**Data from: Evolution of jaw disparity in fishes**

Jennifer J. Hill1, Mark N. Puttick1, Thomas L. Stubbs1, Emily J. Rayfield1\*, and Philip C.J. Donoghue1\*

1School of Earth Sciences, University of Bristol, Life Sciences Building, Tyndall Avenue, Bristol BS8 1TQ, United Kingdom

\*Corresponding authors: phil.donoghue@bristol.ac.uk and emily.rayfield@bristol.ac.uk

**Supplementary Information**

**Table of Contents**

1. **Study materials (taxon information including lateral lower jaw profiles for analysis and additional methodology)**
2. **Percentage of variance per axis: scree plots**
3. **Multivariate analyses**
4. **Disparity**
5. **Statistical tests**
6. **Supplementary references**

**1. Study materials**

*Taxon sample.* The dataset consists of 138 lower jaw specimens (75 fossil taxa and 63 extant taxa) of modern and Palaeozoic aquatic vertebrates. Analysed taxa include 8 acanthodians, 48 actinopterygians, 26 chondrichthyans, 32 placoderms, and 24 sarcopterygians. The same groups or clade divisions were used as justified by the supplementary material provided by Anderson et al (2011). Sampling is at order and family level and when certain orders have 100 or more families (e.g. Perciformes), a range of taxa were included, representative of both generalised morphology and outliers representative of the most extreme jaw anatomies. Only genera that were complete (well-preserved) or with reconstructed lower jaws were used. Any mandibles not found in the institutions listed on the accompanying Excel spreadsheet were obtained from literature (including published photographs, illustrations, figure drawings, and reconstructions) as well as photographs from Anderson et al (2011).

**Table S1:** An excel spreadsheet composed of taxa information for 138 specimens including institutions, museums, and sources where specimen photographs were obtained. \*\*Excel spreadsheet known as Set1\_SpecimenList.xlxs

*Equipment Used to Obtain Specimen Photographs.* Jaw outline data was generatedprimarily from photographs obtained by the lead author or from the literature. Images from Phil Anderson were taken with a Nikon D70 digital SLR with a Nikkor 18-55mm lens. All other images were photographed by Jennifer J. Hill using a Canon EOS 5D Mark II digital SLR with a EF24-70mm, f/2.8L USM lens, Canon Speedlite 580EX II Handheld Flash, and a woodblock-UV + Antiglare museum quality glass rig set up with black velvet fabric as a background. Original photographs included either a millimetre or centimetre unit scale.

*Orientation of Lower Jaw Shapes for Outline Analysis.* Lower jaws have been orientated with the posterior portion (back of the jaw) at the left and the anterior most tip (front portion of the jaw) at the right. As is seen in Figure S1 below:



Anterior

Posterior

**Figure S1:** Lateral profile of the lower jaw of an Angler Monkfish showing the orientation of the mandible used for outline analysis in Momocs [4].

*Panel: Lateral view of Lower Jaw Profiles of 138 specimens.* Each lateral lower jaw shape profile corresponds with the specimen as listed alphabetically in **Table S1** (as read in rows from left to right).



**Figure S2:** Each mandible is colour coded by group/clade: Placoderms (**Dark Orange**), Acanthodians (**Dark Green**), Actinopterygians (**Deep Sky Blue**), Chondrichthyans (**Dark Blue**), and Sarcopterygians (**Chartreuse**).

*Additional Methodology: mandibular outline generation.*

During the digitization process, each lower jaw went through a segmentation process converting the original captured colourscale image into a greyscale image, which was then converted into a binary image based on threshold level. Binary images of each mandible were converted into *coo* class objects in Momocs (Bonhomme*, et al.* 2014) for outline analysis and extended eigenshape analysis was used to quantify variation in lower jaw morphology. After a Procrustes transformation was performed, outlines were subjected to *a priori* normalization (where, norm=FALSE), which means that during elliptical Fourier transformation, the first harmonic is not dropped, but rather included (averaged) in the approximation of form of the total lower jaw shape. Crampton (1995) proposed that shape and variance in a specimen is usually captured sufficiently with eight harmonics, although the total number of harmonics used to describe shape variation depends on the complexity of the curvature of the object.

