Shark mandible evolution reveals patterns of trophic and habitat-mediated diversification Supplementary Information

Faviel A. López-Romero^{1,2*}, Sebastian Stumpf¹, Pepijn Kamminga³, Christine Böhmer^{4,5}, Alan Pradel⁶, Martin D. Brazeau^{7,8}, Jürgen Kriwet¹

¹ University of Vienna, Faculty of Earth Sciences, Geography and Astronomy, Department of Palaeontology, Evolutionary Morphology Research Group, Josef-Holaubek-Platz 2, 1190 Vienna, Austria.

² University of Vienna, Vienna Doctoral School of Ecology and Evolution (VDSEE), Djerassiplatz 1, 1030 Vienna, Austria

³ Naturalis Biodiversity Center, Darwinweg 2, 2333 CR, Leiden, The Netherlands ⁴MECADEV UMR 7179 CNRS/MNHN, Département Adaptations du Vivant, Muséum National d'Histoire Naturelle, CP 55, 57 rue Cuvier, 75231 Paris, France. ⁵Department für Geo- und Umweltwissenschaften und GeoBio-Center, Ludwig-Maximilians-Universität München, Richard-Wagner-Straße 10, 80333 München 6 CR2P, Centre de Recherche en Paléontologie – Paris, Muséum National d'Histoire Naturelle – Sorbonne Université ‐ CNRS, CP 38, 57 rue Cuvier, F75231 Paris Cedex 05, France.

⁷Department of Life Sciences, Imperial College London, Silwood Park Campus, Ascot, SL5 7PY, UK.

⁸Department of Earth Sciences, The Natural History Museum, Cromwell Road, London, SW7 5BD, UK

Supplementary Methods

Landmarking*.*

Six fixed landmarks were digitized on each specimen. In between the landmarks we placed sliding landmarks to capture the full extent of the curvature displayed in the jaws. Landmark 1 correspond to the upper medial symphysis, landmark 2 is the posterior end of the dental groove, landmark 3 is the medial quadratomandibular joint, landmark 4 is the posterior quadratomandibular joint, landmark 5 is the lower symphysis and landmark 6 is the inner lower symphysis. Between landmark 1 and 2 there are 8 sliding landmarks to capture the extent of the dental groove along the jaw. For the landmarks 2 and 3 there are 5 landmarks to describe the curve between the end of the dental groove and the quadratomandibular joint. Two curves encircle landmarks 3 and 4, in the medial view 7 sliding landmarks were used to describe the perimeter comprising the medial quadratomandibular joint and the lateral quadratomandibular joint, while on the ventral view 8 sliding landmarks were used to describe the mandibular knob. For landmarks 4 and 5, 15 sliding landmarks were used to capture the lower curvature of the jaw from the posterior quadratomandibular joint to the lower symphysis. Finally, from landmarks 2 and 6, 8 semilandmarks were used for the curve of the dental sulcus from the posterior end of the dentition to the lower ventral symphysis. This configuration was captured for every specimen. Additionally, in the specimen of *Carcharhinus acronotus*, this configuration was used to also place surface landmarks on the medial view of the jaw, with 43 surface landmarks this configuration was used as a template, for a total of 100 coordinates. Initially, a configuration with 257 coordinates was used, however this was re-estimated with the LASEC function (1), which indicated that around 100 coordinates were sufficient to describe the shape among the number of specimens (Supplementary figure S2). The complete configuration of landmarks and semilandmarks for all the specimens was then imported as a PTS file for subsequent analysis.

Initially with the reference template built in *Carcharhinus acronotus*, we used the function createAtlas in the R package Morpho (2). This function allowed us to have a reference for the rest of the coordinates configurations in the other specimens and designate specific coordinates to be treated as fixed, curve or surface. We then built an array object from the PTS file which was resampled to remove duplicated landmarks (especially in the connection extremes of the curves used), and this array was used with the reference atlas with the placePatch function. This function was used with varying parameters since the scale can make the arrangement of surface landmarks to displace them and thus not capturing the surface shape. The final parameters for the placePatch function designated were: ray = TRUE, inflate = 3.5, tol = 5, relaxpatch = FALSE. The array of coordinates finally was then used for the geometric morphometric analyses.

