

1 **A genomic assessment of the marine-speciation paradox within the toothed whale**  
2 **superfamily Delphinoidea**

3  
4  
5 Michael V Westbury<sup>1\*</sup>, Andrea A. Cabrera<sup>1</sup>, Alba Rey-Iglesia<sup>1</sup>, Binia De Cahsan<sup>1</sup>, David A.  
6 Duchêne<sup>1</sup>, Stefanie Hartmann<sup>2</sup>, Eline D Lorenzen<sup>1\*</sup>

- 7 1. GLOBE Institute, University of Copenhagen, Øster Voldgade 5-7, Copenhagen,  
8 Denmark  
9 2. Institute of Biochemistry and Biology, University of Potsdam, Karl-Liebknecht-Str.  
10 24-25, Potsdam, Germany

11 \* Corresponding authors: m.westbury@sund.ku.dk, elinelorenzen@sund.ku.dk

12  
13 **Abstract**

14 The importance of post-divergence gene flow in speciation has been documented  
15 across a range of taxa in recent years, and may have been especially widespread in highly  
16 mobile, wide-ranging marine species, such as cetaceans. Here, we studied individual  
17 genomes from nine species across the three families of the toothed whale superfamily  
18 Delphinoidea (Delphinidae, Phocoenidae, Monodontidae). To investigate the role of post-  
19 divergence gene flow in the speciation process, we used a multifaceted approach, including:  
20 (i) phylogenomics, (ii) the distribution of shared derived alleles, and (iii) demographic  
21 inference. We found the divergence of lineages within Delphinoidea did not follow a process  
22 of pure bifurcation, but was much more complex. Sliding-window phylogenomics reveal a  
23 high prevalence of discordant topologies within the superfamily, with further analyses  
24 indicating these discordances arose due to both incomplete lineage sorting and gene flow. D-  
25 statistics, D-foil, and *f*-branch analyses supported gene flow between members of  
26 Delphinoidea, with the vast majority of gene flow occurring as ancient interfamilial events.  
27 Demographic analyses provided evidence that introgressive gene flow has likely ceased  
28 between all species pairs tested, despite reports of contemporary interspecific hybrids. Our  
29 study provides the first steps towards resolving the large complexity of speciation within  
30 Delphinoidea; we reveal the prevalence of ancient interfamilial gene flow events prior to the  
31 diversification of each family, and **show** contemporary hybridisation events may be  
32 disadvantageous, as hybrid individuals do not appear to contribute to the parental species'  
33 gene pools.

## 43 Introduction

44

45 The formation of new species involves the divergence of lineages through  
46 reproductive isolation. Isolation can initially occur in allopatry (geographical isolation  
47 without gene flow) or in sympatry (biological/ecological isolation with gene flow). Over  
48 time, isolation can be maintained and strengthened, ultimately leading to the formation of  
49 new species (Norris and Hull, 2012). While allopatric speciation requires geographical  
50 isolation plus time, sympatric speciation often requires a broader and more complicated set of  
51 mechanisms (Turelli et al., 2001). These mechanisms mostly rely on ecologically mediated  
52 natural selection. Parapatric speciation, on the other hand, encompasses intermediate  
53 scenarios of partial, but incomplete, physical restrictions to gene flow leading to speciation.

54

55 Through the analysis of whole-genome datasets, the detection of post-divergence gene  
56 flow in various distinct taxonomic groups is becoming commonplace (Árnason et al., 2018;  
57 Barlow et al., 2018; Westbury et al., 2020), demonstrating that speciation is much more  
58 complex than a simple bifurcating process (Campbell and Poelstra, 2018; Feder et al., 2012).  
59 Speciation is not an instantaneous process, but usually requires tens of thousands to millions  
60 of generations to achieve complete reproductive isolation (Butlin and Smadja, 2018; Coyne  
61 and Orr, 2004; Liu et al., 2014). The duration it takes to reach this isolation may be especially  
62 long in highly mobile marine species, such as cetaceans, due to a relative lack of geographic  
63 barriers in the marine realm, and therefore high potential for secondary contact and gene flow  
64 (Árnason et al., 2018).

65

66 The apparent inability to undergo allopatric speciation in marine species has been  
67 termed the marine-speciation paradox (Bierne et al., 2003). However, over the past decade,  
68 genomic studies have provided insights into how speciation can occur within cetaceans  
69 (Árnason et al., 2018; Moura et al., 2020). For example, initial phases of allopatry among  
70 populations of killer whales (*Orcinus orca*) may have led to the accumulation of ecological  
71 differences between populations, which strengthened population differences even after  
72 secondary contact (Foote et al., 2011; Foote and Morin, 2015). However, whether these initial  
73 phases of allopatry caused the divergence, or whether speciation occurred purely in sympatry,  
74 remains debated (Foote, 2018; Moura et al., 2015). But, these two hypotheses are not  
75 necessarily mutually exclusive. Instead, differentiation in parapatry, encompassing features of  
76 both allopatric and sympatric speciation, may have been key in the evolutionary history of  
77 cetaceans.

78

79 The toothed whale superfamily Delphinoidea represents an interesting opportunity to  
80 further explore speciation in the presence of putative interspecific gene flow. The crown root  
81 of Delphinoidea has been dated at ~19 million years ago (Ma) (95% CI 19.73 - 18.26 Ma)  
82 (McGowen et al., 2020) and has given rise to three families: (i) Delphinidae, the most  
83 species-rich family, which comprises dolphins and ‘black-fish’ (such as killer whales and  
84 pilot whales (*Globicephala spp.*)); (ii) Phocoenidae, commonly known as porpoises; and (iii)  
85 Monodontidae, which comprises two extant lineages, beluga (*Delphinapterus leucas*) and  
86 narwhal (*Monodon monoceros*).

87

88 Delphinoidea is of particular interest, as contemporary interspecific hybrids have been  
89 reported within all three families (Delphinidae (Espada et al., 2019; Miyazaki et al., 1992;  
90 Silva et al., 2005); Phocoenidae (Willis et al., 2004); Monodontidae (Skovrind et al., 2019)).  
91 However, these represent recent hybridization events that occurred long after species  
92 divergence, and their contribution to the parental gene pools is mostly unknown. The  
93 presence of more ancient introgressive hybridization events between families, and during the  
94 early radiations of these families, has yet to be investigated. With the rapid increase of  
95 genomic resources for cetaceans, and in particular for species within Delphinoidea, we are  
96 presented with the ideal opportunity to investigate post-divergence gene flow between  
97 lineages, furthering our understanding of speciation processes in cetaceans.

98

99 Here, we utilise publicly available whole-genome data from nine species of  
100 Delphinoidea, representing all three families, to investigate signs of post-divergence gene  
101 flow across their genomes. Our analyses included five Delphinidae (killer whale, Pacific  
102 white-sided dolphin (*Lagenorhynchus obliquidens*), long-finned pilot whale (*Globicephala*  
103 *melas*), bottlenose dolphin (*Tursiops truncatus*), Indo-Pacific bottlenose dolphin (*T.*  
104 *aduncus*)); two Phocoenidae (harbour porpoise (*Phocoena phocoena*), finless porpoise  
105 (*Neophocaena phocaenoides*)); and two Monodontidae (beluga, narwhal). Moreover, we  
106 compare their species-specific genetic diversity and demographic histories, and explore how  
107 species abundances may have played a role in interspecific hybridisation over the last two  
108 million years.

109

## 110 **Results and discussion**

111

### 112 **Detecting gene flow**

113 To assess the evolutionary relationships across the genomes of the nine Delphinoidea  
114 species investigated, we computed non-overlapping, sliding-window, maximum-likelihood  
115 phylogenies of four different window sizes in RAxML (Stamatakis, 2014). These analyses  
116 resulted in 43,207 trees (50 kilobase (kb) windows), 21,387 trees (100 kb windows), 3,705  
117 trees (500 kb windows), and 1,541 trees (1 megabase (Mb) windows) (Fig. 1, Supplementary  
118 Fig. S1, Supplementary Table S1). The 50 kb windows retrieved 96 unique topologies, 100  
119 kb windows retrieved 47 unique topologies, 500 kb windows retrieved 16 unique topologies,  
120 and 1 Mb windows retrieved 15 unique topologies. Regardless of window size, we retrieved  
121 consensus support for the species tree previously reported using target-sequence capture  
122 (McGowen et al., 2020). However, when considering the smallest window size (50 kb), we  
123 found a considerable proportion of trees (up to 76%) with an alternative topology to the  
124 species tree (Fig. 1A). These alternative topologies may be due to incomplete lineage sorting  
125 (ILS) or interspecific gene flow (Leaché et al., 2014). Moreover, the higher prevalence of this  
126 pattern in the shorter 50 kb windows may indicate that inconsistencies in topology are caused  
127 by ancient, rather than recent, gene flow events, as recombination is expected to break up  
128 longer introgressed regions over time (as a comparison, only 21% of windows in the 1 Mb  
129 dataset do not show the most common topology, Fig. 1B).

130

131 We explored whether the large number of phylogenetic discrepancies in the 50kb  
132 windows could be linked to the GC content (%GC) of the windows as elevated levels of GC  
133 content can result from higher levels of GC-Biased Gene Conversion (gBGC) in regions with  
134 higher levels of recombination (Lartillot, 2013). When binning windows into either high,  
135 medium, or low levels of GC content, the most common topologies were consistent, but with  
136 slight differences in overall values (Supplementary Table S2). This result suggests that the  
137 topological discrepancies are not arising purely due to GC-content linked biases and  
138 recombination rate.

139  
140 To investigate whether the alternative topologies could simply be explained by ILS,  
141 or a combination of ILS and gene flow, we ran Quantifying Introgression via Branch Lengths  
142 (QuIBL) (Edelman et al., 2019) on every twentieth tree from the 50 kb sliding-window  
143 analysis (Supplementary Table S3), as well as on a dataset that contained trees constructed  
144 using 20 kb windows with a 1 Mb slide (Supplementary Table S4). We were only able to  
145 investigate the potential cause of discordances within the Delphinidae family, as we did not  
146 recover any phylogenetic discordances between families, and all families were respectively  
147 monophyletic.

148  
149 When considering the results using 50 kb windows, we found significant evidence of  
150 ILS and gene flow in all species pairwise comparisons within Delphinidae. The only  
151 comparisons that did not show significant results for gene flow were those that contained  
152 both the bottlenose and Indo-Pacific bottlenose dolphins. The lacking evidence of gene flow  
153 when both *Tursiops* species were included, suggests signals of gene flow between either  
154 *Tursiops* species and killer whale, Pacific white-sided dolphin, or pilot whale are likely  
155 remnants of ancestral gene flow events between the ancestral *Tursiops* and the given  
156 comparative species.

157  
158 Similar to the 50 kb windows, the 20 kb window analysis showed a large proportion  
159 of alternative topologies within Delphinidae likely arose due to both ILS gene flow. Again,  
160 we retrieved most non-significant results when both *Tursiops* species were included in the  
161 analysis. Moreover, although we found no evidence of gene flow between killer whale and  
162 pilot whale when either *Tursiops* was included as the triplet outgroup, we found evidence of  
163 gene flow when the Pacific white-sided dolphin was the triplet outgroup. We also found no  
164 evidence for gene flow between the Indo-Pacific bottlenose and Pacific white-sided dolphins,  
165 regardless of triplet outgroup. It is difficult to ascertain why we observe discrepancies  
166 between results based on the triplet outgroup. But, taken together, our QuIBL analyses  
167 suggest a combination of ILS and gene flow played a role in shaping the evolutionary history  
168 of Delphinidae.

169  
170 To further explore potential gene flow while taking ILS into account, we used D-  
171 statistics (Durand et al., 2011; Green et al., 2010). D-statistics uses a four-taxon approach  
172 [[[H1, H2], H3], Outgroup] to uncover the differential distribution of shared derived alleles,  
173 which may represent gene flow between either H1/H3 or H2/H3. Here we used baiji (*Lipotes*  
174 *vexillifer*) as the outgroup, and alternated ingroup positions based on the consensus topology.

175 In congruence with the QuIBL results, we found significant levels of gene flow within  
176 Delphinidae. However, we also found higher levels of gene flow between the killer whale,  
177 pilot whale, and Pacific white-sided dolphin and the Indo-Pacific bottlenose dolphin, relative  
178 to the bottlenose dolphin. **In fact, 85 out of 86 tests showed significant signs of gene flow**  
179 **both within and between families (Supplementary Table S5).** The only comparison that did  
180 not return a significant result was [[[finless porpoise, harbour porpoise], narwhal], outgroup].  
181 This does not necessarily mean there was no gene flow between these species, but could be  
182 caused by equal amounts of gene flow between both porpoise species and narwhal. Such  
183 abundant signs of gene flow suggests the evolutionary history of Delphinoidea was more  
184 complex than a simple bifurcating process. Alternatively, our findings may reflect **limitations**  
185 **of the D-statistic and false positives due to gene flow between ancestral lineages** (Moodley et  
186 al., 2020).

