The dynamic adaptive landscape of cetacean body size

Highlights

- The adaptive landscape of cetacean body size is remarkably flat over deep time

- The largest peak shift is associated with cetaceans adapting to aquatic life

- Local peaks are also present, linked to adaptations for feeding and deep diving

- Adaptive peaks are dynamic and can also evolve, making them harder to detect

Authors

Gustavo Burin, Travis Park,
Tamora D. James, Graham J. Slater,
Natalie Cooper

Correspondence

gustavo.burin@nhm.ac.uk

In brief

Burin, Park, et al. estimate the adaptive landscape of cetacean body size over 50 million years. The landscape is quite flat, with few new peaks arising after cetaceans entered the oceans. However, trends linked to specific adaptations are evident at finer scales. These results confirm that adaptive landscapes are dynamic and peaks evolve over time.

CellPress

Report

The dynamic adaptive landscape
of cetacean body size

Gustavo Burin,[1,4,5,10,*] Travis Park,[1,4,6] Tamora D. James,[2,7] Graham J. Slater,[3,8] and Natalie Cooper[1,9]
[1]Natural History Museum London, Cromwell Road, London SW7 5BD, UK
[2]School of Earth and Environment, University of Leeds, Leeds LS2 9JT, UK
[3]Department of the Geophysical Sciences, University of Chicago, Chicago, IL 60637, USA
[4]These authors contributed equally
[5]Twitter: @gustavoburin
[6]Twitter: @blogozoic
[7]Twitter: @tamorajames
[8]Twitter: @grahamjslater
[9]Twitter: @nhcooper123
[10]Lead contact
*Correspondence: gustavo.burin@nhm.ac.uk
https://doi.org/10.1016/j.cub.2023.03.014

SUMMARY

Adaptive landscapes are central to evolutionary theory, forming a conceptual bridge between micro- and macroevolution.[1–4] Evolution by natural selection across an adaptive landscape should drive lineages toward fitness peaks, shaping the distribution of phenotypic variation within and among clades over evolutionary timescales.[5] The location and breadth of these peaks in phenotypic space can also evolve,[4] but whether phylogenetic comparative methods can detect such patterns has largely remained unexplored.[6] Here, we characterize the global and local adaptive landscape for total body length in cetaceans (whales, dolphins, and relatives), a trait that spans an order of magnitude, across their ~ 53-million-year evolutionary history. Using phylogenetic comparative methods, we analyze shifts in long-term mean body length[7] and directional changes in average trait values[8] for 345 living and fossil cetacean taxa. Remarkably, we find that the global macroevolutionary adaptive landscape of cetacean body length is relatively flat, with very few peak shifts occurring after cetaceans entered the oceans. Local peaks are more numerous and manifest as trends along branches linked to specific adaptations. These results contrast with previous studies using only extant taxa,[9] highlighting the vital role of fossil data for understanding macroevolution.[10–12] Our results indicate that adaptive peaks are dynamic and are associated with subzones of local adaptations, creating moving targets for species adaptation. In addition, we identify limits in our ability to detect some evolutionary patterns and processes and suggest that multiple approaches are required to characterize complex hierarchical patterns of adaptation in deep time.

RESULTS

Total length data imputation

By using phylogenetically informed imputations and a range of proxy measurements (see STAR Methods), along with total length information from the literature and museum specimens, we generated the largest total length dataset to date for Cetacea, comprising 345 taxa (89 living and 256 fossils; Figure 1). Where total lengths were known, imputed values showed a strong correlation with them, regardless of the amount of information used for estimation (see Figure S1), meaning that total lengths could be estimated for even incomplete fossil taxa (Figure S1), allowing their inclusion in downstream analyses. These additional data are critical to robust macroevolutionary inference, though we note that it is impossible to calculate total length values without a degree of uncertainty (see STAR Methods), and this must be considered in any interpretation of the results.

Exploring tempo and mode in body size evolution

As in previous work,[9,14] we find that disparity through time in extant cetaceans is lower than would be expected under a constant-rates process (center of gravity = 0.43, p = 0.005), consistent with early burst adaptive radiation models (Figure 2). However, this signal is entirely erased by the inclusion of fossil taxa, regardless of whether we include all fossil cetaceans or limit our sample to Neoceti only (center of gravity = 0.29, p = 0.294).

The macroevolutionary adaptive landscape is often conceptualized as an *n*-dimensional topographic map, the ruggedness of which depicts the relationship between morphological form and fitness[4] (Figure S2). If the topographic relief of the landscape is sufficient, selection should cause traits to evolve rapidly toward peaks and remain there until selection relaxes or the topography of the landscape is altered, for example, due to environmental disturbance or changes in resource availability. The topography
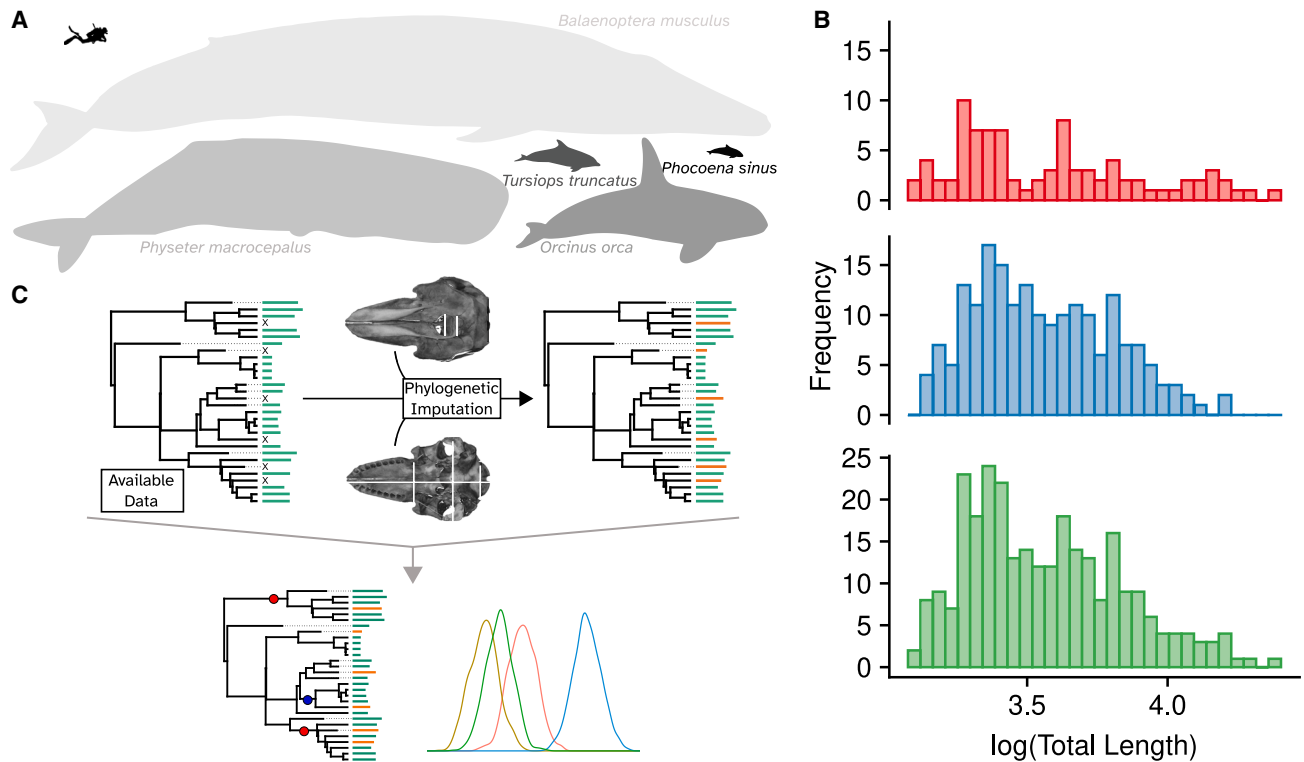
Figure 1. Total length variation in cetaceans and the analytical approach used in the analyses
(A) Silhouettes of five species of cetacean highlighting the variation in total length within the group (diver silhouette to scale).
(B) Histograms showing the distribution of $\log_{10}$-transformed total length (m) for living (red bars/upper panel), fossil (blue bars/middle panel), and all cetacean species (green bars/bottom panel).
(C) Schematic representation of the analytical approach used in the paper. We used data from the literature, new measurements of museum specimens, and the phylogeny of Lloyd and Slater[13] to estimate total length values for species lacking these data. The images in the center of the panel correspond to the dorsal (top) and ventral (bottom) views of the skull of an example odontocete. The imputed dataset and phylogeny were then analyzed using *bayou*[7] and the Fabric model[8] to detect shifts in the evolutionary regimes of total length, generating insights about the adaptive landscape of this trait.
See also Figure S1 for a quality assessment of the imputation approach.