Diet and ecology data

We updated the data from the work of Cortes (3); we first checked the raw values. In some cases, on the original table the values can go up to 100.1 % which is unreliable. In some of the cases this can be caused by an automatic rounding of the values which have decimals as 0.66667 to be equal to 0.7. Therefore, we first corrected such instances to have adjusted values. We then removed the species for which we do not have a specimen with jaws

reconstruction. In some specific cases we replaced a species feeding data with the closest relative for which stomach content studies have been performed. This was only done for a few specimens which were underrepresented within an order. Such instances are highlighted in the supplementary data table 1 with their respective reference. With the prey categories described by Cortes (3) we complemented from the literature for other species and create a table with percentages used for the Index of Relative Importance (Supplementary Data Table 2). The final table consist of 75 species. With this table we performed a clustering with the Bray Curtis Index followed by clustering by UPGMA. However, after the first observation of the cluster dendrogram (Supplementary Figure S3) we notice certain species affect the outcome of the average silhouette width (with the find k function in dendextend R package (4). For this reason, we ran the analysis again excluding these species for which a 100% of diet component is present (Supplementary figure S4). Overall, the defined groups were consistent in the second dendrogram, however, other groups like the Generalists and Crustacean consumers required further assessment. We decided to keep the Generalist group since it shows a wider prey spectrum compared to the Crustacean consumers as shown in the heatmap. This groups were used as categorical variables for other comparisons with the shape and size data (Supplementary Data Table 3). In the case of the trophic levels, we first obtained the information from the Index of Relative Importance of each of the used species as reported [\(http://www.fishbase.org/\)](http://www.fishbase.org/) (5), as in the case of stomach content information, some species do not have an exact estimation of the trophic level. In these particular cases the reported value is estimated from the value on close relatives of the species. This information is also highlighted in the supplementary Data Table 1. The trophic level was then used as a categorical variable as proposed by Froese (6), with values lower than 3.8 as Low-level predator (LP; n = 20), from 3.81-4.2 as Mesopredators (MP; $n = 41$), and values greater than 4.2 as top predators (TP; $n = 29$). But also, the raw values of trophic level were also used with the shape variables for comparison. The information on ecological lifestyle determinate from the ecomorphotypes (7), which were adapted in Dulvy et al (8). In this sense, the reef species are usually found with complex environments, either coral or rocky. The shelf species are found in the intertidal zone and up to 200m depth, but without a highly 3D complex environment. The pelagic species are found mostly in open sea, and less than 200m depth. And finally, the deep sea are found below 200m depth. In total we have reef associated ($n = 28$), shelf ($n = 24$), pelagic (n = 17), and deep-sea (n = 21) (Supplementary Data Table 3).

Biomechanical and shape comparison

We selected from the aligned coordinates those which could extract information related to the in-lever and out-lever, in order to estimate the mechanical advantage of jaw opening. The ratio was estimated by measuring the interlandmark distance between Landmark 1 and Landmark 6 divided by the interlandmark distance between Landmark 1 and Landmark 2. The second comparison was in relation to the torsional resistance. To estimate this measurement we used the interlandmark distance between Landmark 5 and 6 divided by interlandmark distance between Landmark 1 and 5, this value was then log transformed to reduce the difference with large specimens. The mechanical advantage estimated was then plotted against PC1, while the torsional resistance was plotted against PC2. This was done in order to corroborate the shape variation on Procrustes for each PC. We then estimated the correlation between these ratios and their corresponding PC. To measure the interlandmark distance the interlmkdist function in geomorph (v. 4.0.4) (9), and the correlation with ggpubr (v. 0.4.0) (10).

Supplementary Figures

Supplementary Figure S1. Landmarks and sliding landmarks configuration on the lower jaw of *Carcharhinus acronotus*. Red dots represent true landmarks at the (L1) upper symphysis, (L2) posterior boundary of the dental groove, (L3) anterior quadratojugal joint, (L4) posterior quadratojugal joint, (L5) lower symphysis, and (L6) inner symphysis. The blue does denote curves along the defined landmarks. Green dots are surface landmarks over the lateral view of the lower jaw. No surface landmarks were used in the inner part of the jaw since some reconstructions were often incomplete in this area.

Supplementary Figure S2. Sampling curve from LaSEC on the 257 3D landmark coordinates. Grey lines indicate iterations fitted values. Solid line indicates the median fit value.