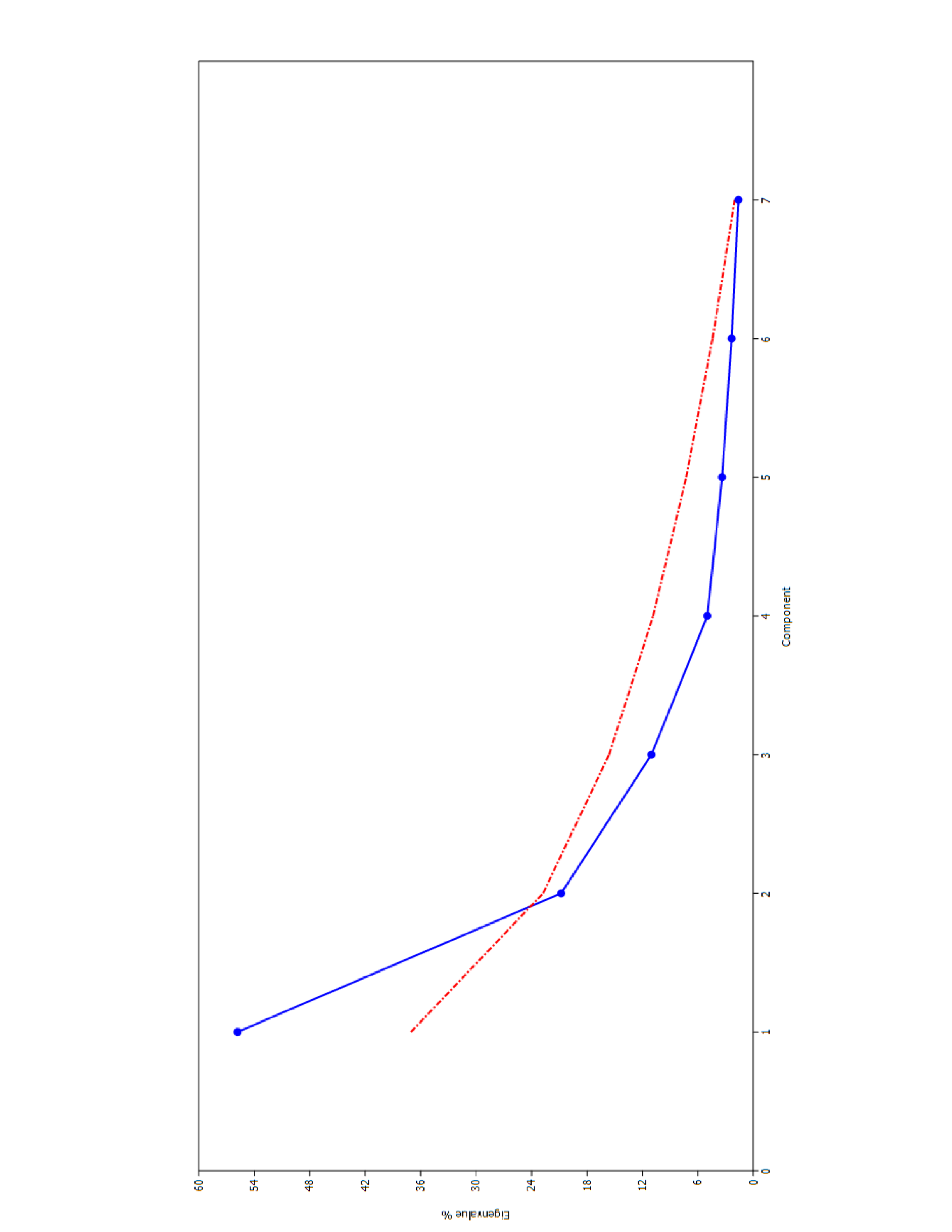
**2. Proportion of Variance per Component**

*Percentage of variance on each axis.* 98.4% of variance is accounted for on 12 component axes. The first 3 components are illustrated in Figures 1a, b, and 2a-2d of the main text; components 4-12 are shown in **Figure S3**.