of the adaptive landscape can be approximated as an Ornstein-Uhlenbeck process, where the positions of peaks on the landscape are represented as a vector of distinct long-term means, $\theta = (\theta_1, \theta_2, \ldots, \theta_n)$, and the ruggedness of the landscape is represented by the magnitude of a parameter, α, which determines the strength of the pull toward the peaks as an exponential function of time.[15] As $\alpha \to 0$, the ruggedness of the adaptive landscape declines and evolution reverts to a random walk. Such a result would not necessarily mean that a trait does not experience strong selection over short (e.g., generational) timescales; rather, it would imply that there are no long-term active trends in trait evolution at the clade level.

As one of the most conspicuous and fundamental features of an organism, body size can improve our understanding of changes in adaptive landscapes over macroevolutionary timescales.[16] Size influences multiple aspects of ecology, physiology, and, ultimately, fitness,[17–19] and has therefore become an obvious proxy for life history in analyses of evolutionary tempo and mode.[20,21] We used reversible jump Markov chain Monte Carlo (MCMC) methods[7] to map the adaptive landscape for cetacean body sizes. An advantage of this approach is that it does not require the user to specify the location of peaks *a priori*;

rather, they are estimated from the data. Estimated peaks differed by an order of magnitude, ranging from approximately 2 up to 12 m, although the 95% credibility interval for the estimated peaks indicated that this variation could be even higher, ranging from 60 cm to 50 m (see Figure S3).

It has long been recognized that secondarily aquatic tetrapods tend to be larger than their terrestrial sister lineages.[22] Diverse explanations have been suggested for this pattern. Life in water frees the limbs from mechanical constraints associated with supporting body mass,[23] which should passively increase the upper limit on potential sizes without imposing a minimum size constraint. In addition, heat is more readily lost in water than in air and, because surface area is reduced relative to volume in larger animals, this should place constraints on the lower size limits of fully aquatic taxa.[24] Our results indicate that the adaptive landscape for cetacean body size was profoundly shaped by the thermoregulatory lower limit on size in marine environments,[22,25,26] with a dramatic shift to a new adaptive peak associated with large (posterior mean = 12 m total length) size during the major transitional stage of adapting to fully aquatic life in ambulocetids, from a small (posterior mean = 2 m) ancestral size.
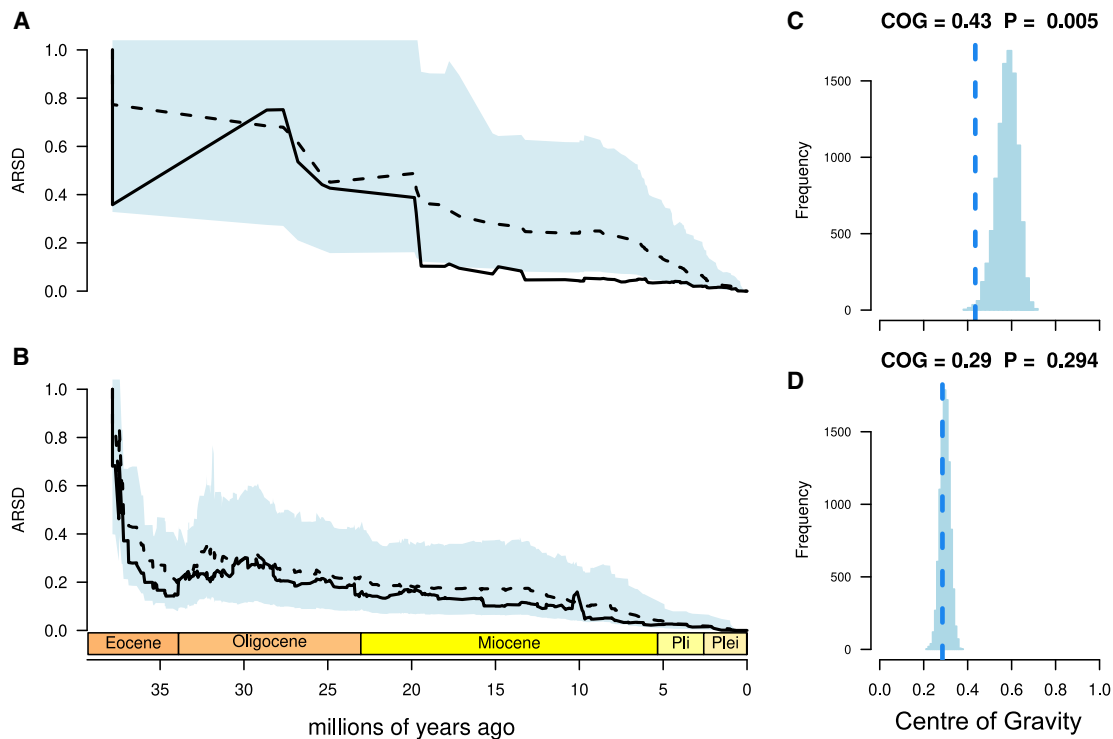
Figure 2. Disparity through time of total length for extant cetaceans

The disparity through time profile for extant cetaceans (A) is consistent with an adaptive radiation, with the average relative subclade disparity (ARSD) dropping substantially below constant rates expectation in the Early Miocene. This is reflected in significantly low center of gravity (C). Adding fossil neocete data (B) (though similar results are obtained for total group Cetacea) results in a subclade disparity profile that is more consistent with, and not significantly different from, (D) the constant rates expectation.

We found limited support for a trophically defined adaptive landscape for body size in cetaceans, *contra* previous work on extant taxa alone.[9] The cetacean body size adaptive landscape remains remarkably stable for ~20 million years after their initial transition to the marine realm. However, a dramatic decrease in long-term mean body size (from ~12.5 to ~2.9 m) in the Late Oligocene, along the branch leading to Delphinida (oceanic dolphins, porpoises, and non-platanistid river dolphins), with a second shift to even smaller sizes (long-term mean ~ 2.1 m) in the Early Miocene, along the branch leading to total group Inioidea + Lipotidae, is consistent with a dietary model. Maneuverability is negatively correlated with body size in marine mammals,[27] and evolution about these smaller long-term means would provide a performance advantage for delphinidans when foraging for small, agile prey in complex environments. It is tempting to speculate that the decrease in long-term mean body size in the Lipotidae + Inioidea clade is an adaptation to living in freshwater river systems, given their extant distributions. However, this modern distribution is convergently derived, and many fossil taxa in the clade are found in marine sediments.[28] Functional studies of *Inia* indicate that its flexible body enhances maneuverability at the expense of speed.[27] Thus, members of this lineage may have become smaller to optimally maneuver in shallow coastal environments, which then served as an exaptation for life in rivers. Delphinoids, in contrast, exhibit a suite of vertebral adaptations that result in a stiffer body,[29] which sacrifices

some of the maneuverability attained by iniids and pontoporiids for increased speed and swimming efficiency.[27] This likely facilitated the diversification of pelagic piscivores during the Late Miocene, which succeeded the archaic stem lineages and platanistoids[30,31] and culminated in the dramatic, rapid radiation of oceanic dolphins that today comprise almost half of cetacean diversity.[9,13,32]

