Supplemental Information

Giant baleen whales emerged from a cold southern cradle

Rule, J.P., Duncan, R.J., Marx, F.G., Pollock, T.I., Evans, A.R., Fitzgerald, E.M.G. 2023. Giant baleen whales emerged from a cold southern cradle. Proceedings of the Royal Society B: Biological Sciences. 10.1098/rspb.2023.2177

Description of NMV P218462

Description refers to both sides unless otherwise stated. Preserved remain include the symphyseal regions of both mandibles, as well as fragments of the left premaxilla and maxilla. The premaxilla and maxilla fragments do not provide any systematic information, and do not preserve the alveolar row, and hence are not described. There is minimal dorsoventral and transverse crushing or shearing, but the anterior margin of the mandible is slightly abraded on the left and notably damaged on the right. The dorsal and ventral edges of the mandible are subparallel, with the anterior edge appearing squared-off with rounded corners. In medial view, the symphyseal groove is distinct but shallow and curves in a gentle arc for about 150 mm posteroventrally.

The transverse thickness of the mandible decreases both ventrally and towards the apex, resulting in an inverted pyriform cross section that becomes asymmetrical posteriorly. In posterior view, the mandibular canal is circular and located dorsolaterally, with a diameter of 11 mm on the left and 8 mm on the right); on the right side, two smaller additional canals occur dorsal to the mandibular canal.

In dorsal view there is a shallow alveolar groove with relictual alveolar foramina (six on the left and three on the right) that are confluent with anteriorly directed sulci. There is a possible shark feeding trace preserved as a shallow trough on the dorsomedial edge of the right mandible, roughly halfway along the preserved surface. It is 14 mm at its shortest edge, and 23 mm at its longest edge. The shortest edge is parallel to the dorsal margin of the mandible, with the longest edge located roughly 60° to the shortest edge.

Extended Methods

Body size of NMV P218462

To estimate the total body length of NMV P218462 we generated predictive regression equations based on mandibular measurements from 20 extant baleen whale specimens, 15 of which had known total body lengths. Our data encompass 7 species (*Balaenoptera acutorostrata*, *B. edeni*, *B. musculus*, *B. physalus*, *Balaenoptera* sp., *Caperea marginata*, *Eubalaena australis*, *Megaptera novaeangliae*,) housed at the following institutions: NMV, Museums Victoria, Melbourne, Australia; NMNZ, Museum of New Zealand Te Papa Tongarewa, Wellington, New Zealand; CNPMAMM: Laboratorio de Mamiferos Marinos, Puerto Madryn, Argentina). Measurements were taken in mm and, where possible, averaged across both sides.

For the extant mysticete mandibles, we took chord (Ch) and curvilinear length (Cu) following [1], as well as three novel measurements from the anterior end of the mandible (Figure 1f): anterior end of the mandible to anterior extent of the symphyseal groove (A-Sg);

dorsoventral height of the mandible (D-V); lateral-medial width of the mandible (L-M). The first D-V and L-M measurements were taken using the anterior end of the symphyseal groove as a landmark. We then established two landmarks posterior to this (double and triple the A-Sg length, respectively) and took a second and third D-V and L-M measurement at these landmarks (Figure 1f). We regressed the six D-V and L-M measurements against each other to establish that the width and height of the mandible had a uniform proportion. This validated calculating a geometric mean of 7 measurements (A-Sg1, D-V1, L-M1, D-V2, L-M2, D-V3, L-M3) to capture the overall size of the anterior end of the mandible. We then log10 transformed all measurements prior to calculating regressions.

We conducted bivariate linear regressions in R (version 4.2.1) for the seven measurements and the geometric mean against chord and curvilinear length. In addition, we regressed the seven measurements, the geometric mean, chord length, and curvilinear length against total body length. This resulted in 16 regressions indirectly predicting total body length (via chord or curvilinear length) and eight regressions predicting total body length directly. We used ggplot2 to plot the regressions and map their respective 95% confidence intervals.

Regressions were assessed using the adjusted R^2 value, P-value, and residual standard error. The slope and intercept of the regression lines were used for predictive equations, the power of which was assessed via percent prediction error (%PE) and percent standard error of the estimate (%SEE). We then selected the best prediction for NMV P218462 based on the above five metrics, with preference given to direct predictive methods.

Comparative dataset

We compiled a composite phylogeny of crown and stem mysticetes using Mesquite and the R-package phytools [2]: the first tree was sourced from [3] (modified from [4]) and included crown mysticetes, *Aglaocetus moreni* and *Mauicetus parki*. Owing to the small sample of stem mysticetes in this source tree, we additionally used the tree of [5] for stem mysticetes and added *Llanocetus* sp. [6] as ancestral (zero-branch length) to *Llanocetus denticrenatus.* NMV P218462 could not be included in the phylogeny because of its fragmentary nature, and so is missing from all phylogenetic comparative analyses. All zero-length branches were assigned a branch length of 0.1 million years to allow analyses to run.