187

188 Due to the inability of the four-taxon D-statistics approach to detect the direction of  
189 gene flow, as well as whether gene flow events may have occurred between ancestral  
190 lineages, we used D-foil (Pease and Hahn, 2015). D-foil enables further characterization of  
191 the D-statistics results, which may be particularly relevant given the complex array of gene  
192 flow putatively present within Delphinoidea. D-foil uses a five-taxon approach [[H1, H2]  
193 [H3, H4], Outgroup] and a system of four independent D-statistics in a sliding-window  
194 fashion, to uncover (i) putative gene flow events, (ii) donor and recipient lineages, and (iii)  
195 whether gene flow events occurred between a distantly related lineage and the ancestor of  
196 two sister lineages, which is indicative of ancestral-lineage gene flow. However, as the input  
197 topology requirements of D-foil are [[H1, H2] [H3, H4], Outgroup], we were only able to  
198 investigate gene flow between families, and not within families, using this analysis. Hence,  
199 we tested for gene flow between Delphinidae/Phocoenidae, Delphinidae/Monodontidae, and  
200 Phocoenidae/Monodontidae.

201

202 The D-foil results underscore the complex pattern of post-divergence gene flow  
203 between families indicated by the D-statistics. We found support for interfamilial gene flow  
204 events between all nine species investigated, to varying extents (Supplementary Table S6).  
205 This could reflect multiple episodes of gene flow between all investigated species.  
206 Alternatively, the pattern could reflect ancient gene flow events between the ancestors of H1-  
207 H2 and H3-H4 (in the topology [[H1, H2] [H3, H4], Outgroup]), with differential inheritance  
208 of the introgressed loci in subsequent lineages. Such ancestral gene flow events have  
209 previously been shown to lead **to false positives between species pairs using D-statistics**  
210 (Moodley et al., 2020). A further putative problem with these results can be seen when  
211 implementing D-foil on the topology [[Delphinidae, Delphinidae], [Monodontidae,  
212 Phocoenidae], Outgroup]. We found the majority of windows support a closer relationship  
213 between Delphinidae (ancestors of H1 and H2) and Monodontidae (H3), as opposed to the  
214 species tree. If this result is correct, it suggests the input topology was incorrect, and the  
215 results reflect more recent common ancestry and not gene flow. This **implies** Delphinidae and  
216 Monodontidae are sister lineages, as opposed to Phocoenidae and Monodontidae. However,  
217 this contrasts with the family topology of [Delphinidae, [Phocoenidae, Monodontidae]]

218 retrieved in our phylogenetic analyses (Fig. 1) and reported by others (McGowen et al., 2020;  
219 Steeman et al., 2009).

220

221 Taken together, it is difficult to ascertain whether our D-statistics and D-foil results of  
222 prevalent gene flow among most species pairs are true, or whether some results may have  
223 arisen due to biases that can occur when attempting to infer gene flow between highly  
224 divergent lineages. False positives and potential biases in D-statistics and D-foil can arise due  
225 to a number of factors including (i) ancestral population structure, (ii) introgression from  
226 unsampled and/or extinct ghost lineages, (iii) differences in relative population size of  
227 lineages or in the timing of gene flow events, (iv) different evolutionary rates or sequencing  
228 errors between H1 and H2, and (v) gene flow between ancestral lineages (Moodley et al.,  
229 2020; Slatkin and Pollack, 2008; Zheng and Janke, 2018). These issues are important to  
230 consider when interpreting our results, as the deep divergences of lineages suggest the  
231 possibility for a number of ancestral gene flow events, as well as gene flow events between  
232 now-extinct lineages, that may bias results.

233

234 Due to the large number of possible D-statistics comparisons, and difficulties  
235 disentangling false positives that may arise due to ancient gene flow events, we performed  
236 the *f*-branch test (Malinsky et al., 2021, 2018). The test takes correlated allele sharing into  
237 account when visualising *f*<sub>4</sub>-ratio (similar to D-statistics) results. The *f*-branch results  
238 suggested several instances of gene flow, many between ancestral lineages with relatively  
239 small values of *f*<sub>b</sub> (<0.04 with the majority being ~0.01) (Fig. 2). This result suggests  
240 widespread gene flow but in small quantities. However, it should be noted that *f*<sub>b</sub> represents  
241 relative quantities of gene flow and likely also decreases the older the introgression event  
242 (Martin et al., 2015) so the values we present here may not fully represent the absolute levels  
243 of gene flow. When considering interfamilial gene flow events, we see excess allele sharing  
244 (*f*<sub>b</sub>) between the ancestral Monodontidae branch and all Delphinidae species, which we  
245 interpret as gene flow between the ancestral lineages of Monodontidae and Delphinidae. We  
246 also uncovered elevated *f*<sub>b</sub> between the ancestor of all Delphinidae (to the exclusion of the  
247 killer whale) and all Phocoenidae and Monodontidae species, which could suggest gene flow  
248 between Delphinidae and the ancestral Phocoenidae/Monodontidae lineage. However, the  
249 exclusion of the killer whale may be due to the inability of the four taxon *f*<sub>4</sub>-ratio test to  
250 calculate gene flow between the killer whale and ancestral Phocoenidae/Monodontidae.  
251 Based on this limitation, we take a conservative approach and suggest this result reflects gene  
252 flow between the ancestral Delphinidae and ancestral Phocoenidae/Monodontidae.

253

254 Further supporting the hypothesis of gene flow between the ancestral Delphinidae and  
255 ancestral Phocoenidae/Monodontidae, we also observed signs of gene flow between the  
256 finless porpoise and all Delphinidae species, which suggests gene flow between the finless  
257 porpoise and ancestral Delphinidae. This seems unreasonable, as the finless porpoise  
258 diverged from the harbour porpoise much more recently (~5 Ma) than the time to the most  
259 recent common ancestor (tMRCA) of all Delphinidae (~10 Ma, (McGowen et al., 2020),  
260 meaning gene flow would have occurred independently between the finless porpoise and  
261 almost every Delphinidae species studied here. Moreover, the *f*-branch showed similar *f*<sub>b</sub>

262 between the Indo-Pacific bottlenose dolphin and all Phocoenidae and Monodontidae, as well  
263 as between the ancestral *Tursiops* and all Phocoenidae and Monodontidae. Similar to the  
264 finless porpoise and ancestral Delphinidae, **this result seems unlikely due to the divergence**  
265 **times of *Tursiops*.**

266  
267 We also found signals of gene flow between beluga and both Phocoenidae species,  
268 but not between narwhal and Phocoenidae. This pattern may be more parsimoniously  
269 explained by an ancestral event between Phocoenidae and Monodontidae, where the narwhal  
270 retained less introgressed alleles. A given *fb* statistic presents the signal of excess gene flow  
271 relative to the ingroup's sister taxa (Malinsky et al., 2021). Hence, not recovering a signal of  
272 gene flow with the sister taxa does not mean it did not occur. Rather, gene flow may have  
273 occurred between taxa, but to a lesser degree. Taking this into account, we suggest our results  
274 may instead be **remnants of ancestral gene flow events between the ancestral Phocoenidae**  
275 **and Monodontidae lineages.** A lack of evidence for more recent, species-specific gene flow  
276 events here is congruent with the sliding-window tree analyses, which consistently showed  
277 Phocoenidae and Monodontidae as monophyletic groups.

278  
279 The *f*-branch test also revealed interspecific gene flow events within Delphinidae may  
280 have been common. We uncovered evidence for gene flow between the Pacific white-sided  
281 dolphin and ancestral *Tursiops*, as well as the killer whale and ancestral *Tursiops*. However,  
282 we are unable to dissect whether there was gene flow between the pilot whale and ancestral  
283 *Tursiops*, due to the limitation of the four-taxon requirement.

284  
285 To investigate whether the X chromosome may have presented a more pronounced  
286 barrier to gene flow relative to the autosomes, we ran the *f*-branch test on scaffolds aligning  
287 to the X chromosome. Results were similar to the genome-wide dataset (Supplementary Fig.  
288 S2). The most obvious difference is that evidence for gene flow between Phocoenidae and  
289 Monodontidae is not as pronounced as in the genome-wide dataset. It is difficult to discern  
290 whether the lack of resolution here is due to the X chromosome constituting a smaller dataset,  
291 or whether parts of the X chromosome were not incorporated into the recipient gene pool due  
292 to the occurrence of more rapid reproductive isolation on the X chromosome (Payseur and  
293 Rieseberg, 2016). The former option appears more probable, due to the consistent evidence  
294 for gene flow between the beluga and both Phocoenidae species, which are likely the  
295 remnants of ancestral gene flow events between Phocoenidae and Monodontidae.

296  
297 By combining results acquired through sliding-window phylogenies, QuIBL, D-  
298 statistics, Dfoil, and *f*-branch, we are able to better decipher the complex evolutionary history  
299 of Delphinidae, and the signatures of interspecific gene-flow events present in most  
300 individuals studied. We found the most probable explanation for such wide-spread signatures  
301 to be the differential inheritance of remnant loci from ancestral gene flow events. However,  
302 as exemplified here and due to the limitations of each method, uncovering the exact lineages  
303 involved in these events is challenging.

304  
305 **Cessation of lineage sorting and/or gene flow**

306 To further elucidate the complexity of interspecific gene flow within Delphinoidea,  
307 we implemented F1 hybrid PSMC (hPSMC) (Cahill et al., 2016) on the autosomes of our  
308 species of interest. This method creates a pseudo-diploid sequence by merging pseudo-  
309 haploid sequences from two different genomes, which in our case represents two different  
310 species. The variation in the interspecific pseudo-F1 hybrid genome cannot coalesce more  
311 recently than the emergence of reproductive isolation between the two parental species. If  
312 some regions within the genomes of two target species are yet to fully diverge, due to ILS or  
313 to gene flow, hybridisation may still be possible. Therefore, we can use this method to infer  
314 when reproductive isolation between two species may have occurred.

315

316 When considering the upper bound of when two target genomes coalesce (equating  
317 the oldest date), and the lower bound of each divergence date (equating the most recent date)  
318 (McGowen et al., 2020), we found the majority of comparisons (29/36) show lineage sorting  
319 and/or gene flow occurred for >50% of the post-divergence branch length (Fig. 3,  
320 Supplementary data - hPSMC). However, we used divergence dates estimated assuming a  
321 fixed tree-like topology without taking gene tree discordances into account, and therefore the  
322 divergence dates may be overestimated due to extended terminal branches from molecular  
323 substitutions of discordant loci needing to be placed somewhere on the tree (Mendes and  
324 Hahn, 2016). Nevertheless, our results suggest that reaching complete reproductive isolation  
325 in Delphinoidea was a slow process, due to ILS and/or gene flow. ILS levels are known to be  
326 proportional to ancestral population sizes, and inversely proportional to time between  
327 speciation events (Pamilo and Nei, 1988). Hence, if ILS was the only explanation for this  
328 phenomenon, this would suggest extremely large ancestral population sizes. We do indeed  
329 see that the species pairs with the highest  $N_e$  prior to the end of lineage sorting/gene flow  
330 (Supplementary table S7) also have the largest discrepancies between divergence date and the  
331 date at which the two genomes coalesce. However, an alternative, and perhaps more likely,  
332 explanation is the occurrence of gene flow after initial divergence, supported by our  
333 phylogenomic, D-statistics, Dfoil, and  $f$ -branch results above. Post-divergence gene flow may  
334 reflect the ability of cetacean species to travel long distances, and the absence of significant  
335 geographical barriers in the marine environment. Alternatively, if geographic barriers did  
336 drive initial divergence, the pattern retrieved in our data may reflect secondary contact prior  
337 to complete reproductive isolation.

