

Divergent evolutionary morphology of the axial skeleton as a potential key innovation in modern cetaceans

Amandine Gillet, Bruno Frédérick and Eric Parmentier

Proceedings of the Royal Society B: Biological Sciences

DOI: 10.1098/rspb.2019.1771

Contents:

Appendix I: *Supplementary methods for L/H ratios*

Appendix II: *Analyses without phylogenetic correction*

Supplementary Figures:

- S1:** *Schematic representation of vertebral shape data collection and standardisation*
- S2:** *Vertebral counts for cetaceans compared to other mammalian lineages*
- S3:** *Relationship between vertebral count, family and habitat*
- S4:** *Biplot of the principal components analysis*
- S5:** *Effect of the habitat on vertebral shape (Delphinidae & Phocoenidae vs. other families)*
- S6:** *L/H ratios of extant and extinct cetaceans*
- S7:** *Evolutionary patterns of the vertebral shape (PC2 to PC8)*
- S8:** *Effect of the habitat on vertebral shape (Delphinoidea vs. non-Delphinoidea)*
- S9:** *Relationship between diversification rate and phenotypic traits*
- S10:** *Effect of the habitat on vertebral shape without accounting for the effect of the phylogeny (All cetaceans)*
- S11:** *Effect of the habitat on vertebral shape without accounting for the effect of phylogeny (Delphinidae & Phocoenidae vs. other families)*
- S12:** *Effect of the habitat on vertebral shape without accounting for the effect of phylogeny (Delphinoidea vs. non-Delphinoidea)*
- S13:** *Evolutionary patterns of the vertebral shape (PC1) without phylogenetic correction*
- S14:** *Evolutionary patterns of the vertebral shape (PC2 to PC8) without phylogenetic correction*

Supplementary Tables:

- S1:** *List of specimens used in this study*
- S2:** *Priors used for the Bayesian multi-regime Ornstein-Uhlenbeck method (Bayou)*
- S3:** *Comparison of intra- and interspecific morphological disparity*
- S4:** *Summary of correlation tests between body size and vertebral count*
- S5:** *Summary of the effect of habitat on vertebral count and shape*
- S6:** *Correlations between diversification rate and morphological traits*
- S7:** *Summary of correlation tests between body size and vertebral count without phylogenetic correction*
- S8:** *Summary of the effect of habitat on vertebral count and shape without phylogenetic correction*
- S9:** *Correlations between diversification rate and morphological traits without phylogenetic correction*

References

Appendix I: Supplementary methods for L/H ratios

Measurements of centrum length and height of 13 fossil specimens were obtained from the literature. Only species that have previously been included in phylogenetic analyses were selected. Data were obtained for four stem cetaceans: *Basilosaurus cetoides* [1], *Dorudon atrox* [2], *Zygorhiza kochii* [1] and *Cyhiacetus peruvianus* [3]. Two extinct mysticetes were also included: *Aetiocetus cotylalveus* [4] and *Thinocetus arthritus* [5]. Data were also obtained for seven extinct odontocetes: *Zarhachis flagellator* [6], *Ninoziphius platyrostris* [7], *Kentriodon pernix* [8], *Pliopontos littoralis* [7], *Atocetus iquensis* [9], *Albireo whistleri* [10] and *Piscolithax longirostris* [7]. Measurements were taken from data tables available in the original description of the specimens except for *Albireo whistleri*. For this later species, no raw measurements were available but ratios were obtained from pictures of the backbone available in the original publication. Phylogenetic relationships were synthesized from various recently published phylogenies [11–17].

For extant species, L/H ratios were calculated for each vertebra of the thoracic, lumbar and caudal regions. The average ratio for the entire backbone was then calculated and used for the analysis. Densities for the violin plot were computed with the R-package *ggplot2* [18]. The number of vertebrae from which measurements were obtained for extinct specimens varied depending on the preservation state of each specimen but it generally consisted of several vertebrae from at least two different regions of the backbone (apart from the cervical region). Some specimens used were reconstructed from a composite of several specimens. Similar to extant species, the mean L/H ratio of each extinct species was then calculated and compared to extant species.

Appendix II: Comparative analyses without phylogenetic correction

1. Material and methods

In order to investigate the raw morphological variation across cetaceans, we repeated all the comparative analyses described in the main text without accounting for phylogenetic information. An identical structure was conserved but all phylogenetically informed analyses were replaced by regular statistics.

(a) Vertebral count and body size.

The effect of habitat on the vertebral count was tested using a regular ANOVA with the function *aov* from the R-package *stats* (v.3.5.1) [19]. We then tested whether Delphinidae and Phocoenidae (i.e. oceanic dolphins and porpoises) differ from the other species in their vertebral count and body size by applying a MANOVA with the *manova* function in R. This analysis was repeated to test such a habitat effect within: (i) Delphinidae and Phocoenidae; (ii) all species except Delphinidae and Phocoenidae; (iii) Delphinoidea; and (iv) non-Delphinoidea. The linear relationship between vertebral count and body length was tested for each group by using a generalized least squares regression (GLS analysis) with the *gls* function from the *nlme* package (v. 3.1-131) [20].

(b) Morphospace of vertebrae.

All linear IMRMs were log₁₀-transformed and were then size-corrected with a generalised least square regression using the function *gls* (figure S1e). The log₁₀ TCL was used as a proxy for body size for size-correction. Angular IMRMs (i.e. inclination of neural processes and inclination of transverse processes) were not correlated to body size and were thus transformed using a cosine function. Species mean regional measurements (SMRMs) for linear and angular values were then computed by calculating the mean value of residuals of each IMRM of all individuals belonging to the same species (figure S1e). In order to explore morphological variation of vertebrae, we performed a principal component analysis (PCA) on SMRMs of all cetacean species using the *prcomp* function in R. PCA was produced on the correlation matrix because the scales of the different morphological traits vary among them. Four distinct PCAs were also run separately for each group: (i) Delphinidae and Phocoenidae; (ii) all species except Delphinidae and Phocoenidae; (iii) Delphinoidea; and (iv) non-Delphinoidea. According to the Jolliffe cut-off, only principal components (PCs) with an eigenvalue equal to or higher than 0.7 were conserved for further analysis. This corresponds to the first eight PCs for each PCA except the 'Delphinoidea' PCA for which the first nine PCs were conserved. Differences in vertebral morphology between the group of dolphins and porpoises and other species were tested by a MANOVA. A second MANOVA was used to test for a difference between Delphinoidea and non-Delphinoidea. We used a multivariate linear regression to test the correlation between vertebral count and vertebral shape (Jolliffe cut-off PCs) for each PCA, using the *procD.lm* in the R-package *geomorph* (v. 3.0.6) [21]. For all cetaceans and for every group, we tested the effect of habitat on the variation of vertebral morphology using

MANOVAs.

(c) Evolutionary shifts of phenotypic traits and relationships with the rate of diversification.

BAMM and Bayou analyses, allowing the detection of evolutionary shifts, were performed on PCs 1 to 8 from the 'all cetaceans' PCA that was calculated on non-phylogenetically corrected residuals. Parameters for these analyses were the same as those used for phylogenetically corrected data.

Similarly, the ES-Sim analyses, used to detect the correlation between speciation rate and phenotypic traits evolutionary rates, were performed on non-phylogenetically corrected PCs 1 to 8 with 1,000 iterations. When a significant relationship was found between speciation and a trait, a regular linear model based on generalized least squares was applied to determine the slope of the regression using the *gls* function from the R-package *nlme*.

2. Results

(a) Vertebral count and shape in relation to body size and ecology.

Generally speaking, results obtained from non-phylogenetically corrected analyses regarding the relationship between vertebral morphology and body size and ecology were very similar to those obtained with phylogenetic comparative methods. The sole difference was that the correlation between the vertebral count and body size for all cetaceans was significant without the phylogenetic correction (GLS: $p = 0.01$, $R^2 = 0.09$) while it was not when accounting for the effect of phylogeny (pGLS: $p = 0.7$). However, the coefficient of determination was very low, reflecting that the linear regression did not fit well these data. Delphinidae and Phocoenidae differ in body size, vertebral count and vertebral shape from the other families (MANOVAs: $p \leq 0.0001$). Similarly, Delphinoidea differ in vertebral shape from non-Delphinoidea (MANOVA on PCs1-8: $p \leq 0.0001$). Projections of the first two PCs for each PCA are in figures S10 to S12 and statistical results are listed in tables S7 and S8.

(b) Evolutionary patterns of phenotypic traits.

When using uncorrected morphological data for phylogenetic information, BAMM found strong support for a shift occurring for the evolutionary rate of PC1 with 72% of simulated trees having one shift. The principal shift sampled in the posteriors was on the branch leading to beaked whales (*Ziphiidae*) (marginal shift probability: 0.50). However, a shift on one of the branches leading to Delphinoidea was still sampled but with a lower marginal shift probability (0.26) (figure S13). Both these shifts are mutually exclusive meaning that they were never sampled together on the same simulated tree.

Bayou also found support for at least one evolutionary shift for PC1 although the effect was weaker than for phylogenetically-corrected data. Three branches with a posterior probability of having a shift greater than 0.12 (*i.e.* 15 times greater than the prior probability of 0.008) were

sampled. The shift on the branch leading to all Delphinidae except the killer whale (*Orcinus orca*) was still sampled (posterior probability = 0.19) but the shift leading to porpoises (*Phocoenidae*) was not detected anymore. Only a shift on the terminal branch of the Dall's porpoise (*Phocoenoides dalli*) was sampled (posterior probability = 0.13). In addition, a shift on the branch supporting river dolphins (*Pontoporia*, *Inia* and *Lipotes* genera) was also detected (posterior probability = 0.16) (figure S12). Results of both BAMM and Bayou for PC2 to 8 obtained from the 'all cetaceans' PCA on non-phylogenetically corrected residuals are shown in figure S14.

The results of the ES-Sim test were very similar to those obtained on phylogenetically corrected data. PC1 scores were significantly correlated to the diversification rate ($p = 0.03$, $R^2 = 0.41$, slope $\pm s.e. = 0.13 \pm 0.02$) while scores on PCs 2 to 8 were not (see table S9).

3. Discussion

Running analyses without accounting for the non-independence of species did not change the main results about the effect of body size and ecology on vertebral count and shape. Analyses on evolutionary patterns (BAMM, Bayou and ES-Sim) still found an evolutionary shift on one of the branches leading to Delphinoidea for PC1. However, the signal was weaker than for analyses on phylogenetically size-corrected data. This is undoubtedly due to the overlap of Delphinoidea and non-Delphinoidea on PC1 that weakens the observed morphological difference between the two groups. Moreover, both methods highlighted other shifts that were less frequently sampled with phylogenetically corrected data which probably also tend to reduce the signal for a shift of the branch leading to Delphinoidea. Nonetheless, the multivariate analysis of variance (MANOVA) ran on PCs 1 to 8 still returned a significant difference between these two groups. This demonstrates that the morphological difference between Delphinoidea and non-Delphinoidea is still present but might be less pronounced on PC1 when data are not phylogenetically corrected.