For our dataset of mysticete body lengths we updated the data of [4] with the balaenid estimates of [7] and added length estimates for any missing taxa missing from the literature [3, 8-19]. For *Chonecetus tomitai* [20] we estimated total body length following [21]. All data were converted to meters. Prior to analysis, we either log10-transformed our data or scaled them to generation time (total body length$^{3/4}$). The latter results in constant evolutionary rates within mammal clades [22] and hence yields a baseline rate that any genuine shift would need to deviate from.

Body size and biogeographical analyses

We generated scatterplots of raw total body length and total body length scaled to generation time for mysticetes using the package ggplot2, with taxa categorised by both feeding morphology (toothed or edentulous) and geographic region (global, Northern Hemisphere, Southern Hemisphere). We then fitted separate trend lines using “geom_smooth” (stats = loess) and their associated confidence intervals to the non-transformed Northern and

Southern Hemisphere data to compare patterns of body size distribution. Because our dataset is dominated by Northern Hemisphere fossils, we assessed the degree of inter-hemispheric fossil sampling bias by downloading occurrence data for Mysticeti from the Paleobiology Database (accessed 31st May 2022). The latter were then grouped by hemisphere and plotted as a pie chart.

We also explored the global distribution of extinct mysticetes through time via a phylogenetically informed biogeographic analysis, performed in R studio using the packages geiger [23] and phytools [2]. After categorising species in the phylogeny as global, Northern Hemisphere or Southern Hemisphere, we performed a Maximum Likelihood ancestral state estimation using the *fitDiscrete* function to test three models (ER: Equal Rates; SYM: symmetrical; ARD: all-rates-different). AIC and log-likelihood values were used to select the best-supported model, the results of which were further visualised as a lineage-through-time plot.

We investigated the deep-time evolution of mysticete body size via a Maximum Likelihood ancestral state estimation based on the body size categories of [7], as well as separate ancestral state estimation and evolutionary rate shift analyses of the log-transformed and scaled data using *RRphylo search.shift* (in *auto.recognize* mode) and *overfitRR* (100 ancestral state estimation regression simulations) functions of RRphylo [24-26].

Table S1. Measurements for left and right mandible tips of NMV P218462. Geometric mean is calculated from A-Sg1, D-V1, L-M1, D-V2, L-M2, D-V3, L-M3. Asterisk (*) indicates incomplete measurement due to preservation.

Measurement	**Left (mm)**	**Right (mm)**
Preserved greatest length (anterior-posterior)	315*	150*
Preserved greatest height (dorsal-ventral)	170	172
Preserved greatest width (medial-lateral)	65	52
Anterior to symphyseal groove (A-Sg1)	52.23	48.72
Dorsal-ventral (D-V1)	157.1	162
Lateral-Medial (L-M1)	47.2	48
Dorsal-ventral (D-V2)	170	170.2
Lateral-Medial (L-M2)	54.2	56.4
Dorsal-ventral (D-V3)	165.3	-
Lateral-Medial (L-M3)	54	-
Geometric Mean	84.92	-

Table S2. Average measurements for extant mysticete specimens for regression equations. Measurements $\log_{10}$ transformed and in mm. Ch = Chord length; Cu = Curvilinear length; A-Sg = Anterior end of the mandible to anterior extent of the symphyseal groove; DV = Dorso-ventral height of the mandible; LM = Lateral-medial width of the mandible; TL = Total Length; GM = Geometric mean.

