frontiers

Supplementary Material

1 Supplementary Materials and Methods

1.1 Data acquisition, model rendering and processing

The coral fragments were placed at 30° in a manual rotating plate with the help of a rubber wedge, to ensure that the corals were comprehensively documented and to decrease the effect of self-shading (Supplementary Figure 1). In this position, both the top (towards where the branches grow) and bottom faces (underneath the table-like shape), were registered over two full-plate rotations and at least three different angles from the rotating plate (0°, 45°, and 90°). In a few cases where the preliminary model (obtained using real-time fusion) displayed missing data (holes) in some complex areas of the coral morphology, additional angles (approx. 25° and 65°) and full-plate rotations were added to improve the quality of the rendered 3D model.

Following Reichert et al. (2016; 2017), the Artec Studio software was used to render and clean up the 3D models. First, fine serial registration was performed using a final mesh size of 0.2 mm. Global registration followed using a minimal distance of 10 and 100 iterations based on the geometry of the objects. Finally, sharp fusion was performed with 0.2 mm resolution, and holes were also filled by radius (max. hole radius=5). The small objects filter was applied to get rid of any artefacts generated by shading or reflection (mode: Leave_biggest_objects). The algorithm of this filter identifies all the outliers obtained after the fusion as all the objects that are not connected to the main object (in this case the coral fragment) and then it removes them, keeping only the object that includes most of the polygons. Additional editing was performed using the rectangular selection eraser when needed.

1.2 Polygon mesh-based estimations

1.2.1 Complexity

The function *vtkMassProperties* estimates the volume (*V*) with the aid of a discrete divergence theorem algorithm, and surface area (*SA*) by taking the sum of each face area of the polygon mesh (Alyassin et al., 1994). Two quantitative measures were extracted directly from *SA* and *V*, the surface-area-to-volume (*S/V*) ratio and the sphericity (φ). Sphericity is a measure to estimate the compactness of an object since it describes the ratio between the surface area of a sphere – the minimum surface area – with a volume equal to *V* to the actual value of *SA* (see Equation 1; Wadell, 1935):

$$\boldsymbol{\varphi} = \frac{\boldsymbol{\pi^{\frac{1}{3}}(6V)^{\frac{2}{3}}}}{\boldsymbol{SA^{-1}}} \quad (1)$$

1.2.2 Curvature

Gaussian curvature (*K*) and mean curvature (*H*) are two common characteristics of surface curvature. The Gaussian curvature of a point (*p*) on a surface is defined as the product of the two principal curvatures at that location:

$$K(p) = k1(p)\, k2(p) \quad (2)$$

where $k1$ is the maximum curvature and $k2$ is the minimum curvature at p. The average of the two principal curvatures is called the mean curvature:

$$H(p) = \frac{1}{2}(k1(p) + k2(p)) \tag{3}$$

However, for obtaining curvature values of a polygon mesh a discrete scheme is required. These schemes do not often derive surface curvatures from the principal curvatures, but rather look at the angle deficit. From the Gauss-Bonnet theorem, a discrete approximation of the Gaussian curvature for vertices in a triangle mesh was derived (Meyer et al., 2003). In this scheme, the curvature for vertex (v) can be obtained as follows:

$$K(v) = \frac{1}{A_{mix}(v)}\left(2\pi - \sum_{f \in F} \alpha f\right) \tag{4}$$

where f are the faces that have v as one of their vertices, α_f is the angle of face f at v, A_{mix} is the surface area associated to the vertex, which is either based on based on the surface area or the Voronoi region of v of each face depending on whether faces are obtuse or non-obtuse, respectively. In addition, Meyer et al. (2003) also proposed a scheme for the mean curvature of triangle meshes. Here, a mean curvature normal operator is derived as follows:

$$\mathbf{H}(v) = \frac{1}{2A_{mix}(v)} \sum_{e \in E} (cot\ \alpha_e + cot\ \beta_e)\, e \tag{5}$$

where E are the edges of v to its neighboring vertices v and α_e and β_e are the opposite angles that correspond to e. The actual mean curvature (H) is equal to:

$$H(v) = \pm\frac{1}{2} \|\mathbf{H}(v)\| \tag{6}$$

where the negative sign is added when $\mathbf{H}(v)$ does not follow the direction of the outward normal vector of the surface at v (Mesmoudi et al., 2010). Normal vectors were derived with *vtkTriangleMeshPointNormals*. Finally, the principal curvatures were obtained from K and H by using the definitions described at the start of this section and stated by the following equations:

$$k1(v) = H(v) + \sqrt{H(v)^2 - K(v)} \tag{7}$$

$$k2(v) = H(v) - \sqrt{H(v)^2 - K(v)} \tag{8}$$

In the rare case that H^2–K is negative, due to errors in the estimation, $k1$ and $k2$ were both set to H.

1.2.3 Medial axis skeleton-derived measures

The voxel thinning was performed with *itkBinaryImageThinningFilter3D* (Homann, 2007), an implementation of the algorithm by Lee et al. (1994) that results in single-voxel thin skeletons. The voxel skeletons were then transformed into graphs (G). To do so, each voxel of the voxel skeleton was

translated to a vertex (v) with coordinates corresponding to the location that the voxel represents and connected by edges using the method described in Reinders et al. (2000). In a graph, neighboring vertices are those connected to each other by edges and the degree of a vertex ($dG(v)$) corresponds to the number of neighboring vertices. All vertices with degrees greater than two ($dG(v)>2$) were considered junction vertices and the vertices with degrees equal to one ($dG(v)=1$) were considered endpoint vertices. A branch (b) was considered to be the set of neighboring vertices with $dG(v)=2$ and the two connected vertices with $dG(v)\neq2$. All the branches that contained an endpoint vertex were considered terminal branches (b_T), except for the branch with the highest average thickness that was considered the root branch (b_R). Terminal branches with less than four vertices were removed from the skeleton as they were likely spurious.

1.3 Global and average values variable assessment

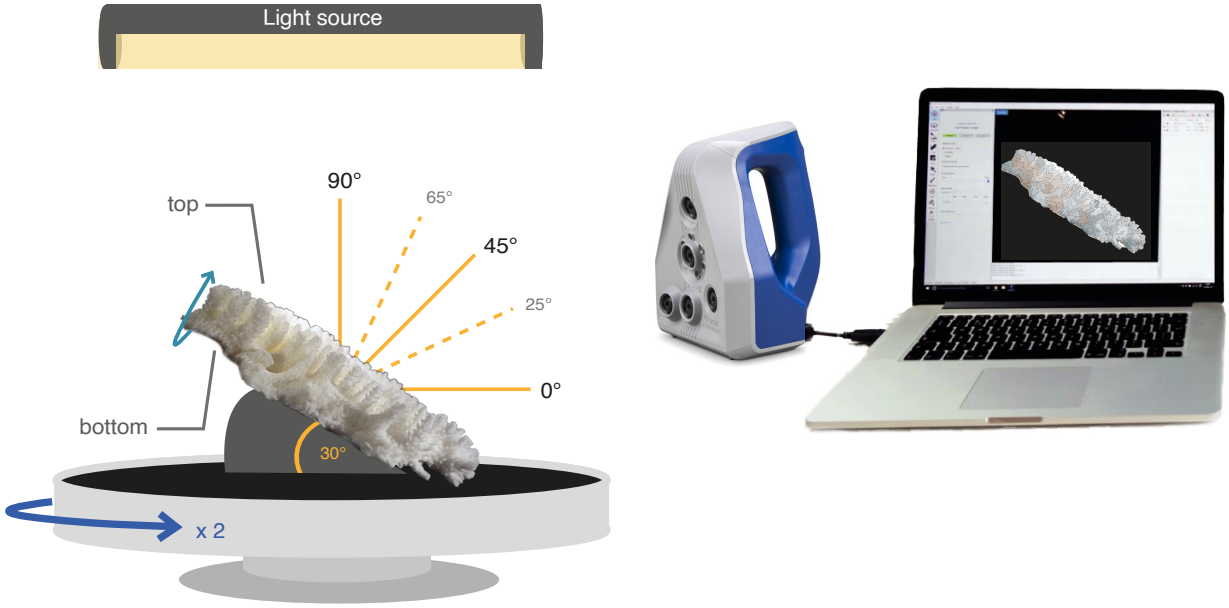
To evaluate the performance of individual features, we first tested the assumptions of normality (q-q plots, ggpubr R package v0.4.0; Kassambara, 2020; Shapiro-Wilk test, stats R package v4.1.0; R Core Team, 2018) and homoscedasticity (Levene's test; car R package 3.0-11; Fox and Weisberg, 2019) in the estimated univariate measures. All variables that deviated from the assumptions were transformed (TR) using the most frequently suggested option after five iterations of the *bestNormalize* function (bestNormalize R package v1.8.1; Peterson and Cavanaugh, 2020). Multivariate normality was then assessed using Mardia's, Henze-Zirkler's, Royston's, energy and Doornik-Hansen's tests (MVN R package v5.9; Korkmaz et al., 2014). Variables that did not meet assumptions of normality and homoscedasticity after transformation, were excluded from downstream analyses.

1.4 Discriminant analysis based on Gaussian mixture modelling

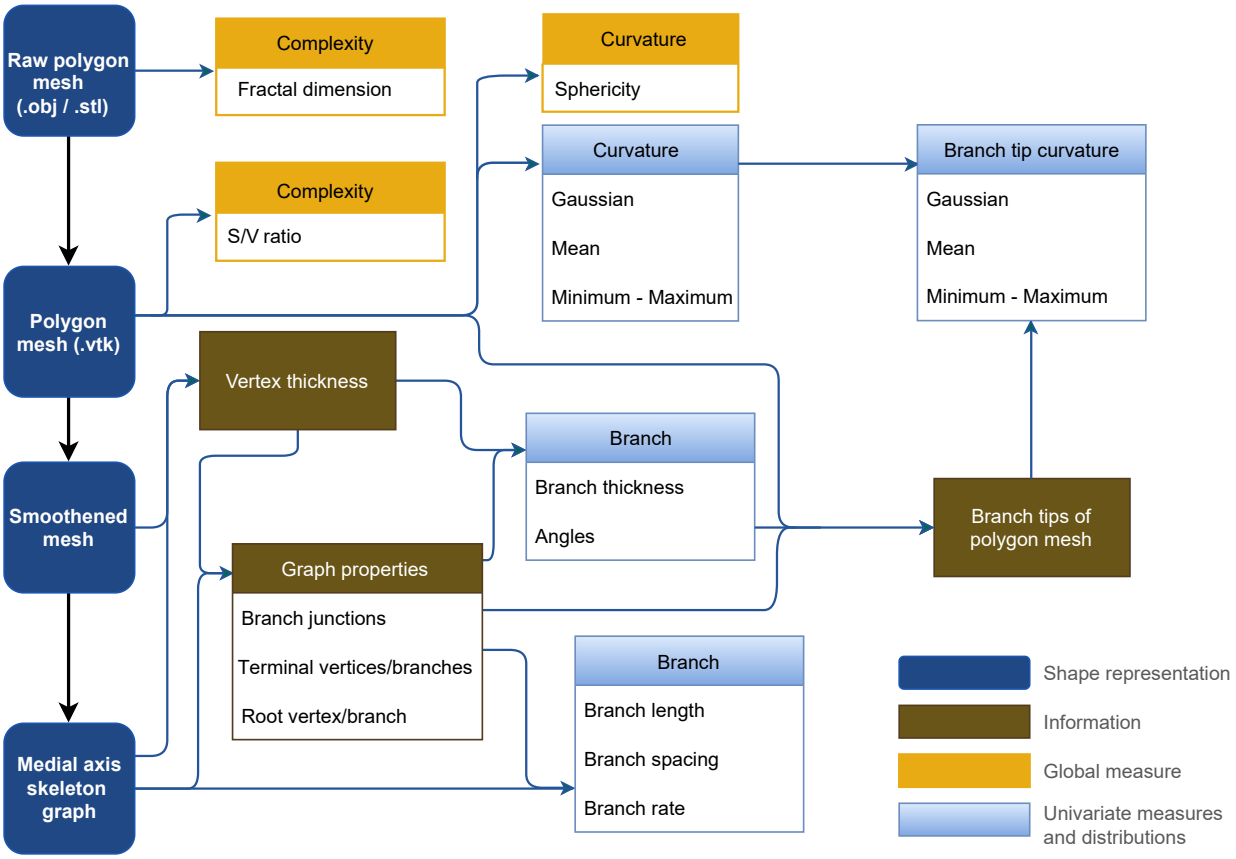
A discriminant analysis based on Gaussian mixture modelling was performed using two different approaches: eigenvalue decomposition discriminant analysis (EDDA), which assumes that the density for each class can be described by a single Gaussian component, and MclustDA, which uses a finite mixture of Gaussian distributions within each class (mclust v5.4.7; Scrucca et al., 2016). The discriminant analysis based on Gaussian finite mixture modelling fitted "EEE" as the mixture model selected according to the Bayesian information criterion (BIC). Following this, the three species distributions were better described by ellipsoidal Gaussian components with equal volume, shape, and orientation (Supplementary Table S10). However, the accuracy of the fitted models to predict the testing subset in this case was only 54.17% (10x cross-validation, error= 0.20, standard error= 0.05) and the prediction accuracy decreased when a finite mixture of Gaussian distributions was used within each species class (50%, 10x cross-validation, error= 0.38, standard error= 0.05). This last approach suggested different mixture models and numbers of components within each species or category: a single component or group for species A (G=1, XXI model= diagonal multivariate normal) and two or more for species B (G=4, EEE model= ellipsoidal, equal volume, shape, and orientation) and C (G=2, EEI= diagonal, equal volume and shape); which suggests higher variation and heterogeneity within these last two groups (Supplementary Table 10).

