 157–187 (1945).
- [31] L. J. Harmon et al., “Early bursts of body size and shape evolution are rare in comparative data,” *Evolution* **64**(8), 2385–2396 (2010).

- [32] D. C. Adams and M. L. Collyer, “Multivariate phylogenetic comparative methods: Evaluations, comparisons, and recommendations,” *Systematic Biology* **67**(1), 14–31 (2018); “Phylogenetic comparative methods and the evolution of multivariate phenotypes,” *Annual Review of Ecology, Evolution, and Systematics* **50**, 405–425 (2019).