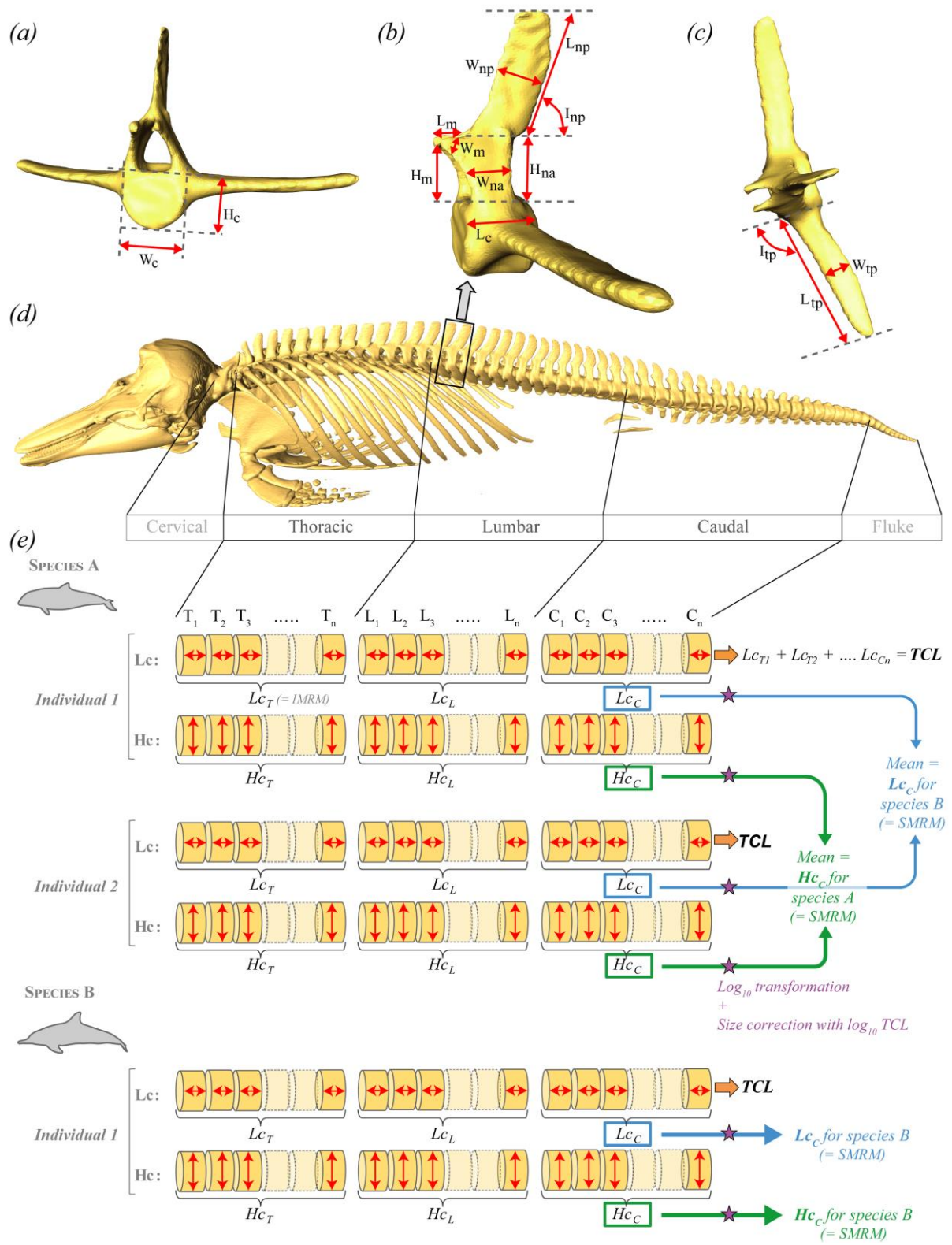


Figure S1. Schematic representation of vertebral shape data collection and standardisation. Twelve linear and two angular measurements (red double arrows) taken on each vertebra are shown on a vertebra in (a) frontal view, (b) lateral view and (c) dorsal view. W_c : centrum width, H_c : centrum height, L_c : centrum length, L_{np} : neural spine height, W_{np} : neural spine width, I_{np} : neural spine inclination, H_{na} : neural arch height, W_{na} : neural arch width, L_m : metapophysis length, W_m : metapophysis width, H_m : metapophysis height, L_{tp} : transverse process length, W_{tp} : transverse process height, I_{tp} : transverse process inclination. (d) 3D model of a harbour porpoise (*Phocoena phocoena*) skeleton based on CT-scan. Measurements were taken only on vertebrae of three regions: thoracic, lumbar and caudal. (e) For each individual, the mean value of each measurement in each region was calculated (= IMRMs). The total centrum length (TCL) was calculated by summing centra lengths of the three regions. All IMRMs were then \log_{10} -transformed and size-corrected using \log_{10} TCL as a proxy for body size. Species values of each measurement of each region (SMRM) were then calculated as the mean of IMRM of all individuals belonging to the species.

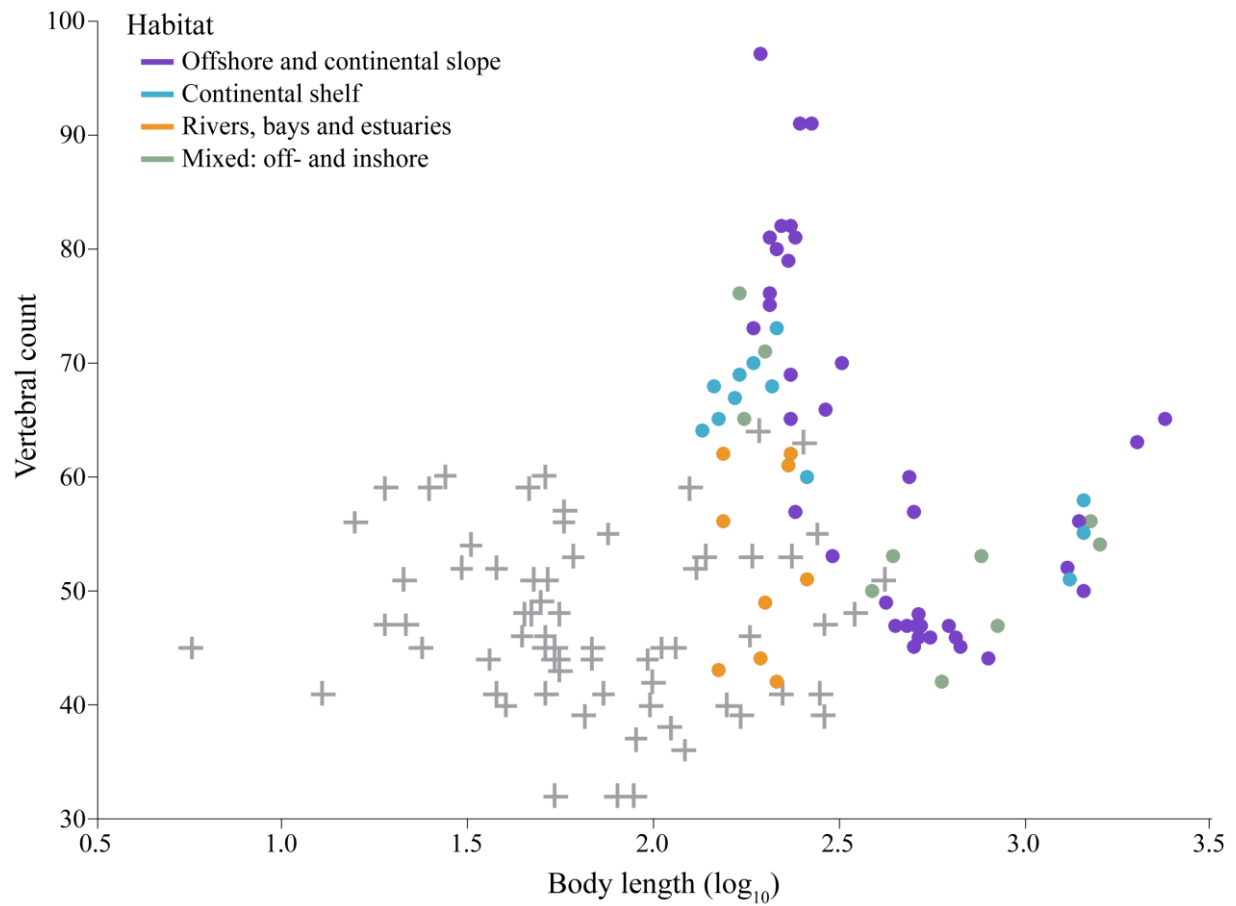


Figure S2. Vertebral counts for cetaceans compared to other mammalian lineages. Vertebral count according to body length in meters (\log_{10} -transformed). Filled circles represent cetaceans; 'plus' signs represent other mammals. For cetaceans, symbol colours correspond to different habitats: orange for rivers, bays and estuaries; light blue for continental shelf; purple for continental slope and offshore; and green for mixed off- and in- shore. For non-cetacean species, vertebral count data are from Narita and Kuratani [22] and body size data are from the panTHERIA database [23].

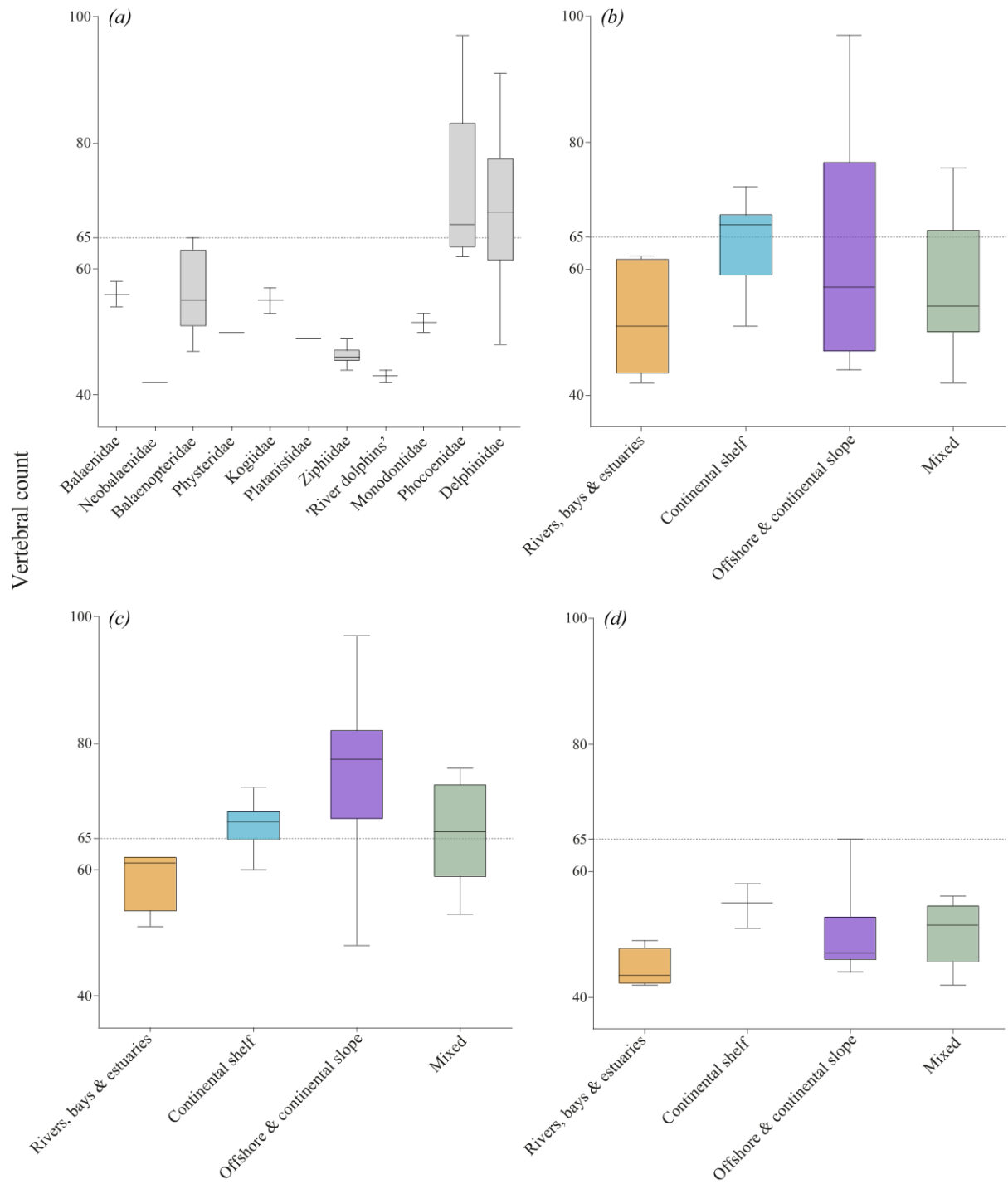


Figure S3. Relationship between vertebral count, family and habitat. Vertebral counts for (a) all cetaceans according to phylogenetic group, (b) all cetaceans according to habitat, (c) Delphinidae and Phocoenidae according to habitat and (d) other cetaceans according to habitat. For each data set, horizontal line represents the median value, lower and upper limits of boxes represent the 25% and 75% quartiles respectively and lower and upper whiskers represent minimum and maximum values respectively.