2 Supplementary Figures and Tables

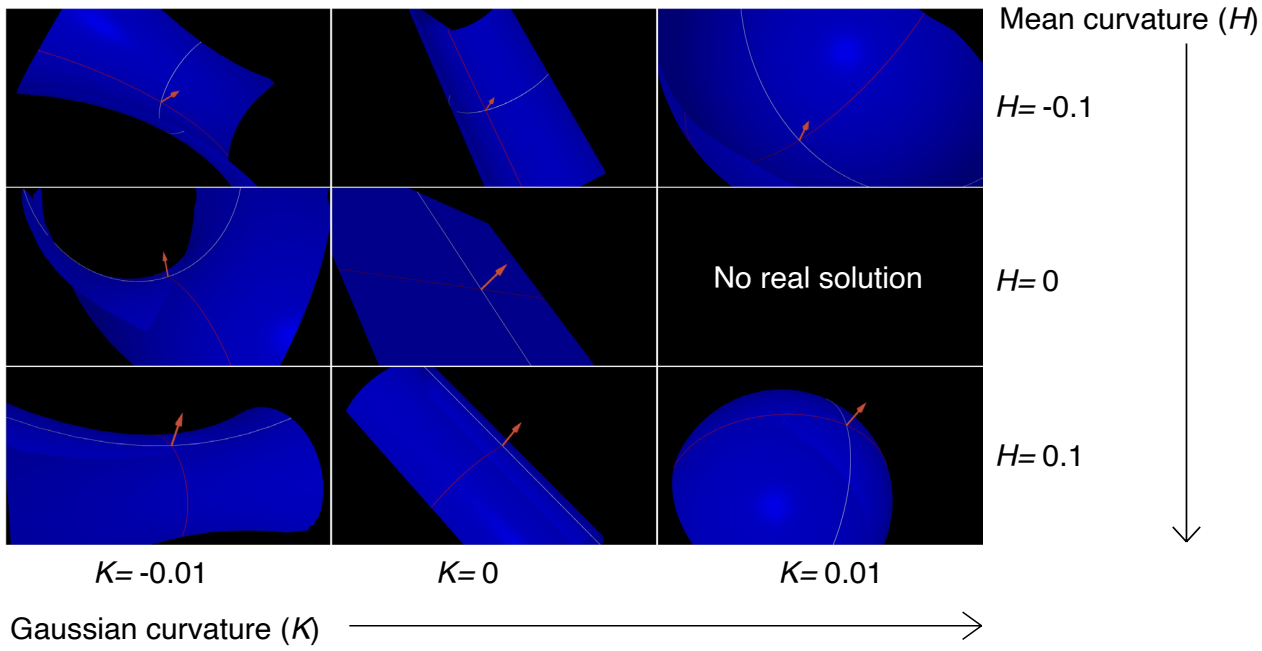
2.1 Supplementary Figures



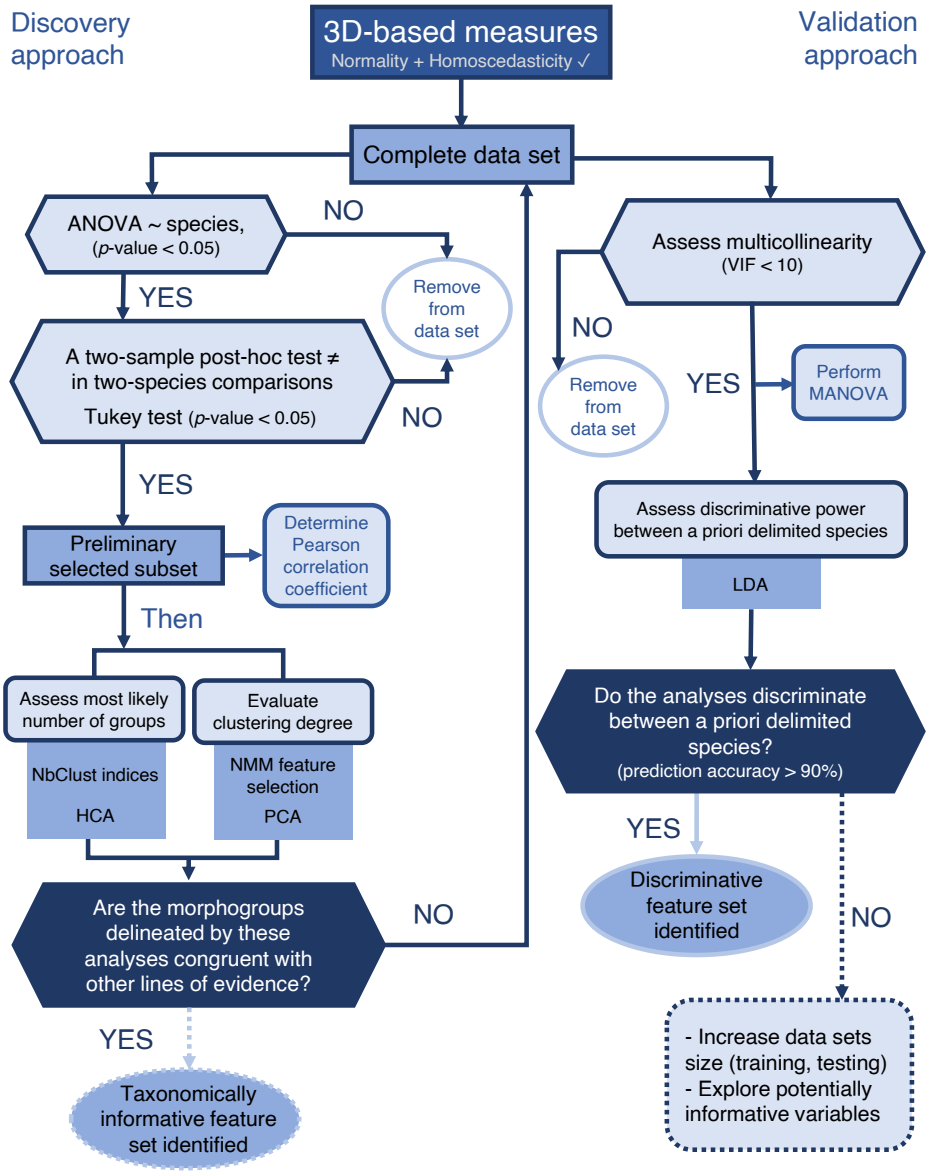
Supplementary Figure S1. Obtaining 3D data from complex-shaped organisms. Schematic representation of how the 3D scanning of the coral specimens was performed, highlighting the primary angles used to best capture the skeleton details (solid lines) and the additional perspectives used when complex branching patterns hampered accurate 3D model rendering (dashed lines). Two full-plate rotations were registered for each one of the angles and both the top and bottom faces of the coral fragments.



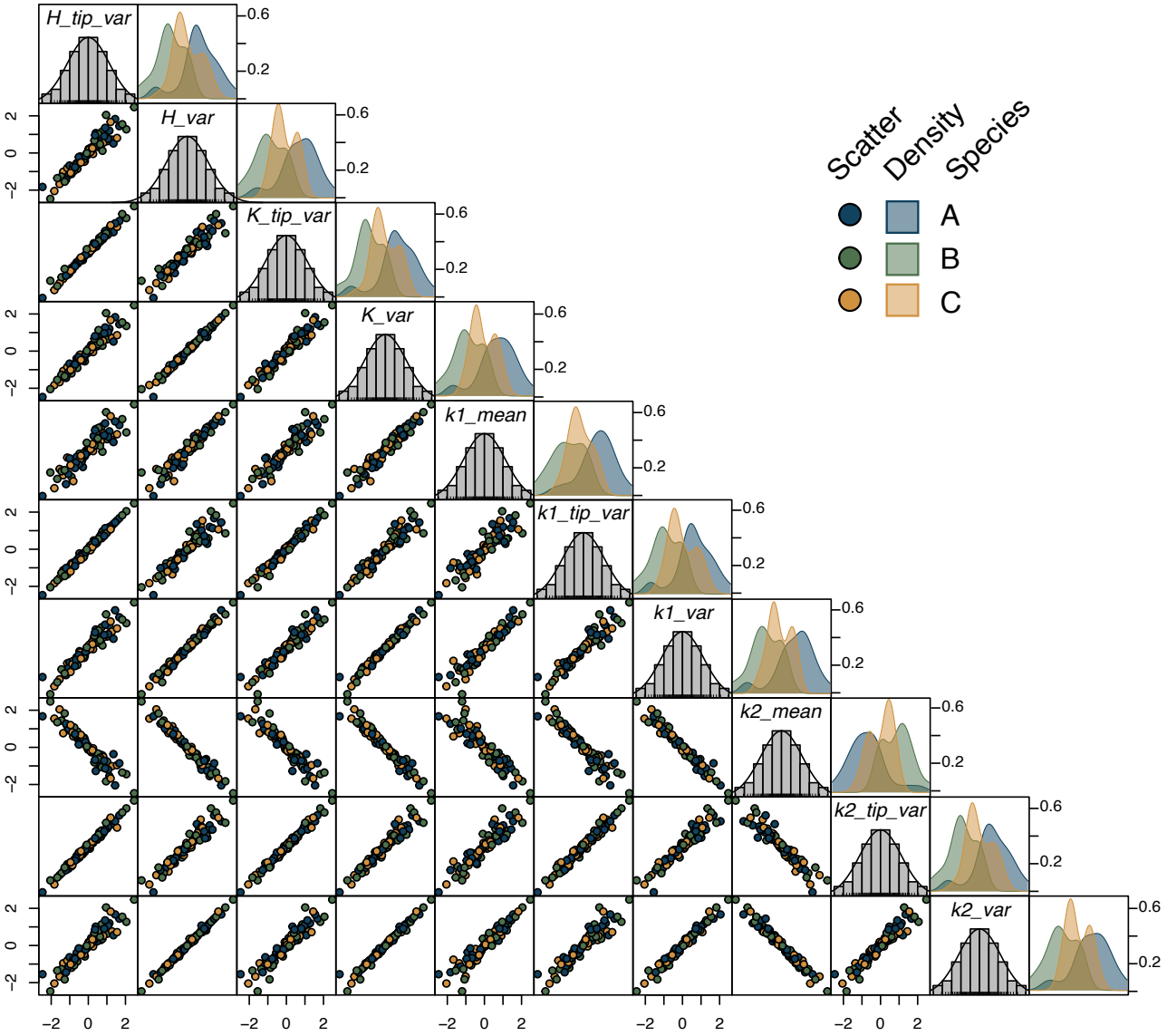
Supplementary Figure S2. Workflow for 3D-based measures estimation. Color-coded representation summarizing the pipeline used to estimate different types of 3D-based measures by extracting and combining information from different shape representations rendered from the light-structured scanning performed of the coral specimens.