Aside from these major shifts, we detected few other peaks on the adaptive landscape that affect large clades or that are associated with substantial size change. Shifts within odontocetes represent increases in body size in relation to the delphinidan mean, and were detected for Kentriodontidae and closely related stem delphinidans, and for *Orcaella* (Figures 3 and S4). Within the baleen whales (Mysticeti), we only observed shifts within the crown group (Figure 3). These shifts indicate localized decreases in body size within balaenids and in "basal thalassotherians" (*sensu* Bisconti et al.[33]), with other lineages of balaenopterids containing the remaining changes (Figures 3 and S4). We found no evidence for shifts toward gigantic sizes in crown Balaenopteridae or Balaenidae, the largest extant cetaceans, likely due to the inference that a large optimal body size is acquired early in cetacean evolutionary history. However, we did find a shift toward smaller size in the minke whale. Virtually all the shifts described above were retained across our sensitivity analyses (see STAR Methods and Burin et al.[34]), regardless of the tree and prior parameters used.[34]
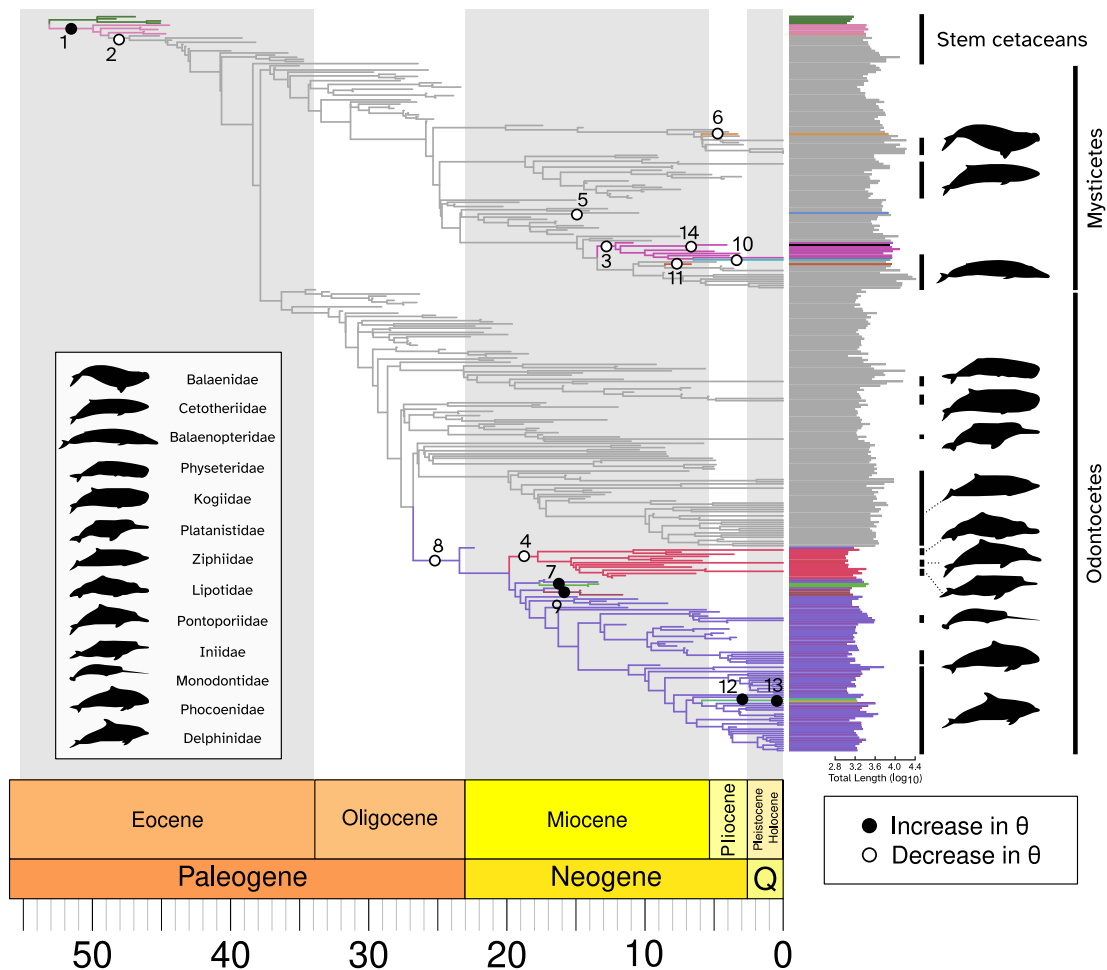
Figure 3. Evolutionary shifts in total length evolution of living and extinct cetaceans

Phylogeny of cetaceans showing the 14 shifts detected by *bayou*.[7] Solid black circles represent increases in average total length (θ) for a given regime; hollow circles represent decreases in average total length (θ) for a given regime. The bars to the right of the phylogeny show the $\log_{10}$-transformed total length values (m) for each species. The colors in the branches and their corresponding bars identify the regime to which each lineage belongs. Silhouettes indicate cetacean families that have at least one living species. In the geological timescale, Q, is the Quaternary.
See also Figures S2–S4.

Although the fit of Ornstein-Uhlenbeck models to comparative data is often interpreted in terms of estimates of θ, the location of the adaptive peak(s), the parameter α that describes the topography of the landscape and thus the historical strength of selection, is of equal importance.[35] Estimates from wild populations indicate that selection can drive incredibly rapid responses in heritable phenotypes within a single generation,[36] though this may represent short-term fluctuations in response to changes in local conditions[37] that do not accumulate to yield large macroevolutionary divergences.[6] Still, for cetacean body size, we estimate an extremely low rate of adaptation (posterior mean $\alpha = 0.0077$, 95% highest posterior density = 0.0001–0.0141), indicating that it would take, on average, approximately 90 million years (or 1.6 times the age of the entire clade) to move halfway toward a new long-term mean body size after a peak shift. Such low estimated rates of adaptation indicate that selection on the global adaptive landscape was not the only factor shaping the distribution of body sizes for living and extinct cetaceans.

A flat adaptive landscape for cetacean body size implies that variance should have increased through time, yielding both larger and smaller taxa in spite of an inferred mean size at the upper end of the body size distribution. However, if our models are insufficient, our low parameter estimates may reflect more complex underlying dynamics rather than weak selection. For example, the surprising lack of peak shifts identified in our bayou analyses, particularly in large-bodied extant taxa, could result from that model's inability to accommodate multiple α values, causing us to miss nested shifts associated with periods of accelerated evolution toward a new local peak, or from small sample sizes (such as where shifts occur along terminal branches) that preclude parameter identifiability. To account for these possibilities, we fitted a simpler phenomenological model, the "Fabric model,"[8] to our data in which phenotypic shifts along individual branches of phylogeny are treated as biased deviations from an otherwise unbiased random walk in which rates are free to vary. Any trends toward larger or smaller
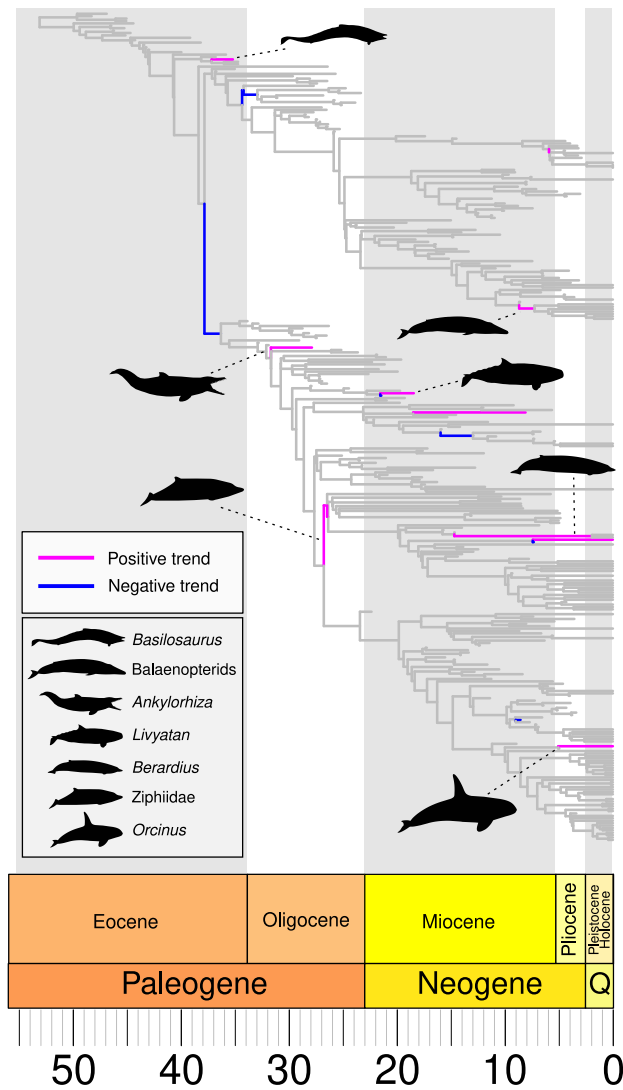
Figure 4. Evolutionary trends in total length evolution of living and extinct cetaceans