**Figure S3:** Approximately, 95% of total lower jaw shape variation is accounted for within the first seven PC-axes, and 83% of the variation is summarised on the first three axes. However, there are a total of 40-PC axes that account for total shape variation.

*Scree Plot of the first 7 principal components.* Scree plots are used in Principal Components Analysis (PCA) to visually assess which components explain the most variance in the data.



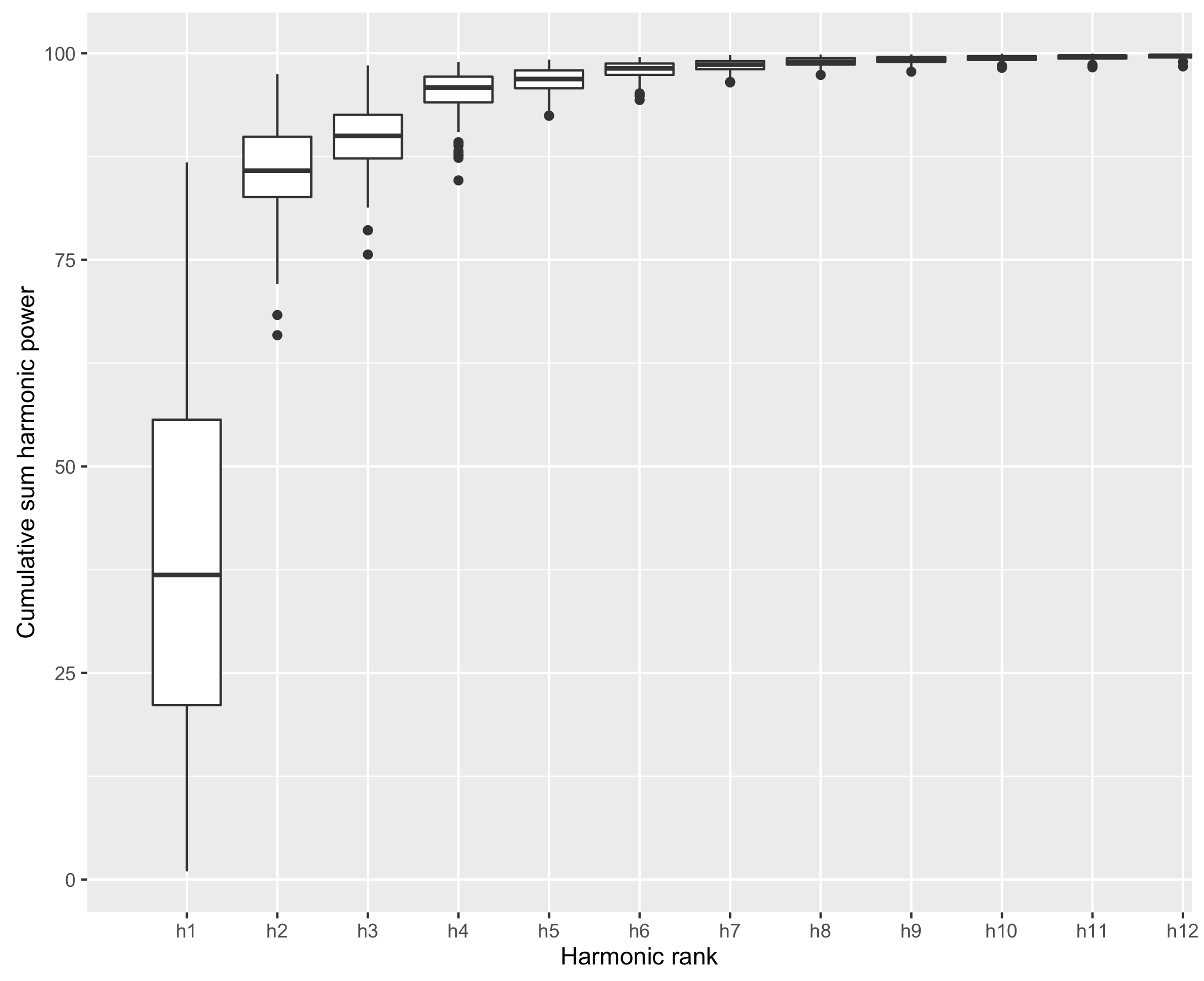
**3. Mean and Standard Deviation of jaw shapes for 7 PC axes**