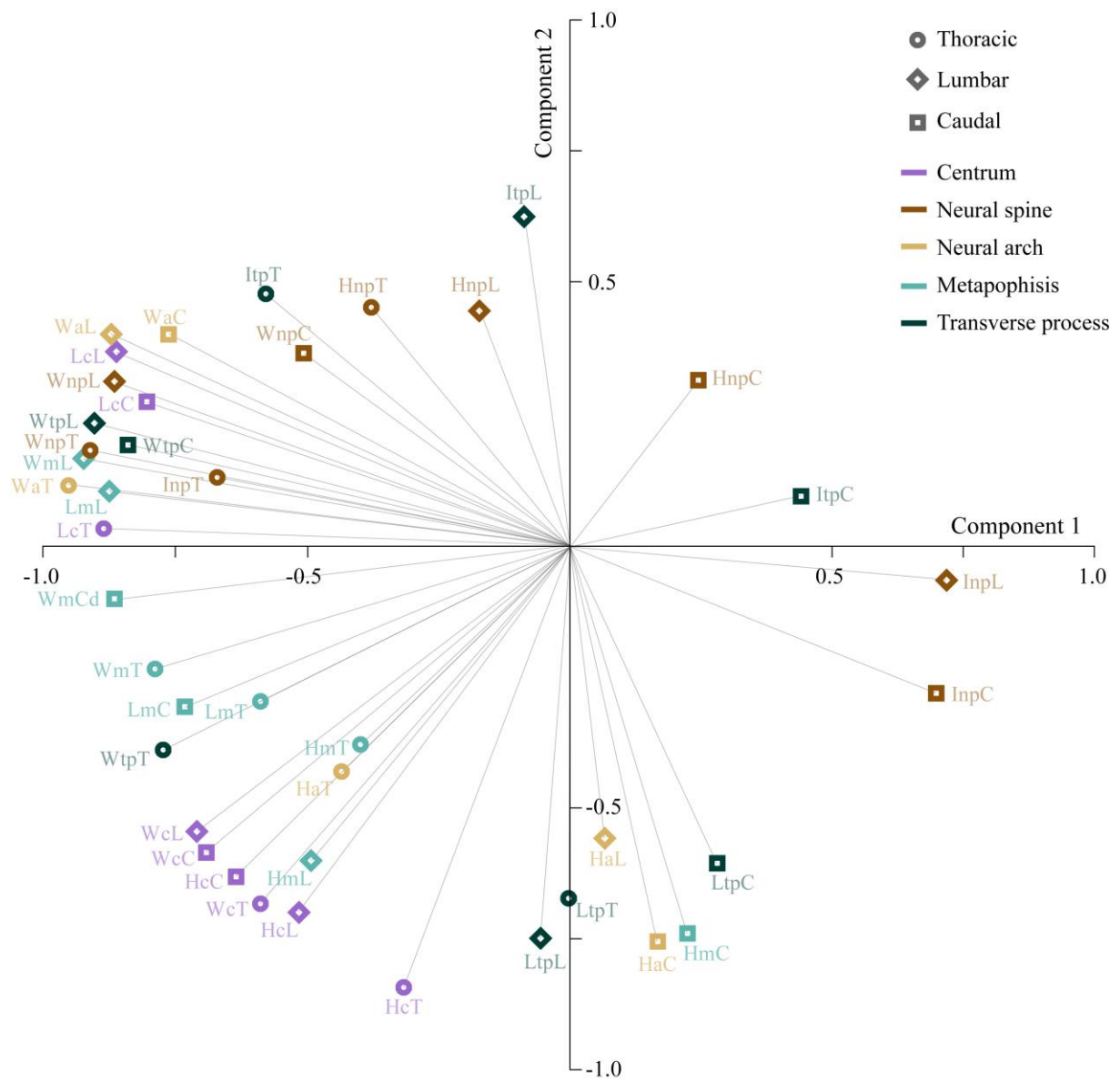


Figure S4. Biplot of the 'all cetaceans' principal components analysis. Loadings of variables on PC1 and PC2 for the PCA calculated on all cetacean species (figure 2). Symbol shapes correspond to the regions of the vertebral column. Each colour corresponds to a different part of the vertebra. Variables abbreviations are explained in figure S1.

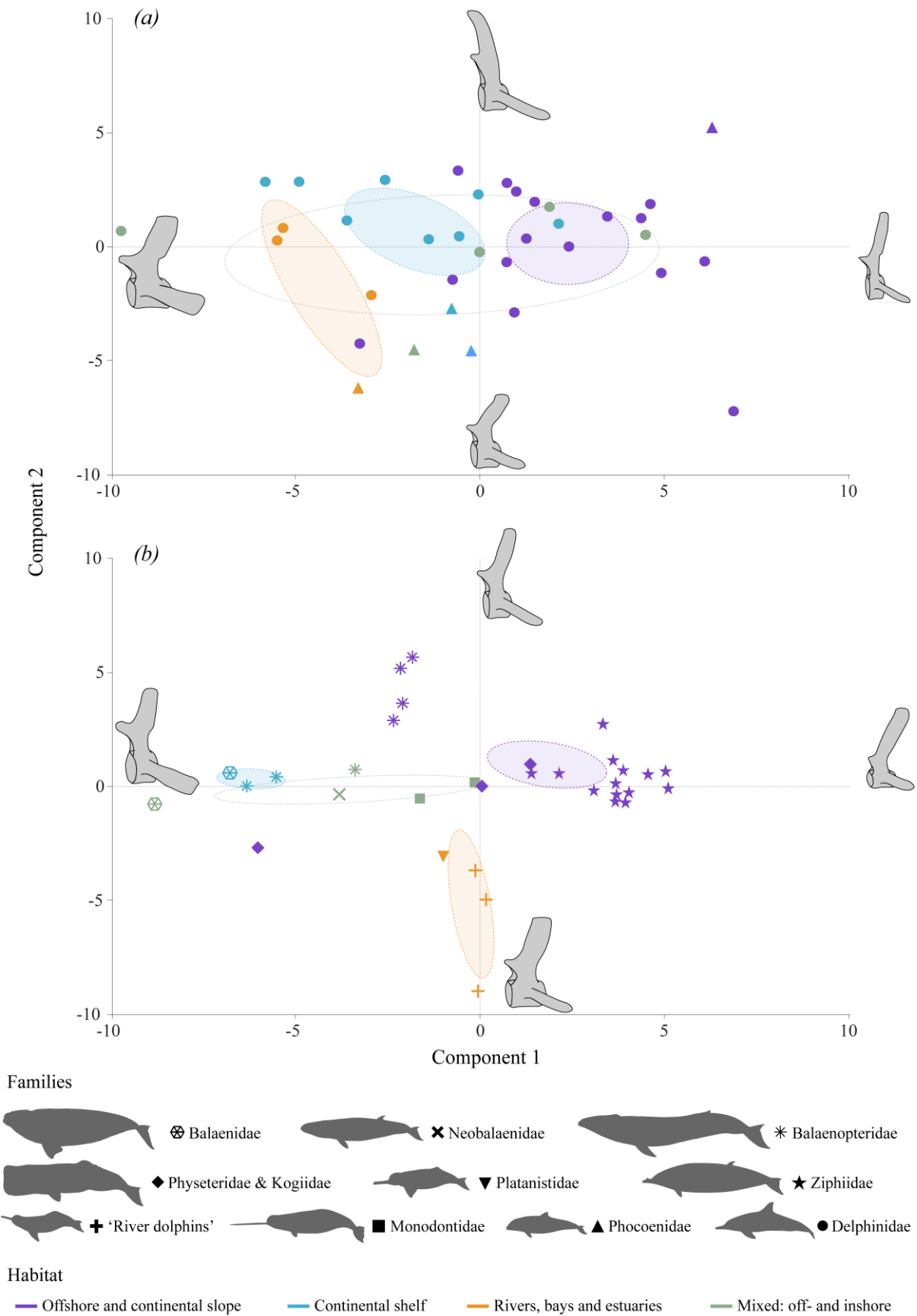


Figure S5. Effect of the habitat on vertebral shape. Principal components analysis plot of PC1 vs. PC2. Symbol shapes correspond to phylogenetic groups. Symbol colours correspond to different habitats. Dashed ellipses represent the 95% confidence intervals of the mean coordinates for each habitat category. (a) Delphinidae and Phocoenidae. PC1: 35% of variance, PC2: 19%. (b) All families excluding Delphinidae and Phocoenidae. PC1: 36%, PC2: 17%.