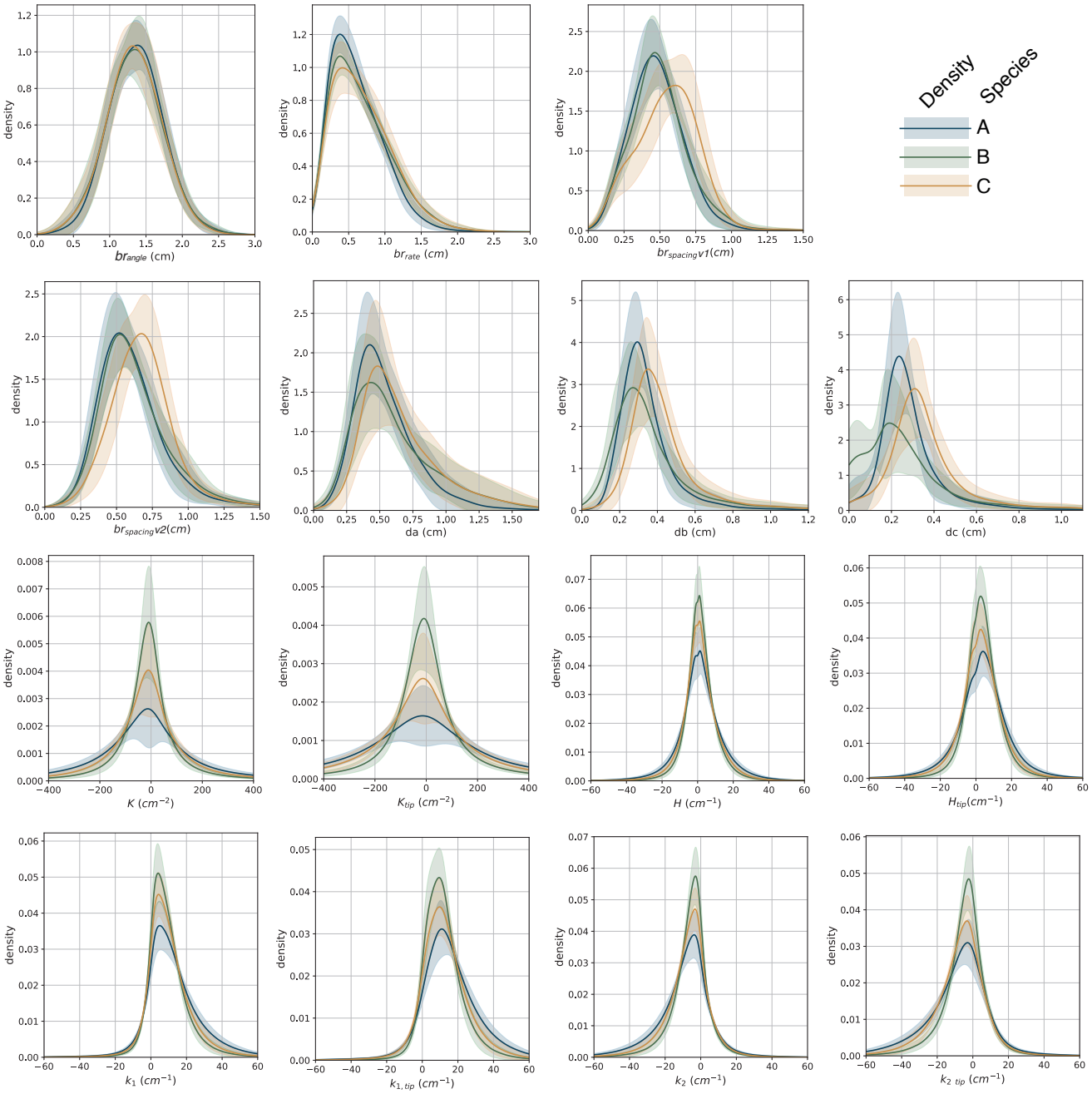
Supplementary Figure S3. Expected shape for different surface curvature combinations. The surfaces for different combinations of *H* (mean curvature) and *K* (Gaussian curvature) are displayed. Orange arrows represent the direction of the outward normal (length=2.5). The red and white lines over the surfaces represent the direction of *k1* (maximum curvature) and *k2* (minimum curvature) respectively.



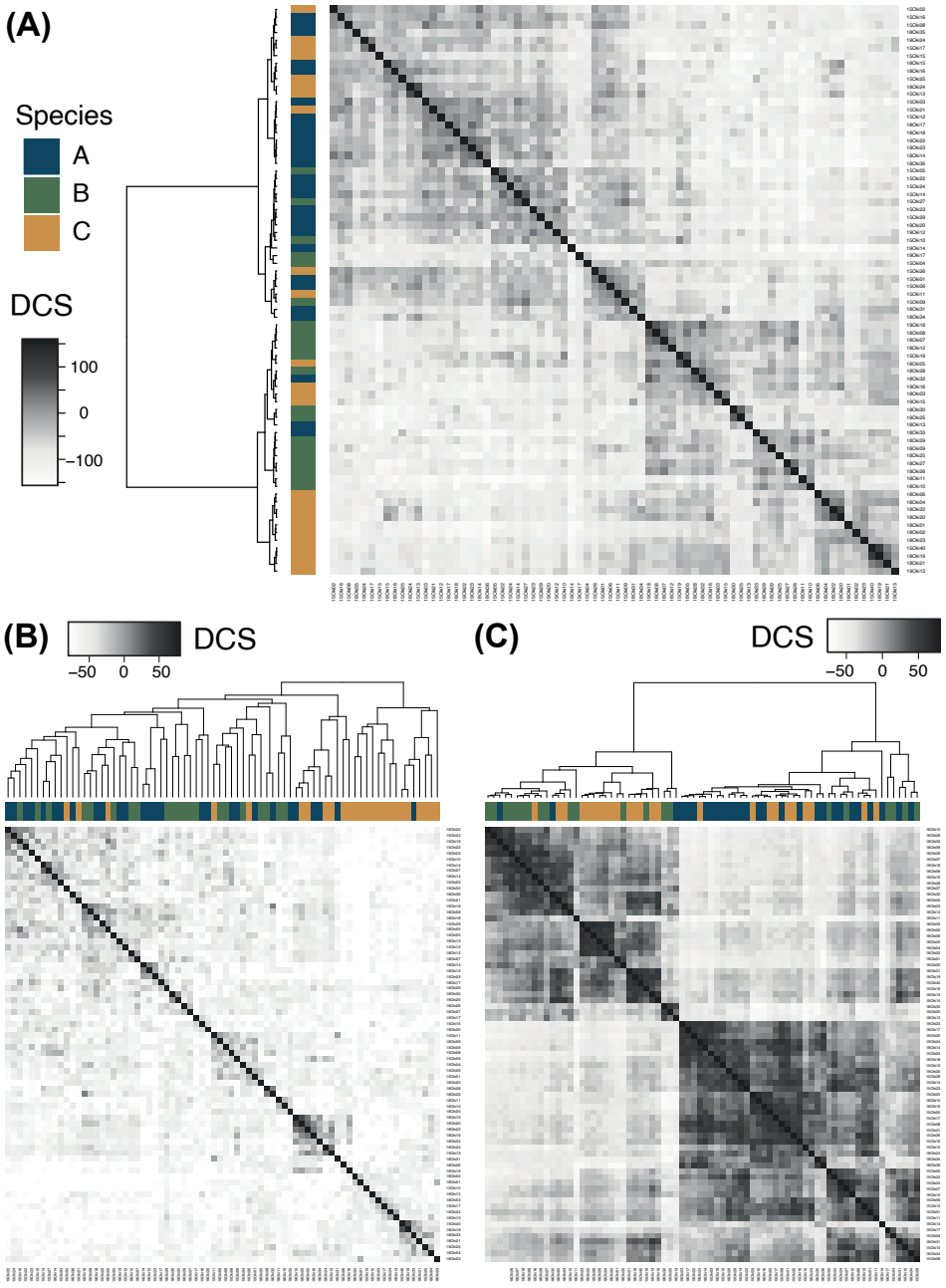
Supplementary Figure S4. Flow chart of 3D-based measures feature screening and selection. To perform feature screening for a prospective combination of representative characters that support interspecific discrimination, 3D-based measures that met the assumptions of normality and homoscedasticity were screened for significant interspecific differences (preliminary selected subset), and clustering was further assessed in this subset to evaluate if morphogroups were congruent with species delineated using other lines of evidence (identify variables useful for discovery approaches) or whether the 3D-morphological analyses enable to discriminate between a priori delimited species (perform a validation approach). Dashed lines and portions within the flow chart represent potential outcomes that were not obtained in this study. Abbreviations: Hierarchical clustering analysis (HCA), normal mixture model (NMM), principal component analysis (PCA), variance inflation factor (VIF), multivariate analysis of variance (MANOVA), linear discriminant analysis (LDA).



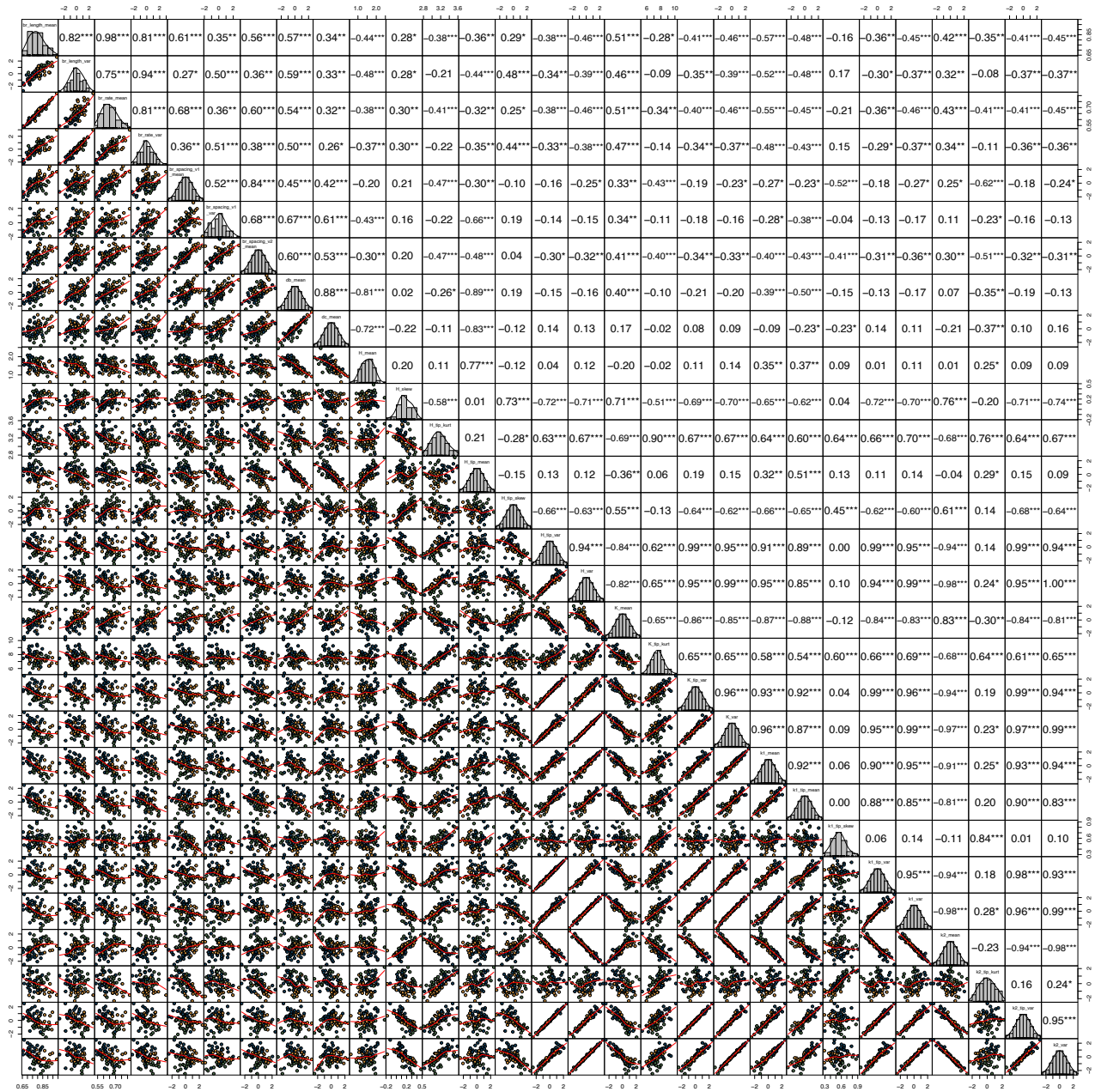
Supplementary Figure S5. Morphospaces depiction using variables with significant differences. Bivariate scatter plots (lower triangle) and univariate density distributions (upper diagonal line) of the ten variables that presented significant interespecific differences were used to assess the degree of overlapping between morphospaces of the three different species and the intra-specific variation. The histograms for each variable are also depicted in the main diagonal.