**Figure S4:** Associated eigenvalues are plotted against the axis order, which can be used as a guide to identify important principal components. The intersection of the broken stick model (dashed red line) with the scree plot (solid blue line) suggests that at least 2 axes should be retained for analysis.

**3. Multivariate analyses**

*Principal Component Analysis.* Forty component axes express total shape variation; 95% of shape variation is accounted for with 7-PC axes alone. Reference **Table S1**, which lists all the specimens and their associated values for each listed component.

*Outline Analysis, Calibration of Harmonic Coefficients, and Elliptical Fourier Transformation Analysis (a priori normalization).* We used outline and Elliptical Fourier Transformation analysis of modern and Palaeozoic aquatic jawed vertebrates to characterize lower jaw shape disparity. Lateral thresholded images of 138 mandibles (as seen in **Figure S2**) were imported into R (Team 2016), where we used Momocs (Bonhomme*, et al.* 2014) to convert binary images into *coo*-class objects. Herein an outline is a list of 200 coordinates, which is classified as a ‘*coo*’ object in R (Team 2016). ‘*Out*’ is used for closed outlines and applied to ‘*coo*’ objects. Refer to Bonhomme *et al.* for more specific information about applied methods and class objects within Momocs (Bonhomme*, et al.* 2014).

A Procrustes transformation was performed where all lower jaw shapes (or outlines) were centred, scaled, and aligned via the x-axis. An origin point (coordinate 0) was redefined for each outline and 200 coordinates or semi-sliding landmarks were assigned per mandible. These coordinates were slid in an easterly direction to best fit the curvature of the jaw. Artefacts may be introduced during the digitization and translation process from a 2D geometric shape towards a vector-based outline; these artefacts can make outlines appear visually jagged or pixellated. Therefore, harmonic coefficients were used to smooth outlines and eliminate any introduced high-frequency noise. Outlines were further analysed using Elliptical Fourier transformation analysis with a priori normalization (where, norm=FALSE); the outlines are decomposed into a series of closed curves known as harmonics, which can be ranked according to the cumulative sum of harmonic power (as seen in Figure S5). 

**Figure S5:** Harmonic power is proportional to harmonic amplitude and is to be considered as a measure of shape information. The power of a given harmonic *n* is calculated as follows for Elliptical Fourier analysis and *n*th harmonic:

(Bonhomme *et al.* 2014).

The total number of harmonics used to describe shape variation varies depending on the curvature of the lower jaw with the caveat that using too many harmonics exposes outlines to high-frequency noise and using too few harmonics risks loss of morphological detail. The number of harmonics is evaluated and selected so that their cumulative power amasses to 99% of the total shape variation (or a total of 10 harmonics as stated in the main text); however, a total of twenty harmonics could have been used to describe an additional 0.9% of shape variation. After the appropriate number of harmonics were selected, Elliptical Fourier analysis is performed on the ‘*coo*’ object using the e-Fourier method. A matrix composed of 138 outlines (one outline per specimen) x [4 coefficients x 10 harmonics (the number selected whose cumulative power accounts for 99% of total shape variation)] is created with group (clade or group) factors, specimen names, and other relevant information as a ‘*coe*’ class object. A PCA was calculated on the ‘*coe*’ class object and the resulting scores were used to plot clade/group-defined empirical morphospaces (Figures 1, 4 of the main text).