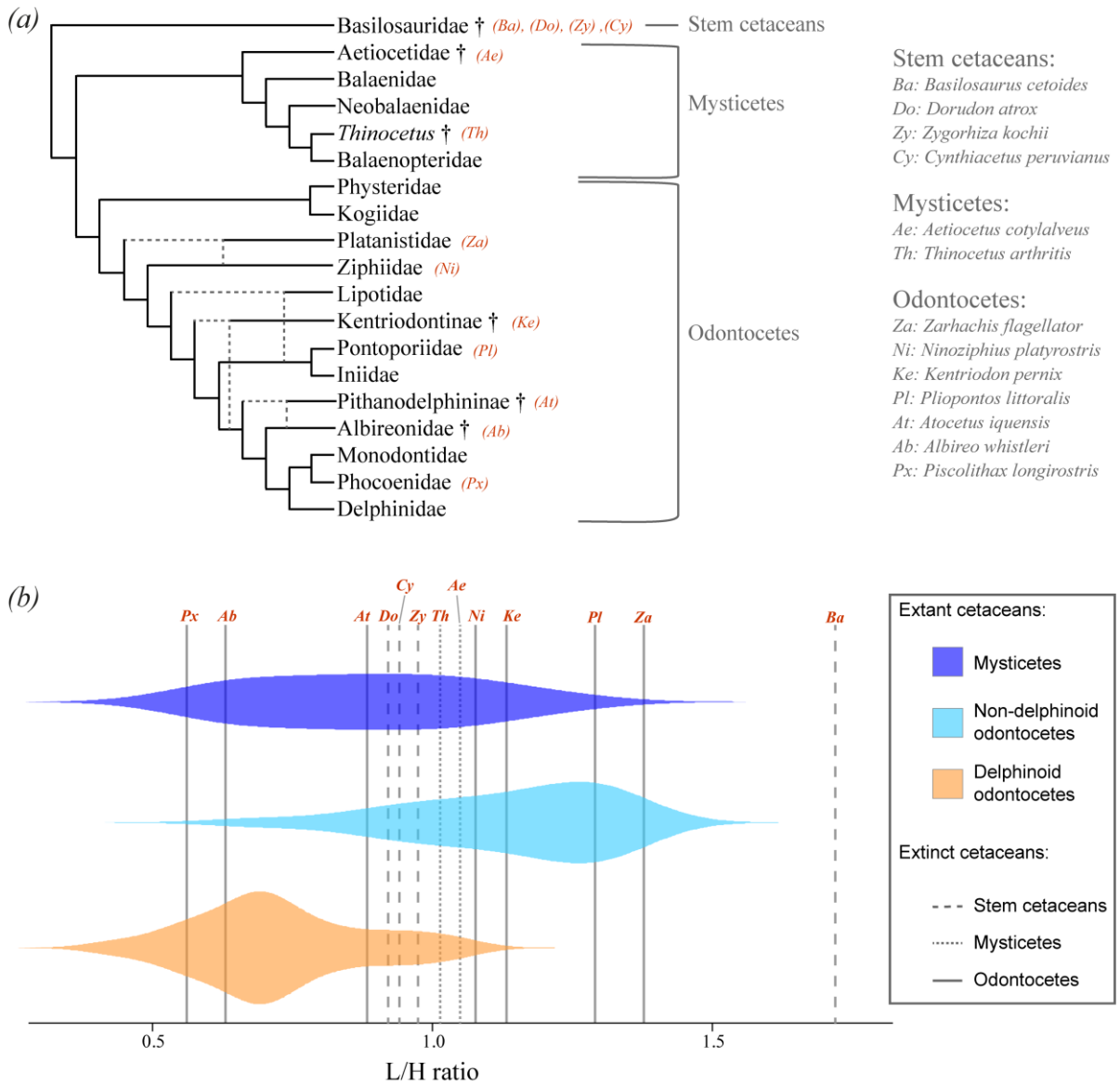


Figure S6. L/H ratios of extant and extinct cetaceans. (a) Simplified phylogenetic tree of cetaceans including some extinct families and genera (indicated by the dagger symbol) adapted from Marx *et al.* [24]. Dotted lines represent the phylogenetic uncertainty of some lineages. Fossil specimens included in the analysis are annotated in red beside their respective lineage. (b) Violin plot of L/H ratios of extant species. L/H ratios of extinct species are represented by vertical lines. The name of the species is annotated in red above its corresponding line. Lower ratios correspond to more discoidal vertebrae and higher ratios correspond to more spool-shaped vertebrae.

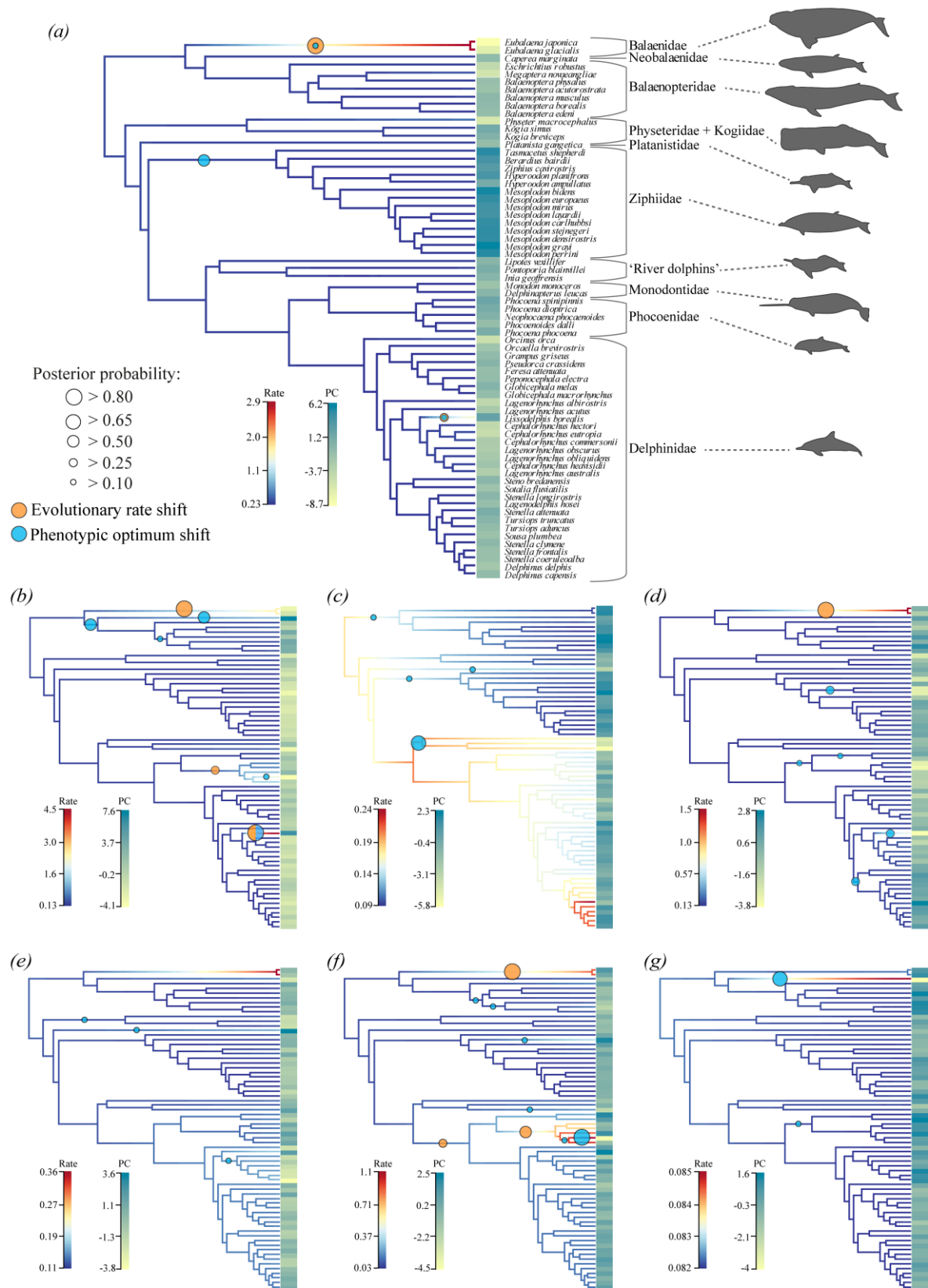


Figure S7. Evolutionary patterns of the vertebral shape (PC2 to PC8). Phylogenetic tree of cetaceans from Steeman [17] with branches coloured according to the evolutionary rates of each principal component from the 'all cetaceans' PCA based on phylogenetically corrected residuals. (a) PC2, (b) PC3, (c) PC4, (d) PC5, (e) PC6, (f) PC7 and (g) PC8. Coloured rectangles at trees tips represent species PC score. Orange circles show shifts in evolutionary rates and were calculated from a Bayesian multi-rate approach (BAMM). Blue circles represent shifts in phenotypic optima and were obtained from a Bayesian multi-regime Ornstein-Uhlenbeck approach (Bayou).

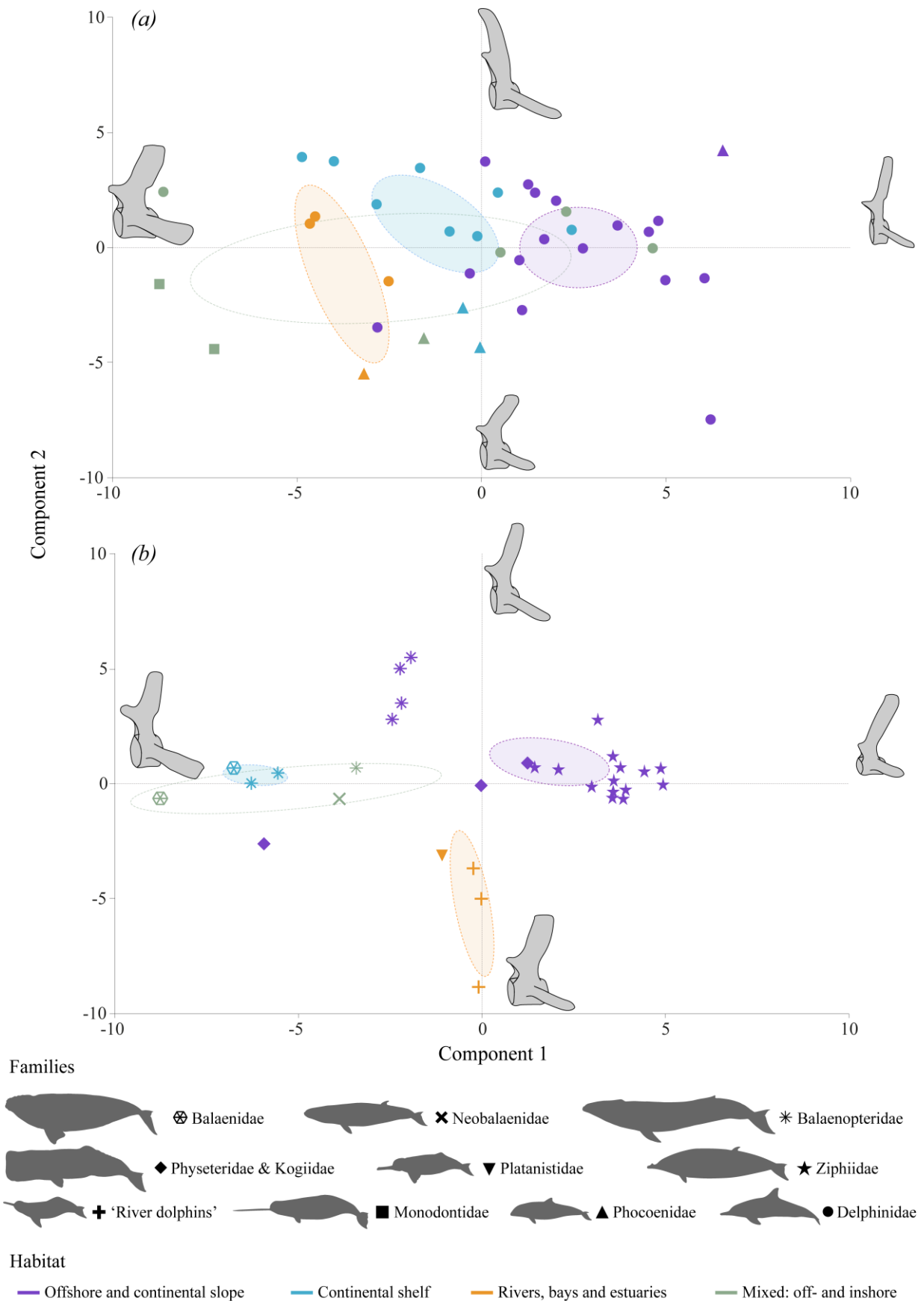


Figure S8. Effect of the habitat on vertebral shape. Principal components analysis plot of PC1 vs. PC2. Symbol shapes correspond to phylogenetic groups. Symbol colours correspond to different habitats. Dashed ellipses represent the 95% confidence intervals of the mean coordinates for each habitat category. (a) Delphinoidea. PC1: 37% of variance, PC2: 19%. (b) Non-Delphinoidea. PC1: 37%, PC2: 18%.