338

339 Our hPSMC results showed an almost simultaneous cessation of lineage sorting/gene  
340 flow regardless of species pair within the Delphinidae family (Fig 3A), as well as  
341 comparisons between families (Fig 3B). Based on our D-statistic/D-foil/ $f$ -branch results  
342 showing many of the signals of gene flow may be remnants of ancestral gene flow events, we  
343 hypothesise that our deep-time hPSMC results may also be produced by ILS of ancestrally  
344 introgressed regions. If we assume the divergence dates are correct, this hypothesis also  
345 offers an explanation regarding why the end of interfamilial ILS/gene flow occurs after the  
346 tMRCA of the family in many cases. For example, the tMRCA of Phocoenidae is ~6Ma, and  
347 the tMRCA of Monodontidae is ~7Ma but our hPSMC suggests that ILS/gene flow did not  
348 stop between Phocoenidae and Monodontidae until ~5Ma. Superficially, this implies that  
349 interfamilial gene flow occurred uniquely between beluga/finless porpoise, beluga/harbour

350 porpoise, narwhal/finless porpoise, and narwhal/harbour porpoise, and ceased for all species  
351 pairs at the same time. While this may have been the case, a more likely explanation is that  
352 lineage sorting of introgressed regions from an ancestral gene flow event was not complete  
353 until the time periods that our hPSMC results recovered.

354

355 Despite our hPSMC results of long-term lineage sorting/gene flow in the majority of  
356 species comparisons, they also suggested that lineage sorting is complete and gene flow has  
357 ceased between all lineages in our dataset. This finding is in contrast with confirmed reports  
358 of fertile contemporary hybrids between several of our target species, and may reflect the  
359 inability of hPSMC to detect low levels of migration. For example, viable offspring have  
360 been reported between bottlenose dolphins and Indo-Pacific bottlenose dolphins (Gridley et  
361 al., 2018) and between bottlenose dolphins and Pacific white-sided dolphins (Crossman et al.,  
362 2016; Miyazaki et al., 1992). Simulations have shown that in the presence of as few as  
363 1/10,000 migrants per generation, hPSMC will suggest continued gene flow. However, this is  
364 not the case with a rate  $< 1/100,000$  migrants per generation. Rather, in the latter case, the  
365 exponential increase in effective population size ( $N_e$ ) of the pseudo-hybrid genome, which  
366 can be used to infer the date at which gene flow ceased between the parental species,  
367 becomes a more gradual transition, leading to a larger estimated time interval of gene flow  
368 (Cahill et al., 2016). Within Delphinidae, we observe a less pronounced increase in  $N_e$  in the  
369 pseudo-hybrids, suggesting continued, but very low migration rates (Supplementary results -  
370 hPSMC). This finding suggests that gene flow within Delphinidae may have continued for  
371 longer than shown by hPSMC, which may not be sensitive enough to detect low rates of  
372 recent gene flow. Either way, our hPSMC results within and between all three families  
373 showed a consistent pattern of long periods of lineage sorting/gene flow in Delphinoidea,  
374 some lasting more than ten million years post divergence.

375

376 We further assessed the robustness of our hPSMC results to the inclusion or exclusion  
377 of repeat regions in the pseudodiploid genome. We compared the hPSMC results when  
378 including and removing repeat regions for three independent species pairs of varying  
379 phylogenetic distance. These included a shallow divergence (bottlenose and Indo-Pacific  
380 bottlenose dolphins), medium divergence (beluga and narwhal), and deep divergence  
381 (bottlenose dolphin and beluga) (Supplementary Figs. S3 - S5). For all species pairs, results  
382 showed that pre-divergence  $N_e$  is almost identical, and the exponential increase in  $N_e$  is just  
383 slightly more recent when removing repeat regions, compared to when repeat regions are  
384 included. This gives us confidence that the inclusion of repeats did not greatly alter our  
385 results.

386

387 To add independent evidence for continued lineage sorting/gene flow for an extended  
388 period after initial divergence, we compared relative divergence time between killer whale,  
389 Pacific white-sided dolphin, and long-finned pilot whale based on the species tree and a set of  
390 alternative topologies (Supplementary Fig. S6). We focused on Delphinidae, due to the large  
391 number of loci per alternative topology (Supplementary Tables S1, S2, S3, and S4). By  
392 assuming ILS and gene flow are the dominant forces behind gene-tree discordance, we can  
393 uncover information about the timing of ILS and gene flow events among lineages, by

394 isolating the loci that produce each topology (Mendes and Hahn, 2016). In agreement with  
395 our hPSMC results, this analysis showed that ILS/gene flow continued for a long time after  
396 initial divergence. For example, we observed that the killer whale diverged from all other  
397 Delphinidae at a relative divergence time of 0.45 (45% of the divergence time of  
398 Delphinoidea and the baiji) in the consensus topology (Supplementary Fig. S6A). In an  
399 alternative topology, the killer whale was placed as sister to the Pacific white-sided dolphin  
400 (Supplementary Fig. S6B); despite still diverging from the remaining Delphinidae at  
401 approximately the same relative timing (0.42), it diverged from the Pacific white-sided  
402 dolphin at a relative divergence time of 0.25. As we assumed the alternative topologies only  
403 arose due to ILS and/or gene flow, this suggested lineage sorting and/or gene flow continued  
404 along ~40% of the post-divergence branch length. This estimate was qualitatively equivalent  
405 to that made using hPSMC (minimally 43%). Similarly, long periods of post-divergence  
406 lineage sorting/gene flow were observed when investigating topologies with the killer whale  
407 and long-finned pilot whale as sister species (Supplementary Fig. S6C, ~43%), and with the  
408 Pacific white-sided dolphin and long-finned pilot whale as sister species (Supplementary Fig.  
409 S6D, ~37%). As the results here included alternative topologies that likely arose due to both  
410 ILS and gene flow, we propose that the numbers present a more conservative estimate. One  
411 would expect ILS to be a more prevalent force behind discordances shortly after the species'  
412 divergence, whereas gene flow can occur after many generations. Therefore, if we could  
413 more confidently disentangle alternative topologies arising due to ILS from those arising due  
414 to gene flow, we would expect much more recent relative divergence times for loci that  
415 underwent gene flow.

416

417 In summary, by combining findings from several analyses, and with the knowledge  
418 that interspecific hybridisation is still ongoing between many of the lineages studied here, we  
419 suggest that both ILS and gene flow played a major role over extended periods of time, in the  
420 speciation of Delphinoidea.

421

### 422 **Interspecific hybridisation**

423

424 Making inferences as to what biological factors lead to interspecific hybridisation is  
425 challenging, as many variables may play a role. One hypothesis is that interspecific  
426 hybridization may occur at a higher rate during periods of low abundance, when a given  
427 species encounters only a limited number of conspecifics (Crossman et al., 2016; Edwards et  
428 al., 2011; Westbury et al., 2019). When considering species that have not yet undergone  
429 sufficient divergence preventing their ability to hybridise, individuals may mate with a  
430 related species, instead of investing energy in finding a relatively rarer conspecific mate.

431

432 To explore the relationship between susceptibility to interspecific hybridisation and  
433 population size, we calculated the level of genome-wide genetic diversity for each species, as  
434 a proxy for their  $N_e$  (Fig. 4A). Narwhal, killer whale, beluga, and long-finned pilot whale had  
435 the lowest diversity levels, respectively, and should therefore be more susceptible to  
436 interspecific hybridization events. A beluga/narwhal hybrid has been reported (Skovrind et  
437 al., 2019), as has hybridisation between long-finned and short-finned pilot whales (Miralles et

438 al., 2016). However, hybrids between species with high genetic diversity, including harbour  
439 porpoise (Willis et al., 2004), Indo-Pacific bottlenose dolphin (Baird et al., 2012), and  
440 bottlenose dolphin (Espada et al., 2019; Herzog and Johnson, 1997), have also been  
441 reported, suggesting genetic diversity alone is not a good proxy for susceptibility to  
442 hybridisation.

443

444 To investigate the effect of interspecific gene flow on  $N_e$ , we estimated changes in  
445 intraspecific genetic diversity through time (Fig. 4B-D). The modelled demographic  
446 trajectories, using a Pairwise Sequentially Markovian Coalescent model (PSMC), span the  
447 past two million years. We could therefore assess the relationship for the three species pairs,  
448 where the putative interval for the cessation of lineage sorting/gene flow was contained  
449 within this period: harbour/finless porpoise (Phocoenidae), beluga/narwhal (Monodontidae),  
450 and bottlenose/Indo-Pacific bottlenose dolphin (Delphinidae) (Fig. 3).

451

452 In the harbour porpoise, we observed an increase in  $N_e$  beginning ~1 Ma, the rate of  
453 which increased further ~0.5 Ma (Fig. 4C). We observed a similar pattern in belugas; an  
454 increase in  $N_e$  ~1 Ma, relatively soon after the proposed cessation of gene flow with narwhals  
455 ~1.8 - 1.2 Ma (Fig. 4D). Although  $N_e$  may reflect abundance, it is also influenced by several  
456 other factors, including population connectivity and gene flow. If gene flow explained our  
457 changes in  $N_e$ , we would therefore expect a decrease in  $N_e$  after gene flow ceased, but  
458 instead we observed an increase. An increase in  $N_e$  may coincide with an increase in relative  
459 abundance, which would increase the number of potential conspecific mates, and in turn  
460 reduce the level of interspecific gene flow. However, this is difficult to say for certain  
461 without more information on abundances through time.

462

463 We observed a different pattern in the bottlenose/Indo-Pacific bottlenose dolphins.  
464 We found a relatively high population size during the period of gene flow in both species;  $N_e$   
465 declines ~1 - 0.5 Ma, coinciding with the putative end of gene flow ~1.2 - 0.4 Ma. The  
466 decline in  $N_e$  could either reflect a decline in abundance, or a loss of connectivity between  
467 the two species. In the latter, we expect levels of intraspecific diversity (and thereby inferred  
468  $N_e$ ) to decline with the cessation of gene flow, even if absolute abundances did not change.  
469 This is indeed suggested by our data, which showed both species undergoing the decline  
470 simultaneously, indicative of a common cause.

471

472 Seven of the nine Delphinoidea genomes investigated showed a similar pattern of a  
473 rapid decline in  $N_e$  starting ~150-100 thousands of years ago (kya) (Fig. 4B-D; the  
474 exceptions are Pacific white-sided dolphin and narwhal). This concurrent decline could  
475 represent actual population declines across species, or, alternatively, simultaneous reductions  
476 in connectivity among populations within each species. Based on similar PSMC analyses, a  
477 decline in  $N_e$  at this time has also been reported in four baleen whale species (Árnason et al.,  
478 2018). Therefore, the species-wide pattern may reflect climate-driven environmental change.  
479 The period of 150-100 kya overlaps with the onset of the last interglacial, when sea levels  
480 increased to levels as high, if not higher, than at present (Polyak et al., 2018), and which may  
481 have had a marine-wide effect on both population connectivity and sizes. The unique life

482 histories, distribution, and ecology of the cetacean species suggests that a combination of  
483 both decreased population connectivity and population sizes across the different studied  
484 species. A similar marine-wide effect has been observed among baleen whales and their prey  
485 species in the Southern and North Atlantic Oceans during the Pleistocene-Holocene climate  
486 transition (12-7 kya) (Cabrera et al., 2018). These results indicate that past marine-wide  
487 environmental shifts have driven demographic changes in population across multiple marine  
488 species.

489

490 Although speculative, we suggest that recent species-wide declines associated with  
491 the onset of the last glacial period, may have facilitated the resurgence of hybridization  
492 between some of the nine Delphinoidea species analysed. If interspecific hybridisation has  
493 increased after these declines, species may already be sufficiently differentiated that offspring  
494 fertility is reduced. Even if offspring are fertile, the high level of differentiation between  
495 species may mean hybrids are unable to occupy either parental niche (Skovrind et al., 2019)  
496 and are strongly selected against. A lack of significant contribution from recent hybrids to the  
497 parental gene pools may be why we observe contemporary hybrids, but do not find evidence  
498 of this in our analyses.

499

## 500 **Conclusions**

501

502 Allopatric speciation is generally considered the most common mode of speciation, as  
503 the absence of gene flow due to geographic isolation can most easily explain the evolution of  
504 ecological, behavioural, morphological, or genetic differences between populations (Norris  
505 and Hull, 2012). However, our findings suggest that within Delphinoidea, speciation in the  
506 presence of gene flow was commonplace, consistent with sympatric/parapatric speciation, or  
507 allopatric speciation and secondary contact.

508

509 The ability for gene flow events to occur long after initial divergence may also  
510 explain the presence of contemporaneous hybrids between several species. In parapatric  
511 speciation, genetic isolation is achieved relatively early due to geographical and biological  
512 isolation, but species develop complete reproductive isolation relatively slowly, through low  
513 levels of migration or secondary contact events (Norris and Hull, 2012). The prevalence of  
514 this mode of speciation in cetaceans, as suggested by our study and previous genomic  
515 analyses (Árnason et al., 2018; Moura et al., 2020), may reflect the low energetic costs of  
516 dispersing across large distances in the marine realm (Fish et al., 2008; Williams, 1999) and  
517 the relative absence of geographic barriers preventing such dispersal events (Palumbi, 1994).  
518 Both factors are believed to be important in facilitating long-distance (including inter-  
519 hemispheric and inter-oceanic) movements in many cetacean species (Stone et al., 1990).