Taxa	Specimen	**Ch**	**Cu**	**ASG1**	**DV1**	**LM1**	**DV2**	**LM2**	**DV3**	**LM3**	**TL**	**GM**
Balaenoptera acutorostrata	NMV C23488	**2.88**	**2.89**	**1.52**	**1.67**	**1.31**	**1.71**	**1.39**	**1.73**	**1.39**	**3.53**	**1.53**
Balaenoptera acutorostrata	NMV C24935	**3.02**	**3.04**	**1.59**	**1.75**	**1.50**	**1.79**	**1.53**	**1.79**	**1.54**	**3.66**	**1.64**
Balaenoptera acutorostrata	NMV C24968	**2.91**	**2.92**	**1.62**	**1.67**	**1.41**	**1.71**	**1.45**	**1.72**	**1.51**	**3.55**	**1.59**
Balaenoptera acutorostrata	NMV C37480	**2.93**	**2.94**	**1.41**	**1.64**	**1.39**	**1.68**	**1.42**	**1.69**	**1.43**	**NA**	**1.52**
Balaenoptera acutorostrata	NMV C24936	**2.69**	**2.70**	**1.28**	**1.45**	**1.20**	**1.50**	**1.23**	**1.51**	**1.28**	**NA**	**1.35**
Balaenoptera edeni	NMV C8441	**3.10**	**3.11**	**1.66**	**1.80**	**1.57**	**1.84**	**1.64**	**1.85**	**1.64**	**3.81**	**1.72**
Balaenoptera edeni	NMV C28703	**3.48**	**3.49**	**1.96**	**2.10**	**1.87**	**2.13**	**1.91**	**2.17**	**1.93**	**4.15**	**2.01**
Balaenoptera edeni	NMV C28758	**3.22**	**3.23**	**1.72**	**1.81**	**1.65**	**1.87**	**1.71**	**1.89**	**1.72**	**3.90**	**1.77**
Balaenoptera edeni	NMNZ MM002240	**3.41**	**3.42**	**1.90**	**2.13**	**1.80**	**2.14**	**1.85**	**2.15**	**1.88**	**3.95**	**1.98**
Balaenoptera musculus	NMV C23571	**3.68**	**3.69**	**1.90**	**2.36**	**2.04**	**2.42**	**2.06**	**2.43**	**2.09**	**4.27**	**2.19**
Balaenoptera musculus	NMV C38098	**3.65**	**3.68**	**1.95**	**2.36**	**2.00**	**2.43**	**2.03**	**2.44**	**2.09**	**4.28**	**2.19**
Balaenoptera physalus	NMV C27802	**3.53**	**3.55**	**2.05**	**2.26**	**1.97**	**2.25**	**2.02**	**2.25**	**2.04**	**4.18**	**2.12**
Balaenoptera sp	NMV C35964	**3.48**	**3.51**	**1.71**	**2.21**	**1.88**	**2.27**	**1.91**	**2.27**	**1.92**	**NA**	**2.02**
Caperea marginata	NMNZ MM002235	**2.82**	**2.83**	**1.41**	**1.58**	**1.11**	**1.68**	**1.27**	**1.72**	**1.20**	**3.51**	**1.42**
Caperea marginata	NMV C37350	**3.07**	**3.09**	**1.60**	**1.86**	**1.46**	**1.99**	**1.51**	**2.02**	**1.53**	**3.76**	**1.71**
Megaptera novaeangliae	NMNZ MM003068	**3.55**	**3.57**	**2.14**	**2.25**	**2.05**	**2.28**	**2.11**	**2.29**	**2.10**	**4.12**	**2.18**
Megaptera novaeangliae	CNPMAMM-774	**3.38**	**3.40**	**1.98**	**2.17**	**1.98**	**2.22**	**2.05**	**2.23**	**2.06**	**4.08**	**2.10**
Eubalaena australis	CNPMAMM-747	**3.55**	**3.57**	**2.00**	**2.05**	**1.83**	**2.22**	**1.85**	**2.32**	**1.87**	**NA**	**2.02**
Eubalaena australis	NMNZ MM002239	**3.09**	**3.10**	**1.70**	**1.82**	**1.68**	**1.89**	**1.66**	**1.88**	**1.64**	**NA**	**1.75**
Eubalaena australis	NMV C28663	**3.33**	**3.36**	**1.78**	**1.93**	**1.80**	**2.09**	**1.87**	**2.15**	**1.90**	**3.96**	**1.93**

Table S3. Predictive regression equations for measurements from the anterior end of the mandible for chord, curvilinear, and total length.

Equation	Regression equation	Adjusted R2	p-value	%PE	%SEE
Ch~Asg	y=1.166x+1.203	0.83	<0.001	21.97	34.08
Ch~DV1	y=1.07x+1.158	0.94	<0.001	13.33	19.47
Ch~LM1	y=1.003x+1.557	0.92	<0.001	15.28	22.22
Ch~DV2	y=1.081x+1.069	0.95	<0.001	11.26	16.69
Ch~LM2	y=1.004x+1.438	0.92	<0.001	14.34	21.38
Ch~DV3	y=1.067x+1.077	0.95	<0.001	11.19	17.05
Ch~LM3	y=1.02x+1.465	0.92	<0.001	14.45	21.61
Ch~GM	y=1.111x+1.197	0.96	<0.001	9.7	14.29
Cu~Asg	y=1.179x+1.197	0.82	<0.001	22.76	35.06
Cu~DV1	y=1.085x+1.145	0.94	<0.001	13.37	19.61
Cu~LM1	y=1.016x+1.552	0.92	<0.001	15.44	22.8
Cu~DV2	y=1.098x+1.051	0.96	<0.001	11.13	16.14
Cu~LM2	y=1.057x+1.432	0.92	<0.001	14.84	21.92
Cu~DV3	y=1.084x+1.059	0.96	<0.001	10.53	16.32
Cu~LM3	y=1.032x+1.459	0.92	<0.001	14.9	22.09
Cu~GM	y=1.126x+1.186	0.96	<0.001	9.76	14.44
TL~Ch	y=0.9237x+0.8933	0.97	<0.001	7.83	10.54
TL~Cu	y=0.9131x+0.9139	0.97	<0.001	8.04	10.76
TL~Asg	y=1.125x+1.903	0.8	<0.001	21.06	32.27
TL~DV1	y=0.9701x+1.9912	0.91	<0.001	13.99	20.11
TL~LM1	y=0.8734x+2.4256	0.92	<0.001	15.22	19.62
TL~DV2	y=1.002x+1.872	0.92	<0.001	13.83	19.65
TL~LM2	y=0.9263x+2.2822	0.91	<0.001	15.19	19.96
TL~DV3	y=1.009x+1.838	0.91	<0.001	14.79	20.63
TL~LM3	y=0.8886x+2.3351	0.91	<0.001	15.49	20.01
TL~GM	y=1.007x+2.029	0.94	<0.001	12.11	16.09