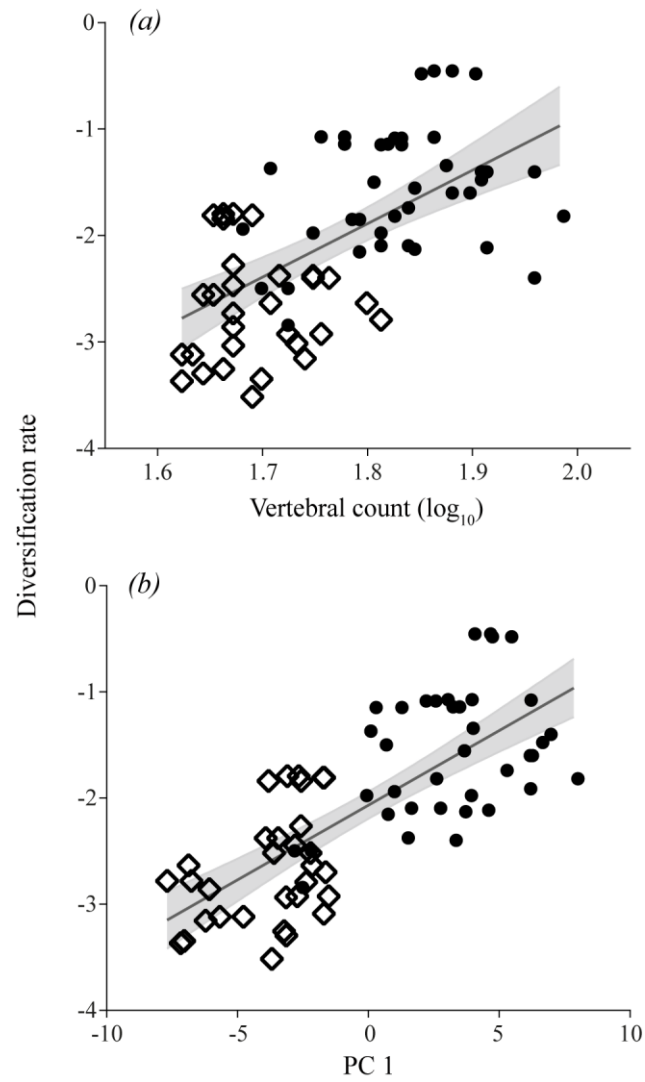


Figure S9. Relationship between diversification rate and phenotypic traits. (a) \log_{10} -transformed vertebral count. (b) PC1 scores from the 'all cetaceans' PCA'. Higher PC1 values correspond to more discoidal vertebrae while lower values correspond to more elongated vertebrae. Filled circles represent Delphinoidea while empty diamonds are for non-Delphinoidea. Diversification rates are the log-transformed speciation rates based on equal splits measure as described by Harvey and Rabosky [47]. For both traits, we found a significant correlation with diversification rates. Solid grey lines represent the linear best fit and grey-shaded areas correspond to the 95% confidence intervals.

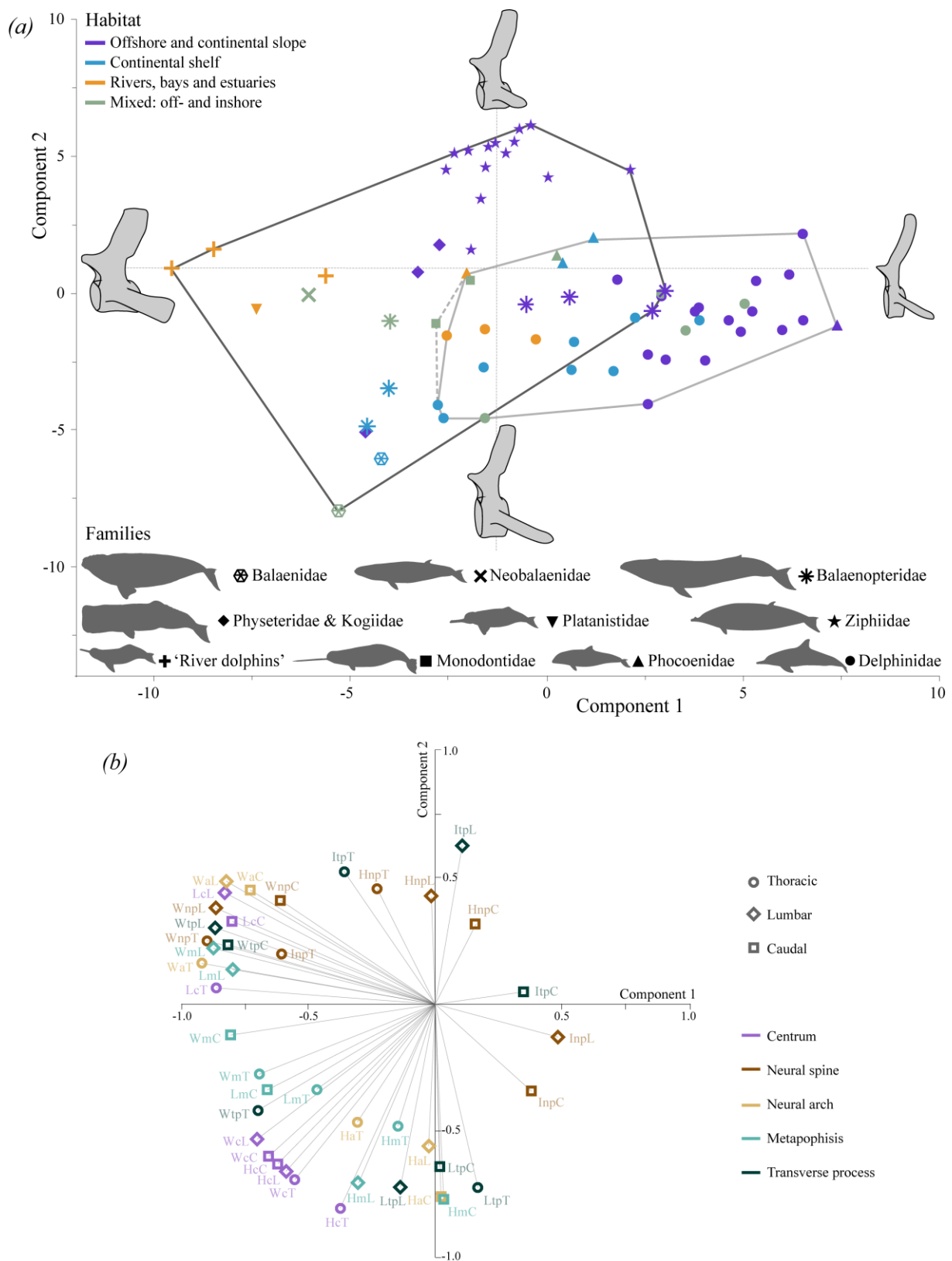


Figure S10. Effect of the habitat on vertebral shape without accounting for the effect of the phylogeny. (a) Projection of species scores on PC1 vs. PC2 for the 'all cetaceans' PCA calculated from non-phylogenetically size-corrected residuals. PC1 represents 35% of the variance and PC2 23%. Typical vertebral shapes are shown at each extremity of the axes. Symbol shapes correspond to phylogenetic groups and colours correspond to habitats. Convex hulls represent (1) Delphinidae and Phocoenidae (grey lines) and (2) non-delphinoidean cetaceans (black lines). Dotted grey lines show the inclusion of Monodontidae with Phocoenidae and Delphinidae. (b) Projection of variable loadings on PC1 and PC2. Symbol shapes correspond to regions of the backbone and colours to different parts of the vertebrae.

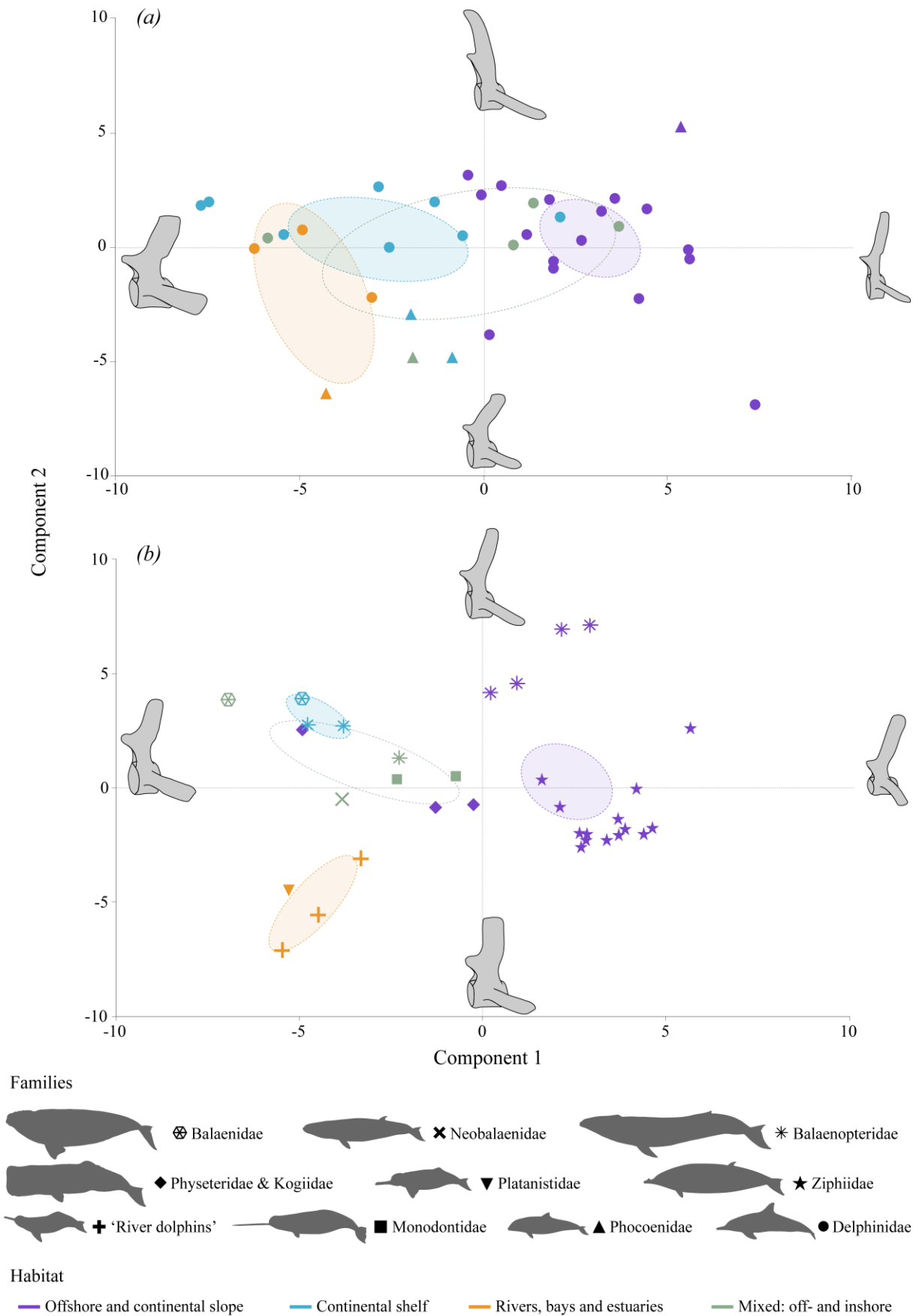


Figure S11. Effect of the habitat on vertebral shape without accounting for the effect of phylogeny. Principal components analysis plot of PC1 vs. PC2. Symbol shapes correspond to phylogenetic groups. Symbol colours correspond to different habitats. Dashed ellipses represent the 95% confidence intervals of the mean coordinates for each habitat category. (a) Delphinidae and Phocoenidae. PC1: 37% of variance, PC2: 18%. (b) All families excluding Delphinidae and Phocoenidae. PC1: 33%, PC2: 27%.