Supplementary Figure S3. Hierarchical cluster analysis of the dissimilarity of diet preferences among the sharks studied. Barplots on the right indicate the %IRI for each prey category reported for each species. Complete data with monospecific diets. Colour code indicates BIR = Bird; CEPH = Cephalopod; CHON = Chondrichthyan; CR = Crustacean; FISH = Bony fish; INV = Invertebrate; MAM = Mammal; MOLL = Hard shelled mollusc; PL = Plant; REP = Reptile and ZOO = Zooplancton.

Supplementary Figure S4. Hierarchical cluster analysis of the dissimilarity of diet preferences for the species without mono-specific diet to observe arrangements in the clustering for larger groups.

Supplementary Figure S5. Trace plots of two independent runs for the BayesTraits analysis, results from the Gelman-Rubin's convergence diagnostic and effective sample size in tables S1 and S2.

Supplementary Figure S6. Jaw shape variation along Principal Component 1 and 2. Coloured circles show the order as indicated in Figure 1 of main manuscript. Labels on each point indicate the species for each specimen.

Supplementary Figure S7. Jaw shape variation along Principal Component 2 and 3. Coloured circles indicate the order.

Supplementary Figure S8. Scree plots of the variance explained (95%) by the components for the whole set of specimens (n = 145) (A), and the phylogenetic aligned PCA (89 species) (B).

Supplementary Figure S9. Log(centroid size) against the multivariate regression of the Procrustes coordinates (regression score) of shape for each species in the four analysed groups. (A) Orders (Lamn = Lamniformes; Carc = Carcharhiniformes; Orec = Orectolobiformes; Squa = Squaliformes; Hexa = Hexanchiformes; Echi = Echinorhiniformes; Hete = Heterodontiformes; Pris = Pristiophoriformes; Squt = Squatiniformes), (B) trophic levels (TP = Top Predator; MP = Mesopredator; LP = Low-level Predator), (C) diet groups, and (D) habitat (CEPH = Cephalopods; GEN = Generalist; FISH = Bony fishes; BP = Big Predator; ZOO = Zooplankton; CR = Crustaceans; INV = Invertebrates; MOLL = Hard shelled molluscs).

Supplementary Figure S10. Evolutionary rate per landmark for the main orders of sharks analysed. Landmarks and semilandmarks configurations on the jaws of representative orders in lateral (top row), dorsal (middle row) and buccal (bottom raw) view. Colder colours indicate low evolutionary rate and warmer colour indicate faster evolutionary rate.

Supplementary Figure S11. Posterior distribution of parameter estimates using ratematrix. Upper left and lower right panels show the variances of each trait, Principal component 1 and2 respectively. Upper right panel shows pairwise covariation of traits. Lower left panel, samples of the posterior distribution within the 95% of credible interval of the distribution.

Supplementary Figure S12. Ancestral shape reconstruction with all the species, including *Callorhinchus milii*. Displayed in lateral (upper) and dorsal (lower) view.

Supplementary Figure S13. Scatter plot of the mechanical advantage (MA) against PC1 scores in figure 1 of the main text. Pearson correlation estimated and P value indicated.

Supplementary Figure S14. Scatter plot of the logarithm of Torsional Resistance (log(TR)) against PC2 scores in figure 1 of the main text. Pearson correlation estimated and P value indicated.

Supplementary Tables

Supplementary Table S1. Summary of mean values of the estimation of the transition rate matrix for discrete trait evolution (Diet, Order, Habitat, Trophic Level) over 100 trees to account for phylogenetic uncertainty. Selected models based on the Akaike Information Criterion (bold). Models tested: All Rates Different (ARD), Symmetric (SYM), and Equal Rates (ER).

Supplementary Table S2. Effective sample size of the runs that converged from BayesTraits.

Supplementary Table S3. Gelman and Rubin's convergence diagnostic results from the two runs that displayed convergence in the trace plots. Multivariate psfr 1

Supplementary Table S4. Effective sample size of the runs that converged from the ratematrix analysis. Regimes are the habitat (deep, pelagic, reef, shelf); diet (BP, CEPH, CR, FISH, GEN), and trophic level (LP, MP, TP).

Supplementary Table S5. Mean evolutionary rates obtained from the estimation on 100 random trees extracted from the posterior distribution. Each group value is indicated along with the median, standard deviation and standard error.