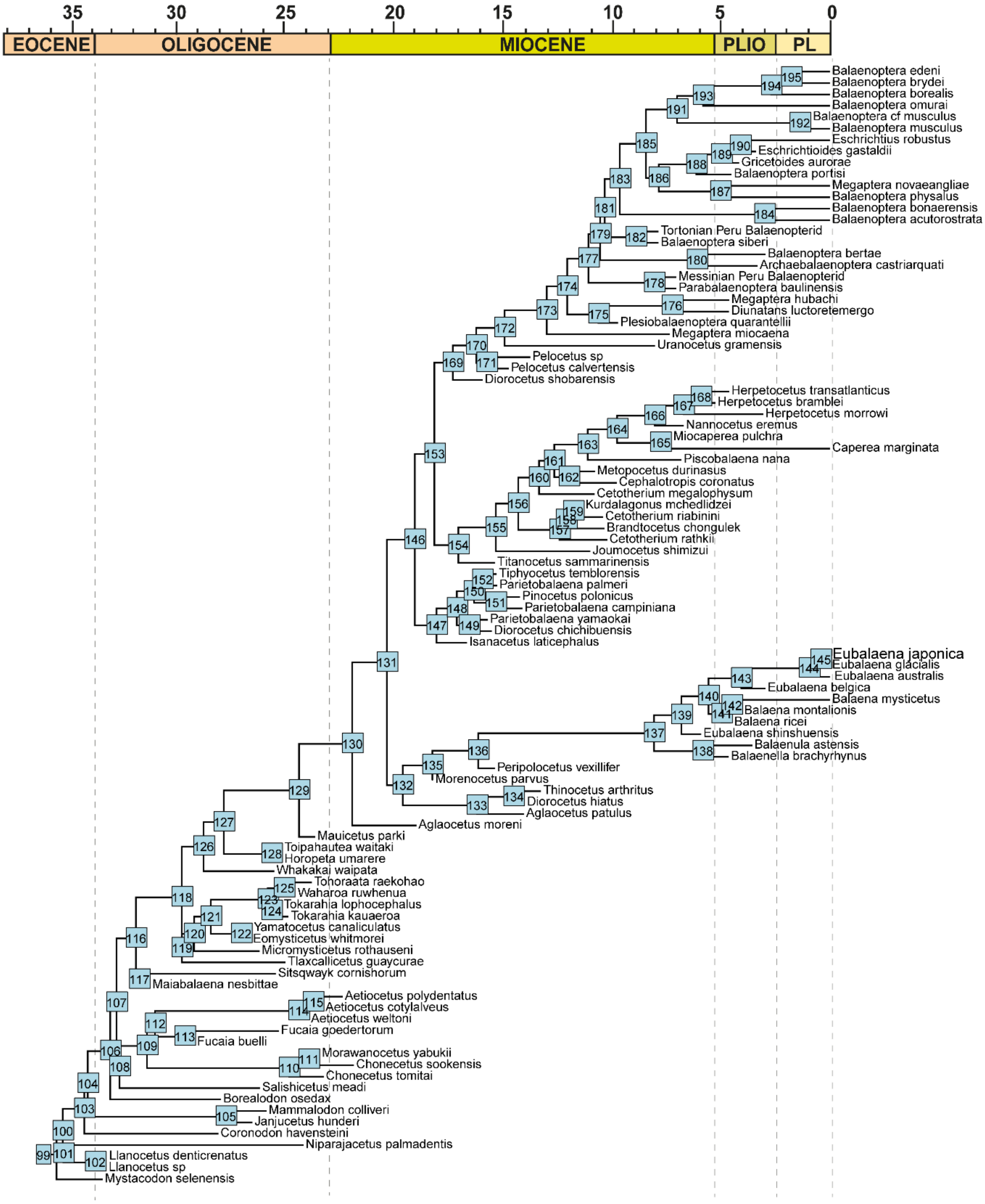

Figure S1. Phylogenetic tree of Mysticeti for evolutionary analyses. Crown Mysticeti from Bianucci et al. 2019 (modified from Slater et al. 2017); stem-Mysticeti from Lloyd and Slater 2021.

Table S4. Evaluation values of categorical Maximum Likelihood ancestral state estimation models.