520

521 Our study shows that speciation in Delphinoidea was a complex process and involved  
522 multiple ecological and evolutionary factors. Our results take a step towards resolving the  
523 enormous complexity of speciation within this superfamily, through a multifaceted analysis  
524 of nuclear genomes. Our study underscores the challenges of accurately interpreting some  
525 results, due to the high levels of divergence between the target species. Moreover, while we

526 make inferences based on a genome-wide dataset, certain regions of the genome may have a  
527 greater contribution to reproductive isolation than others, e.g. sex chromosomes and regions  
528 of reduced recombination (Payseur and Rieseberg, 2016). By using the hypotheses we form  
529 about general patterns and major processes of gene flow and speciation uncovered in our  
530 data, we hope that future studies may be able to build on our results to make more specific  
531 inferences as to the genomics of speciation in Delphinoidea, as additional genomic data and  
532 new methodologies for data analysis become available.

533

## 534 **Methods**

535

### 536 **Data collection**

537 We downloaded the assembled genomes and raw sequencing reads from nine toothed  
538 whales from the superfamily Delphinoidea. The data included five Delphinidae: Pacific  
539 white-sided dolphin (NCBI Biosample: SAMN09386610), Indo-Pacific bottlenose dolphin  
540 (NCBI Biosample: SAMN06289676), bottlenose dolphin (NCBI Biosample:  
541 SAMN09426418), killer whale (NCBI Biosample: SAMN01180276), and long-finned pilot  
542 whale (NCBI Biosample: SAMN11083132); two Phocoenidae: harbour porpoise (Autenrieth  
543 et al., 2018) and finless porpoise (NCBI Biosample: SAMN02192673); and two  
544 Monodontidae: beluga (NCBI Biosample: SAMN06216270) and narwhal (NCBI Biosample:  
545 SAMN10519625). To avoid artificially inflating signals of genetic similarities between a  
546 highly divergent outgroup and an ingroup species used as mapping reference (Liu et al.,  
547 2021), we downloaded the assembled outgroup baiji genome (Genbank accession code:  
548 GCF\_000442215.1) as mapping reference in the gene flow analyses. Delphinoidea and the  
549 baiji diverged ~24.6 Ma (95% CI 25.2 - 23.8 Ma) (McGowen et al., 2020).

550

### 551 **Initial data filtering**

552 To determine which scaffolds were most likely autosomal in origin, we identified  
553 putative sex chromosome scaffolds for each genome through synteny, and omitted them from  
554 further analysis. We found putative sex chromosome scaffolds in all ten assemblies by  
555 aligning them to the Cow X (Genbank accession: CM008168.2) and Human Y (Genbank  
556 accession: NC\_000024.10) chromosomes. Alignments were performed using satsuma  
557 synteny v2.1 (Grabherr et al., 2010) with default parameters. Since short scaffolds have a  
558 higher likelihood of including assembly errors, we also removed scaffolds smaller than 100  
559 kb from all downstream analyses.

560

### 561 **Mapping**

562 We trimmed adapter sequences from all raw reads using skewer v0.2.2 (Jiang et al.,  
563 2014). We mapped the trimmed reads to the baiji for downstream gene flow analyses, and to  
564 the species-specific reference genome for downstream demographic history and genetic  
565 diversity analyses using BWA v0.7.15 (Li and Durbin, 2009) and the mem algorithm. We  
566 parsed the output and removed duplicates and reads with a mapping quality lower than 30  
567 with SAMtools v1.6 (Li et al., 2009). Mapping statistics can be found in supplementary tables  
568 S8 and S9.

569

## 570 **Sliding-window phylogeny**

571 For the sliding-window phylogenetic analysis, we created fasta files for all individuals  
572 mapped to the baiji genome using a consensus base call (-dofasta 2) approach in ANGSD  
573 v0.921 (Korneliussen et al., 2014), and specifying the following filters: minimum read depth  
574 of 5 (-mininddepth 5), minimum mapping quality of 30 (-minmapq 30), minimum base  
575 quality (-minq 30), only consider reads that map to one location uniquely (-uniqueonly 1),  
576 and only include reads where both mates map (-only\_proper\_pairs 1). All resultant fasta files,  
577 together with the assembled baiji genome, were aligned, and sites where any individual had  
578 more than 50% missing data were filtered before performing maximum likelihood  
579 phylogenetic analyses in a non-overlapping sliding-window approach using RAxML v8.2.10  
580 (Stamatakis, 2014). We performed this analysis four times independently, specifying a  
581 different window size each time (50 kb, 100 kb, 500 kb, and 1 Mb). We used RAxML with  
582 default parameters, specifying baiji as the outgroup, and a GTR+G substitution model. We  
583 computed the genome-wide majority rule consensus tree for each window size in PHYLIP  
584 (Felsenstein, 2005), with branch support represented by the proportion of trees displaying the  
585 same topology. We simultaneously visualised all trees of equal sized windows using  
586 DensiTree (Bouckaert, 2010).

587

588 We tested whether discordant phylogenetic topologies may be linked to GC content in  
589 the 50kb windows. To do this, we calculated the GC content for each window and binned the  
590 windows into three bins: The 33% with the lowest levels of GC content, the 33% with  
591 intermediate levels, and the 33% with the highest levels of GC content.

592

## 593 **Quantifying Introgression via Branch Lengths (QuIBL)**

594 To test hypotheses of whether phylogenetic discordance between all possible triplets  
595 can be explained by incomplete lineage sorting (ILS) alone, or by a combination of ILS and  
596 gene flow, we implemented QuIBL (Edelman et al., 2019) in two different datasets. The first  
597 dataset leveraged the results of the above 50 kb-window analysis, by taking every twentieth  
598 tree from the 50kb sliding-window analysis and running it through QuIBL. The second  
599 dataset was created specifically for this test, and contained topologies generated from 20 kb  
600 windows with a 1 Mb slide using the phylogenetic methods mentioned above. We ran QuIBL  
601 specifying the baiji as the overall outgroup (totaloutgroup), to test either ILS or ILS with gene  
602 flow (numdistributions 2), the number of total EM steps as 50 (numsteps), and a likelihood  
603 threshold of 0.01. We determined the significance of gene flow by comparing the BIC1 (ILS  
604 alone) and BIC2 (ILS and gene flow). When BIC2 was lower than BIC1, with a difference of  
605  $> 10$ , we assumed incongruent topologies arose due to both ILS and gene flow. Triplet  
606 topologies supporting the species tree, and those that had  $< 5$  alternative topologies, were  
607 excluded from interpretations.

608

## 609 **D-statistics**

610 To test for signs of gene flow in the face of ILS, we ran D-statistics (Durand et al.,  
611 2011; Green et al., 2010) using all individuals mapped to the baiji genome in ANGSD, and  
612 using a consensus base call approach (-doabbababa 2), specifying the baiji sequence as the  
613 ancestral outgroup sequence, and the same filtering as for the fasta file construction with the

614 addition of setting the block size as 1Mb (-blocksize). Significance of the results was  
615 evaluated using a block jackknife approach with the Rscript provided in the ANGSD  
616 package.  $|Z| > 3$  was deemed significant.

617

### 618 **D-foil**

619 As D-statistics only tests for the presence and not the direction of gene flow, we ran  
620 D-foil (Pease and Hahn, 2015), an extended version of the D-statistic, which is a five-taxon  
621 test for gene flow, making use of all four combinations of the potential D-statistics  
622 topologies. For this analysis, we used the same fasta files constructed above, which we  
623 converted into an mvf file using MVFtools (Pease and Rosenzweig, 2018). We specified the  
624 5-taxa [[H1, H2], [H3, H4], baiji], for all possible combinations, following the species tree  
625 (Fig. 1) and a 100 kb window size. All scaffolds were trimmed to the nearest 100 kb to avoid  
626 the inclusion of windows shorter than 100 kb. The significance of each window was  
627 separately assessed by a chi-squared goodness-of-fit test within the software.

628

### 629 **The *f*-branch statistic**

630 To aid in the interpretation of the multitude of D-statistics comparisons, we  
631 implemented the *f*-branch test (Malinsky et al., 2021, 2018) to uncover correlations between  
632 results that may indicate ancestral gene flow events. For this analysis, we needed a variant  
633 call file (VCF). However, the raw sequencing reads for the baiji are not available. To  
634 overcome this, we simulated 100 million 150 bp reads from the assembled genome using  
635 SAMtools wgsim, which we mapped back to the baiji assembly using the same mapping  
636 parameters specified above. We constructed a multi-individual VCF of all individuals  
637 mapped to the baiji using bcftools mpileup, and filtered said VCF file to only include SNPs  
638 using BCFtools call and the -mv parameter, resulting in 138,715,767 sites for downstream  
639 analyses. We ran the multi-individual VCF through Dtrios in **Dsuite v0.4** (Malinsky et al.,  
640 2021) and specified the species tree as the most common topology from our sliding window  
641 analyses, and otherwise default parameters. **We ran the output from Dtrios through *f*-branch**  
642 **and visualised the output using the dtools.py script from Dsuite. To assess whether sex**  
643 **chromosomes may support a different scenario of gene flow events, we also ran the *f*-branch**  
644 **on scaffolds >1 Mb aligning to the X chromosome which gave us 3,728,572 sites.**

645

### 646 **Mutation rate estimation**

647 For use in the downstream demographic analyses, we computed the mutation rate per  
648 generation for each species. To do this, we estimated the pairwise distances between all  
649 ingroup species mapped to the baiji, using a consensus base call in ANGSD (-doIBS 2), and  
650 applying the same filters as above, with the addition of only considering sites in which all  
651 individuals were covered (-minInd). The pairwise distances used in this calculation were  
652 those from the closest lineage to the species of interest (Supplementary Tables S10 and S11).  
653 The mutation rates per generation were calculated using the resultant pairwise distance as  
654 follows: mutation rate = pairwise distance x generation time / 2 x divergence time.  
655 Divergence times were taken from the full dataset 10-partition auto-correlated rate (mean)  
656 values from **McGowen et al. (McGowen et al., 2020)** (Supplementary Table S11). Generation  
657 times were taken from previously published data (Supplementary Table S12).

658

### 659 **Cessation of lineage sorting and/or gene flow**

660 To estimate when lineage sorting and/or gene flow may have ceased between each  
661 species pair, we used the F1-hybrid PSMC (hPSMC) approach (Cahill et al., 2016). As input  
662 we used the haploid consensus sequences mapped to the baiji that were created for the  
663 phylogenetic analyses. Despite the possibility of producing consensus sequences when  
664 mapping to conspecific reference genomes, we chose the baiji for all comparisons, as  
665 previous analyses have shown the choice of reference genome does not influence hPSMC  
666 results (Moodley et al., 2020; Westbury et al., 2019). We merged the haploid sequences from  
667 each possible species pair into pseudo-diploid sequences using the scripts available in the  
668 hPSMC toolsuite. We independently ran each resultant species pair pseudo-diploid sequences  
669 through PSMC, specifying atomic intervals 4+25\*2+4+6. We plotted the results using the  
670 average (i) mutation rate per generation and (ii) generation time for each species pair being  
671 tested. From the output of this analysis, we visually estimated the pre-divergence  $N_e$  of each  
672 hPSMC plot (i.e.  $N_e$  prior to the point of asymptotic increase in  $N_e$ ) to be used as input for  
673 downstream simulations. Based on these empirical results, we ran simulations in ms (Hudson,  
674 2002) using the estimated pre-divergence  $N_e$ , and various predefined divergence times, to  
675 find the interval in which gene flow may have ceased between a given species pair. The time  
676 intervals and pre-divergence  $N_e$  for each species pair used for the simulations can be seen in  
677 supplementary table S7. The ms commands were produced using the scripts available in the  
678 hPSMC toolsuite. We plotted the simulated and empirical hPSMC results to find the  
679 simulations with an asymptotic increase in  $N_e$  closest to, but not overlapping with, the  
680 empirical data. The predefined divergence times of the simulations showing this pattern  
681 within 1.5x and 10x of the pre-divergence  $N_e$  were taken as the time interval in which gene  
682 flow ceased.

683

684 We repeated the above analysis for three species pairs: bottlenose/Indo-Pacific  
685 bottlenose dolphins, beluga/narwhal, and beluga/bottlenose dolphin, but with an additional  
686 step, where we masked repeat elements of the haploid genomes using bedtools v2.26.0  
687 (Quinlan, 2014) and the repeat annotations available on Genbank. Once we masked the repeat  
688 elements, we re-ran the hPSMC analysis as above.