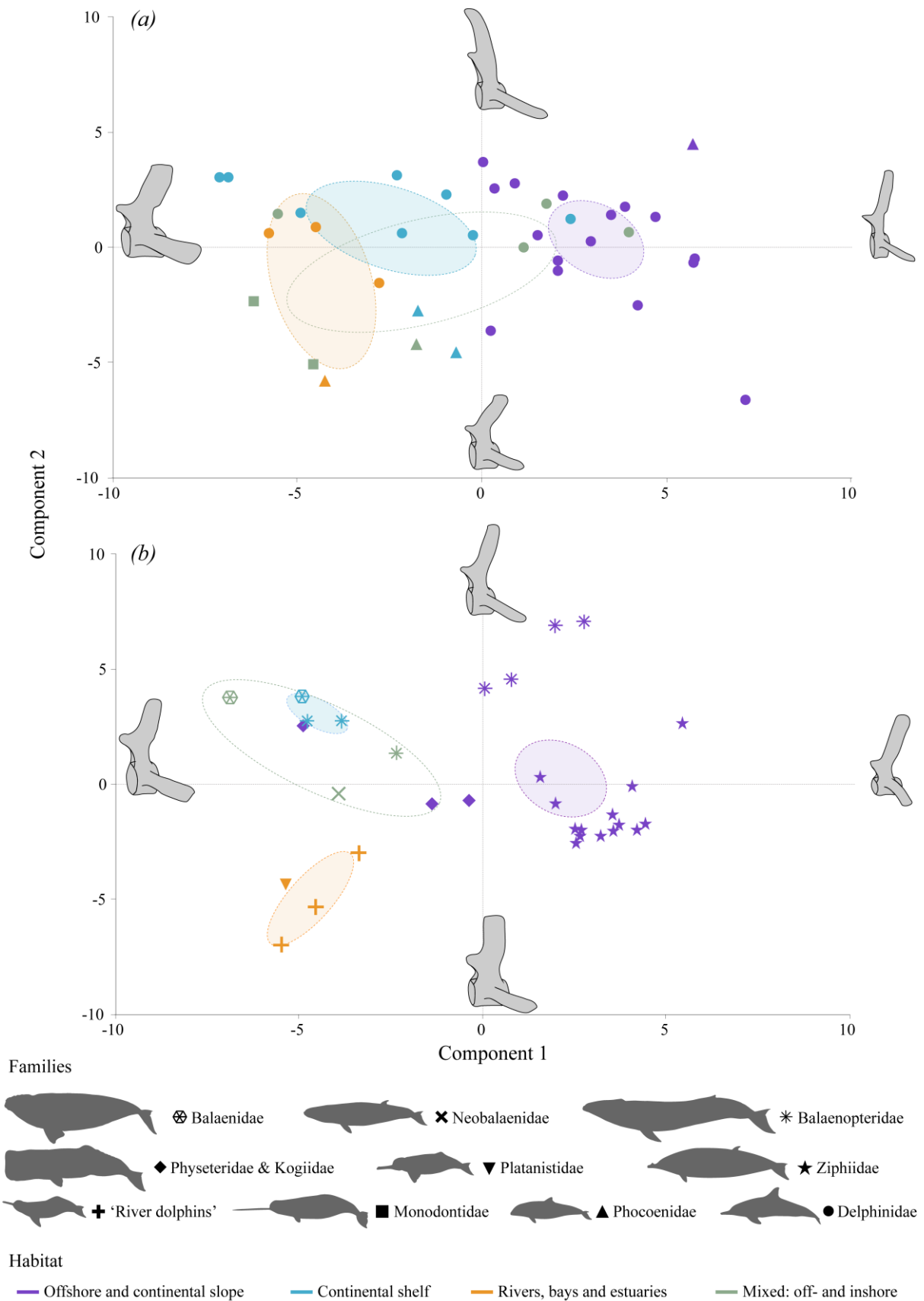


Figure S12. Effect of the habitat on vertebral shape without accounting for the effect of phylogeny. Principal components analysis plot of PC1 vs. PC2. Symbol shapes correspond to phylogenetic groups. Symbol colours correspond to different habitats. Dashed ellipses represent the 95% confidence intervals of the mean coordinates for each habitat category. (a) Delphinoidea. PC1: 37% of variance, PC2: 18%, (b) non-Delphinoidea. PC1: 34%, PC2: 27%.

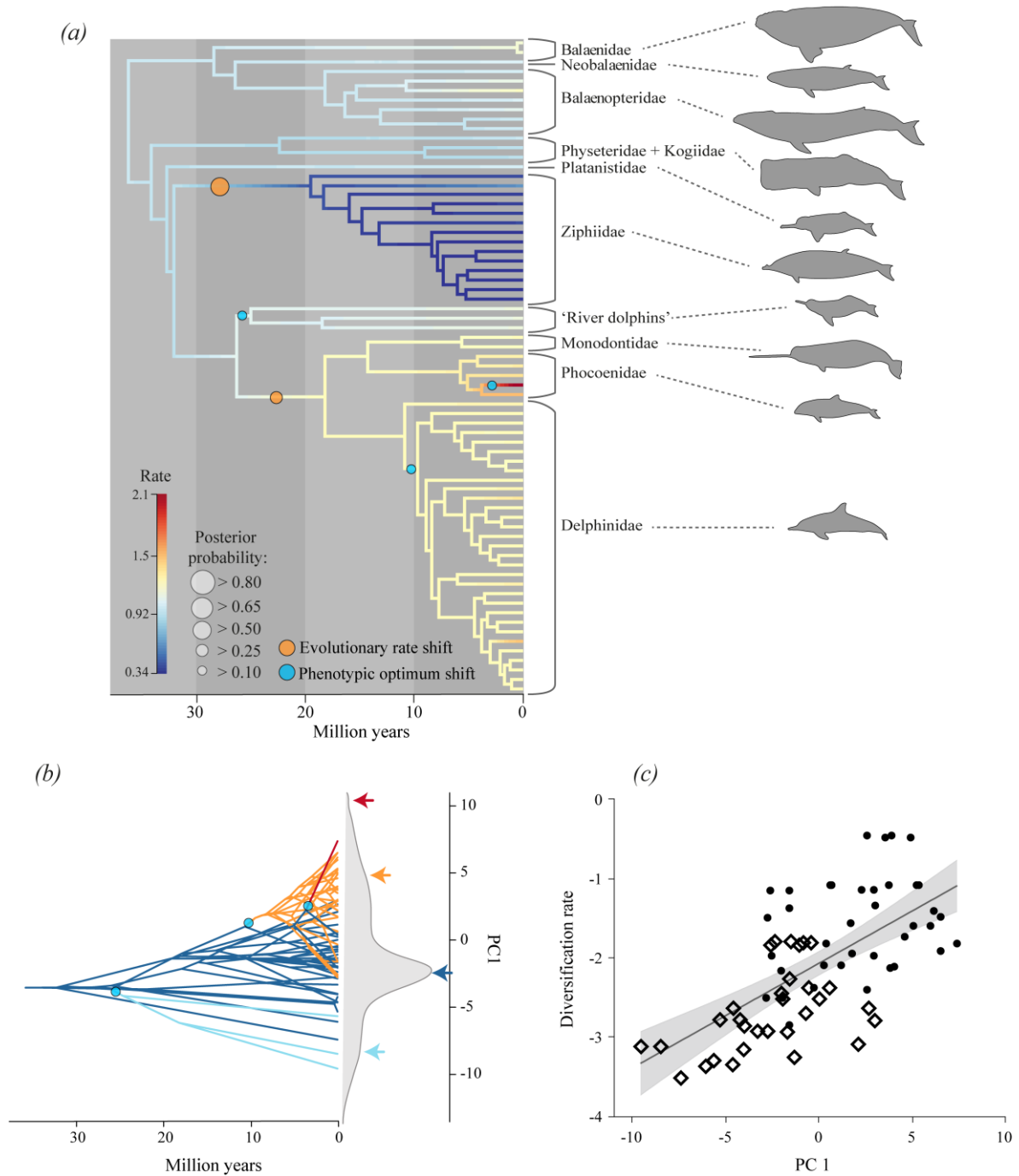


Figure S13. Evolutionary patterns of the vertebral shape (PC1) without phylogenetic correction. (a) Phylogenetic tree of cetaceans from Steeman [17] with branches coloured according to the evolutionary rates of the first principal component (PC1) from the 'all cetaceans' PCA based on non-phylogenetically corrected residuals. Evolutionary rates were obtained from a Bayesian multi-rate approach (BAMM). Grey bands represent 10 million year intervals. (b) Phenogram showing pattern of vertebral shape (PC1) evolution calculated from a Bayesian multi-regime Ornstein-Uhlenbeck approach (Bayou). Phylogenetic tree tips and internal nodes are plotted in function of their vertebral count value. Branches colours show clades with different evolutionary regimes identified by Bayou and coloured arrows correspond to their respective phenotypic optima. Posterior distribution of traits optima is represented by the grey shaded area and their mean value by the coloured arrows. (c) Regression between PC1 and the diversification rate for each species obtained from the ES-Sim test. The solid grey line represents the linear best fit of the statistically significant regression. Grey-shaded area corresponds to the 95% confidence intervals. Orange and blue circles in (a) and (b) show significant shifts of the evolutionary rate (BAMM) and the phenotypic optima (Bayou) respectively. Circles relative sizes correspond to the posterior probability of the shift. Both evolutionary rate shifts identified on (a) are mutually exclusive.

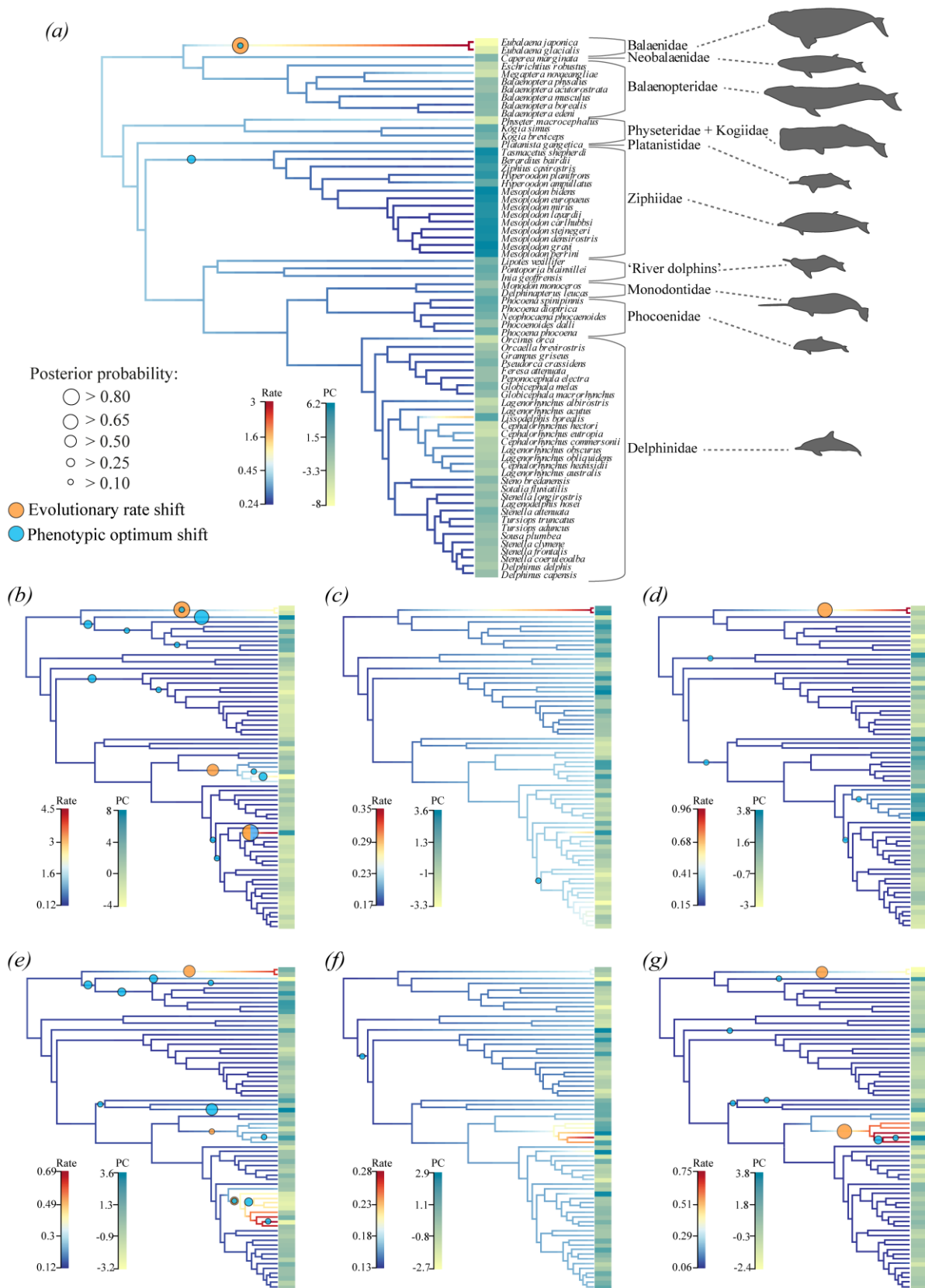


Figure S14. Evolutionary patterns of the vertebral shape (PC2 to PC8) without phylogenetic correction. Phylogenetic tree of cetaceans from Steeman [17] with branches coloured according to the evolutionary rates of each principal component from the 'all cetaceans' PCA based on non-phylogenetically corrected residuals. (a) PC2, (b) PC3, (c) PC4, (d) PC5, (e) PC6, (f) PC7 and (g) PC8. Coloured rectangles at trees tips represent species PC score. Orange circles show shifts in evolutionary rates and were calculated from a Bayesian multi-rate approach (BAMM). Blue circles represent shifts in phenotypic optima and were obtained from a Bayesian multi-regime Ornstein-Uhlenbeck approach (Bayou).