Phylogeny of cetaceans showing the branches with trends toward increasing or decreasing total length detected by the Fabric model.[8] Magenta lines represent branches with evidence of positive trends in average total length; blue lines represent branches with evidence of negative trends in average total length. Silhouettes indicate the main groups for which significant trends were found. In the geological timescale, Q, Quaternary.
See also Figure S2.

the branch leading to crown group balaenopterids (excluding minkes). Additional increases were detected in crown balaenids and the deep-diving ziphiids. Seven trends toward smaller size were also recovered. For example, the ancestor of total group odontocetes underwent a substantial (35%) decrease in size relative to the neocete ancestor, while smaller decreases occurred in the common ancestor of the pygmy sperm whale family Kogiidae, several odontocete terminal lineages, and a pair of nested trends toward smaller size in the toothed mysticete lineage Aetiocetidae.

DISCUSSION

The evolution of morphological diversity demands explanation. Models of gradual evolution predict a simple increase in phenotypic variance through time, with more closely related lineages exhibiting greater similarity to one another, on average, than they do to more distantly related lineages.[38,39] Where selection acts to drive lineages toward distinct phenotypes, it should mold and shape the distribution of phenotypic variation within and across clades, yielding patterns that deviate from the expectations of simple gradualistic models.[5,40] Characterizing these "ghosts of selection past" is a central challenge of comparative evolutionary biology. Our analyses suggest fundamental limits in their ability to detect some evolutionary patterns and processes and that a diversity of approaches is required to characterize these complex hierarchical patterns of phenotypic adaptation.

Comparative biologists increasingly model phenotypic macroevolutionary landscapes using the mathematics of the Ornstein-Uhlenbeck process, which connects microevolutionary processes[5] to patterns of variance and covariance over phylogeny.[41] In this framework, peaks on the landscape act as phenotypic attractors while selection strength reflects the steepness of the peaks. However, these models are data-hungry, and phylogenies with massive numbers of tips are required for accurate parameter estimation.[42] This, in turn, can result in a coarse, global view of the adaptive landscape, such as the identification of a single shift of large effect associated with the transition to the marine realm, or a shift to a much smaller mean size associated with increased maneuverability in piscivorous delphinidan odontocetes. What is missed, then, are more localized changes in the topography of the adaptive landscape, which may be particularly difficult to detect when different lineages experience change in different ways.[4]

These hidden dynamics are evident as multiple trends in body size evolution that appear on more local scales than the global peak shifts identified by our bayou analyses. In some cases, they occur along branches leading to individual species that can be characterized as possessing similar ecologies, such as trends toward larger size along branches leading to the macropredators (see also Coombs et al.[43]). In other cases, trends are clade-dependent, such as a recent trend toward larger sizes in balaenopterids and, to a lesser extent, balaenids, which is associated with bulk feeding on dense prey patches,[26,44–46] and ziphiids, which is associated with deep-diving behaviors.[47] Trends toward smaller sizes associated with the origin of odontocetes and kogiids may reflect similar responses to foraging on small, mobile prey relative to ancestral states.[48]

sizes can be thought of as evolutionary change that falls outside of the range of possibilities permitted by an otherwise unbiased process, though no explicit mechanistic explanation can be invoked.

Using the phenomenological model, we identified 18 branches along which evolution was biased away from an expected change of 0, none of which were found by bayou[7] (Figure 4). 11 of these shifts were toward larger sizes. Notably, some of these involved macroraptorial predators *Basilosaurus*, *Ankylorhiza tiedemani*, *Livyatan melvillei*, and the modern orca *Orcinus orca*. A massive (175%) increase in size was also detected along

The incongruence between a low number of global shifts and multiple conspicuous, more taxonomically restricted local body size trends suggests a much more dynamic adaptive landscape than would be inferred from any single analysis. How could this be the case? Simpson[2,3] conceptualized adaptive zones in terms of phenotypic solutions to functional problems, suggesting that traits associated with ecological behaviors like feeding and locomotion might be associated with rugged adaptive landscapes.[49] Thus the global body size adaptive landscape may be fairly flat when averaged over time, but there could still be distinct peaks on the cranial morphology adaptive landscape, for instance, associated with bulk feeding, piscivory, etc.[43] This means that there are likely multiple phenotypic dimensions to the landscape of which we are only exploring one. However, body size is a complex and multifaceted trait, interconnected with almost every aspect of a species' biology.[17,50] This, in turn, means that body size is under multiple selection pressures, each with its own adaptive mean, resulting in a locally rugged adaptive landscape (i.e., the principle of frustration, *sensu* Marshall[51]). We argue that the close proximity of these long-term means facilitates movement between peaks on the adaptive landscape in relatively small increments, which are not identified as discrete shifts in our bayou analyses but appear as trends in the analysis using the Fabric model.[8] Therefore, we describe the adaptive landscape of body size in cetaceans over macroevolutionary timescales by comparing it to a turbulent ocean surface rather than to the more stable topography of terrestrial landscapes (*sensu* Bell[52]). The few shifts we identified are analogous to large waves, with the smaller peaks representing unique local conditions of adaptation,[4] more ephemeral in nature, akin to ripples on the surface of the ocean that change rapidly across both space and time. These dynamics likely further increase the difficulty of peaks being detected as discrete shifts in comparative analyses. We note that a major reason our analyses detect these complex dynamics and previous studies do not is that we include fossil taxa. Fossils flatten the adaptive landscape and reveal different peaks. Together, these results highlight the need to include fossils and use multiple comparative approaches to fully characterize adaptive landscapes in deep time, especially in traits as labile and integrated as body size.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Total length data
  - Phylogeny
- METHOD DETAILS
  - Total length data imputation
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Exploring tempo and mode in body size evolution
  - Sensitivity analyses

- Removing taxa with zero-length branches
- Evaluating per-branch trends in body size evolution

REFERENCES

1. Wright, S. (1932). The roles of mutation, inbreeding, crossbreeding, and selection in evolution. Proceedings of the Sixth International Congress on Genetics *1*, 356–366.

2. Simpson, G.G. (1944). Tempo and Mode in Evolution (Columbia University Press).

3. Simpson, G.G. (1953). Major Features of Evolution (Columbia University Press).

4. Arnold, S.J., Pfrender, M.E., and Jones, A.G. (2001). The adaptive landscape as a conceptual bridge between micro-and macroevolution. Genetica *112*, 9–32.

5. Lande, R. (1985). Expected time for random genetic drift of a population between stable phenotypic states. Proc. Natl. Acad. Sci. USA *82*, 7641–7645.

6. Voje, K.L. (2016). Tempo does not correlate with mode in the fossil record. Evolution *70*, 2678–2689.

7. Uyeda, J.C., and Harmon, L.J. (2014). A novel Bayesian method for inferring and interpreting the dynamics of adaptive landscapes from phylogenetic comparative data. Syst. Biol. *63*, 902–918.

8. Pagel, M., O'Donovan, C., and Meade, A. (2022). General statistical model shows that macroevolutionary patterns and processes are consistent with Darwinian gradualism. Nat. Commun. *13*, 1113.

9. Slater, G.J., Price, S.A., Santini, F., and Alfaro, M.E. (2010). Diversity versus disparity and the radiation of modern cetaceans. Proc. Biol. Sci. *277*, 3097–3104.

10. Slater, G.J., Harmon, L.J., and Alfaro, M.E. (2012). Integrating fossils with molecular phylogenies improves inference of trait evolution. Evolution *66*, 3931–3944.

11. Hunt, G., and Slater, G. (2016). Integrating paleontological and phylogenetic approaches to macroevolution. Annu. Rev. Ecol. Evol. Syst. *47*, 189–213.

12. Louca, S., and Pennell, M.W. (2020). Extant timetrees are consistent with a myriad of diversification histories. Nature *580*, 502–505.