Trait	Model	Log-likelihood	AIC	AICc	free parameters
Biogeography	ER	-68.76	139.34	139.38	1
	SYM	-63.89	133.78	134.03	3
	ARD	-60.66	133.32	134.24	6
Body size	ER	-72.04	146.08	146.12	1
	SYM	-62.01	136.01	136.93	6
	ARD	-57.07	138.13	141.8	12

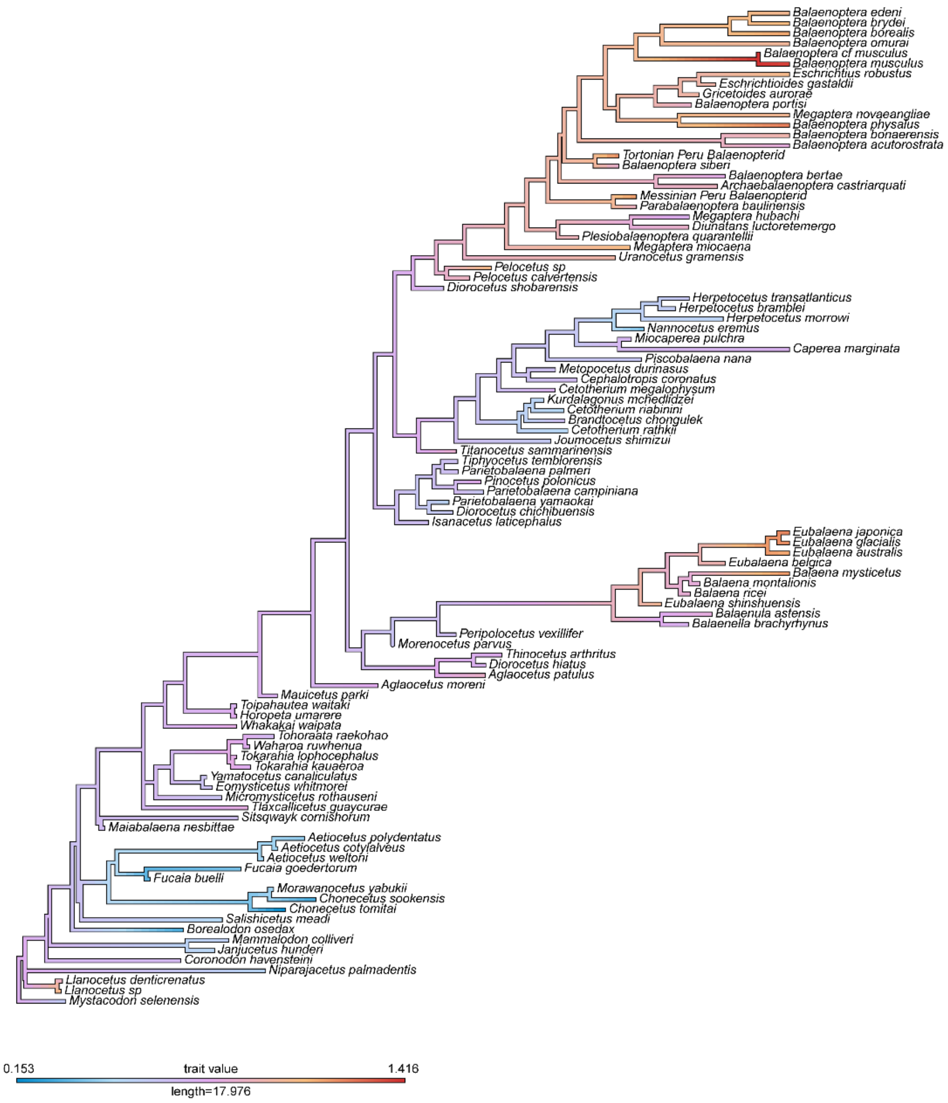

Figure S2. RRphylo maximum likelihood ancestral state estimation for Log_{10} total body length for Mysticeti.