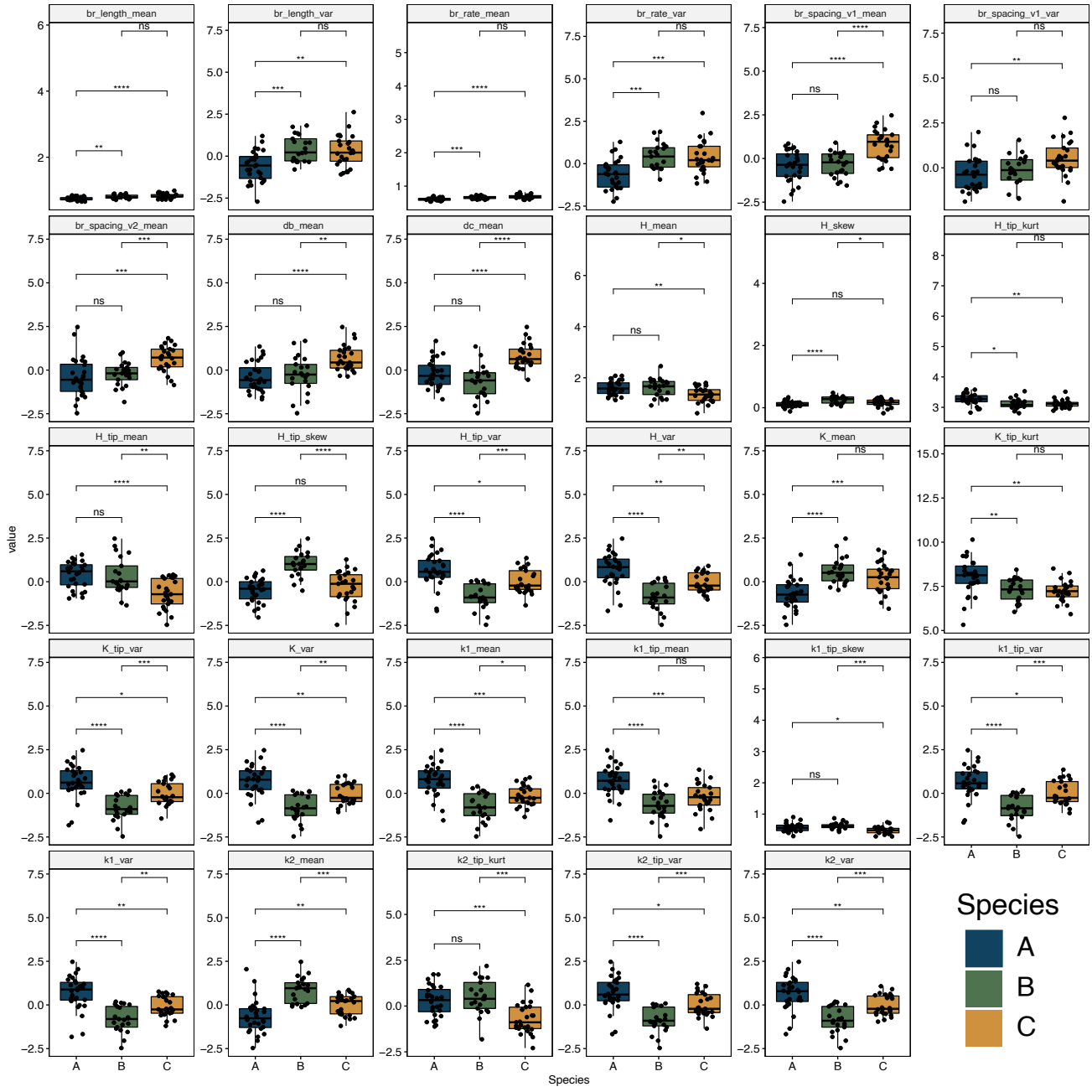
Supplementary Figure S6. Assessment of variables using Kernel density. Gaussian kernel densities (KD) were estimated for seven branch-derived variables variables (first two rows) and eight curvature variables (last two rows). The corresponding probability density function profiles (pdf) are depicted for each of the three species.



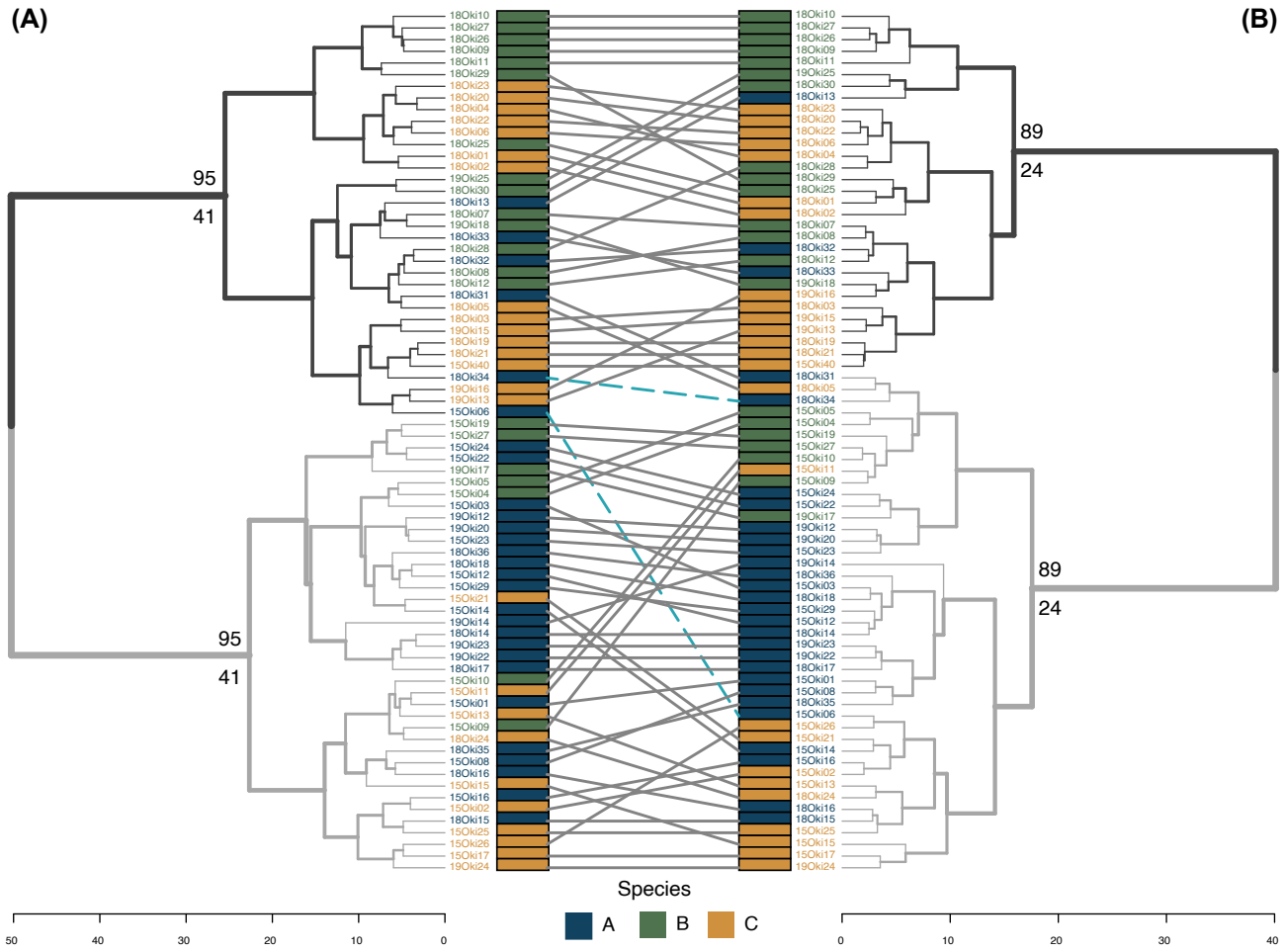
Supplementary Figure S7. Assessment of significant interspecific differences in distributions. Heatmaps depicting recoded and summarized p-values obtained from two-sample Mann-Whitney U tests for different sets of variables: (A) all variables, (B) branch-derived variables, and (C) curvature variables. Samples were re-organized using hierarchical clustering according to similarity in pairwise comparisons, calculated as the distribution comparison score (DCS) or the cumulative value of similarity according to the p-value significance cut-off (α= 0.05).