13. Lloyd, G.T., and Slater, G.J. (2021). A total-group phylogenetic metatree for Cetacea and the importance of fossil data in diversification analyses. Syst. Biol. *70*, 922–939.

14. Slater, G.J., and Pennell, M.W. (2014). Robust regression and posterior predictive simulation increase power to detect early bursts of trait evolution. Syst. Biol. *63*, 293–308.

15. Hansen, T.F. (2012). Adaptive landscapes and macroevolutionary dynamics. In The Adaptive Landscape in Evolutionary Biology, *205* (Oxford University Press), p. 26.

16. LaBarbera, M. (1989). Analyzing body size as a factor in ecology and evolution. Annu. Rev. Ecol. Syst. *20*, 97–117.

17. Calder, W.A. (1996). Size, Function, and Life History (Courier Corporation).

18. Jablonski, D. (1996). Body size and macroevolution. In Evolutionary Paleobiology, D. Jablonski, D. Erwin, and J. Lipps, eds. (University of Chicago Press), pp. 256–289.

19. Uyeda, J.C., Pennell, M.W., Miller, E.T., Maia, R., and McClain, C.R. (2017). The evolution of energetic scaling across the vertebrate tree of life. Am. Nat. *190*, 185–199.

20. Cooper, N., and Purvis, A. (2010). Body size evolution in mammals: complexity in tempo and mode. Am. Nat. *175*, 727–738.

21. Venditti, C., Meade, A., and Pagel, M. (2011). Multiple routes to mammalian diversity. Nature *479*, 393–396.

22. Gearty, W., McClain, C.R., and Payne, J.L. (2018). Energetic tradeoffs control the size distribution of aquatic mammals. Proc. Natl. Acad. Sci. USA *115*, 4194–4199.

23. Biewener, A.A. (1989). Scaling body support in mammals: limb posture and muscle mechanics. Science *245*, 45–48.

24. Ahlborn, B.K., and Blake, R.W. (1999). Lower size limit of aquatic mammals. Am. J. Phys. *67*, 920–922.

25. Downhower, J.F., and Blumer, L.S. (1988). Calculating just how small a whale can be. Nature *335*, 675.

26. Clauset, A. (2013). How large should whales be? PLoS One *8*, e53967.

27. Fish, F.E. (1997). Biological designs for enhanced maneuverability: analysis of marine mammal performance. In Tenth International Symposium on Unmanned Untethered Submersible Technology (Autonomous Undersea Systems Institute), pp. 109–117.

28. Pyenson, N.D., Vélez-Juarbe, J., Gutstein, C.S., Little, H., Vigil, D., and O'Dea, A. (2015). Isthminia panamensis, a new fossil inioid (Mammalia, Cetacea) from the Chagres formation of panama and the evolution of 'river dolphins' in the Americas. PeerJ *3*, e1227.

29. Gillet, A., Frédérich, B., and Parmentier, E. (2019). Divergent evolutionary morphology of the axial skeleton as a potential key innovation in modern cetaceans. Proc. Biol. Sci. *286*, 20191771.

30. Fordyce, R.E. (2018). Cetacean evolution. In Encyclopedia of Marine Mammals (Elsevier), pp. 180–185.

31. Marx, F.G., Lambert, O., and Uhen, M.D. (2016). Cetacean Paleobiology (John Wiley & Sons).

32. Steeman, M.E., Hebsgaard, M.B., Fordyce, R.E., Ho, S.Y.W., Rabosky, D.L., Nielsen, R., Rahbek, C., Glenner, H., Sørensen, M.V., and Willerslev, E. (2009). Radiation of extant cetaceans driven by restructuring of the oceans. Syst. Biol. *58*, 573–585.

33. Bisconti, M., Lambert, O., and Bosselaers, M. (2013). Taxonomic revision of isocetus depauwi (Mammalia, Cetacea, Mysticeti) and the phylogenetic relationships of archaic 'cetothere'mysticetes. Palaeontology *56*, 95–127.

34. Burin, G., Park, T., James, T.D., Slater, G.J., and Cooper, N. (2023). The dynamic adaptive landscape of cetacean body size: submitted code and data. https://doi.org/10.5281/zenodo.7563162.

35. Cooper, N., Thomas, G.H., Venditti, C., Meade, A., and Freckleton, R.P. (2016). A cautionary note on the use of Ornstein Uhlenbeck models in macroevolutionary studies. Biol. J. Linn. Soc. Lond. *118*, 64–77.

36. Bonnet, T., Morrissey, M.B., de Villemereuil, P., Alberts, S.C., Arcese, P., Bailey, L.D., Boutin, S., Brekke, P., Brent, L.J.N., Camenisch, G., et al. (2022). Genetic variance in fitness indicates rapid contemporary adaptive evolution in wild animals. Science *376*, 1012–1016.

37. Uyeda, J.C., Hansen, T.F., Arnold, S.J., and Pienaar, J. (2011). The million-year wait for macroevolutionary bursts. Proc. Natl. Acad. Sci. USA *108*, 15908–15913.

38. Raup, D.M., and Gould, S.J. (1974). Stochastic simulation and evolution of morphology-towards a nomothetic paleontology. Syst. Biol. *23*, 305–322.

39. Felsenstein, J. (1985). Phylogenies and the comparative method. Am. Nat. *125*, 1–15.

40. Revell, L.J., Harmon, L.J., and Collar, D.C. (2008). Phylogenetic signal, evolutionary process, and rate. Syst. Biol. *57*, 591–601.

41. Hansen, T.F. (1997). Stabilizing selection and the comparative analysis of adaptation. Evolution *51*, 1341–1351.

42. Beaulieu, J.M., Jhwueng, D.C., Boettiger, C., and O'Meara, B.C. (2012). Modeling stabilizing selection: expanding the Ornstein-Uhlenbeck model of adaptive evolution. Evolution *66*, 2369–2383.

43. Coombs, E.J., Felice, R.N., Clavel, J., Park, T., Bennion, R.F., Churchill, M., Geisler, J.H., Beatty, B., and Goswami, A. (2022). The tempo of cetacean cranial evolution. Curr. Biol. *32*. 2233–2247.e4.

44. Slater, G.J., Goldbogen, J.A., and Pyenson, N.D. (2017). Independent evolution of baleen whale gigantism linked to Plio-Pleistocene ocean dynamics. Proc. Biol. Sci. *284*, 20170546.

45. Goldbogen, J.A., Cade, D.E., Wisniewska, D.M., Potvin, J., Segre, P.S., Savoca, M.S., Hazen, E.L., Czapanskiy, M.F., Kahane-Rapport, S.R., DeRuiter, S.L., et al. (2019). Why whales are big but not bigger: physiological drivers and ecological limits in the age of ocean giants. Science *366*, 1367–1372.

46. Pyenson, N.D., and Vermeij, G.J. (2016). The rise of ocean giants: maximum body size in Cenozoic marine mammals as an indicator for productivity in the pacific and Atlantic Oceans. Biol. Lett. *12*, 20160186.

47. Mori, Y. (2002). Optimal diving behaviour for foraging in relation to body size. J. Evol. Biol. *15*, 269–276.

48. Boessenecker, R.W., Fraser, D., Churchill, M., and Geisler, J.H. (2017). A toothless dwarf dolphin (Odontoceti: Xenorophidae) points to explosive feeding diversification of modern whales (Neoceti). Proc. Biol. Sci. *284*, 20170531.

49. Slater, G.J. (2022). Topographically distinct adaptive landscapes for teeth, skeletons, and size explain the adaptive radiation of Carnivora (Mammalia). Evolution *76*, 2049–2066.

50. Peters, R.H. (1986). The Ecological Implications of Body Size (Cambridge University Press).

51. Marshall, C.R. (2014). The evolution of morphogenetic fitness landscapes: conceptualising the interplay between the developmental and ecological drivers of morphological innovation. Aust. J. Zool. *62*, 3–17.

52. Bell, M.A. (2012). Adaptive landscapes, evolution, and the fossil record. In The Adaptive Landscape in Evolutionary Biology (Oxford Academic), pp. 243–256.