689

### 690 **Relative divergence times in Delphinidae**

691 To further examine the timing of the ending of lineage sorting and/or gene flow, we  
692 performed phylogenetic inferences to uncover the relative divergence times on subsets of  
693 genomic loci showing alternative topologies in Delphinidae. To do this, we masked repeats in  
694 the same fasta files used for our other phylogenetic analyses using the baiji Genbank  
695 annotation and bedtools (Quinlan, 2014). We extracted 1 kb windows with a 1 Mb slide from  
696 the aligned fasta files and only kept loci containing less than 50% missing data for any  
697 individual. We separated our data set into the loci that supported each of four sets of  
698 relationships. These included loci that supported (i) the consensus species tree ( $n = 109$ ), (ii)  
699 the Pacific white-sided dolphin as sister to the killer-whale ( $n = 84$ ), (iii) the Pacific white-  
700 sided dolphin as sister to the clade of bottlenose dolphins, with the long-finned pilot and

701 killer whales in a monophyletic clade as sisters to this group (n = 48), and (iv) the Pacific  
702 white-sided dolphin as sister to the long-finned pilot whale (n = 59).

703

704 As focal species, we selected to test the Pacific white-sided dolphin, killer whale, and  
705 long-finned pilot whale, as they showed the highest number of discordances, allowing for a  
706 more balanced comparison of divergence-time estimates among different topologies. For  
707 each of the four sets of loci, we inferred the relative divergence times across our samples of  
708 Delphinidae, also including the beluga and the baiji in the taxon set. We analysed each data  
709 set independently, constrained the tree topology to that of the corresponding set of loci, and  
710 constrained the age of the root to 1. We performed Bayesian dating using a GTR+ $\Gamma$   
711 substitution model and an uncorrelated-gamma relaxed clock model in MCMCtree, as  
712 implemented in PAML v4.8 (Yang, 2007). The posterior distribution was approximated using  
713 Markov chain Monte Carlo (MCMC) sampling, with samples drawn every  $10^3$  MCMC steps  
714 over  $10^7$  steps, after discarding a burn-in phase of  $10^5$  steps. Convergence to the stationary  
715 distribution was verified by comparing parameter estimates from two independent analyses,  
716 and confirming that effective sample sizes were above 200 for all sampled parameters.

717

### 718 **Heterozygosity**

719 As a proxy for species-level genetic diversity, we estimated autosome-wide  
720 heterozygosity for each of the nine Delphinoidea species. We estimated autosomal  
721 heterozygosity using allele frequencies (-doSaf 1) in ANGSD (Korneliussen et al., 2014),  
722 taking genotype likelihoods into account (-GL 2) and specifying the same filters as for the  
723 fasta file construction, with the addition of adjusting quality scores around indels (-baq 1). To  
724 ensure comparability between genomes of differing coverage, we uniquely set the subsample  
725 filter (-downSample) for each individual to result in a 20x genome-wide coverage.  
726 Heterozygosity was computed from the output of this using realSFS from the ANGSD  
727 toolsuite and specifying 20 Mb windows of covered sites (-nSites).

728

### 729 **Demographic reconstruction**

730 To determine the demographic histories of all nine species over a two million year  
731 time scale, we ran a Pairwise Sequentially Markovian Coalescent model (PSMC) (Li and  
732 Durbin, 2011) on each diploid genome independently. We called diploid genome sequences  
733 using SAMtools and BCFtools v1.6 (Narasimhan et al., 2016), specifying a minimum quality  
734 score of 20 and minimum coverage of 10. We ran PSMC specifying atomic intervals  
735  $4+25*2+4+6$  and performed 100 bootstrap replicates to investigate support for the resultant  
736 demographic trajectories. PSMC outputs were plotted using species-specific mutation rates  
737 and generation times (Supplementary Table S12).

738

### 739 **Figure legends:**

740

741 **Figure 1: Sliding-Window Maximum likelihood trees of nine Delphinoidea species and**  
742 **the baiji.** The trees were constructed using non-overlapping sliding windows of (A) 50 kb in  
743 length and (B) 1 Mb in length. Black lines show the consensus tree, grey lines show  
744 individual trees. Numbers on branches show the proportion of windows supporting the node.

745 Branches without numbers had 100% support. Bottlenose dolphin silhouette: license Public  
746 Domain Dedication 1.0; remaining Delphinoidea silhouettes: Chris huh, license CC-BY-SA-  
747 3.0 (<https://creativecommons.org/licenses/by-sa/3.0/>).

748

749 **Figure 2: Genome-wide *f*-branch results.** (A) Species tree; (B) and (C) Species tree in  
750 expanded form, with internal branches as dotted lines. The values in the matrix refer to  
751 excess allele sharing between the expanded tree branch (relative to its sister branch) and the  
752 species on the *x*-axis. Lines connecting branches show: (B) gene flow events inferred directly  
753 from the *f*-branch results; (C) gene flow events hypothesised from the *f*-branch results, while  
754 accounting for (i) the inability to detect gene flow between sister lineages, and (ii) a lack of a  
755 positive means less gene flow relative to the sister lineage, rather than no gene flow.

756

757 **Figure 3: Estimated divergence times and time intervals during which gene flow ceased**  
758 **between species (A) within families and (B) between families.** Estimated time intervals of  
759 when gene flow ceased between species pairs are based on hPSMC results and simulated  
760 data. Divergence time estimates are taken from McGowen et al 2020.

761

762 **Figure 4: Autosome-wide heterozygosity and demographic histories over the past two**  
763 **million years.** (A) Autosome-wide levels of heterozygosity calculated in 20 Mb sliding  
764 windows. (B-D) Demographic history of all studied species within (B) Delphinidae, (C)  
765 Phocoenidae, and (D) Monodontidae, estimated using PSMC. Thick coloured lines show  
766 estimated demographic trajectory, faded lines show bootstrap support values. Colours of B-D  
767 correspond to species' colour from A.

768

## 769 Acknowledgements

770 The work was supported by the Independent Research Fund Denmark | Natural Sciences,  
771 Forskningsprojekt 1, grant no. 8021-00218B and the Villum Fonden Young Investigator  
772 Programme, grant no. 13151 to EDL. AAC was funded by the Rubicon-NWO grant (project  
773 019.183EN.005). We would like to thank all those contributing to the ever-increasing  
774 abundance of publicly available genomic resources. Without the availability of such data, our  
775 study would not have been possible. We would also like to thank Michael Fontaine,  
776 Christelle Fraïsse, Camille Roux, Andrew Foote, and Simon Martin for their helpful input to  
777 previous versions of this manuscript.

778

## 779 Author contributions

780 Conceptualization, MVW; Formal analysis, MVW, AAC, AR-I, BDC, DAD, SH; Writing –  
781 Original Draft MVW; Writing – Review & Editing, All authors; Supervision, MVW, EDL;  
782 Funding Acquisition, EDL

783

## 784 References:

785 Árnason Ú, Lammers F, Kumar V, Nilsson MA, Janke A. 2018. Whole-genome sequencing  
786 of the blue whale and other rorquals finds signatures for introgressive gene flow. *Sci Adv*

787 4:eaap9873.

788 Autenrieth M, Hartmann S, Lah L, Roos A, Dennis AB, Tiedemann R. 2018. High-quality  
789 whole-genome sequence of an abundant Holarctic odontocete, the harbour porpoise  
790 (*Phocoena phocoena*). *Mol Ecol Resour* **18**:1469–1481.

791 Baird RW, Gorgone AM, McSweeney DJ, Ligon AD, Deakos MH, Webster DL, Schorr GS,  
792 Martien KK, Salden DR, Mahaffy SD. 2012. Population structure of island-associated  
793 dolphins: Evidence from mitochondrial and microsatellite markers for common  
794 bottlenose dolphins (*Tursiops truncatus*) in the main Hawaiian Islands. *Mar Mamm Sci*.  
795 Barlow A, Cahill JA, Hartmann S, Theunert C, Xenikoudakis G, Fortes GG, Paijmans JLA,  
796 Rabeder G, Frischauf C, Grandal-d’Anglade A, García-Vázquez A, Murtskhvaladze M,  
797 Saarma U, Anijalg P, Skrbinšek T, Bertorelle G, Gasparian B, Bar-Oz G, Pinhasi R,  
798 Slatkin M, Dalén L, Shapiro B, Hofreiter M. 2018. Partial genomic survival of cave  
799 bears in living brown bears. *Nat Ecol Evol* **2**:1563–1570.

800 Bierne N, Bonhomme F, David P. 2003. Habitat preference and the marine-speciation  
801 paradox. *Proc Biol Sci* **270**:1399–1406.

802 Bouckaert RR. 2010. DensiTree: making sense of sets of phylogenetic trees. *Bioinformatics*  
803 **26**:1372–1373.

804 Butlin RK, Smadja CM. 2018. Coupling, Reinforcement, and Speciation. *Am Nat* **191**:155–  
805 172.

806 Cabrera AA, Schall E, Bérubé M, Bachmann L, Berrow S, Best PB, Clapham PJ, Cunha HA,  
807 Rosa LD, Dias C, Findlay KP, Haug T, Heide-Jørgensen MP, Kovacs KM, Landry S,  
808 Larsen F, Lopes XM, Lydersen C, Mattila DK, Oosting T, Pace RM, Papetti C, Paspatis  
809 A, Pastene LA, Prieto R, Ramp C, Robbins J, Ryan C, Sears R, Secchi ER, Silva MA,  
810 Víkingsson G, Wiig Ø, Øien N, Palsbøll PJ. 2018. Strong and lasting impacts of past  
811 global warming on baleen whale and prey abundance. *bioRxiv*.

812 Cahill JA, Soares AER, Green RE, Shapiro B. 2016. Inferring species divergence times using  
813 pairwise sequential Markovian coalescent modelling and low-coverage genomic data.  
814 *Philos Trans R Soc Lond B Biol Sci* **371**. doi:10.1098/rstb.2015.0138

815 Campbell CR, Poelstra JW. 2018. What is Speciation Genomics? The roles of ecology, gene  
816 flow, and genomic architecture in the formation of species. *Biol J Linn Soc Lond*  
817 **124**:561–583.

818 Coyne JA, Orr HA. 2004. Speciation. Sinauer Associates Sunderland, MA.

819 Crossman CA, Taylor EB, Barrett-Lennard LG. 2016. Hybridization in the Cetacea:  
820 widespread occurrence and associated morphological, behavioral, and ecological factors.  
821 *Ecol Evol* **6**:1293–1303.

822 Durand EY, Patterson N, Reich D, Slatkin M. 2011. Testing for ancient admixture between  
823 closely related populations. *Mol Biol Evol* **28**:2239–2252.

824 Edelman NB, Frandsen PB, Miyagi M, Clavijo B, Davey J, Dikow RB, García-Accinelli G,  
825 Van Belleghem SM, Patterson N, Neafsey DE, Challis R, Kumar S, Moreira GRP,  
826 Salazar C, Chouteau M, Counterman BA, Papa R, Blaxter M, Reed RD, Dasmahapatra  
827 KK, Kronforst M, Joron M, Jiggins CD, McMillan WO, Di Palma F, Blumberg AJ,  
828 Wakeley J, Jaffe D, Mallet J. 2019. Genomic architecture and introgression shape a  
829 butterfly radiation. *Science* **366**:594–599.

830 Edwards CJ, Suchard MA, Lemey P, Welch JJ, Barnes I, Fulton TL, Barnett R, O’Connell  
831 TC, Coxon P, Monaghan N, Valdiosera CE, Lorenzen ED, Willerslev E, Baryshnikov  
832 GF, Rambaut A, Thomas MG, Bradley DG, Shapiro B. 2011. Ancient hybridization and  
833 an Irish origin for the modern polar bear matriline. *Curr Biol* **21**:1251–1258.

834 Espada R, Olaya-Ponzzone L, Haasova L, Martín E, García-Gómez JC. 2019. Hybridization in  
835 the wild between *Tursiops truncatus* (Montagu 1821) and *Delphinus delphis* (Linnaeus  
836 1758). *PLoS One* **14**:e0215020.

837 Feder JL, Egan SP, Nosil P. 2012. The genomics of speciation-with-gene-flow. *Trends Genet*  
838 **28**:342–350.

839 Felsenstein J. 2005. PHYLIP (Phylogeny Inference Package) version 3.6.

840 Fish FE, Howle LE, Murray MM. 2008. Hydrodynamic flow control in marine mammals.  
841 *Integr Comp Biol* **48**:788–800.