**4. Disparity**

*Sum of Variances and Partial Disparity.* These particular metrics were used to assess the statistical significance of shape variation between living and fossil specimens and between taxonomic groups in morphospace. The total variance of the dataset is defined by the sum of variances (SOV) of the individual components from a total of 40-PC axes. Partial disparity was calculated using a subset of 7-PC axes (accounting for 95% of the total variation) to examine the relative contribution of both extant and fossil taxa for all taxa shown in Figure 1 the initial radiation of gnathostomes (Sil-Dev) morphospace in Figure 4A, B and the Recent/Extant morphospace in Figure 4C, D of the main text. Calculations were performed using the MDA (Navarro 2003) software package in Matlab.

**Table S2:** Partial disparity of extant and fossil gnathostomes within each group for all taxa (138 specimens) found in Figure 1a, b. Actinopterygians account for a total of 60% shape variation, sarcopterygians (5%), acanthodians (5%), chondrichthyans (18%), and placoderms (12%).

|  |  |  |  |
| --- | --- | --- | --- |
| **Group/Clade** | **Fossil** | **Extant/Living** | **Total** |
| Acanthodians | 5% | NA | 5% |
| Actinopterygians | 2% | 58% | 60% |
| Chondrichthyans | 5% | 13% | 18% |
| Placoderms | 12% | NA | 12% |
| Sarcopterygians | 4% | 1% | 5% |

|  |  |
| --- | --- |
| **Group/Clade** | **Fossil** |
| Acanthodians | 18% |
| Actinopterygians | 9% |
| Chondrichthyans | 18% |
| Placoderms | 36% |
| Sarcopterygians | 19% |

**Table S3:** Partial disparity of each group within the Sil-Dev dataset, which consists of 75 fossil specimens, found in Figure 2a, b. Fossil actinopterygians account for 9% of shape variation, fossil non-tetrapodomorph sarcopterygian fish (19%), acanthodians (18%), fossil chondrichthyans (18%), and placoderms (36%).

|  |  |
| --- | --- |
| **Group/Clade** | **Extant/Living** |
| Actinopterygians | 79% |
| Chondrichthyans | 20% |
| Sarcopterygians | 1% |

**Table S4:** Partial disparity of each group within the Extant/Living dataset, which consists of 63 lower jaw specimens found in Figure 2c, d. Actinopterygians account for 79% of the total shape variation, sarcopterygians (1%), and chondrichthyans (20%).

**6. Statistical Tests**

*Statistical methods.* Significance differences in shape variation between living and fossil specimens between taxonomic groups of morphospace were tested by using non-parametric multivariate analyses of variance (NPMANOVA) for All Taxa, in addition to both the Sil-Dev and Extant datasets using the *Adonis* function in the R package Vegan (Team 2016; Oksanen *et al.* 2013). *Adonis* is a function used for the analysis and partitioning sum of squares using semi-metric and metric distance matrices. This function allows one to partition sums of squared deviations from the centroid.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **All Taxa [by clade/ group]** | **F. Model** | **R2** | **P-value** | **P-adjusted** |
| Actinopterygians vs Placoderms | 8.202333305 | 0.095152103 | 0.001 | 0.0025 |
| Actinopterygians vs Acanthodians | 1.311900145 | 0.023718226 | 0.23 | 0.23 |
| Actinopterygians vs Chondrichthyans | 6.864801035 | 0.087045183 | 0.001 | 0.0025 |
| Actinopterygians vs Sarcopterygian Fish | 1.270493919 | 0.017826366 | 0.23 | 0.23 |
| Placoderms vs Acanthodians | 4.454266825 | 0.104919179 | 0.007 | 0.01 |
| Placoderms vs Chondrichthyans | 4.297605111 | 0.071273231 | 0.016 | 0.02 |
| Placoderms vs Sarcopterygian Fish | 8.300353086 | 0.133231237 | 0.001 | 0.0025 |
| Acanthodians vs Chondrichthyans | 5.332923833 | 0.142847741 | 0.001 | 0.0025 |
| Acanthodians vs Sarcopterygian Fish | 3.804734202 | 0.112550336 | 0.005 | 0.008333333 |
| Chondrichthyans vs Sarcopterygian Fish | 6.39368582 | 0.117544633 | 0.002 | 0.004 |