Table S1. List of specimens used in this study. Specimens belong to nine different museums: the American Museum of Natural History, New York (AMNH); the French National Museum of Natural History, Paris (MNHN); the Swedish Royal Museum of Natural History, Stockholm (NRM); the Bayworld Port Elizabeth Museum, Port Elizabeth (PEM); the Queensland Museum, Brisbane (QM); the Royal Belgian Institute of Natural Sciences, Brussels (RBINS); the Iziko South African Museum, Cape Town (SAM); the State Museum of Natural History, Stuttgart (SMNS) and the National Museum of Natural History, Washington D.C. (USNM). For most specimens, both vertebral count and shape data were collected. Specimens on which only vertebral count was collected are marked by a ^{'C'}, specimens on which only vertebral shape was collected are marked with an ^{'S'}.

<u>Balaenidae</u>				<i>Hyperoodon planifrons</i>	NRM	558395
	<i>Balaena mysticetus</i>	NRM	558409 ^C		SAM	ZM41892
	<i>Eubalaena australis</i>	PEM	N0019 ^C	<i>Mesoplodon bidens</i>	MNHN	A14519
	<i>Eubalaena glacialis</i>	NRM	558386 ^S		NRM	558398
		NRM	558389		USNM	594220
		USNM	593893	<i>Mesoplodon carlhubbsi</i>	USNM	504128
	<i>Eubalaena japonica</i>	USNM	339990 ^S	<i>Mesoplodon densirostris</i>	USNM	504217
<u>Neobalaenidae</u>					USNM	550754
	<i>Caperea marginata</i>	RBINS	1536		USNM	550951
<u>Balaenopteridae</u>				<i>Mesoplodon europaeus</i>	USNM	550824
	<i>Balaenoptera acutorostrata</i>	NRM	558397 ^S		USNM	572952
		USNM	49775		USNM	593439
	<i>Balaenoptera borealis</i>	NRM	558432	<i>Mesoplodon grayi</i>	PEM	N0021
		USNM	236680	<i>Mesoplodon layardii</i>	PEM	N0020
	<i>Balaenoptera edeni</i>	NRM	558399		SAM	ZM19931
		SAM	ZM12962 ^S	<i>Mesoplodon mirus</i>	SAM	ZM36844
		USNM	572922		USNM	504612
	<i>Balaenoptera musculus</i>	NRM	558430 ^C		USNM	504724
		USNM	124326	<i>Mesoplodon perrini</i>	USNM	504260
	<i>Balaenoptera physalus</i>	NRM	558431	<i>Mesoplodon stejnegeri</i>	USNM	504731 ^S
		NRM	558434		USNM	550113
	<i>Eschrichtius robustus</i>	NRM	558391 ^C	<i>Tasmacetus shepherdi</i>	SAM	ZM40484
		USNM	593558 ^S		USNM	484878
	<i>Megaptera novaeangliae</i>	NRM	558433	<i>Ziphius cavirostris</i>	RBINS	1504
		SAM	ZM02288		USNM	347645
<u>Physeteridae</u>					USNM	49599
	<i>Physeter macrocephalus</i>	NRM	558211 ^S	<u>'River dolphins'</u>		
		NRM	558400 ^C	<i>Inia geoffrensis</i>	MNHN	A61
		SMNS	26429		SMNS	45662
		USNM	301634		USNM	395614
<u>Kogiidae</u>					USNM	49582
	<i>Kogia breviceps</i>	PEM	N1862	<i>Lipotes vexillifer</i>	AMNH	57333
		PEM	N989	<i>Pontoporia blainvillei</i>	SMNS	45597
		SMNS	7618 ^S		USNM	501168
		USNM	504737		USNM	501179
		USNM	572932		USNM	504920
	<i>Kogia sima</i>	PEM	N1564	<u>Monodontidae</u>		
		PEM	N3554	<i>Delphinapterus leucas</i>	MNHN	A3246 ^S
		USNM	504221		NRM	558404
		USNM	593890		RBINS	1508
<u>Platanistidae</u>					USNM	571021
	<i>Platanista gangetica</i>	MNHN	A7945	<i>Monodon monoceros</i>	MNHN	A3235 ^S
		NRM	608417 ^C		NRM	558407
		SMNS	45652 ^S		USNM	594407
		SMNS	45648	<u>Phocoenidae</u>		
		SMNS	45651 ^S	<i>Neophocaena phocaenoides</i>	SMNS	45679
		SMNS	45653		SMNS	45680
<u>Ziphiidae</u>					SMNS	45681
	<i>Berardius bairdii</i>	USNM	49726 ^S		USNM	240002
	<i>Hyperoodon ampullatus</i>	NRM	558402	<i>Phocoena dioptrica</i>	USNM	571485
		RBINS	1503		USNM	571486 ^S

<i>Phocoena phocoena</i>	NRM 558322	<i>Lagenorhynchus australis</i>	USNM 395347
	NRM 805026		USNM 395350
	NRM 815072	<i>Lagenorhynchus obliquidens</i>	USNM 504412 ^S
	NRM 835011		USNM 504413
	NRM 845002		USNM 504415
	NRM 855083	<i>Lagenorhynchus obscurus</i>	SAM ZM41890 ^S
	NRM 855196		SAM ZM35681
	NRM 865039 ^C	<i>Lissodelphis borealis</i>	USNM 484929
	NRM 865044		USNM 550026
	NRM 875045	<i>Lissodelphis peronii</i>	NRM 558419 ^C
	NRM 875216	<i>Orcaella brevirostris</i>	RBINS 1512
	NRM 875358 ^C	<i>Orcaella heinssohni</i>	QM JM511 ^C
	NRM 895156	<i>Orcinus orca</i>	MNHN A3231 ^S
	NRM 20065226		NRM 558250 ^S
	RBINS 16233 ^C		NRM 558251
	USNM 550312		NRM 558401
	USNM 571709	<i>Peponocephala electra</i>	SAM ZM38245
<i>Phocoena spinipinnis</i>	USNM 395751		USNM 550399
	USNM 550782		USNM 593799
	USNM 550785 ^S		USNM 593941 ^S
<i>Phocoenoides dalli</i>	USNM 396304	<i>Pseudorca crassidens</i>	NRM 558271 ^S
	USNM 504417		NRM 558405
	USNM 504969		QM J14210
Delphinidae			SMNS 7617
<i>Cephalorhynchus commersonii</i>	SAM ZM40555	<i>Sotalia fluviatilis</i>	RBINS 1516
	USNM 550154		RBINS 20137
	USNM 550156		USNM 571558
<i>Cephalorhynchus eutropia</i>	NRM 616647	<i>Sousa plumbea</i>	PEM N1179
	USNM 395374		PEM N1266
<i>Cephalorhynchus heavisidii</i>	SAM ZM 0014		PEM N1582 ^S
	SAM ZM19943		PEM N1593
	SAM ZM36717 ^S		USNM 550939
<i>Cephalorhynchus hectori</i>	SAM ZM36182	<i>Stenella attenuata</i>	USNM 395390
	USNM 500864		USNM 396028
<i>Delphinus capensis</i>	PEM N1649		USNM 500122
	SMNS 45763 ^S	<i>Stenella clymene</i>	USNM 550501
<i>Delphinus delphis</i>	NRM 805172		USNM 550511
	RBINS 1519B		USNM 550532
	USNM 500273	<i>Stenella coeruleoalba</i>	PEM N289 ^S
	USNM 593770		USNM 504350
<i>Feresa attenuata</i>	PEM N4762 ^S		USNM 504384
	PEM N4763	<i>Stenella frontalis</i>	USNM 21915
	SMNS 8841		USNM 22017
	USNM 571268		USNM 504321
<i>Globicephala macrorhynchus</i>	USNM 22561 ^S	<i>Stenella longirostris</i>	PEM N1278 ^S
	USNM 593641		USNM 395414
<i>Globicephala melas</i>	NRM 558264		USNM 500017
	USNM 21118	<i>Steno bredanensis</i>	SAM ZM41124
<i>Grampus griseus</i>	MNHN A3248		USNM 504462
	NRM 558392 ^C		USNM 504468
	PEM N117 ^S	<i>Tursiops aduncus</i>	SAM ZM38240
	USNM 347613		SMNS 45711 ^S
	USNM 504328	<i>Tursiops truncatus</i>	SAM ZM35678 ^S
<i>Lagenodelphis hosei</i>	PEM N395		USNM 484529
	PEM N827 ^S		USNM 504618
	USNM 571619		USNM 504726 ^S
<i>Lagenorhynchus acutus</i>	USNM 504153		USNM 504906 ^S
	USNM 504154		USNM 550225
	USNM 504164		USNM 550364
<i>Lagenorhynchus albirostris</i>	NRM 20065395		USNM 550422
	SMNS 7591		USNM 550852
	USNM 550208		USNM 571388
			USNM 572831

Table S2. Priors used for the Bayesian multi-regime Ornstein-Uhlenbeck method (Bayou). $1/4$ *Ntips*: a quarter of the total number of species in the dataset, *Data mean*: mean value of the phenotypic trait of interest for all species, 2 *Data SD*: twice the standard deviation of the phenotypic trait of interest for all species.

Prior	Distribution function	Function parameters	Param. value
Constrain parameter (α)	Half-Cauchy	Scale:	1
Evolutionary rate (σ^2)	Half-Cauchy	Scale:	1
Shift position on branches (sb)	Uniform	Maximum number of shift per branch:	1
Expected number of shifts (k)	Conditional Poisson	Total number (λ):	3
		Maximum number:	$1/4$ Ntips
Phenotypic optimum (θ)	Normal	Mean:	Data mean
		Standard deviation:	2 Data SD

Table S3. Comparison of intra- and interspecific morphological disparity. Interspecific disparity is reported for the entire cetacean clade and intraspecific disparity is reported for two species (*P. phocoena* and *T. truncatus*). The vertebral shape disparity was calculated on size-corrected shape measurements. The absolute disparity of each group is expressed as Procrustes variance in the first part of the table. The second part of the table shows the results of pairwise comparisons of disparity. Significant *p*-values are in bold. *Diff.*: absolute pairwise differences between variances.