842 Foote AD. 2018. Sympatric Speciation in the Genomic Era. *Trends Ecol Evol* **33**:85–95.

843 Foote AD, Morin PA. 2015. Sympatric speciation in killer whales? *Heredity* **114**:537–538.

844 Foote AD, Morin PA, Durban JW, Willerslev E. 2011. Out of the Pacific and back again:  
845 insights into the matrilineal history of Pacific killer whale ecotypes. *PLoS*.

846 Grabherr MG, Russell P, Meyer M, Mauceli E, Alföldi J, Di Palma F, Lindblad-Toh K. 2010.  
847 Genome-wide synteny through highly sensitive sequence alignment: Satsuma.  
848 *Bioinformatics* **26**:1145–1151.

849 Green RE, Krause J, Briggs AW, Maricic T, Stenzel U, Kircher M, Patterson N, Li H, Zhai  
850 W, Fritz MH-Y, Hansen NF, Durand EY, Malaspinas A-S, Jensen JD, Marques-Bonet  
851 T, Alkan C, Prüfer K, Meyer M, Burbano HA, Good JM, Schultz R, Aximu-Petri A,  
852 Butthof A, Höber B, Höffner B, Siegemund M, Weihmann A, Nusbaum C, Lander ES,  
853 Russ C, Novod N, Affourtit J, Egholm M, Verna C, Rudan P, Brajkovic D, Kucan Ž,  
854 Gušić I, Doronichev VB, Golovanova LV, Lalueva-Fox C, de la Rasilla M, Fortea J,  
855 Rosas A, Schmitz RW, Johnson PLF, Eichler EE, Falush D, Birney E, Mullikin JC,  
856 Slatkin M, Nielsen R, Kelso J, Lachmann M, Reich D, Pääbo S. 2010. A draft sequence  
857 of the Neandertal genome. *Science* **328**:710–722.

858 Gridley T, Elwen SH, Harris G, Moore DM, Hoelzel AR, Lampen F. 2018. Hybridization in  
859 bottlenose dolphins—A case study of *Tursiops aduncus* × *T. truncatus* hybrids and  
860 successful backcross hybridization events. *PLoS One* **13**:e0201722.

861 Herzog DL, Johnsonz CM. 1997. Interspecific interactions between Atlantic spotted  
862 dolphins (*Stenella frontalis*) and bottlenose dolphins (*Tursiops truncatus*) in the  
863 Bahamas 1985-1995. *Aquat Mamm*.

864 Hudson RR. 2002. Generating samples under a Wright–Fisher neutral model of genetic  
865 variation. *Bioinformatics* **18**:337–338.

866 Jiang H, Lei R, Ding S-W, Zhu S. 2014. Skewer: a fast and accurate adapter trimmer for  
867 next-generation sequencing paired-end reads. *BMC Bioinformatics* **15**:182.

868 Korneliussen TS, Albrechtsen A, Nielsen R. 2014. ANGSD: Analysis of Next Generation  
869 Sequencing Data. *BMC Bioinformatics* **15**:356.

870 Lartillot N. 2013. Phylogenetic patterns of GC-biased gene conversion in placental mammals  
871 and the evolutionary dynamics of recombination landscapes. *Mol Biol Evol* **30**:489–502.

872 Leaché AD, Harris RB, Rannala B, Yang Z. 2014. The influence of gene flow on species tree  
873 estimation: a simulation study. *Syst Biol* **63**:17–30.

874 Li H, Durbin R. 2011. Inference of human population history from individual whole-genome  
875 sequences. *Nature* **475**:493–496.

876 Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows–Wheeler  
877 transform. *Bioinformatics* **25**:1754–1760.

878 Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin  
879 R, 1000 Genome Project Data Processing Subgroup. 2009. The Sequence  
880 Alignment/Map format and SAMtools. *Bioinformatics* **25**:2078–2079.

881 Liu S, Lorenzen ED, Fumagalli M, Li B, Harris K, Xiong Z, Zhou L, Korneliussen TS, Somel  
882 M, Babbitt C, Wray G, Li J, He W, Wang Z, Fu W, Xiang X, Morgan CC, Doherty A,  
883 O’Connell MJ, McInerney JO, Born EW, Dalén L, Dietz R, Orlando L, Sonne C, Zhang  
884 G, Nielsen R, Willerslev E, Wang J. 2014. Population genomics reveal recent speciation  
885 and rapid evolutionary adaptation in polar bears. *Cell* **157**:785–794.

886 Liu S, Westbury MV, Dussex N, Mitchell KJ, Sinding M-HS, Heintzman PD, Duchêne DA,

887 Kapp JD, von Seth J, Heiniger H, Sánchez-Barreiro F, Margaryan A, André-Olsen R, De  
888 Cahsan B, Meng G, Yang C, Chen L, van der Valk T, Moodley Y, Rookmaaker K,  
889 Bruford MW, Ryder O, Steiner C, Bruins-van Sonsbeek LGR, Vartanyan S, Guo C,  
890 Cooper A, Kosintsev P, Kirillova I, Lister AM, Marques-Bonet T, Gopalakrishnan S,  
891 Dunn RR, Lorenzen ED, Shapiro B, Zhang G, Antoine P-O, Dalén L, Gilbert MTP.  
892 2021. Ancient and modern genomes unravel the evolutionary history of the rhinoceros  
893 family. *Cell* **184**:4874–4885.e16.

894 Malinsky M, Matschiner M, Svardal H. 2021. Dsuite - Fast D-statistics and related admixture  
895 evidence from VCF files. *Mol Ecol Resour* **21**:584–595.

896 Malinsky M, Svardal H, Tyers AM, Miska EA, Genner MJ, Turner GF, Durbin R. 2018.  
897 Whole-genome sequences of Malawi cichlids reveal multiple radiations interconnected  
898 by gene flow. *Nat Ecol Evol* **2**:1940–1955.

899 Martin SH, Davey JW, Jiggins CD. 2015. Evaluating the use of ABBA–BABA statistics to  
900 locate introgressed loci. *Mol Biol.*

901 McGowen MR, Tsagkogeorga G, Álvarez-Carretero S, Dos Reis M, Struebig M, Deaville R,  
902 Jepson PD, Jarman S, Polanowski A, Morin PA, Rossiter SJ. 2020. Phylogenomic  
903 Resolution of the Cetacean Tree of Life Using Target Sequence Capture. *Syst Biol*  
904 **69**:479–501.

905 Mendes FK, Hahn MW. 2016. Gene Tree Discordance Causes Apparent Substitution Rate  
906 Variation. *Syst Biol* **65**:711–721.

907 Miralles L, Oremus M, Silva MA, Planes S, Garcia-Vazquez E. 2016. Interspecific  
908 Hybridization in Pilot Whales and Asymmetric Genetic Introgression in Northern  
909 Globicephala melas under the Scenario of Global Warming. *PLoS One* **11**:e0160080.

910 Miyazaki N, Hirosaki Y, Kinuta T, Omura H. 1992. Osteological study of a hybrid between  
911 *Tursiops truncatus* and *Grampus griseus*. *Bull Natl Mus Nat Sci Ser B Bot* **18**:79–94.

912 Moodley Y, Westbury MV, Russo I-RM, Gopalakrishnan S, Rakotoarivelo A, Olsen R-A,  
913 Prost S, Tunstall T, Ryder OA, Dalén L, Bruford MW. 2020. Interspecific gene flow and  
914 the evolution of specialisation in black and white rhinoceros. *Mol Biol Evol.*  
915 doi:10.1093/molbev/msaa148

916 Moura AE, Kenny JG, Chaudhuri RR, Hughes MA. 2015. Phylogenomics of the killer whale  
917 indicates ecotype divergence in sympatry. *Heredity* **114**:48–55.

918 Moura AE, Shreves K, Pilot M, Andrews KR, Moore DM, Kishida T, Möller L, Natoli A,  
919 Gaspari S, McGowen M, Chen I, Gray H, Gore M, Culloch RM, Kiani MS, Willson MS,  
920 Bulushi A, Collins T, Baldwin R, Willson A, Minton G, Ponnampalam L, Hoelzel AR.  
921 2020. Phylogenomics of the genus *Tursiops* and closely related Delphininae reveals  
922 extensive reticulation among lineages and provides inference about eco-evolutionary  
923 drivers. *Mol Phylogenet Evol* **146**:106756.

924 Narasimhan V, Danecek P, Scally A, Xue Y, Tyler-Smith C, Durbin R. 2016. BCFtools/RoH:  
925 a hidden Markov model approach for detecting autozygosity from next-generation  
926 sequencing data. *Bioinformatics* **32**:1749–1751.

927 Norris RD, Hull PM. 2012. The temporal dimension of marine speciation. *Evol Ecol* **26**:393–  
928 415.

929 Palumbi SR. 1994. Genetic divergence, reproductive isolation, and marine speciation. *Annu*  
930 *Rev Ecol Syst* **25**:547–572.

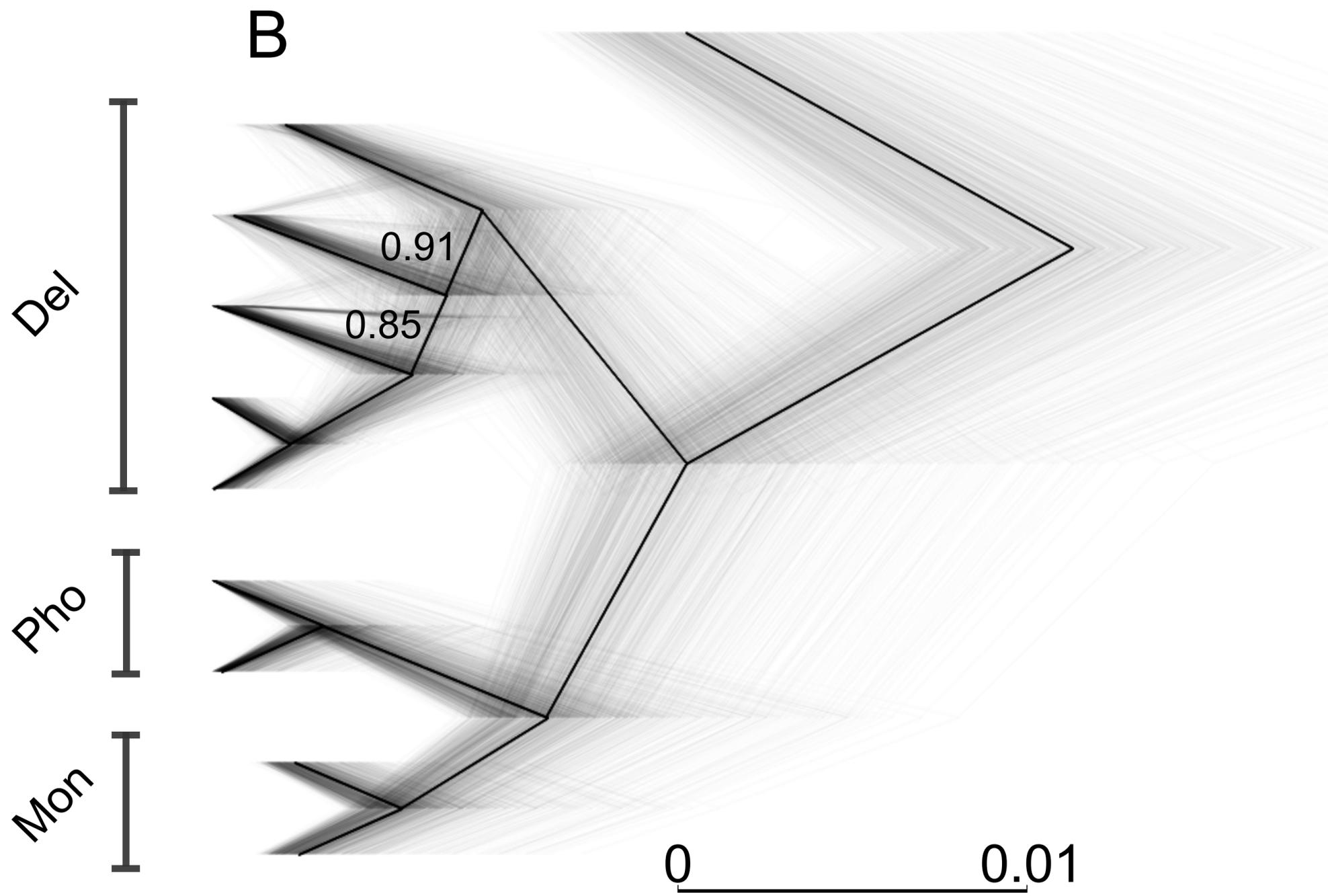
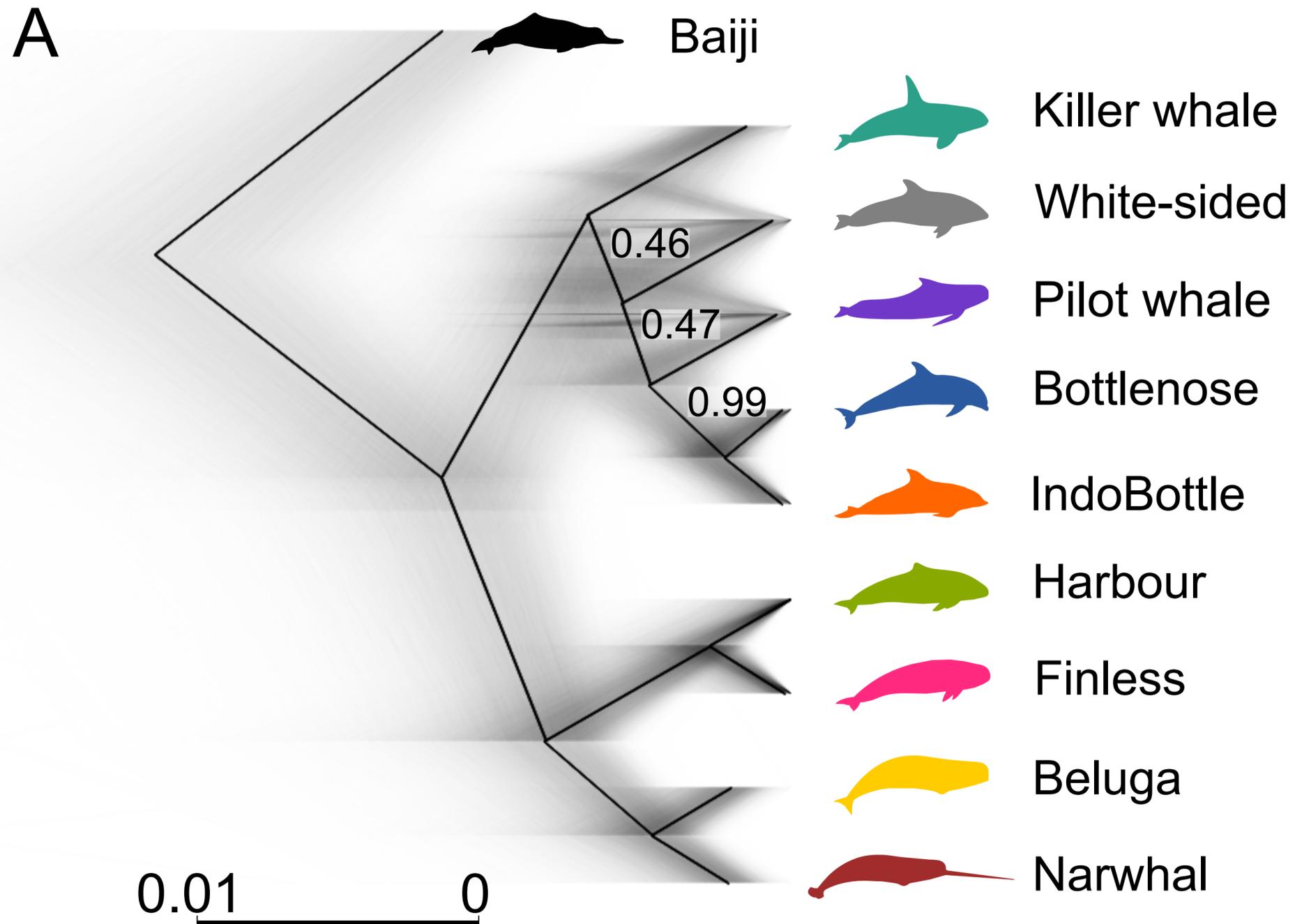
931 Pamilo P, Nei M. 1988. Relationships between gene trees and species trees. *Mol Biol Evol*  
932 **5**:568–583.