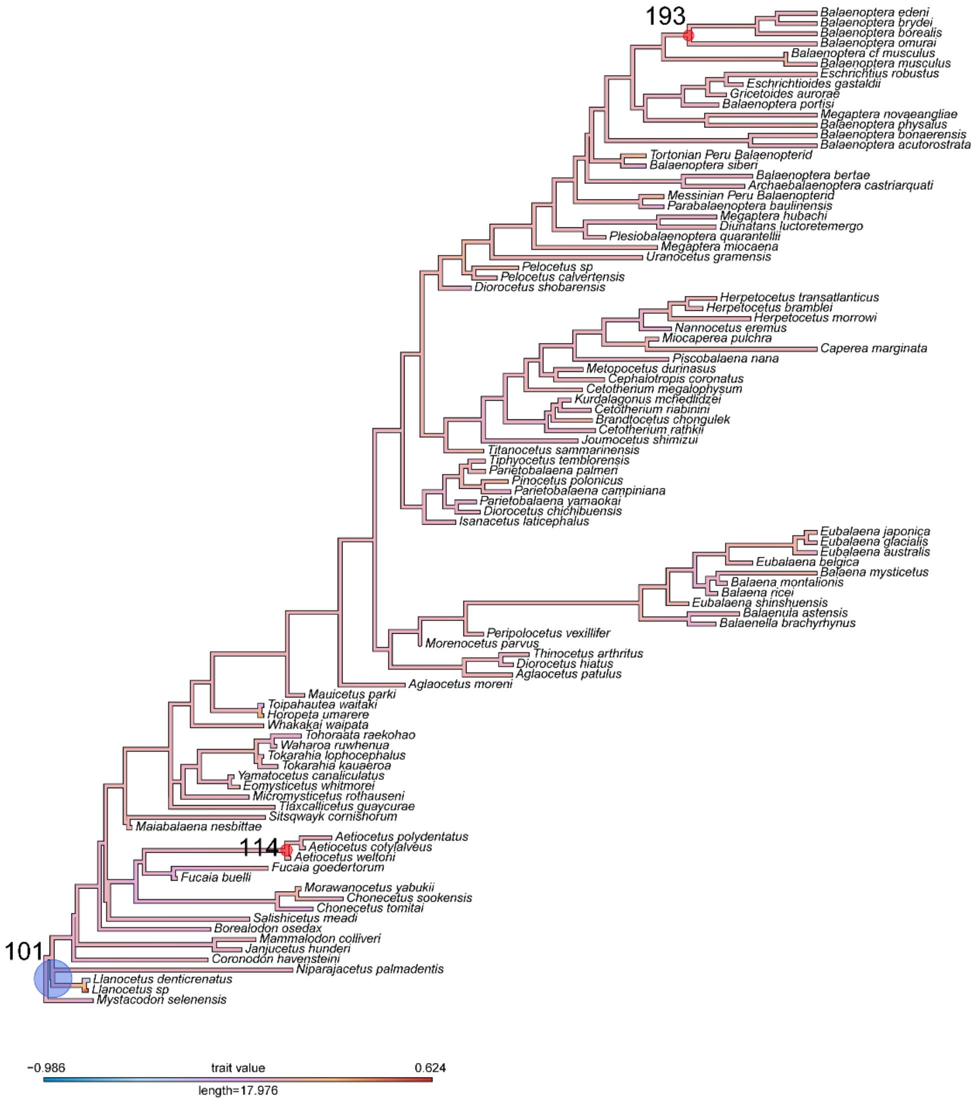

Figure S3. RRphylo evolutionary rates for Log$_{10}$ total body length for Mysticeti. Rate shifts from the function search.shift indicated by circles, where red indicates a decrease in evolutionary rate and blue indicates an increase in evolutionary rate. Numbers indicate clade name.

Table S5. Results of the search.shift function of RRphylo. Results for clades with a significant (P = <0.05, indicated in bold) shift in at least one of the datasets analysed are shown.

	Log10		**GT**	
Clade	**rate.difference**	**p.value**	**rate.difference**	**p.value**
109	-0.007018244	0.542	-0.21159119	**0.027**
112	-0.024321284	0.144	-0.250450449	**0.013**
114	-0.038855857	**0.007**	-0.315423592	**0.001**
153	-0.026431275	**0.001**	-0.157611475	**0.03**
154	-0.021068562	**0.061**	-0.237587854	**0.001**
155	-0.023549356	**0.021**	-0.254027454	**0.001**
156	-0.023053964	**0.03**	-0.252785898	**0.001**
160	-0.023454707	0.056	-0.250859539	**0.001**
161	-0.022772044	0.08	-0.24880167	**0.002**
163	-0.017261494	0.277	-0.223900189	**0.031**
167	-0.028317854	0.163	-0.26956564	**0.044**
169	-0.017713883	**0.047**	-0.030119638	0.432
170	-0.020367873	**0.029**	-0.046128693	0.384
172	-0.021465016	**0.028**	-0.048874796	0.41
173	-0.020746389	**0.022**	-0.040887003	0.41
174	-0.019743787	**0.043**	-0.030913262	0.467
179	-0.023038078	**0.01**	-0.069578853	0.389
181	-0.021000145	**0.044**	-0.046461152	0.458
183	-0.030916341	**0.001**	-0.153219416	0.1
185	-0.030091706	**0.002**	-0.139861131	0.193
189	-0.034833346	**0.041**	-0.237343047	0.109
191	-0.028485709	**0.028**	-0.10758478	0.444
193	-0.032701247	**0.021**	-0.199511237	0.165