53. Goolsby, E.W., Bruggeman, J., and Ané, C. (2017). Rphylopars: fast multivariate phylogenetic comparative methods for missing data and within-species variation. Methods Ecol. Evol. *8*, 22–27.

54. Park, T., Burin, G., Slater, G.J., and Cooper, N. (2022). Data from the "Back to the Water" project (Natural History Museum). https://doi.org/10.5519/vmbrpkuq.

55. Pyenson, N.D., and Sponberg, S.N. (2011). Reconstructing body size in extinct crown Cetacea (Neoceti) using allometry, phylogenetic methods and tests from the fossil record. J. Mammal. Evol. *18*, 269–288.

56. The Commitee on Marine Mammals American Society of Mammalogists (1961). Standardized methods for measuring and recording data on the smaller cetaceans. J. Mammal. *42*, 471–476.

57. Swenson, N.G. (2014). Phylogenetic imputation of plant functional trait databases. Ecography *37*, 105–110.

58. Oba, S., Sato, M.A., Takemasa, I., Monden, M., Matsubara, K., and Ishii, S. (2003). A bayesian missing value estimation method for gene expression profile data. Bioinformatics *19*, 2088–2096.

59. Poyatos, R., Sus, O., Badiella, L., Mencuccini, M., and Martínez-Vilalta, J. (2018). Gap-filling a spatially explicit plant trait database: comparing imputation methods and different levels of environmental information. Biogeosciences *15*, 2601–2617.

60. Jardim, L., Bini, L.M., Diniz-Filho, J.A.F., and Villalobos, F. (2021). A cautionary note on phylogenetic signal estimation from imputed databases. Evol. Biol. *48*, 246–258.

61. James, T.D., Salguero-Gómez, R., Jones, O.R., Childs, D.Z., and Beckerman, A.P. (2021). Bridging gaps in demographic analysis with phylogenetic imputation. Conserv. Biol. *35*, 1210–1221.

62. Harmon, L.J., Schulte, J.A., Larson, A., and Losos, J.B. (2003). Tempo and mode of evolutionary radiation in iguanian lizards. Science *301*, 961–964.

63. Pennell, M.W., Eastman, J.M., Slater, G.J., Brown, J.W., Uyeda, J.C., FitzJohn, R.G., Alfaro, M.E., and Harmon, L.J. (2014). geiger v2. 0: an expanded suite of methods for fitting macroevolutionary models to phylogenetic trees. Bioinformatics *30*, 2216–2218.

64. R Core Team (2022). R: a language and environment for statistical computing (R Foundation for Statistical Computing). https://www.R-project.org/.

65. Meade, A., and Pagel, M. (2022). BayesTraits (University of Reading). http://www.evolution.reading.ac.uk/BayesTraitsV4.0.0/BayesTraitsV4.0.0.html.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
Tree for all phylogenetic analyses	https://doi.org/10.5519/vmbrpkuq	https://doi.org/10.5519/vmbrpkuq
	Lloyd and Slater[13]	https://doi.org/10.1093/sysbio/syab002
Software and algorithms		
R v.4.1.2	R Core Team.[64]	N/A

RESOURCE AVAILABILITY

Lead contact
Further information and requests for resources should be directed to and will be fulfilled by the lead contact, Gustavo Burin (gustavo.burin@nhm.ac.uk).

Materials availability
The collated dataset generated in this study has been deposited in the NHM Data Portal.[54]

Data and code availability

- All original code is available at https://doi.org/10.5519/vmbrpkuq Zenodo: https://doi.org/10.5281/zenodo.7563162 and is publicly available as of the date of publication. DOIs are listed in the key resources table. All R package version details available at Github Repository: https://github.com/gburin/bodysize-evolution-cetacea.
- Data are available from the NHM Data Portal[54]
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Total length data
We collated total length (m) for living species primarily from stranding records in the literature (n = 1,327 individuals), and the strandings database at the Natural History Museum London (n = 1,659 individuals). If no other data were available we used data from aboriginal subsistence harvests (n = 80 individuals). We only included values within the range of adult lengths for each species, and, where possible, we recorded a minimum of 10 total length values for each species. This resulted in a total length dataset of 2,986 individuals of living species. We collated data on total lengths of fossils from the literature; some were true total length measurements, others were based on estimates (a total of n = 273). 87 species in our phylogeny (see below), two living and 85 fossil, had no available total length data. To estimate total length for these taxa we collated data for nine proxy measurements[55,56] from 913 adult specimens (451 extinct and 462 living) from the published literature, or by measuring specimens using either Mitutoyo Absolute Digimatic CD-8"AX calipers, large calipers, or tape measures depending on the size of the feature. The proxy measurements were (1) width of antorbital notches, (2) bizygomatic width, (3) exoccipital width, (4) occipital condyle breadth, (5) condylobasal length, (6) maximum narial fossa width, (7) maximum nasal width, (8) atlas articular facet width, and (9) humerus length. We obtained a minimum of two proxy measurements from each specimen, as few fossils possessed all features. Prior to the analyses we removed species with uncertain taxonomy and those not present in the phylogeny.

Phylogeny
We used the "Safe" analysis phylogeny from Lloyd and Slater[13] and removed all taxa without any total length proxy measurements from the phylogeny leaving 345 taxa. 60 of these remaining taxa were singletons, forming terminals with branch lengths of zero. Because our analyses required branch lengths, we added 0.001 MY to all branches to remove the zero-length branches. To ensure this did not substantially influence our results we repeated all downstream analyses (and sensitivity analyses; see supplemental information) removing the species with zero length branches from the phylogeny (n = 285 taxa; see supplemental information; Burin et al.[34]).

METHOD DETAILS

Total length data imputation

To impute total lengths for the 87 species without available total length data, we used a phylogenetically-informed imputation approach. These methods use the eigenvectors of the phylogenetic distance matrix to estimate trait values for species without the focal variable[53,57] while accounting for intraspecific variation.[53,57] We used the phylogenetic imputation algorithm in the R package *Rphylopars*,[53,57] using a Brownian Motion model for trait evolution, and including specimens with at least one proxy measurement to incorporate intraspecific variation. All measurements were log_{10}-transformed prior to the imputation. The final total length dataset contained 345 taxa; 89 living and 256 fossil.

To assess the quality of total length estimates we generated datasets with different percentages of missing total length data for species with available empirical total length data as follows. We randomly removed data for 10% (n = 25), 20% (n = 51) and 50% (n = 129) of all species with existing total length data, leaving 233, 207 and 129 species respectively. Note that we removed total length data for all specimens belonging to these species, so in some cases the percentage of missing total length data *records* is larger than the percentage of missing *species*. We also generated a dataset with empirical total length data from extant species only (n = 89). We then assessed the accuracy of the approach by plotting the imputed values against the empirical average values for the species whose total length data was removed from each dataset (Figure S1). Imputed total length data were strongly correlated with the empirical values, regardless of the amount of information used for the estimation of missing values.

Finally, we calculated the Normalized Root Mean Squared Error for the main dataset (NRMSE[58]), which represents the mean squared errors for the imputed values standardized by the range of empirical values (Equation 1; see also Table S1).

$$NRMSE = \frac{\sqrt{\frac{\sum (x_i^* - x_i)^2}{n}}}{max(X) - min(X)} \qquad \text{(Equation 1)}$$

Where X represents the variable "total length", X_i^* is the imputed value of total length for species i, X_i is the empirical value of total length for species i, and n is the number of taxa. Mean squared errors were calculated for estimates using all possible combinations of proxy measurements used for the imputation (ranging from using just one measurement up to all nine). NRMSE values for all four datasets were low (Table S1), and within acceptable values found in other papers.[59–61]

QUANTIFICATION AND STATISTICAL ANALYSIS

Exploring tempo and mode in body size evolution

We performed disparity through time analyses[62] using the dtt function in the R package geiger,[63] with 10,000 simulated datasets used to generate a null, constant-rates distribution. The centre of gravity was computed as the weighted sum of the product of standardised average subclade disparity and relative node age, with significance determined as the proportion of simulated centres of gravity that were lower than the empirical value.[49]

We used *bayou*[7] to detect evolutionary shifts in total length evolution across cetaceans. *bayou* uses reversible-jump MCMC algorithms to fit multi-peak OU models of trait evolution, identifying the number, magnitude and location of shifts in θ, the evolutionary regime mean of total length. We used a cumulative Poisson distribution as the prior distribution for the number of shifts, setting the prior on the expected number of shifts (λ) as 2.5% of the total number of branches in the corresponding tree, i.e. $\lambda = 15$. We allowed an unlimited number of shifts to occur on each branch, while setting the prior probability for each branch to contain a shift to be proportional to the branch's length. For all other parameters, we used the recommended default distributions found in (https://github.com/uyedaj/bayou/blob/master/tutorial.md). We ran the MCMC chains for 1 million generations, sampling every 1,000 generations, yielding a posterior sample of 700 after discarding the first 30% as burn-in. We assessed the convergence of each chain using Gelman and Rubin's R statistic, and confirmed that all parameters had an effective sample size greater than 200.