**Table S5:** MANOVA\_PW results for All Taxa by clade/group. This table shows dissimilarities between both fossil and living groups using Euclidean distances.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Sil-Dev [by clade/ group]** | **F. Model** | **R2** | **P-value** | **P-adjusted** |
| Placoderms vs Acanthodians | 4.45394504 | 0.104912395 | 0.01 | 0.033333333 |
| Placoderms vs Sarcopterygian Fish | 7.488767893 | 0.125885409 | 0.001 | 0.01 |
| Placoderms vs Chondrichthyans | 1.249116048 | 0.034459214 | 0.261 | 0.421428571 |
| Placoderms vs Actinopterygians | 4.928154595 | 0.112186698 | 0.014 | 0.035 |
| Acanthodians vs Sarcopterygian Fish | 3.582138183 | 0.113422915 | 0.009 | 0.033333333 |
| Acanthodians vs Chondrichthyans | 1.105889129 | 0.091351335 | 0.349 | 0.43625 |
| Acanthodians vs Actinopterygians | 1.488028657 | 0.090249034 | 0.189 | 0.378 |
| Sarcopterygian Fish vs Chondrichthyans | 1.156464391 | 0.04421333 | 0.295 | 0.421428571 |
| Sarcopterygian Fish vs Actinopterygians | 0.949070717 | 0.031689488 | 0.393 | 0.436666667 |
| Chondrichthyans vs Actinopterygians | 0.632740802 | 0.050087373 | 0.621 | 0.621 |

**Table S6:** MANOVA\_PW results for Sil-Dev dataset by clade/group. This table shows dissimilarities between fossil groups using Euclidean distances.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Extant/Living [by clade/ group]** | **F. Model** | **R2** | **P-value** | **P-adjusted** |
| Actinopterygians vs Chondrichthyans | 6.678758009 | 0.103260455 | 0.001 | 0.003 |
| Actinopterygians vs Sarcopterygian Fish | 0.369089443 | 0.009375107 | 0.644 | 0.644 |
| Chondrichthyans vs Sarcopterygian Fish | 2.092308794 | 0.090606306 | 0.117 | 0.1755 |

**Table S7:** MANOVA\_PW results for Extant/Living dataset by clade/group. This table shows dissimilarities between extant/living groups using Euclidean distances.

**7. Supplementary references**

ANDERSON, P. S. L., FRIEDMAN, M., BRAZEAU, M. D. and RAYFIELD, E. J. 2011. Initial radiation of jaws demonstrated stability despite faunal and environmental change. *Nature*, **476**, 206-209.

BONHOMME, V., PICQ, S., GAUCHEREL, C. and CLAUDE, J. 2014. Momocs: Outline Analysis Using R. *Journal of Statistical Software*, **56**.

CRAMPTON, J. S. 1995. Elliptic Fourier shape analysis of fossil bivalves: some practical considerations. *Lethaia*, **28**, 179-186.

NAVARRO, N. 2003. MDA: a MATLAB-based program for morphospace-disparity analysis. *Computing in Geosciences*, **29**, 655–664.

OKSANEN, J., BLANCHET, F. G., KINDT, R., LEGENDRE, P., MINCHIN, P. R., O’HARA, R. B., SIMPSON, G. L., SOLYMOS, P., STEVENS, M. H. H. and WAGNER, H. 2013. vegan: Community Ecology Package. R package version 2.0-7. *Available at:* [*http://CRAN.R-project.org/package=vegan*](http://CRAN.R-project.org/package=vegan).

TEAM, R. C. 2016. R: A language and environment for statistical computing.