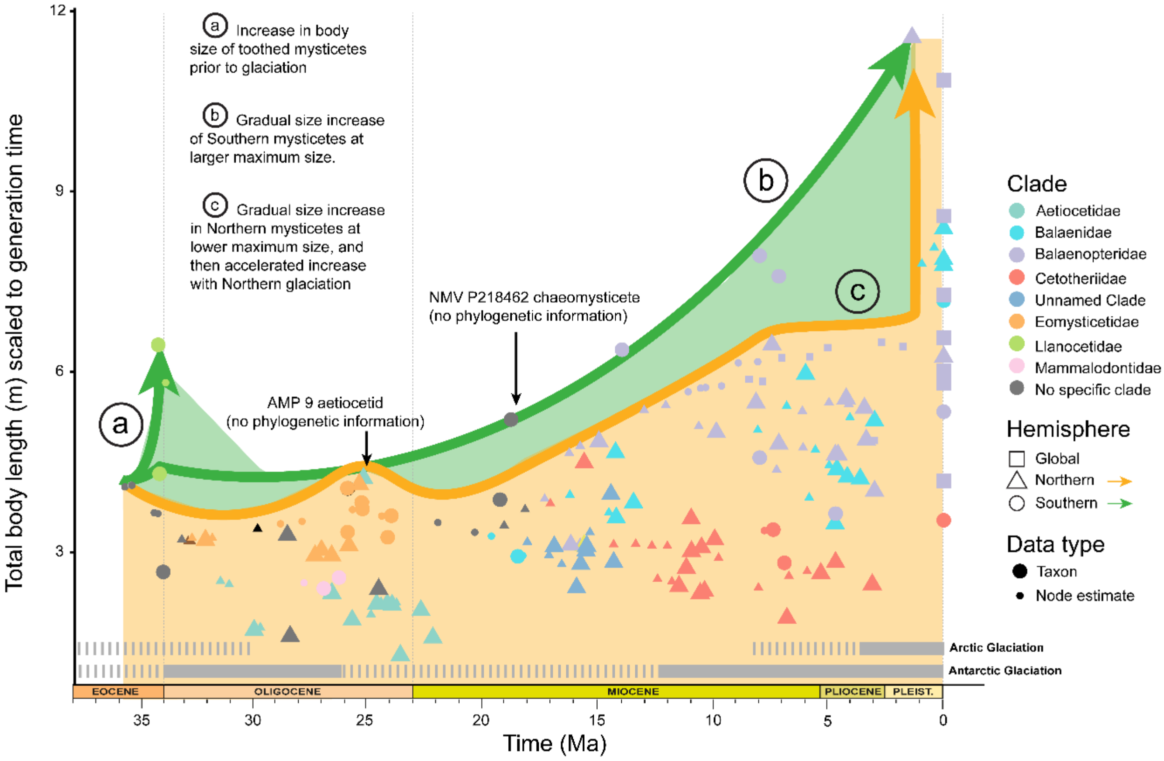

Figure S4. Body size of mysticetes through time. Datapoints are tip and estimated node data from the ancestral state estimations of biogeography and total body length (in metres) scaled to generation time. Polar glaciation data from [27]: solid lines represent permanent glaciation; dashed lines represent ephemeral glaciation. Green and orange arrows are inferred progression of body size through time, and not calculated trendlines.

References

1. Pyenson N.D., Goldbogen J.A., Shadwick R.E. 2013 Mandible allometry in extant and fossil Balaenopteridae (Cetacea: Mammalia): the largest vertebrate skeletal element and its role in rorqual lunge feeding. *Biological Journal of the Linnean Society* **108**(3), 586-599.
2. Revell L.J. 2012 phytools: an R package for phylogenetic comparative biology (and other things). *Methods in ecology and evolution* (2), 217-223.
3. Bianucci G., Marx F.G., Collareta A., Di Stefano A., Landini W., Morigi C., Varola A. 2019 Rise of the titans: baleen whales became giants earlier than thought. *Biology Letters* **15**(5), 20190175.
4. Slater G.J., Goldbogen J.A., Pyenson N.D. 2017 Independent evolution of baleen whale gigantism linked to Plio-Pleistocene ocean dynamics. *Proceedings of the Royal Society B: Biological Sciences* **284**(1855), 20170546.
5. Lloyd G.T., Slater G.J. 2021 A total-group phylogenetic metatree for Cetacea and the importance of fossil data in diversification analyses. *Systematic Biology* **70**(5), 922-939.
6. Marx F.G., Buono M.R., Evans A.R., Fordyce R.E., Reguero M., Hocking D.P. 2019 Gigantic mysticete predators roamed the Eocene Southern Ocean. *Antarctic Science* **31**(2), 98-104.
7. Bisconti M., Pellegrino L., Carnevale G. 2021 Evolution of gigantism in right and bowhead whales (Cetacea: Mysticeti: Balaenidae). *Biological Journal of the Linnean Society* **134**(2), 498-524.
8. Boessenecker R.W., Fordyce R.E. 2015 A new Eomysticetid (Mammalia: Cetacea) from the Late Oligocene of New Zealand and a re-evaluation of 'Mauicetus' waitakiensis. *Papers in Palaeontology* **1**(2), 107-140.
9. Fordyce R.E., Marx F.G. 2018 Gigantism precedes filter feeding in baleen whale evolution. *Current Biology* **28**(10), 1670-1676. e1672.