To assess the relative importance of the shifts we followed Uyeda and Harmon's[7] criteria: we selected shifts that had posterior probabilities greater than 0.1 (corresponding to approximately an eightfold increase in the posterior probability compared to the prior probability), with only three of these shifts leading to singleton taxa. For the selected shifts, we calculated the posterior-to-prior ratio as a measure of how much information our data is providing on the location and magnitude of changes in evolutionary regime. All bayou analyses were performed in R.[64]

Sensitivity analyses

To test the sensitivity of our results to both tree topology and the prior parameter for expected number of shifts (λ), we replicated the *bayou* protocol described above for the original tree (*Full* tree) and six additional trees, with five different values of λ.

The six trees were as follows:

1. *Full*. The most comprehensive phylogeny possible that included all species with total length data (n = 345). This is the the "Safe" analysis phylogeny from Lloyd and Slater[13,14] with species without total length data (empirical or imputed) removed, and is the tree used for our main analyses.
2. *No Archaeoceti*. The Full tree after dropping all species that are classed as archaeocetes (n = 322).

3. ***No Extant***. The Full tree for fossil species only (n = 256).
4. ***No Imput***. The *Full* tree after dropping all species that lacked direct total length data (n = 258).
5. ***Extant***. The *Full* tree for living species only (n = 89). Note that this tree does not contain any taxa with zero-length branches (see below).
6. ***Baleen***. The *Full* tree for baleen whales (Mysticeti) only (n = 105).
7. ***Toothed***. The *Full* tree for toothed whales (Odontoceti) only (n = 217).

For each of these seven phylogenies in turn, we replicated the same *bayou* analysis pipeline described above using five different values for the prior parameter on the expected number of shifts (λ), namely 5, 15, 50, 250 and 500 (the last two values were not used for the *Extant*, *Baleen* or *Toothed* trees because there were too few branches in those trees). Note that the prior on the expected number of shifts for the main analyses was set as 2.5% of the total number of branches in the corresponding tree, i.e. λ = 15. The results were evaluated the same way for all analyses; we considered only the shifts with posterior probability greater than 0.1 and also assessed the posterior-to-prior ratio for the selected shifts.

Removing taxa with zero-length branches

As described above, 60 of the taxa in the *Full* tree were singletons, forming terminals with branch lengths of zero. Because our analyses required branch lengths, we added 0.001 MY to all branches to remove the zero-length branches, and used these branch lengths in our main analyses. To ensure this did not substantially influence our results we repeated all our analyses removing the species with zero length branches from the phylogeny, leaving n = 285 taxa in the *Full* tree. We repeated this for the other four trees containing singletons: *No Archaeoceti* (n = 267), *No Extant* (n = 207), *No Imput* (n = 216), *Baleen* (n = 83), *Toothed* (n = 184). Note that the *Extant* tree does not contain any taxa with zero-length branches.

Overall, our results using different trees and values of λ were congruent with the main results, as were analyses removing taxa with zero length branches.[34]

Evaluating per-branch trends in body size evolution

We used the Fabric model[8] as implemented in BayesTraits v4.0[65] to evaluate the presence, location, and strength of linear trends in cetacean total length evolution. This model assumes that evolution follows a random walk in which the variance parameter (= evolutionary rate) can vary across branches of the phylogeny. However, it also uses a reversible jump MCMC to sample trend parameters along branches that result in biases towards larger (+ve) or smaller (-ve) trait values, accepting these if they increase the probability of the model.

We ran three independent chains with 5 million steps of burn-in and 20 million steps of sampling. After visually checking that all three chains converged on the same target distribution and that effective sample sizes were larger than 200, we used the Fabric post-processing software (included in the suite of Meade and Pagel[65]) to summarise the output of each chain, ultimately retaining all branch trends that were significant in all three chains. The processed output file can be downloaded from the NHM Data Portal.[54]

Supplemental Information

The dynamic adaptive landscape

of cetacean body size

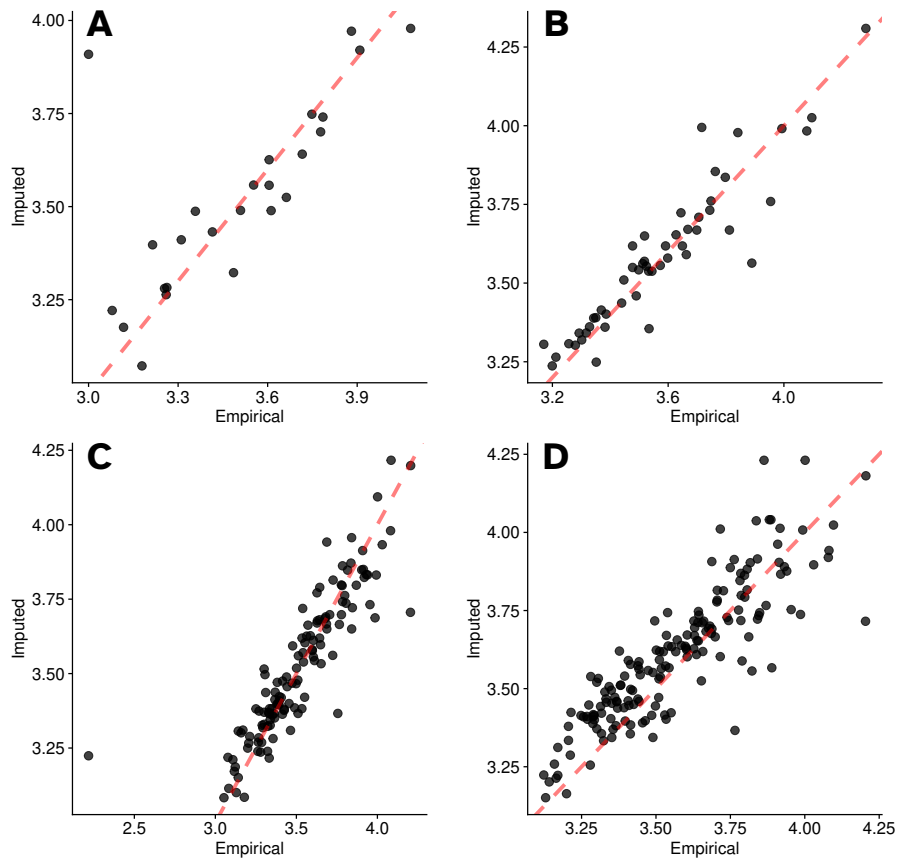Gustavo Burin, Travis Park, Tamora D. James, Graham J. Slater, and Natalie Cooper