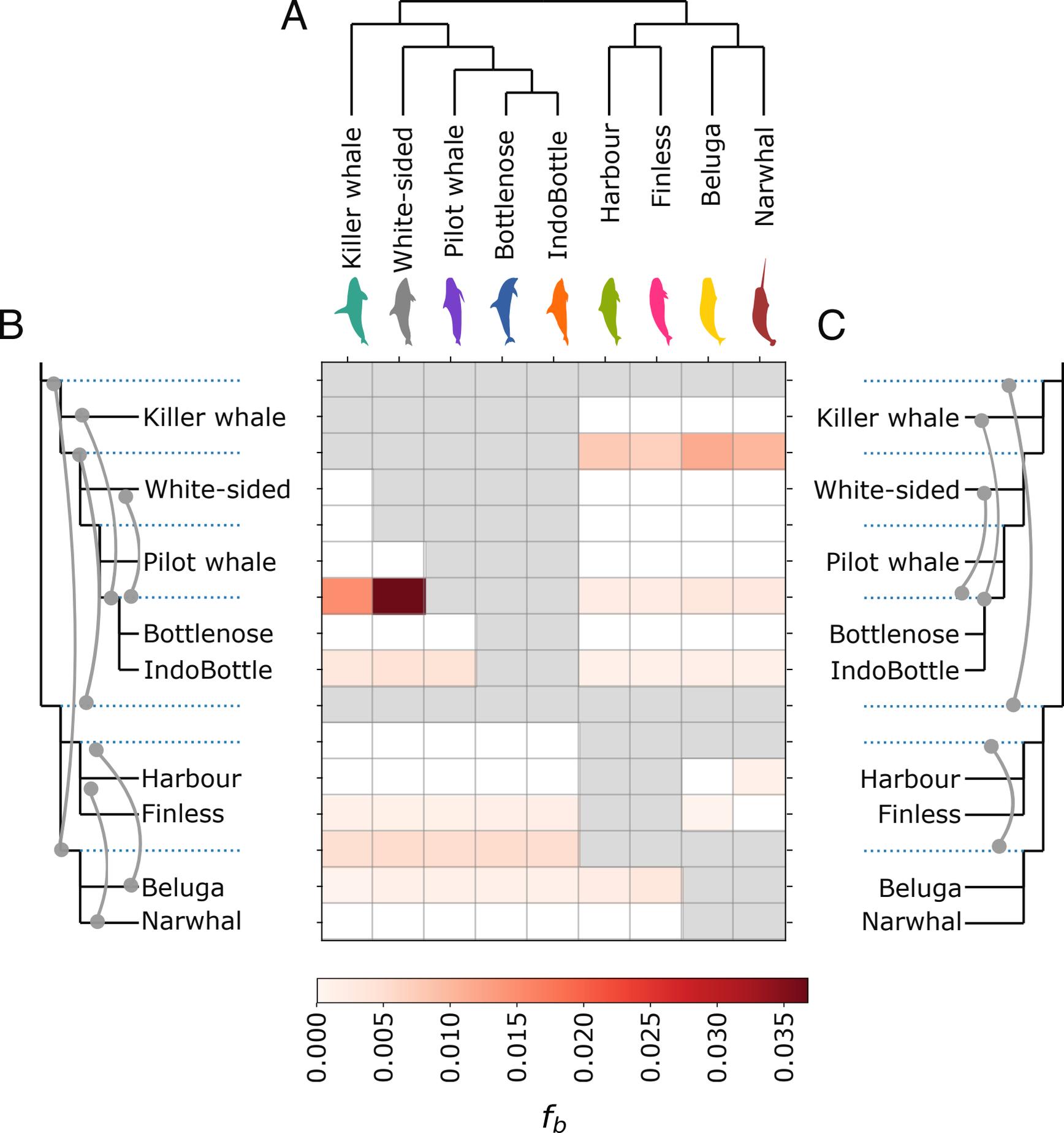
933 Payseur BA, Rieseberg LH. 2016. A genomic perspective on hybridization and speciation.  
934 *Mol Ecol* **25**:2337–2360.

935 Pease JB, Hahn MW. 2015. Detection and Polarization of Introgression in a Five-Taxon  
936 Phylogeny. *Syst Biol* **64**:651–662.

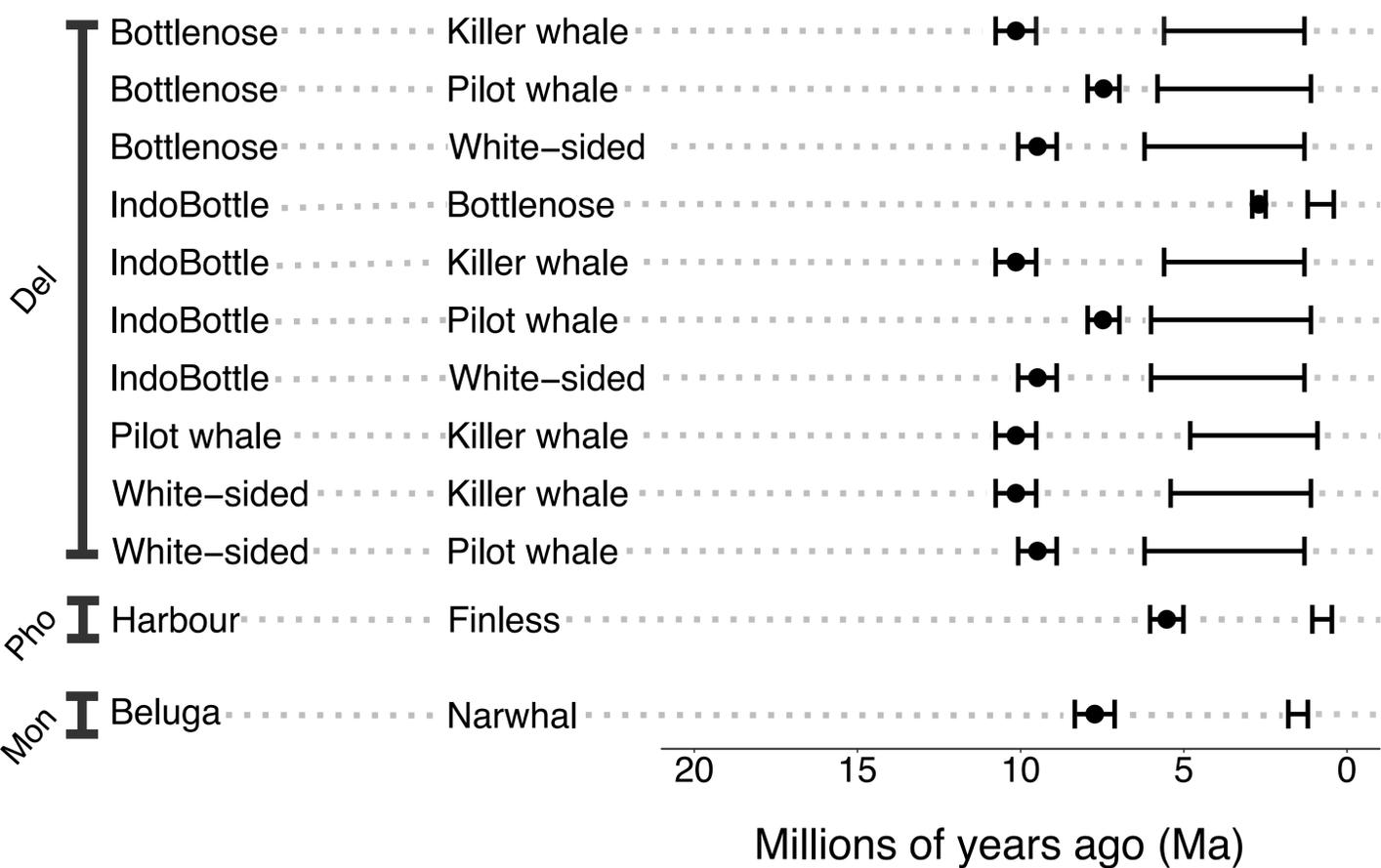
- 937 Pease JB, Rosenzweig BK. 2018. Encoding Data Using Biological Principles: The  
938 Multisample Variant Format for Phylogenomics and Population Genomics. *IEEE/ACM*  
939 *Trans Comput Biol Bioinform* **15**:1231–1238.
- 940 Polyak VJ, Onac BP, Fornós JJ, Hay C, Asmerom Y, Dorale JA, Ginés J, Tuccimei P, Ginés  
941 A. 2018. A highly resolved record of relative sea level in the western Mediterranean Sea  
942 during the last interglacial period. *Nat Geosci* **11**:860–864.
- 943 Quinlan AR. 2014. BEDTools: The Swiss-Army Tool for Genome Feature Analysis. *Curr*  
944 *Protoc Bioinformatics* **47**:11.12.1–34.
- 945 Silva JM, Silva FJL, Sazima I. 2005. Two presumed interspecific hybrids in the genus  
946 *Stenella* (Delphinidae) in the Tropical West Atlantic. *Aquat Mamm* **31**:468.
- 947 Skovrind M, Castruita JAS, Haile J, Treadaway EC, Gopalakrishnan S, Westbury MV,  
948 Heide-Jørgensen MP, Szpak P, Lorenzen ED. 2019. Hybridization between two high  
949 Arctic cetaceans confirmed by genomic analysis. *Sci Rep* **9**:7729.
- 950 Slatkin M, Pollack JL. 2008. Subdivision in an ancestral species creates asymmetry in gene  
951 trees. *Mol Biol Evol* **25**:2241–2246.
- 952 Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of  
953 large phylogenies. *Bioinformatics* **30**:1312–1313.
- 954 Steeman ME, Hebsgaard MB, Fordyce RE, Ho SYW, Rabosky DL, Nielsen R, Rahbek C,  
955 Glenner H, Sørensen MV, Willerslev E. 2009. Radiation of extant cetaceans driven by  
956 restructuring of the oceans. *Syst Biol* **58**:573–585.
- 957 Stone G, Florez-Gonzalez L, Katona S. 1990. Whale migration record. *Nature* **346**:705–705.
- 958 Turelli M, Barton NH, Coyne JA. 2001. Theory and speciation. *Trends Ecol Evol* **16**:330–  
959 343.
- 960 Westbury MV, Hartmann S, Barlow A, Preick M, Ridush B, Nagel D, Rathgeber T, Ziegler  
961 R, Baryshnikov G, Sheng G, Ludwig A, Wiesel I, Dalen L, Bibi F, Werdelin L, Heller  
962 R, Hofreiter M. 2020. Hyena paleogenomes reveal a complex evolutionary history of  
963 cross-continental gene flow between spotted and cave hyena. *Science Advances*  
964 **6**:eaay0456.
- 965 Westbury MV, Petersen B, Lorenzen ED. 2019. Genomic analyses reveal an absence of  
966 contemporary introgressive admixture between fin whales and blue whales, despite  
967 known hybrids. *PLoS One* **14**:e0222004.
- 968 Williams TM. 1999. The evolution of cost efficient swimming in marine mammals: limits to  
969 energetic optimization. *Philosophical Transactions of the Royal Society of London*  
970 *Series B: Biological Sciences* **354**:193–201.
- 971 Willis PM, Crespi BJ, Dill LM, Baird RW, Hanson MB. 2004. Natural hybridization between  
972 Dall’s porpoises (*Phocoenoides dalli*) and harbour porpoises (*Phocoena phocoena*). *Can*  
973 *J Zool* **82**:828–834.
- 974 Yang Z. 2007. PAML 4: phylogenetic analysis by maximum likelihood. *Mol Biol Evol*  
975 **24**:1586–1591.
- 976 Zheng Y, Janke A. 2018. Gene flow analysis method, the D-statistic, is robust in a wide  
977 parameter space. *BMC Bioinformatics* **19**:10.