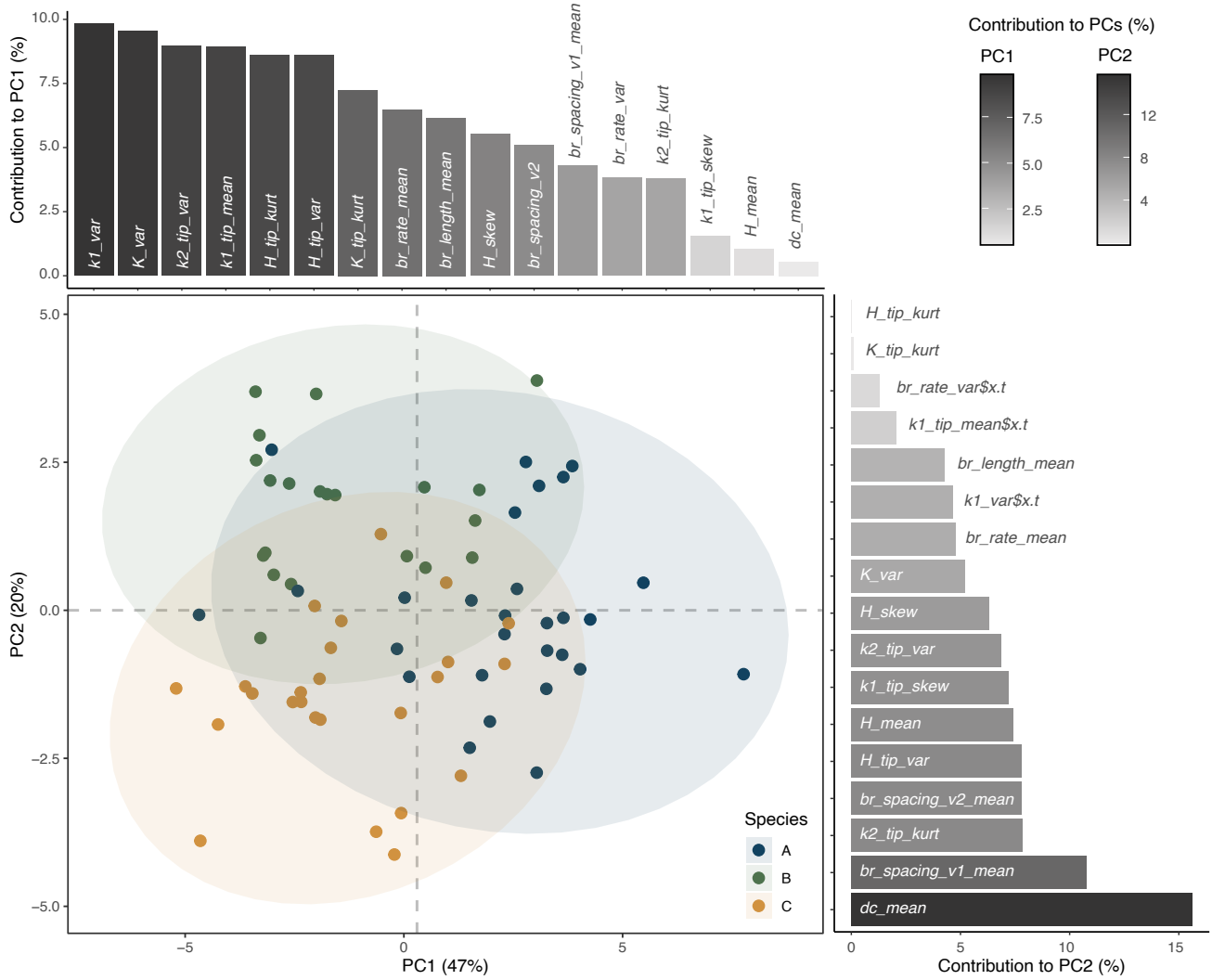
Supplementary Figure S8. Correlation assessment between preliminary selected variables. Correlation plot depicting the pairwise relationships between all the preliminary selected variables. The panels in the diagonal depict histograms of each feature. The graphs in the upper triangle depict Pearson's r correlation coefficients (ρ) and the corresponding asterisks indicate their significance (p-value: * ≤ 0.05, ** ≤ 0.01, *** ≤0.001). The graphs below the diagonal show the scatter plots for each pairwise comparison.



Supplementary Figure S9. Significant interespecific differences in preliminary selected variables. ox plots highlighting significant interespecific differences in the preliminary selected variables. The corresponding significance is displayed by asterisks according to the results of the test (*p*-value: ns = not significant, * ≤ 0.05, ** ≤ 0.01, *** ≤0.001, **** ≤ 0.0001). This subset of 29 variables exhibited significant differences in the ANOVA and in at least two-species comparisons of the two-sample post-hoc Tukey test.



Supplementary Figure S10. Tanglegram of exploratory clustering analyses. Comparison of hierarchical clustering analyses (HCA) using Euclidean distance and Ward linkage for the complete set of variables (A, agglomerative coefficient= 0.8978) and the preliminary selected set (B, agglomerative coefficient= 0.9190). Values on the nodes correspond to the percentage (%) of 1000 bootstrap replications from multiscale re-sampling for approximately unbiased (AU) probability values (top), and bootstrap probability (BP) values (bottom) for the most likely number of groups indicated by the different branch colors (k=2, black and grey). Lines connecting the dendrograms' leaves indicate the topological differences between them, with grey solid lines displaying specimens with positions within the same cluster in A and B, and turquoise dashed lines depicting specimens that belong to different clusters in each dendrogram. Line-width of the branches highlights the height of each dendrogram at a certain point.



Supplementary Figure S11. Exploratory principal component analysis. Biplot of the first two principal components of a principal component analysis (PCA) using a subset of the preliminarily selected variables after performing unbiased variable and selection through Gaussian model-based clustering and the greedy algorithm both (forward and backward directions) according to the Bayesian information criterion (BIC; n= 17). The contribution (%) of each of the 17 variables to the principal components (PCs) is depicted along the axes along with the corresponding color scale.

2.2 Supplementary Tables

2.2.1 SUPPLEMENTARY TABLE S1. Information of specimens used in this study.

Sample ID, year of collection and taxonomy of specimens used for this study are displayed. Open nomenclature qualifiers are used according to the field identification and taxonomic identity of the *Acropora* morphospecies described by Ramírez-Portilla et al. (2022). Briefly, the open nomenclature "aff." is used to indicate that the specimens from species "C" have affinities with *Acropora hyacinthus* but belong most likely to a distinct species. In the case of species "A" and "B" evidence suggest these specimens are similar to *A. bifurcata* and *A. cytherea* respectively, but further information is required to confirm these assessments. For simplicity, species ID are used for this study.

Sample ID	Collection year	Genus	Open nomenclature	Species	Species ID
15Oki01	2015	*Acropora*	cf.	*bifurcata*	A
15Oki02	2015	*Acropora*	aff.	*hyacinthus*	C
15Oki03	2015	*Acropora*	cf.	*bifurcata*	A
15Oki04	2015	*Acropora*	cf.	*cytherea*	B
15Oki05	2015	*Acropora*	cf.	*cytherea*	B
15Oki06	2015	*Acropora*	cf.	*bifurcata*	A
15Oki08	2015	*Acropora*	cf.	*bifurcata*	A
15Oki09	2015	*Acropora*	cf.	*cytherea*	B
15Oki10	2015	*Acropora*	cf.	*cytherea*	B
15Oki11	2015	*Acropora*	aff.	*hyacinthus*	C
15Oki12	2015	*Acropora*	cf.	*bifurcata*	A
15Oki13	2015	*Acropora*	aff.	*hyacinthus*	C
15Oki14	2015	*Acropora*	cf.	*bifurcata*	A
15Oki15	2015	*Acropora*	aff.	*hyacinthus*	C
15Oki16	2015	*Acropora*	cf.	*bifurcata*	A
15Oki17	2015	*Acropora*	aff.	*hyacinthus*	C
15Oki19	2015	*Acropora*	cf.	*cytherea*	B
15Oki21	2015	*Acropora*	aff.	*hyacinthus*	C
15Oki22	2015	*Acropora*	cf.	*bifurcata*	A
15Oki23	2015	*Acropora*	cf.	*bifurcata*	A
15Oki24	2015	*Acropora*	cf.	*bifurcata*	A
15Oki25	2015	*Acropora*	aff.	*hyacinthus*	C
15Oki26	2015	*Acropora*	aff.	*hyacinthus*	C
15Oki27	2015	*Acropora*	cf.	*cytherea*	B
15Oki29	2015	*Acropora*	cf.	*bifurcata*	A
15Oki40	2015	*Acropora*	aff.	*hyacinthus*	C
18Oki01	2018	*Acropora*	aff.	*hyacinthus*	C
18Oki02	2018	*Acropora*	aff.	*hyacinthus*	C
18Oki03	2018	*Acropora*	aff.	*hyacinthus*	C
18Oki04	2018	*Acropora*	aff.	*hyacinthus*	C
18Oki05	2018	*Acropora*	aff.	*hyacinthus*	C
18Oki06	2018	*Acropora*	aff.	*hyacinthus*	C
18Oki07	2018	*Acropora*	cf.	*cytherea*	B

18Oki08	2018	*Acropora*	cf.	*cytherea*	B
18Oki09	2018	*Acropora*	cf.	*cytherea*	B
18Oki10	2018	*Acropora*	cf.	*cytherea*	B
18Oki11	2018	*Acropora*	cf.	*cytherea*	B
18Oki12	2018	*Acropora*	cf.	*cytherea*	B
18Oki13	2018	*Acropora*	cf.	*bifurcata*	A
18Oki14	2018	*Acropora*	cf.	*bifurcata*	A
18Oki15	2018	*Acropora*	cf.	*bifurcata*	A
18Oki16	2018	*Acropora*	cf.	*bifurcata*	A
18Oki17	2018	*Acropora*	cf.	*bifurcata*	A
18Oki18	2018	*Acropora*	cf.	*bifurcata*	A
18Oki19	2018	*Acropora*	aff.	*hyacinthus*	C
18Oki20	2018	*Acropora*	aff.	*hyacinthus*	C
18Oki21	2018	*Acropora*	aff.	*hyacinthus*	C
18Oki22	2018	*Acropora*	aff.	*hyacinthus*	C
18Oki23	2018	*Acropora*	aff.	*hyacinthus*	C
18Oki24	2018	*Acropora*	aff.	*hyacinthus*	C
18Oki25	2018	*Acropora*	cf.	*cytherea*	B
18Oki26	2018	*Acropora*	cf.	*cytherea*	B
18Oki27	2018	*Acropora*	cf.	*cytherea*	B
18Oki28	2018	*Acropora*	cf.	*cytherea*	B
18Oki29	2018	*Acropora*	cf.	*cytherea*	B
18Oki30	2018	*Acropora*	cf.	*cytherea*	B
18Oki31	2018	*Acropora*	cf.	*bifurcata*	A
18Oki32	2018	*Acropora*	cf.	*bifurcata*	A
18Oki33	2018	*Acropora*	cf.	*bifurcata*	A
18Oki34	2018	*Acropora*	cf.	*bifurcata*	A
18Oki35	2018	*Acropora*	cf.	*bifurcata*	A
18Oki36	2018	*Acropora*	cf.	*bifurcata*	A
19Oki12	2019	*Acropora*	cf.	*bifurcata*	A
19Oki13	2019	*Acropora*	aff.	*hyacinthus*	C
19Oki14	2019	*Acropora*	cf.	*bifurcata*	A
19Oki15	2019	*Acropora*	aff.	*hyacinthus*	C
19Oki16	2019	*Acropora*	aff.	*hyacinthus*	C
19Oki17	2019	*Acropora*	cf.	*cytherea*	B
19Oki18	2019	*Acropora*	cf.	*cytherea*	B
19Oki20	2019	*Acropora*	cf.	*bifurcata*	A
19Oki22	2019	*Acropora*	cf.	*bifurcata*	A
19Oki23	2019	*Acropora*	cf.	*bifurcata*	A
19Oki24	2019	*Acropora*	aff.	*hyacinthus*	C
19Oki25	2019	*Acropora*	cf.	*cytherea*	B

2.2.2 SUPPLEMENTARY TABLE S2. Statistics summary of variables assessed in this study.

Significance values (*p*-values) from the normality assumption tested (Shapiro-Wilk test, α = 0.05) for all the estimated univariate measures either jointly (S-W) or per species (S-W ~ spp) are displayed: "all", if they met the assumption, or highlighting the species that do not (i.e., A, B, C). All variables that deviated from the assumption were transformed (TR) using the most frequently suggested option after five iterations of the *bestNormalize* function (TR bestNormalize). Homogeneity of variance was also tested on the TR data set (Levene's test, α = 0.05) and corresponding *p*-values are displayed. Variables that did not meet assumptions of normality and homoscedasticity after transformation, were excluded from downstream analyses.