Supplementary Table S6. Procrustes variance estimated for each order (with 3 or more species) with observed disparity and bootstrapped median (100 bootstrap replicates). Lower triangle indicates the pairwise comparison with p value after Wilcoxon test with Bonferroni correction. Hexa = Hexanchiformes; Squa = Squaliformes; Squt = Squatiniformes; Orec = Orectolobiformes; Lamn = Lamniformes; Carc = Carcharhiniformes.

Supplementary Table S7. Procrustes variance estimated diet groups (with more than 5 species) with observed disparity and bootstrapped median (100 bootstrap replicates). Lower triangle indicates the pairwise comparison with p value after Wilcoxon test with Bonferroni correction. BP = Big predator; CEPH = Cephalopod consumer; CR = Crustacean consumer; FISH =Fish consumer; GEN = Generalist

Supplementary Table S8. Procrustes variance estimated for each trophic level group with observed disparity and bootstrapped median (100 bootstrap replicates). Lower triangle indicates the pairwise comparison with p value after Wilcoxon test with Bonferroni correction. TP = Top predator; MP = Mesopredator; LP = Low level predator.

Supplementary Table S9. Procrustes variance estimated for habitat groups with observed disparity and bootstrapped median (100 bootstrap replicates). Lower triangle indicates the pairwise comparison with p value after Wilcoxon test with Bonferroni correction.

Supplementary Table S10. Procrustes MANCOVA results for each category covarying with the logarithm of centroid size. Bold indicates significant at p < 0.05. Df = Degrees of freedom; SS = Sum of squares; MS = Mean squares; Rsq = r squared; F = F value; Z = Z-score; log(CS) = logarithm of centroid size; TR = Trophic Level.

Supplementary Table S11. Comparison of evolutionary models by penalized likelihood with generalized information criterion (GIG). Bold indicates the best supported model. BM = Brownian Motion; OU = Ornstein-Uhlenbeck; EB = Early Burst; lambda = Pagel's lambda transformation.

Supplementary Table S12. Summary statistics of the posterior distribution overlap among the rates of evolution and correlations between traits. The regimes are the four habitat groups. Bold indicates differences between posterior parameters estimates between regimes overlap.

Supplementary Table S13. Summary statistics of the posterior distribution overlap among the rates of evolution and correlations between traits. The regimes are the main five diet groups. BP = Big predator; CEPH = Cephalopod consumer; CR = Crustacean consumer; FISH =Fish consumer; GEN = Generalist

Supplementary Table S14. Summary statistics of the posterior distribution overlap among the rates of evolution and correlations between traits. The regimes are the three trophic level groups. TP = Top predator; MP = Mesopredator; LP = Low level predator.

Supplementary References

- 1. Watanabe, A. (2018). How many landmarks are enough to characterize shape and size variation?. *PloS one*. **13**, e0198341 (2018).
- 2. Schlager, S. Morpho and Rvcg–Shape Analysis in R: R-Packages for geometric morphometrics, shape analysis and surface manipulations. in *Statistical Shape and Deformation Analysis* (eds Zheng, G. et al.) 217–256 (Academic Press, 2017).
- 3. Cortés, E. Standardized diet compositions and trophic levels of sharks. *ICES J. Mar. Sci.* **56**, 707–717 (1999).
- 4. Galili, T. dendextend: an R package for visualizing, adjusting and comparing trees of hierarchical clustering. *Bioinformatics*. **31**, 3718-3720 (2015).
- 5. Froese, R. & Pauly, D. FishBase. World Wide Web electronic publication www.fishbase.org (2022).
- 6. Froese, R. *Life-History Strategies of Recent Fishes: A Meta-Analysis.* PhD thesis, Christian-Albrecht Universität (2005).
- 7. Compagno, L. J. Alternative life-history styles of cartilaginous fishes in time and space. Environmental Biology of Fishes, **28**, 33-75 (1990).
- 8. Dulvy, N. K. et al. Extinction risk and conservation of the world's sharks and rays. *eLife* **3**, e00590 (2014).
- 9. Adams, D. C. & Otárola‐Castillo, E. geomorph: an R package for the collection and analysis of geometric morphometric shape data. *Methods Ecol. Evol.* **4**, 393-399 (2013).
- 10. Kassambara, A. & Kassambara, M. A. Package 'ggpubr'. R package version 0.1 6 (R Foundation for Statistical Computing, 2020).