	Vertebral count		Vertebral shape	
	Variance		Variance	
Cetacea	163.27		0.916	
<i>P. phocoena</i>	18.07		0.293	
<i>T. truncatus</i>	4.26		0.379	
	Diff.	<i>P</i> -value	Diff.	<i>P</i> -value
Cetacea vs <i>P. phocoena</i>	145.19	0.009	0.622	0.005
Cetacea vs <i>T. truncatus</i>	159.01	0.023	0.537	0.017
<i>P. phocoena</i> vs <i>T. truncatus</i>	13.81	0.863	0.086	0.731

Table S4. Summary of correlation tests between body size and vertebral count. The table shows the comparison of the effect of body size on vertebral count when excluding or including Monodontidae with Delphinidae and Phocoenidae. Significant values are indicated in bold.

Test	n	P-value	R ²	Slope \pm s.e.
Effect of body size on vertebral count (PGLS)				
All cetaceans	71	0.7	-0.35	-0.27 \pm 0.69
Delphinidae and Phocoenidae	38	0.04	0.14	-4.36 \pm 2.00
Other cetacean families	33	0.0001	0.56	0.80 \pm 0.18
Delphinoidea	40	0.03	0.19	-4.51 \pm 1.93
Non-Delphinoidea	31	0.0001	0.61	0.80 \pm 0.17
Effect of body size on vertebral shape (multivariate PGLS)				
All cetaceans	71	0.06	0.11	/
Delphinidae and Phocoenidae	38	0.13	0.05	/
Other cetacean families	33	0.03	0.11	/
Delphinoidea	40	0.07	0.06	/
Non-Delphinoidea	31	0.04	0.11	/

Table S5. Summary of the effect of habitat on vertebral count and shape. The table shows the comparison of the analysis of variance tests when excluding or including Monodontidae with Delphinidae and Phocoenidae. Significant values are indicated in bold. *n*: number of species, *df*: degrees of freedom, *F*: F-value, *P*: *P*-value, η^2 : effect size (eta-squared), ω^2 : effect size (omega-squared).

Test	n	df	F	P	η^2	ω^2
Effect of habitat on vertebral count (pANOVA)						
All cetaceans	71	3,67	1.87	0.39	0.08	0.04
Delphinidae and Phocoenidae	38	3,34	4.86	0.01	0.30	0.24
Other cetacean families	33	3,29	2.05	0.52	0.17	0.09
Delphinoidea	40	3,36	5.8	0.02	0.33	0.27
Non-Delphinoidea	31	3,27	1.9	0.53	0.17	0.08
Effect of habitat on vertebral shape (pMANOVA)						
All cetaceans	69	3,65	4.96	0.001	0.35	/
Delphinidae and Phocoenidae	36	3,32	3.03	0.001	0.46	/
Other cetacean families	33	3,29	6.15	0.001	0.62	/
Delphinoidea	38	3,34	2.44	0.009	0.40	/
Non-Delphinoidea	31	3,27	6.99	0.001	0.62	/

Table S6. Correlations between diversification rate and morphological traits. The ES-sim test was run with 1,000 iterations on vertebral count and on each principal component (PC) of the PCA applied on all cetacean species. Significant values are indicated in bold. Slope was only calculated for significant correlations.

Variable	n	<i>P</i> -value	R ²	Slope \pm <i>s.e.</i>
Vertebral count	71	0.030	0.367	5.02 \pm 0.79
PC1	69	0.002	0.553	0.14 \pm 0.02
PC2	69	0.923	0.002	/
PC3	69	0.675	0.033	/
PC4	69	0.701	0.026	/
PC5	69	0.697	0.043	/
PC6	69	0.777	0.019	/
PC7	69	0.813	0.016	/
PC8	69	0.999	0.0001	/

Table S7. Summary of correlation tests between body size and vertebral count without phylogenetic correction. The table shows the comparison of the effect of body size on vertebral count when excluding or including Monodontidae with Delphinidae and Phocoenidae. Significant values are indicated in bold.

Test	n	P-value	R ²	Slope \pm s.e.
Correlation between vertebral count and body size (GLS)				
All cetaceans	71	0.01	0.09	-0.81 \pm 0.32
Delphinidae and Phocoenidae	38	0.01	0.15	-3.47 \pm 1.35
Other cetacean families	33	< 0.0001	0.57	0.75 \pm 0.12
Delphinoidea	40	0.003	0.21	-4.08 \pm 1.30
Non-Delphinoidea	31	< 0.0001	0.61	0.79 \pm 0.12

Table S8. Summary of the effect of habitat on vertebral count and shape without phylogenetic correction. The table shows the comparison of the analysis of variance tests when excluding or including Monodontidae with Delphinidae and Phocoenidae. Significant values are indicated in bold.

Test	n	df	F	P	η^2	ω^2
Effect of habitat on vertebral count (ANOVA)						
All cetaceans	71	3,67	1.87	0.143	0.08	0.04
Delphinidae and Phocoenidae	38	3,34	4.86	0.006	0.30	0.24
Other cetacean families	33	3,29	2.05	0.129	0.17	0.09
Delphinoidea	40	3,36	5.80	0.002	0.33	0.27
Non-Delphinoidea	31	3,27	1.90	0.153	0.17	0.08
Effect of habitat on vertebral shape (MANOVA)						
All cetaceans	69	3,65	4.52	< 0.001		/
Delphinidae and Phocoenidae	36	3,32	2.84	< 0.001		/
Other cetacean families	33	3,29	6.15	< 0.001		/
Delphinoidea	38	3,34	2.44	0.002		/
Non-Delphinoidea	31	3,27	6.99	< 0.001		/

Table S9. Correlations between diversification rate and morphological traits without phylogenetic correction. The ES-sim test was run with 1,000 iterations on each principal component (PC) of the 'all cetaceans' PCA applied on non-phylogenetically corrected residuals. Significant values are indicated in bold. Slope was only calculated for significant correlations.

Variable	n	<i>P</i> -value	R ²	Slope \pm <i>s.e.</i>
PC1	69	0.030	0.408	0.13 \pm 0.02
PC2	69	0.757	0.014	/
PC3	69	0.470	0.085	/
PC4	69	0.448	0.107	/
PC5	69	0.857	0.015	/
PC6	69	0.685	0.042	/
PC7	69	0.903	0.0002	/
PC8	69	0.929	0.001	/

References:

1. Kellogg R. 1936 A review of the archaeoceti. *Carnegie Institution Washingt. Publ.* **482**, 1–366.
2. Uhen MD. 2004 Form, function, and anatomy of *Dorudon atrox* (Mammalia, Cetacea): an archaeocete from the middle to late Eocene of Egypt. *Univ. Michigan Pap. Paleontol.* **34**, 1–222.
3. Martínez-Cáceres M, Lambert O, de Muizon C. 2017 The anatomy and phylogenetic affinities of *Cynthiacetus peruvianus*, a large *Dorudon*-like basilosaurid (Cetacea, Mammalia) from the late Eocene of Peru. *Geodiversitas* **39**, 7–163. (doi:10.5252/g2017n1a1)
4. Emlong D. 1966 A new archaic cetacean from the Oligocene of northwest Oregon. *Bull. Museum Nat. Hist. Univ. Oregon* **3**, 1–51.
5. Kellogg R. 1969 Cetothere skeletons from the Miocene Choptank Formation of Maryland and Virginia. *United States Natl. Museum Bull.* **294**, 1–40. (doi:10.5479/si.03629236.294.1)
6. Kellogg R. 1924 A fossil porpoise from the Calvert formation of Maryland. *Proc. United States Natl. Museum* **63**, 1–39. (doi:10.5479/si.00963801.63-2482.1)
7. de Muizon C. 1984 Les vertébrés fossiles de la Formation Pisco (Pérou). Deuxième partie: les odontocètes (Cetacea, Mammalia) du Pliocène inférieur de Sud-Sacaco. *Trav. l'Institut Français d'Etudes Andin.* **50**, 1–188.
8. Kellogg R. 1927 *Kentriodon pernix*, a Miocene porpoise from Maryland. *Proc. United States Natl. Museum* **69**, 1–55. (doi:10.5479/si.00963801.69-2645.1)
9. de Muizon C. 1988 Les vertébrés fossiles de la Formation Pisco (Pérou). Troisième partie: Les Odontocètes (Cetacea, Mammalia) du Miocène. *Trav. l'Institut Français d'Etudes Andin.* **78**, 1–244.
10. Barnes LG. 2008 Miocene and Pliocene Albireonidae (Cetacea, Odontoceti), rare and unusual fossil dolphins from the eastern North Pacific Ocean. *Nat. Hist. Museum Los Angeles Cty. Sci. Ser.* , 99–152.
11. Aguirre-Fernández G, Fordyce RE. 2014 *Papahu taitapu*, gen. et sp. nov., an early Miocene stem odontocete (Cetacea) from New Zealand. *J. Vertebr. Paleontol.* **34**, 195–210. (doi:10.1080/02724634.2013.799069)
12. Geisler JH, Godfrey SJ, Lambert O. 2012 A new genus and species of late Miocene inioid (Cetacea , Odontoceti) from the Meherrin River, North Caroline, USA. *J. Vertebr. Paleontol.* **32**, 198–211. (doi:10.1080/02724634.2012.629016)
13. Geisler J, McGowen M, Yang G, Gatesy J. 2011 A supermatrix analysis of genomic, morphological, and paleontological data from crown Cetacea. *BMC Evol. Biol.* **11**, 1–33. (doi:10.1017/CBO9781107415324.004)
14. McGowen MR, Spaulding M, Gatesy J. 2009 Divergence date estimation and a comprehensive molecular tree of extant cetaceans. *Mol. Phylogenet. Evol.* **53**, 891–906. (doi:10.1016/j.ympev.2009.08.018)
15. Murakami M, Shimada C, Hikida Y, Soeda Y, Hirano H. 2014 *Eodelphis kabatensis*, a new name for the oldest true dolphin *Stenella kabatensis* Horikawa, 1977 (Cetacea , Odontoceti , Delphinidae), from the upper Miocene of Japan, and the phylogeny and paleobiogeography of Delphinoidea. *J. Vertebr. Paleontol.* **34**, 491–511. (doi:10.1080/02724634.2013.816720)
16. Slater GJ, Price SA, Santini F, Alfaro ME. 2010 Diversity versus disparity and the radiation of modern cetaceans. *Proc. R. Soc. B Biol. Sci.* **277**, 3097–3104. (doi:10.1098/rspb.2010.0408)

17. Steeman ME *et al.* 2009 Radiation of extant cetaceans driven by restructuring of the oceans. *Syst. Biol.* **58**, 573–585. (doi:10.1093/sysbio/syp060)
18. Wickham H. 2016 *ggplot2: Elegant graphics for data analysis*. New York: Springer-Verlag. See <https://ggplot2.tidyverse.org>.
19. R Core Team. 2017 R: A language and environment for statistical computing. *R Fund. Stat. Comput.* See <https://www.r-project.org/>.
20. Pinheiro J, Bates D, DebRoy S, Sarkar D, R Core Team. 2017 nlme: Linear and nonlinear mixed effects models. *R Packag. version 3.1-131*. See <https://cran.r-project.org/package=nlme>.
21. Adams DC, Collyer ML, Kaliontzopoulou A. 2018 Geomorph: Software for geometric morphometric analyses. *R Packag. version 3.0.6*. See <https://cran.r-project.org/package=geomorph>.
22. Narita Y, Kuratani S. 2005 Evolution of the vertebral formulae in mammals: A perspective on developmental constraints. *J. Exp. Zool. Part B Mol. Dev. Evol.* **304**, 91–106. (doi:10.1002/jez.b.21029)
23. Jones KE *et al.* 2009 PanTHERIA: a species-level database of life history, ecology, and geography of extant and recently extinct mammals. *Ecology* **90**, 2648–2648. (doi:10.1890/08-1494.1)
24. Marx FG, Lambert O, Uhen MD. 2016 *Cetacean paleobiology*. Chichester, UK: John Wiley & Sons.