Variable	Type	S-W	S-W ~ spp	TR bestNormalize	S-W TR	S-W ~ spp TR	Levene TR
b_angle_mean	Branch	0.6383	all	N/A	0.6383	all	0.4786
b_angle_var	Branch	0.0020	all	log_x	0.2250	all	0.1917
br_length_mean	Branch	0.1426	all	N/A	0.1426	all	0.4813
br_length_var	Branch	0.0004	B, C	boxcox	0.9758	all	0.8372
br_rate_mean	Branch	0.2714	all	N/A	0.2714	all	0.0981
br_rate_var	Branch	0.0013	C	log_x	0.8669	all	0.9174
br_spacing_v1_mean	Branch	0.0139	all	orderNorm	1.0000	all	0.6831
br_spacing_v1_var	Branch	0.0001	A, C	log_x	0.4921	all	0.7742
br_spacing_v2_mean	Branch	0.0557	all	orderNorm	1.0000	all	0.0941
br_spacing_v2_var	Branch	0.0000	A, C	log_x	0.4481	all	0.8609
d_avg_var	Branch	0.0000	A, B, C	boxcox	0.5823	all	0.4331
da_avg_mean	Branch	0.0000	A, B, C	boxcox	0.2069	all	0.3737
da_mean	Branch	0.0000	A, B, C	boxcox	0.4611	all	0.3714
da_var	Branch	0.0000	A, B, C	boxcox	0.2992	all	0.2950
db_mean	Branch	0.0001	A, C	orderNorm	1.0000	all	0.4155
db_var	Branch	0.0000	A, B, C	orderNorm	1.0000	all	0.5417
dc_mean	Branch	0.0100	A, C	orderNorm	1.0000	all	0.4190
dc_var	Branch	0.0000	A, B, C	log_x	0.0920	all	0.6926
FD	Complexity	0.0622	A	orderNorm	1.0000	all	0.5942
H_kurt	Curvature	0.1187	C	yeojohnson	0.8190	all	0.7095
H_mean	Curvature	0.8538	all	N/A	0.8538	all	0.5621
H_skew	Curvature	0.5515	all	N/A	0.5515	all	0.3555
H_tip_kurt	Curvature	0.5158	all	N/A	0.5158	all	0.3148
H_tip_mean	Curvature	0.3281	B	orderNorm	1.0000	all	0.6047
H_tip_skew	Curvature	0.0081	all	orderNorm	1.0000	all	0.4073
H_tip_var	Curvature	0.0005	B, C	orderNorm	1.0000	all	0.5702
H_var	Curvature	0.0001	B, C	orderNorm	1.0000	all	0.1768
K_kurt	Curvature	0.2770	all	N/A	0.2770	all	0.1579
K_mean	Curvature	0.0001	B, C	orderNorm	1.0000	all	0.7620
K_skew	Curvature	0.0002	A	orderNorm	1.0000	all	0.6720
K_tip_kurt	Curvature	0.8411	all	N/A	0.8411	all	0.0955
K_tip_mean	Curvature	0.1056	C	center_scale	0.1056	all	0.0435
K_tip_skew	Curvature	0.0509	all	log_x	0.7247	all	0.0117

K_tip_var	Curvature	0.0000	A, B, C	orderNorm	1.0000	all	0.3229
K_var	Curvature	0.0000	A, B, C	orderNorm	1.0000	all	0.2573
k1_kurt	Curvature	0.0356	C	N/A	1.0000	all	0.2257
k1_mean	Curvature	0.0031	B, C	orderNorm	1.0000	all	0.1942
k1_skew	Curvature	0.0209	all	boxcox	0.7109	all	0.6118
k1_tip_kurt	Curvature	0.0001	C	orderNorm	1.0000	all	0.2341
k1_tip_mean	Curvature	0.0039	A	orderNorm	1.0000	all	0.8684
k1_tip_skew	Curvature	0.6146	all	N/A	0.6146	all	0.3792
k1_tip_var	Curvature	0.0008	B, C	orderNorm	1.0000	all	0.5904
k1_var	Curvature	0.0001	B, C	orderNorm	1.0000	all	0.2058
k2_kurt	Curvature	0.0874	C	boxcox	0.8008	all	0.8111
k2_mean	Curvature	0.0008	A, B, C	orderNorm	1.0000	all	0.2245
k2_skew	Curvature	0.0072	A, B	orderNorm	1.0000	all	0.2228
k2_tip_kurt	Curvature	0.0472	all	yeojohnson	0.6917	all	0.8340
k2_tip_mean	Curvature	0.2727	C	orderNorm	1.0000	C	0.8376
k2_tip_skew	Curvature	0.5219	all	N/A	0.5219	all	0.0073
k2_tip_var	Curvature	0.0004	B, C	orderNorm	1.0000	all	0.4173
k2_var	Curvature	0.0001	B, C	orderNorm	1.0000	all	0.2253
S	Complexity	0.0127	all	boxcox	0.9444	all	0.9506
SV_ratio	Complexity	0.1301	all	N/A	0.1301	all	0.8204

2.2.3 SUPPLEMENTARY TABLE S3. Results of the analysis of variance (ANOVA).

Variable	*p*-value	Significance
b_angle_mean	4.92E-01	ns
b_angle_var	2.10E-01	ns
br_length_mean	1.34E-05	****
br_length_var	9.07E-05	****
br_rate_mean	3.46E-06	****
br_rate_var	2.40E-05	****
br_spacing_v1_mean	6.96E-07	****
br_spacing_v1_var	0.003352	**
br_spacing_v2_mean	7.29E-05	****
br_spacing_v2_var	4.05E-02	*
d_avg_var	0.17765	ns
da_avg_mean	2.75E-02	*
da_mean	9.33E-03	**
da_var	6.82E-02	ns
db_mean	6.56E-05	****
db_var	1.48E-02	*
dc_mean	9.86E-08	****
dc_var	0.000694	***
FD	5.19E-01	ns
H_kurt	3.89E-01	ns
H_mean	5.61E-03	**
H_skew	8.84E-05	****
H_tip_kurt	0.00126	**
H_tip_mean	2.23E-05	****
H_tip_skew	5.48E-09	****
H_tip_var	5.02E-08	****
H_var	1.29E-08	****
K_kurt	0.003158	**
K_mean	1.75E-06	****
K_skew	0.251868	ns
K_tip_kurt	4.13E-04	***
K_tip_mean	2.18E-08	****
K_tip_skew	7.00E-01	ns
K_tip_var	5.44E-08	****
K_var	3.24E-08	****
k1_kurt	1.37E-01	ns
k1_mean	1.47E-08	****
k1_skew	1.51E-01	ns
k1_tip_kurt	1.33E-02	*
k1_tip_mean	2.92E-07	****
k1_tip_skew	0.000443	***
k1_tip_var	1.19E-07	****
k1_var	5.10E-08	****
k2_kurt	4.97E-02	*
k2_mean	2.27E-08	****
k2_skew	0.109138	ns
k2_tip_kurt	1.37E-05	****
k2_tip_mean	2.11E-08	****
k2_tip_skew	1.34E-03	**
k2_tip_var	3.02E-08	****
k2_var	2.21E-08	****
S	4.16E-02	*
SV_ratio	0.01883	*

Significance (*p*-values):

(<0.0001) ****

(<=0.001)***

(<=0.01)**

(<0.05)*

(>=0.05) ns

2.2.4 SUPPLEMENTARY TABLE S4. Results of the post-hoc Tukey test for interspecific differences.

Significance values (*p*-values) from post-hoc Tukey tests for significant interspecific differences between A (*n*=28), B (*n*=21) and C (*n*=25) are displayed for each variable in the corresponding order: A vs. B, A vs. C and finally B vs. C. The corresponding significance levels according to an α of 0.05 for each pairwise comparison are also shown so that: (<0.0001) ****, (≤0.001)***, (≤0.01)**, (<0.05)* and (≥0.05) ns.

Variable	*p*-value	*p*-value Sig.	Sig. species pair
b_angle_mean	0.6048	ns	ns
b_angle_mean	0.5345	ns	ns
b_angle_mean	0.9980	ns	ns
b_angle_var	0.1819	ns	ns
b_angle_var	0.7149	ns	ns
b_angle_var	0.5709	ns	ns
br_length_mean	0.0032	**	AvsB
br_length_mean	0.0000	****	AvsC
br_length_mean	0.3914	ns	ns
br_length_var	0.0003	***	AvsB
br_length_var	0.0010	**	AvsC
br_length_var	0.8698	ns	ns
br_rate_mean	0.0029	**	AvsB
br_rate_mean	0.0000	****	AvsC
br_rate_mean	0.2305	ns	ns
br_rate_var	0.0002	***	AvsB
br_rate_var	0.0002	***	AvsC
br_rate_var	0.9694	ns	ns
br_spacing_v1_mean	0.9259	ns	ns
br_spacing_v1_mean	0.0000	****	AvsC
br_spacing_v1_mean	0.0000	****	BvsC
br_spacing_v1_var	0.7149	ns	ns
br_spacing_v1_var	0.0030	**	AvsC
br_spacing_v1_var	0.0481	*	BvsC
br_spacing_v2_mean	0.6938	ns	ns
br_spacing_v2_mean	0.0001	****	AvsC
br_spacing_v2_mean	0.0035	**	BvsC
br_spacing_v2_var	0.0605	ns	ns
br_spacing_v2_var	1.0000	ns	ns
br_spacing_v2_var	0.0686	ns	ns
d_avg_var	0.1797	ns	ns
d_avg_var	0.9541	ns	ns
d_avg_var	0.3085	ns	ns
da_avg_mean	0.2608	ns	ns
da_avg_mean	0.0220	*	AvsC
da_avg_mean	0.5862	ns	ns
da_mean	0.3941	ns	ns

da_mean	0.0065	**	AvsC
da_mean	0.2285	ns	ns
da_var	0.0558	ns	ns
da_var	0.4062	ns	ns
da_var	0.5224	ns	ns
db_mean	0.8719	ns	ns
db_mean	0.0001	***	AvsC
db_mean	0.0016	**	BvsC
db_var	0.0113	*	AvsB
db_var	0.2189	ns	ns
db_var	0.3848	ns	ns
dc_mean	0.1209	ns	ns
dc_mean	0.0001	****	AvsC
dc_mean	0.0000	****	BvsC
dc_var	0.0004	***	AvsB
dc_var	0.1007	ns	ns
dc_var	0.1270	ns	ns
FD	0.9150	ns	ns
FD	0.7147	ns	ns
FD	0.5053	ns	ns
H_kurt	0.5145	ns	ns
H_kurt	0.4338	ns	ns
H_kurt	0.9968	ns	ns
H_mean	0.9944	ns	ns
H_mean	0.0123	*	AvsC
H_mean	0.0165	*	BvsC
H_skew	0.0001	****	AvsB
H_skew	0.1999	ns	ns
H_skew	0.0144	*	BvsC
H_tip_kurt	0.0062	**	AvsB
H_tip_kurt	0.0037	**	AvsC
H_tip_kurt	0.9999	ns	ns
H_tip_mean	0.9709	ns	ns
H_tip_mean	0.0001	****	AvsC
H_tip_mean	0.0004	***	BvsC
H_tip_skew	0.0000	****	AvsB
H_tip_skew	0.5614	ns	ns
H_tip_skew	0.0000	****	BvsC
H_tip_var	0.0000	****	AvsB
H_tip_var	0.0163	*	AvsC
H_tip_var	0.0010	***	BvsC
H_var	0.0000	****	AvsB
H_var	0.0021	**	AvsC