Figure S1: **Quality assessment of imputation procedure. Related to Figure 1**. Relationship between imputed and empirical average total length values after randomly dropping: (A) 10% of species with total length data; (B) 20% of species with total length data; (C) 50% of species with total length data; and (D) all fossil species with total length data. Each point represents a species with total length data removed prior to the imputation (n = 25). The red dotted line is the 1:1 line.
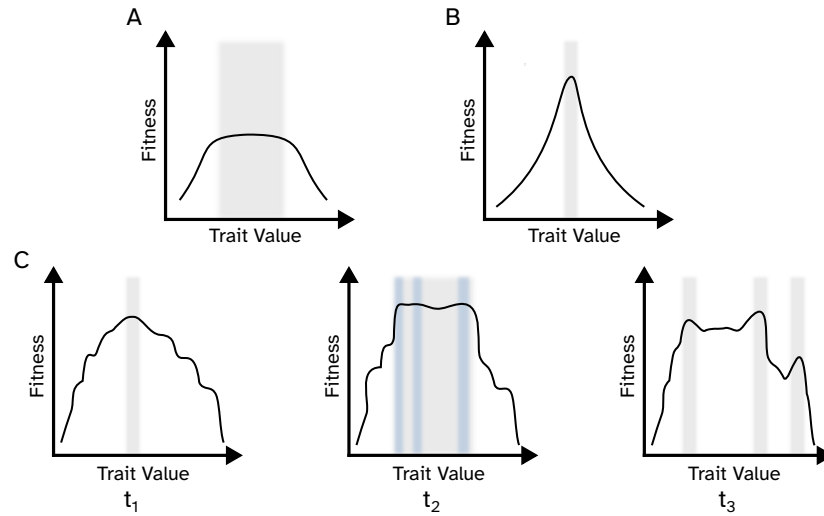
Figure S2: **Conceptual representation of adaptive landscapes. Related to Figures 3 and 4.** Conceptual representations of the relationship between a given one-dimensional trait and its fitness values in a (A) flat and (B) peaky adaptive landscape. (C): Temporal snapshots of a hypothetical adaptive landscape showing the changes in its peaks through time. In all panels the grey shaded areas indicate trait values that provide high fitness, and in the central panel the blue shaded areas indicate regions within high-fitness trait values that represent local optima.
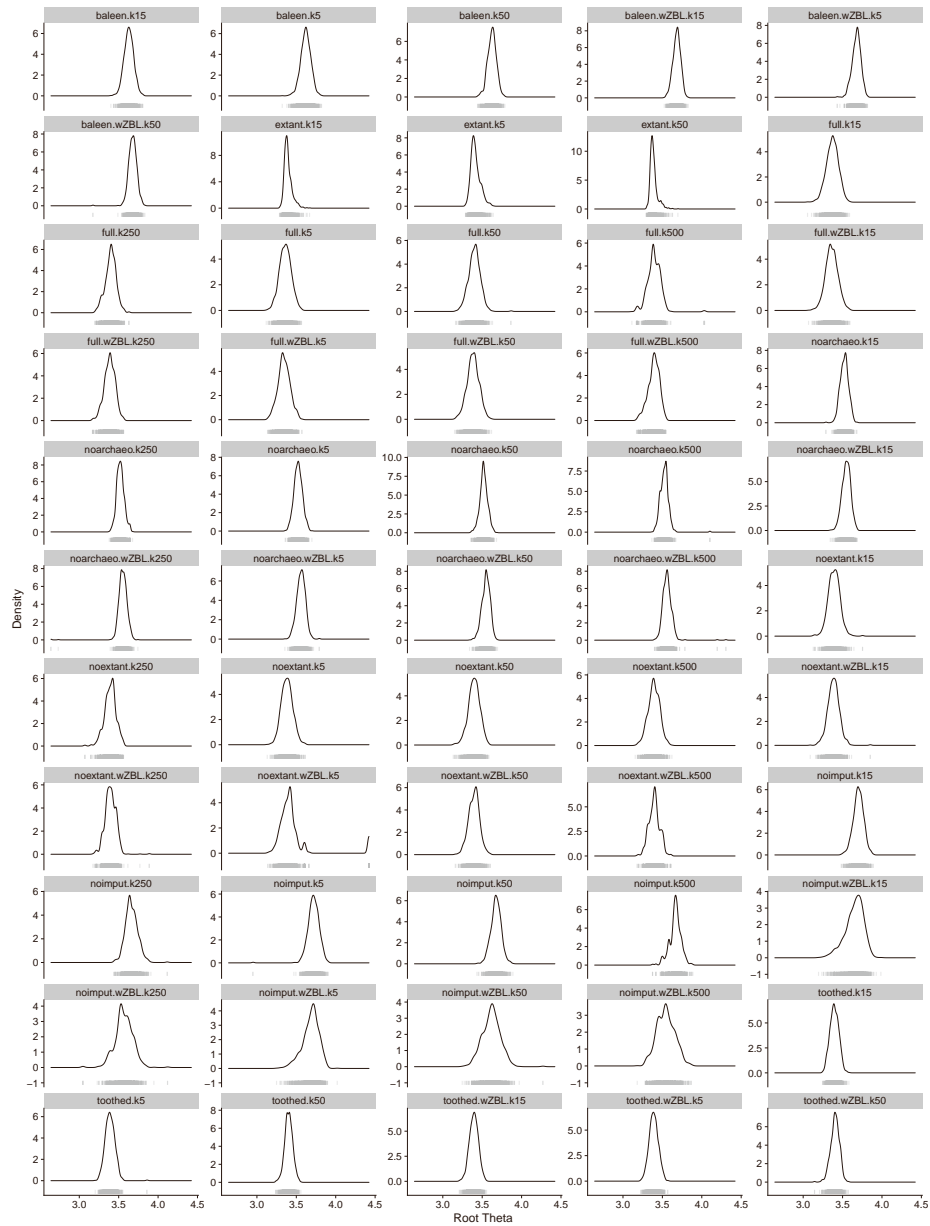
Figure S3: **Rooth theta estimated by *bayou*. Related to Figure 3.** Posterior distribution of estimated θ values at the root for each scenario of the bayou analyses. In each panel, the grey dashes below the x axis represent the raw values of the posterior, that are summarised as density curves in the plots.
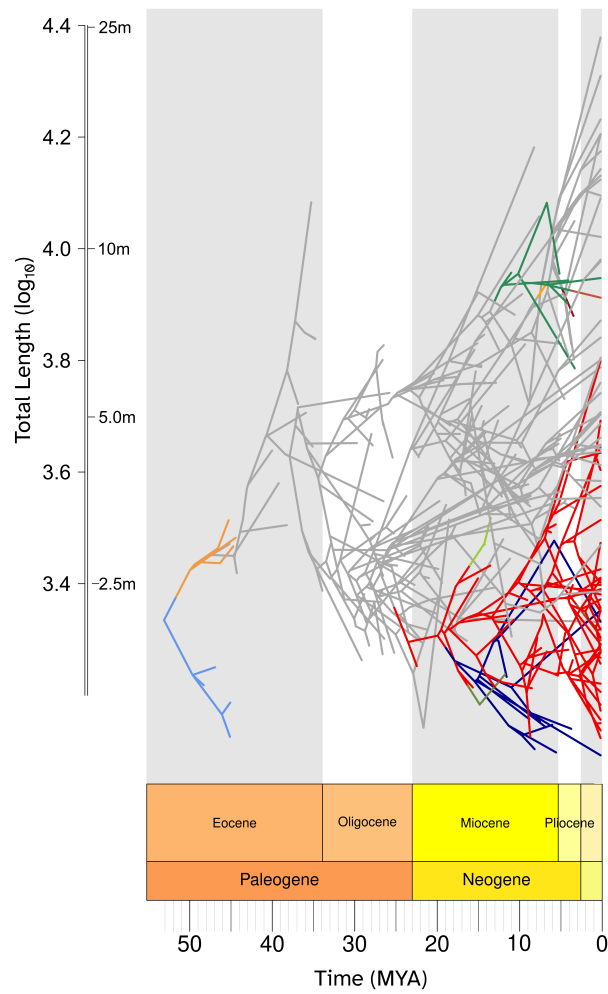
Figure S4: **Phenogram of total length evolution in cetaceans. Related to Figure 3**. The tips, branches and nodes of the phylogeny are re-positioned according to the trait value, and are coloured according to the regime shifts in Figure 2). This figure shows that there is little overlap in total length between the major regimes, and that the trait space is homogeneously occupied, which highlights the flat adaptive landscape found for cetaceans.

Removed variables	1	2	3	4	5	6	7	8
NRMSE	0.087	0.176	0.259	0.321	0.322	0.259	0.167	0.082

Table S1: **Normalized Root Mean Square Errors (NRMSE) values calculated for each manipulated dataset. Related to STAR methods.** Each column represents the NRMSE value calculated over all possible combinations of measurements after removing a given number of measurements prior to the imputation. Mean $\pm$ standard deviation of combined NRMSE values $= 0.2091 \pm 0.0959$.