10. Marx F.G., Tsai C.-H., Fordyce R.E. 2015 A new Early Oligocene toothed 'baleen'whale (Mysticeti: Aetiocetidae) from western North America: one of the oldest and the smallest. *Royal Society Open Science* **2**(12), 150476.
11. Tsai C.-H., Fordyce R.E. 2018 A new archaic baleen whale Toipahautea waitaki (early Late Oligocene, New Zealand) and the origins of crown Mysticeti. *Royal Society Open Science* **5**(4), 172453.
12. Peredo C.M., Pyenson N.D. 2018 Salishicetus meadi, a new aetiocetid from the late Oligocene of Washington State and implications for feeding transitions in early mysticete evolution. *Royal Society Open Science* **5**(4), 172336.
13. Peredo C.M., Pyenson N.D., Marshall C.D., Uhen M.D. 2018 Tooth loss precedes the origin of baleen in whales. *Current Biology* **28**(24), 3992-4000. e3992.
14. Peredo C.M., Uhen M.D. 2016 A new basal chaeomysticete (Mammalia: Cetacea) from the late Oligocene Pysht Formation of Washington, USA. *Papers in Palaeontology* **2**(4), 533-554.
15. Shipps B., Peredo C.M., Pyenson N.D. 2019 Borealodon osedax, a new stem mysticete (Mammalia, Cetacea) from the Oligocene of Washington State and its implications for fossil whale-fall communities. *Royal Society open science* **6**(7), 182168.
16. Geisler J.H., Boessenecker R.W., Brown M., Beatty B.L. 2017 The origin of filter feeding in whales. *Current Biology* **27**(13), 2036-2042. e2032.
17. Lambert O., Martínez-Cáceres M., Bianucci G., Di Celma C., Salas-Gismondi R., Steurbaut E., Urbina M., De Muizon C. 2017 Earliest mysticete from the Late Eocene of Peru sheds new light on the origin of baleen whales. *Current Biology* **27**(10), 1535-1541. e1532.
18. Hernandez Cisneros A.E. 2018 A new group of late Oligocene mysticetes from México. *Palaeontologia Electronica* **21**(1).
19. Solis-Añorve A., González-Barba G., Hernández-Rivera R. 2019 Description of a new toothed mysticete from the Late Oligocene of San Juan de La Costa, BCS, México. *Journal of South American Earth Sciences* **89**, 337-346.
20. Barnes L.G., Kimura M., Furusawa H., Sawamura H. 1994 Classification and distribution of oligocene aetiocetidae (mammalia; cetacea; mysticeti) from western north America and Japan. *Island Arc* **3**(4), 392-431.
21. Pyenson N.D., Sponberg S.N. 2011 Reconstructing body size in extinct crown Cetacea (Neoceti) using allometry, phylogenetic methods and tests from the fossil record. *Journal of Mammalian Evolution* **18**(4), 269-288.
22. Evans A.R., Jones D., Boyer A.G., Brown J.H., Costa D.P., Ernest S.M., Fitzgerald E.M., Fortelius M., Gittleman J.L., Hamilton M.J. 2012 The maximum rate of mammal evolution. *Proceedings of the National Academy of Sciences* **109**(11), 4187-4190.
23. Pennell M.W., Eastman J.M., Slater G.J., Brown J.W., Uyeda J.C., FitzJohn R.G., Alfaro M.E., Harmon L.J. 2014 geiger v2. 0: an expanded suite of methods for fitting macroevolutionary models to phylogenetic trees. *Bioinformatics* **30**(15), 2216-2218.
24. Castiglione S., Serio C., Mondanaro A., Di Febbraro M., Profico A., Girardi G., Raia P. 2019 Simultaneous detection of macroevolutionary patterns in phenotypic means and rate of change with and within phylogenetic trees including extinct species. *PloS one* **14**(1), e0210101.
25. Castiglione S., Serio C., Mondanaro A., Melchionna M., Carotenuto F., Di Febbraro M., Profico A., Tamagnini D., Raia P. 2020 Ancestral state estimation with phylogenetic ridge regression. *Evolutionary Biology* **47**(3), 220-232.
26. Castiglione S., Tesone G., Piccolo M., Melchionna M., Mondanaro A., Serio C., Di Febbraro M., Raia P. 2018 A new method for testing evolutionary rate variation and shifts in phenotypic evolution. *Methods in Ecology and Evolution* **9**(4), 974-983.
27. St. John K. 2008 Cenozoic ice‐rafting history of the central Arctic Ocean: Terrigenous sands on the Lomonosov Ridge. *Paleoceanography* **23**(1).