H_var	0.0030	**	BvsC
K_kurt	0.0694	ns	ns
K_kurt	0.0027	**	AvsC
K_kurt	0.5745	ns	ns
K_mean	0.0000	****	AvsB
K_mean	0.0004	***	AvsC
K_mean	0.2889	ns	ns
K_skew	0.8401	ns	ns
K_skew	0.4822	ns	ns
K_skew	0.2418	ns	ns
K_tip_kurt	0.0045	**	AvsB
K_tip_kurt	0.0009	***	AvsC
K_tip_kurt	0.9490	ns	ns
K_tip_mean	0.0000	****	AvsB
K_tip_mean	0.0889	ns	ns
K_tip_mean	0.0001	****	BvsC
K_tip_skew	0.6798	ns	ns
K_tip_skew	0.9588	ns	ns
K_tip_skew	0.8407	ns	ns
K_tip_var	0.0000	****	AvsB
K_tip_var	0.0070	**	AvsC
K_tip_var	0.0026	**	BvsC
K_var	0.0000	****	AvsB
K_var	0.0039	**	AvsC
K_var	0.0032	**	BvsC
k1_kurt	0.2606	ns	ns
k1_kurt	0.1647	ns	ns
k1_kurt	0.9847	ns	ns
k1_mean	0.0000	****	AvsB
k1_mean	0.0003	***	AvsC
k1_mean	0.0233	*	BvsC
k1_skew	0.9818	ns	ns
k1_skew	0.2242	ns	ns
k1_skew	0.2014	ns	ns
k1_tip_kurt	0.1902	ns	ns
k1_tip_kurt	0.0105	*	AvsC
k1_tip_kurt	0.5382	ns	ns
k1_tip_mean	0.0000	****	AvsB
k1_tip_mean	0.0002	***	AvsC
k1_tip_mean	0.1681	ns	ns
k1_tip_skew	0.2907	ns	ns
k1_tip_skew	0.0191	*	AvsC
k1_tip_skew	0.0004	***	BvsC
k1_tip_var	0.0000	****	AvsB
k1_tip_var	0.0170	*	AvsC

k1_tip_var	0.0019	**	BvsC
k1_var	0.0000	****	AvsB
k1_var	0.0017	**	AvsC
k1_var	0.0104	*	BvsC
k2_kurt	0.9683	ns	ns
k2_kurt	0.0575	ns	ns
k2_kurt	0.1367	ns	ns
k2_mean	0.0000	****	AvsB
k2_mean	0.0055	**	AvsC
k2_mean	0.0017	**	BvsC
k2_skew	0.1527	ns	ns
k2_skew	0.1913	ns	ns
k2_skew	0.9789	ns	ns
k2_tip_kurt	0.5843	ns	ns
k2_tip_kurt	0.0004	***	AvsC
k2_tip_kurt	0.0000	****	BvsC
k2_tip_mean	0.0000	****	AvsB
k2_tip_mean	0.1749	ns	ns
k2_tip_mean	0.0000	****	BvsC
k2_tip_skew	0.0046	**	AvsB
k2_tip_skew	0.0055	**	AvsC
k2_tip_skew	0.9762	ns	ns
k2_tip_var	0.0000	****	AvsB
k2_tip_var	0.0102	*	AvsC
k2_tip_var	0.0011	**	BvsC
k2_var	0.0000	****	AvsB
k2_var	0.0049	**	AvsC
k2_var	0.0019	**	BvsC
S	0.2920	ns	ns
S	0.0345	*	AvsC
S	0.6461	ns	ns
SV_ratio	0.1068	ns	ns
SV_ratio	0.0204	*	AvsC
SV_ratio	0.8452	ns	ns

2.2.5 SUPPLEMENTARY TABLE S5. Pearson correlation coefficients (ρ) for the preliminary selected variables.

Pearson correlation coefficients (ρ) and the corresponding p-values for pairwise comparisons of preliminary selected variables with $|\rho| > 0.05$ and p-value ≥ 0.05 are depicted.

Variable 1	Variable 2	ρ	p-value
br_length_mean	br_length_var	0.821971	0.00E+00
br_length_mean	br_rate_mean	0.977405	0
br_length_mean	br_rate_var	0.813567	0.00E+00
br_length_mean	br_spacing_v1_mean	0.612135	6.83E-09
br_length_mean	br_spacing_v2_mean	0.564328	1.64E-07
br_length_mean	db_mean	0.573809	9.07E-08
br_length_mean	K_mean	0.50865	3.71E-06
br_length_var	br_rate_mean	0.753196	9.77E-15
br_length_var	br_rate_var	0.940212	0.00E+00
br_length_var	db_mean	0.585441	4.29E-08
br_rate_mean	br_rate_var	0.805754	0.00E+00
br_rate_mean	br_spacing_v1_mean	0.681958	2.24E-11
br_rate_mean	br_spacing_v2_mean	0.595146	2.24E-08
br_rate_mean	db_mean	0.535541	8.81E-07
br_rate_mean	K_mean	0.510711	3.34E-06
br_rate_var	br_spacing_v1_var	0.508633	3.71E-06
br_rate_var	db_mean	0.500248	5.67E-06
br_spacing_v1_mean	br_spacing_v1_var	0.515338	2.62E-06
br_spacing_v1_mean	br_spacing_v2_mean	0.84211	0.00E+00
br_spacing_v1_var	br_spacing_v2_mean	0.684745	1.73E-11
br_spacing_v1_var	db_mean	0.667435	8.35E-11
br_spacing_v1_var	dc_mean	0.610329	7.78E-09
br_spacing_v2_mean	db_mean	0.599635	1.65E-08
br_spacing_v2_mean	dc_mean	0.525311	1.54E-06
db_mean	dc_mean	0.882935	0.00E+00
H_mean	H_tip_mean	0.769637	1.11E-15
H_skew	H_tip_skew	0.728279	1.94E-13
H_skew	K_mean	0.706734	1.98E-12
H_skew	k2_mean	0.756884	6.22E-15
H_tip_kurt	H_tip_var	0.627013	2.27E-09
H_tip_kurt	H_var	0.671125	6.02E-11
H_tip_kurt	K_tip_kurt	0.8954	0
H_tip_kurt	K_tip_var	0.667175	8.54E-11
H_tip_kurt	K_var	0.669626	6.88E-11
H_tip_kurt	k1_mean	0.639956	8.32E-10
H_tip_kurt	k1_tip_mean	0.603323	1.28E-08
H_tip_kurt	k1_tip_skew	0.644084	5.98E-10
H_tip_kurt	k1_tip_var	0.663498	1.18E-10
H_tip_kurt	k1_var	0.699778	4.01E-12
H_tip_kurt	k2_tip_kurt	0.755303	7.55E-15
H_tip_kurt	k2_tip_var	0.63515	1.22E-09
H_tip_kurt	k2_var	0.673067	5.06E-11
H_tip_mean	k1_tip_mean	0.514985	2.67E-06
H_tip_skew	K_mean	0.550019	3.85E-07

H_tip_skew	k2_mean	0.607618	9.43E-09
H_tip_var	H_var	0.942736	0.00E+00
H_tip_var	K_tip_kurt	0.615587	5.32E-09
H_tip_var	K_tip_var	0.991478	0
H_tip_var	K_var	0.954129	0.00E+00
H_tip_var	k1_mean	0.90962	0.00E+00
H_tip_var	k1_tip_mean	0.893528	0.00E+00
H_tip_var	k1_tip_var	0.994328	0.00E+00
H_tip_var	k1_var	0.949508	0.00E+00
H_tip_var	k2_tip_var	0.991872	0.00E+00
H_tip_var	k2_var	0.938409	0.00E+00
H_var	K_tip_kurt	0.653429	2.78E-10
H_var	K_tip_var	0.946437	0.00E+00
H_var	K_var	0.99193	0.00E+00
H_var	k1_mean	0.952507	0.00E+00
H_var	k1_tip_mean	0.845339	0.00E+00
H_var	k1_tip_var	0.936956	0.00E+00
H_var	k1_var	0.992149	0.00E+00
H_var	k2_tip_var	0.954843	0.00E+00
H_var	k2_var	0.995836	0.00E+00
K_mean	k2_mean	0.829018	0.00E+00
K_tip_kurt	K_tip_var	0.645418	5.37E-10
K_tip_kurt	K_var	0.645409	5.37E-10
K_tip_kurt	k1_mean	0.584395	4.59E-08
K_tip_kurt	k1_tip_mean	0.537042	8.10E-07
K_tip_kurt	k1_tip_skew	0.598583	1.77E-08
K_tip_kurt	k1_tip_var	0.655142	2.40E-10
K_tip_kurt	k1_var	0.690724	9.76E-12
K_tip_kurt	k2_tip_kurt	0.636663	1.08E-09
K_tip_kurt	k2_tip_var	0.606152	1.05E-08
K_tip_kurt	k2_var	0.65368	2.72E-10
K_tip_var	K_var	0.961587	0.00E+00
K_tip_var	k1_mean	0.930562	0
K_tip_var	k1_tip_mean	0.91759	0
K_tip_var	k1_tip_var	0.986314	0.00E+00
K_tip_var	k1_var	0.95547	0.00E+00
K_tip_var	k2_tip_var	0.992263	0.00E+00
K_tip_var	k2_var	0.940085	0.00E+00
K_var	k1_mean	0.962754	0
K_var	k1_tip_mean	0.870886	0
K_var	k1_tip_var	0.947507	0.00E+00
K_var	k1_var	0.985157	0.00E+00
K_var	k2_tip_var	0.967258	0.00E+00
K_var	k2_var	0.98694	0.00E+00
k1_mean	k1_tip_mean	0.918782	0
k1_mean	k1_tip_var	0.89504	0.00E+00
k1_mean	k1_var	0.945341	0.00E+00
k1_mean	k2_tip_var	0.930863	0.00E+00
k1_mean	k2_var	0.942417	0
k1_tip_mean	k1_tip_var	0.882499	0.00E+00

k1_tip_mean	k1_var	0.851984	0.00E+00
k1_tip_mean	k2_tip_var	0.899742	0.00E+00
k1_tip_mean	k2_var	0.830025	0
k1_tip_skew	k2_tip_kurt	0.843069	0
k1_tip_var	k1_var	0.947465	0
k1_tip_var	k2_tip_var	0.981884	0.00E+00
k1_tip_var	k2_var	0.932699	0
k1_var	k2_tip_var	0.956563	0.00E+00
k1_var	k2_var	0.986052	0
k2_tip_var	k2_var	0.952764	0
br_length_var	k1_mean	-0.52341	1.71E-06
br_rate_mean	k1_mean	-0.54885	4.13E-07
br_spacing_v1_mean	k1_tip_skew	-0.52138	1.91E-06
br_spacing_v1_mean	k2_tip_kurt	-0.61531	5.43E-09
br_spacing_v1_var	H_tip_mean	-0.65794	1.90E-10
br_spacing_v2_mean	k2_tip_kurt	-0.5085	3.74E-06
db_mean	H_mean	-0.81004	0.00E+00
db_mean	H_tip_mean	-0.89419	0.00E+00
db_mean	k1_tip_mean	-0.50484	4.50E-06
dc_mean	H_mean	-0.71907	5.37E-13
dc_mean	H_tip_mean	-0.82991	0.00E+00
H_skew	H_tip_kurt	-0.58341	4.90E-08
H_skew	H_tip_var	-0.71609	7.40E-13
H_skew	H_var	-0.70968	1.46E-12
H_skew	K_tip_kurt	-0.50939	3.57E-06
H_skew	K_tip_var	-0.69266	8.09E-12
H_skew	K_var	-0.69669	5.45E-12
H_skew	k1_mean	-0.64516	5.48E-10
H_skew	k1_tip_mean	-0.61844	4.32E-09
H_skew	k1_tip_var	-0.72457	2.93E-13
H_skew	k1_var	-0.69521	6.30E-12
H_skew	k2_tip_var	-0.71348	9.78E-13
H_skew	k2_var	-0.73637	7.64E-14
H_tip_kurt	K_mean	-0.68855	1.20E-11
H_tip_kurt	k2_mean	-0.6794	2.84E-11
H_tip_skew	H_tip_var	-0.65541	2.35E-10
H_tip_skew	H_var	-0.62957	1.87E-09
H_tip_skew	K_tip_var	-0.63931	8.76E-10
H_tip_skew	K_var	-0.62048	3.71E-09
H_tip_skew	k1_mean	-0.65588	2.26E-10
H_tip_skew	k1_tip_mean	-0.64812	4.31E-10
H_tip_skew	k1_tip_var	-0.62316	3.04E-09
H_tip_skew	k1_var	-0.59978	1.63E-08
H_tip_skew	k2_tip_var	-0.67606	3.86E-11
H_tip_skew	k2_var	-0.63867	9.22E-10
H_tip_var	K_mean	-0.83974	0.00E+00
H_tip_var	k2_mean	-0.93963	0.00E+00
H_var	K_mean	-0.82041	0.00E+00
H_var	k2_mean	-0.98205	0.00E+00
K_mean	K_tip_kurt	-0.6516	3.23E-10
K_mean	K_tip_var	-0.85764	0.00E+00

K_mean	K_var	-0.84883	0
K_mean	k1_mean	-0.8688	0.00E+00
K_mean	k1_tip_mean	-0.88048	0.00E+00
K_mean	k1_tip_var	-0.84009	0.00E+00
K_mean	k1_var	-0.83269	0.00E+00
K_mean	k2_tip_var	-0.84385	0.00E+00
K_mean	k2_var	-0.81363	0.00E+00
K_tip_kurt	k2_mean	-0.67552	4.05E-11
K_tip_var	k2_mean	-0.93646	0.00E+00
K_var	k2_mean	-0.97427	0.00E+00
k1_mean	k2_mean	-0.91009	0
k1_tip_mean	k2_mean	-0.8053	0.00E+00
k1_tip_var	k2_mean	-0.93839	0.00E+00
k1_var	k2_mean	-0.97668	0.00E+00
k2_mean	k2_tip_var	-0.94423	0.00E+00
k2_mean	k2_var	-0.98329	0.00E+00

2.2.6 SUPPLEMENTARY TABLE S6. Variable selection for dimensionality reduction using Gaussian model-based clustering.

The set of variables that best discriminated groups using normal mixture models (NMMs) defined using the greedy algorithm both in forward and backward directions is summarized. The Mclust model, the supported number of components or groups (G) detected in each case, the corresponding Log-likelihood, degrees of freedom (df), Bayesian information criterion (BIC) and integrated complete likelihood (ICL) are also depicted along with the lists of variables.

	Direction	
	Backward	**Forward**
Mclust model	EEE (ellipsoidal, equal volume, shape and orientation)	EEI (diagonal, equal volume and shape)
Log-likelihood	340.7631	111.8223
df	188	13
BIC	-127.638	167.6918
ICL	-127.638	161.9822
G (# components)	2	2
Initial # variables	29	29
Selected # variables	17	4
Variable subsets	br_length_mean br_rate_mean br_rate_var br_spacing_v1_mean br_spacing_v2_mean dc_mean H_mean H_skew H_tip_var K_tip_kurt K_var k1_tip_mean k1_tip_skew k1_var k2_tip_kurt k2_tip_var H_tip_kurt	br_length_mean H_skew k1_tip_mean H_tip_kurt

2.2.7 SUPPLEMENTARY TABLE S7. Results of variance inflation factor (VIF) assessment.

Variables that did not present collinearity (Variance Inflation Factor (VIF) <10) from the complete data set are listed below. Those that were also present in the preliminary selected subset depicting significant interspecific differentiation are indicated accordingly.

Variables	VIF	Preliminary selected subset
H_mean	6.128335	YES
H_skew	7.747286	YES
K_tip_kurt	9.83418	YES
b_angle_mean	3.59385	
b_angle_var	2.772566	
br_rate_var	4.074808	YES
br_spacing_v1_mean	3.920421	YES
br_spacing_v1_var	5.065652	YES
br_spacing_v2_var	2.643434	
d_avg_var	6.434551	
da_avg_mean	4.359019	
da_var	8.221089	
FD	4.806128	
H_tip_mean	5.538008	YES
H_tip_skew	6.452885	YES
H_kurt	5.397675	
K_skew	3.116537	
k1_tip_kurt	5.612972	
k2_mean	6.399289	YES
k2_tip_kurt	6.550996	YES
S	5.967713	

2.2.8 SUPPLEMENTARY TABLE S8. Results of the multiple analysis of variance (MANOVA).

The summarized results of a multivariate analysis of variance (MANOVA) using three different tests (i.e., Pillai, Wilks, and Roy) and the variables set are presented. Degrees of freedom (Df) along with *p*-values and MANOVA statistics are reported. Significance levels correspond to: (<0.001)***, (≤0.01)**, (≤0.01)*, (≤0.05). and (≤0.1)‘ ’.

	Df	**Pillai**	**Approx. F**	**num Df**	**den Df**	***p*-value**	**Significance**
(Intercept)	1	0.99879	2012.83	21	51	< 2.2e-16	***
Species	2	1.47443	6.95	42	104	5.337e-16	***
Residuals	71						
	Df	**Wilks**	**Approx. F**	**num Df**	**den Df**	***p*-value**	**Significance**
(Intercept)	1	0.001205	2012.83	21	51	< 2.2e-16	***
Species	2	0.060179	7.47	42	102	< 2.2e-16	***
Residuals	71						
	Df	**Roy**	**Approx. F**	**num Df**	**den Df**	***p*-value**	**Significance**
(Intercept)	1	828.81	2012.83	21	51	< 2.2e-16	***
Species	2	4.93	12.21	21	52	1.982e-13	***
Residuals	71						

2.2.9 SUPPLEMENTARY TABLE S9. Linear discriminant analysis (LDA) prediction accuracy.

The summarized results of the linear discriminant analysis (LDA) are reported for the complete data set and the randomly selected training (68% of data) and testing (32% of data) subsets using prediction accuracy or confusion matrices. The central diagonal represents the number of specimens that were correctly identified by the linear model in each one of the cases. The accuracy percentage was calculated for each subset based on the total number of specimens in each set.

Complete (*n*= 74)	Is species A	Is species B	Is species C
Allocated to species A	26	0	0
Allocated to species B	1	21	0
Allocated to species C	1	0	25
		Accuracy %	97.30%

Training (*n*= 50)	Is species A	Is species B	Is species C
Allocated to species A	21	0	0
Allocated to species B	0	12	0
Allocated to species C	0	0	17
		Accuracy %	100%

Testing (*n*= 24)	Is species A	Is species B	Is species C
Allocated to species A	5	1	3
Allocated to species B	1	8	0
Allocated to species C	1	0	5
		Accuracy %	75%

2.2.10 SUPPLEMENTARY TABLE S10. Prediction accuracy of discriminant analysis based on Gaussian mixture modelling.

The summarized results of the discriminant analysis based on Gaussian mixture modelling (Supplementary Materials and Methods) are reported for the randomly selected training (68% of data) and testing (32% of data) subsets using two different approaches: the eigenvalue decomposition discriminant analysis (EDDA), which assumes that the density for each class can be described by a single Gaussian component; and the MclustDA approach that uses a finite mixture of Gaussian distributions within each class (mclust R package v5.4.7). Log-likelihood (Log-L), number of specimens used (n), degrees of freedom (df), Bayesian information criterion (BIC), number of components in each class (G, groups), accuracy (AC%), and the cross-validation classification error (CE) and standard error (SE) of the cross-validated statistic are also reported for each approach.

EDDA					**MclustDA**				
Log-L	**n**		**df**	**BIC**	**Log-L**	**n**		**df**	**BIC**
-303.8129	50		99	-994.9162	-284.1052	50		169	-1229.342
Classes	**n**	**%**	**Model**	**G**	**Classes**	**n**	**%**	**Model**	**G**
A	17	34	EEE	1	A	17	34	XXI	1
B	15	30	EEE	1	B	15	30	EEE	4
C	18	36	EEE	1	C	18	36	EEI	2
Training (*n*= 50)	Allocated to species A	Allocated to species B	Allocated to species C		**Training (*n*= 50)**	Allocated to species A	Allocated to species B	Allocated to species C	
Is species A	17	0	0		Is species A	16	0	1	
Is species B	2	13	0		Is species B	0	15	0	
Is species C	1	1	16		Is species C	2	0	16	
		AC%	92%				**AC %**	94%	
Testing (*n*= 24)	Allocated to species A	Allocated to species B	Allocated to species C		**Testing (*n*= 24)**	Allocated to species A	Allocated to species B	Allocated to species C	
Is species A	4	4	3		Is species A	5	0	6	
Is species B	2	4	0		Is species B	3	0	3	
Is species C	0	2	5		Is species C	0	0	7	
		AC %	54.17%				**AC %**	50%	
CE	0.20	**SE**	0. 0.0495		**CE**	0.380	**SE**	0.0526	

3 Supplementary References

Alyassin, A. M., Lancaster, J. L., Downs, J. H., and Fox, P. T. (1994). Evaluation of new algorithms for the interactive measurement of surface area and volume. *Med. Phys.* 21, 741–752. doi:10.1118/1.597333.

Fox, J., and Weisberg, S. (2019). *An R companion to applied regression*. 3rd ed. Thousand Oaks, CA: Sage Available at: https://socialsciences.mcmaster.ca/jfox/Books/Companion/.

Homann, H. (2007). Implementation of a 3D thinning algorithm. *Insight J.*, 4. doi:10.54294/xjdr5f.

Kassambara, A. (2020). ggpubr: "ggplot2" based publication ready plots. Available at: https://cran.r-project.org/package=ggpubr.

Korkmaz, S., Goksuluk, D., and Zararsiz, G. (2014). MVN: an R package for assessing multivariate normality. *R J.* 6, 151–162. Available at: https://journal.r-project.org/archive/2014-2/korkmaz-goksuluk-zararsiz.pdf.

Lee, T. C., Kashyap, R. L., and Chu, C. N. (1994). Building skeleton models via 3D medial surface axis thinning algorithms. *CVGIP Graph. Model. Image Process.* 56, 462–478. doi:10.1006/cgip.1994.1042.

Mesmoudi, M. M., De Floriani, L., and Magillo, P. (2010). A geometric approach to curvature estimation on triangulated 3D shapes. in *Proceedings of the International Conference on Computer Graphics Theory and Applications* (SciTePress - Science and and Technology Publications), 90–95. doi:10.5220/0002825900900095.

Meyer, M., Desbrun, M., Schröder, P., and Barr, A. H. (2003). "Discrete differential-geometry operators for triangulated 2-manifolds," in *Visualization and Mathematics III*, eds. H. C. Hege and K. Polthier (Berlin-Heidelberg: Springer), 35–57. doi:10.1007/978-3-662-05105-4_2.

Peterson, R. A., and Cavanaugh, J. E. (2020). Ordered quantile normalization: a semiparametric transformation built for the cross-validation era. *J. Appl. Stat.* 47, 2312–2327. doi:10.1080/02664763.2019.1630372.

R Core Team (2018). R: A language and environment for statistical computing. Available at: https://www.r-project.org/.

Ramírez-Portilla, C., Baird, A. H., Cowman, P. F., Quattrini, A. M., Harii, S., Sinniger, F., et al. (2022). Solving the coral species delimitation conundrum. *Syst. Biol.* 71, 461–475. doi:10.1093/sysbio/syab077.

Reichert, J., Backes, A. R., Schubert, P., and Wilke, T. (2017). The power of 3D fractal dimensions for comparative shape and structural complexity analyses of irregularly shaped organisms. *Methods Ecol. Evol.* 8, 1650–1658. doi:10.1111/2041-210X.12829.

Reichert, J., Schellenberg, J., Schubert, P., and Wilke, T. (2016). 3D scanning as a highly precise,

reproducible, and minimally invasive method for surface area and volume measurements of scleractinian corals. *Limnol. Oceanogr. Methods* 14, 518–526. doi:10.1002/lom3.10109.

Reinders, F., Jacobson, M. E. D., and Post, F. H. (2000). Skeleton graph generation for feature shape description. in *Data Visualization 2000*, eds. W. Cornelis de Leeuw and R. van Liere (Amsterdam: Springer Netherlands), 73–82. doi:10.1007/978-3-7091-6783-0_8.

Scrucca, L., Fop, M., Murphy, T. B., and Raftery, A. E. (2016). mclust 5: clustering, classification and density estimation using Gaussian finite mixture models. *R J.* 8, 289–317. Available at: http://www.ncbi.nlm.nih.gov/pubmed/27818791.

Wadell, H. (1935). Volume, shape, and roundness of quartz particles. *J. Geol.* 43, 250–280.