978

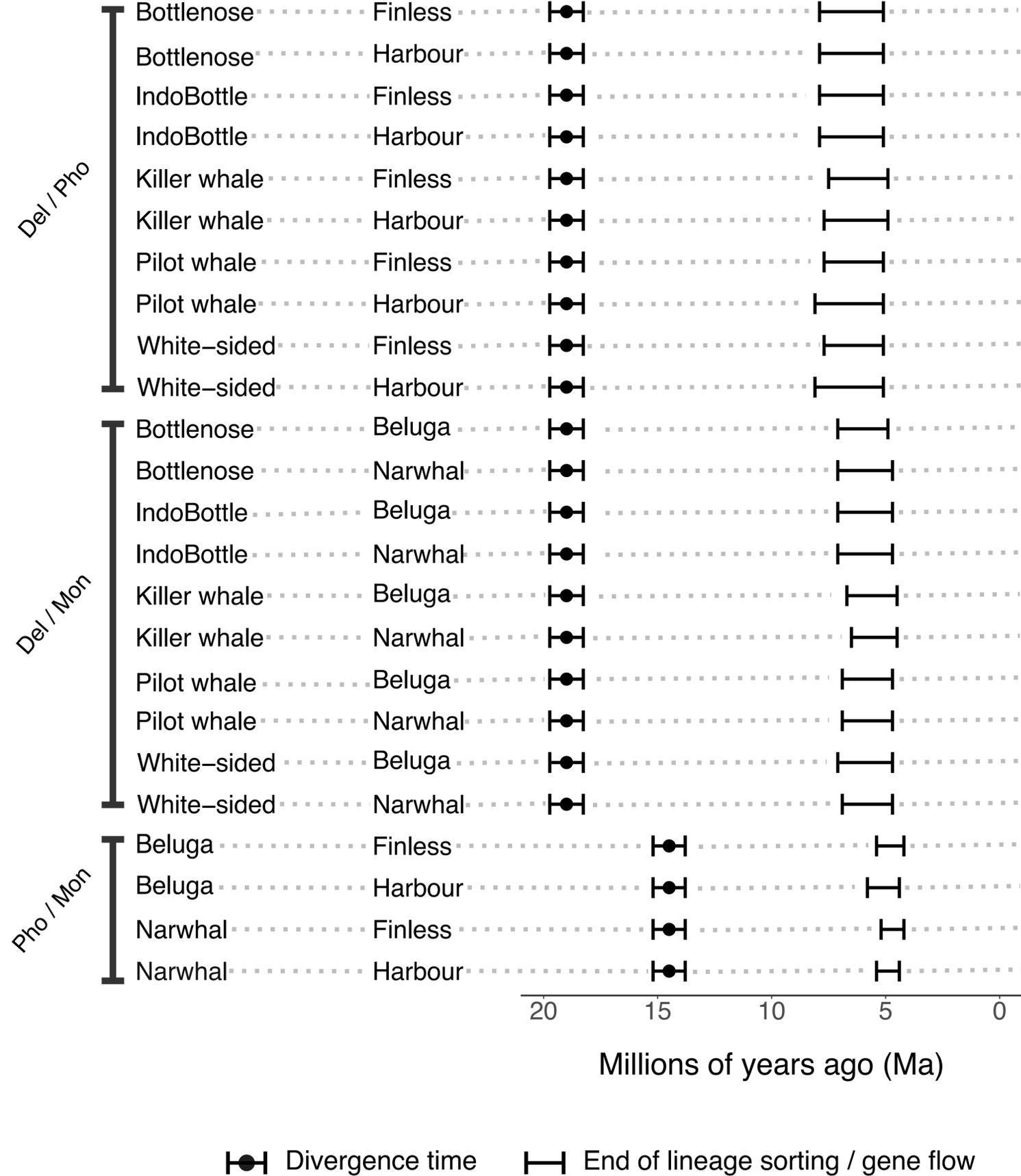




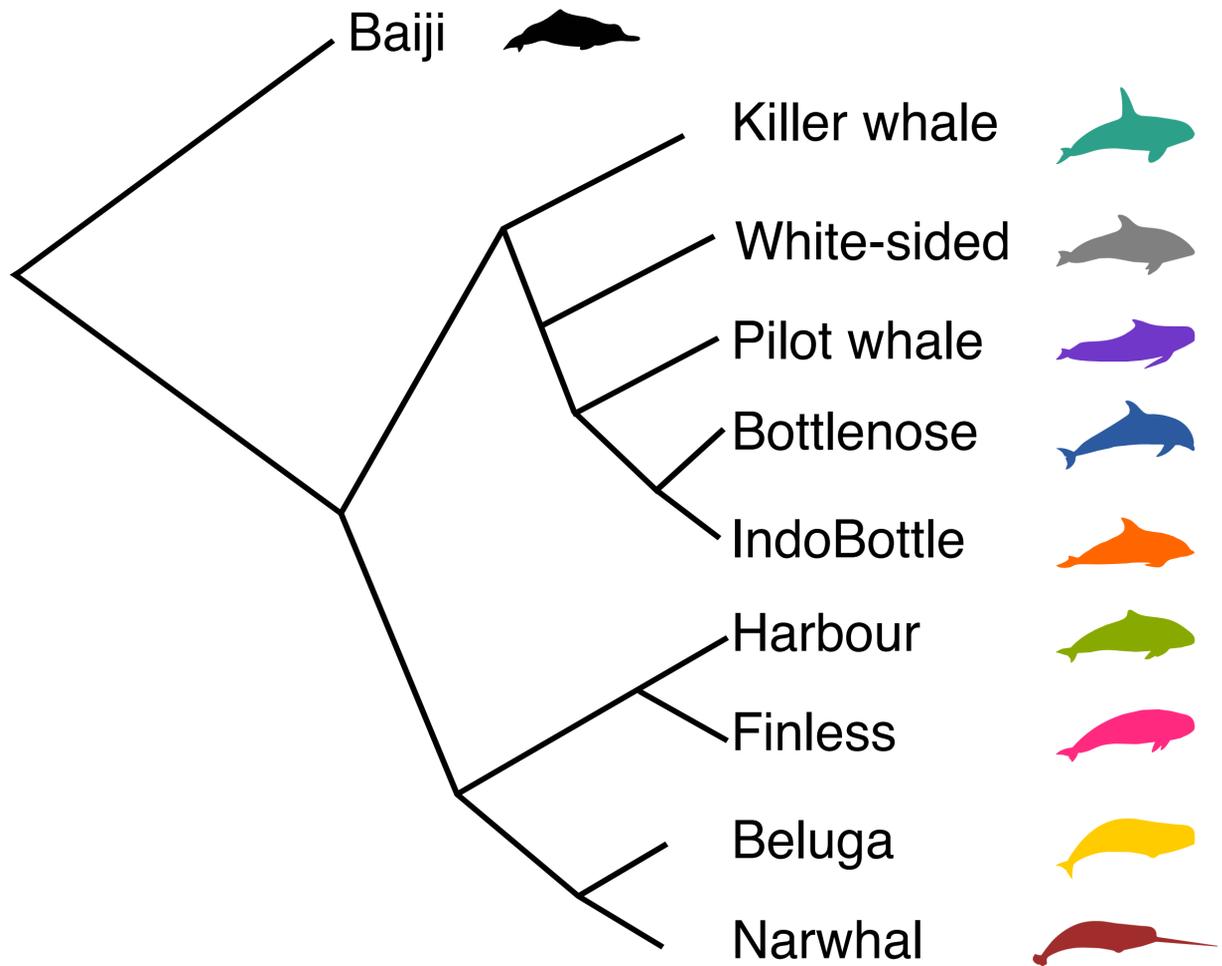
### A Within families

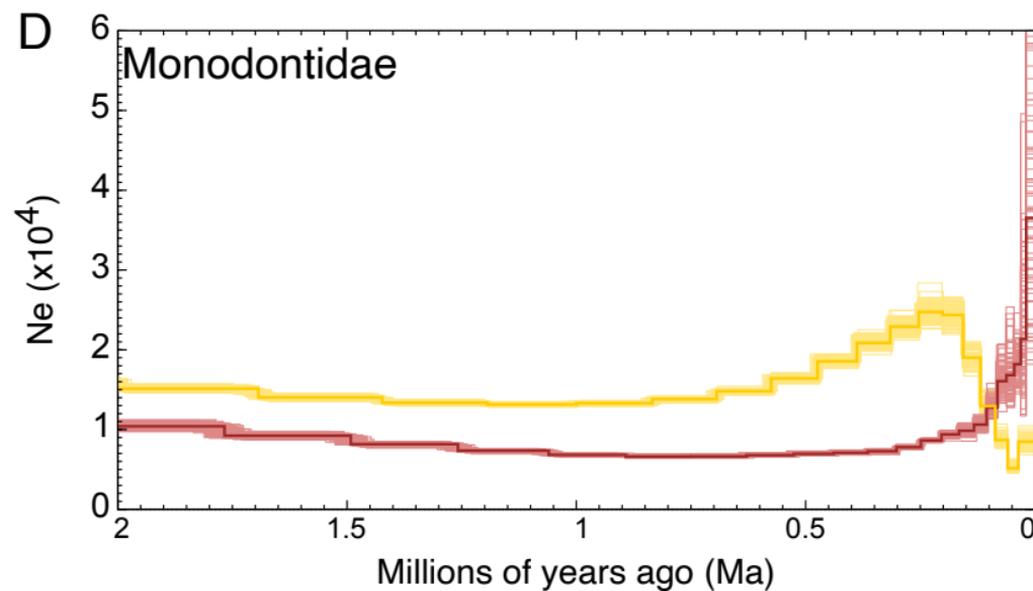
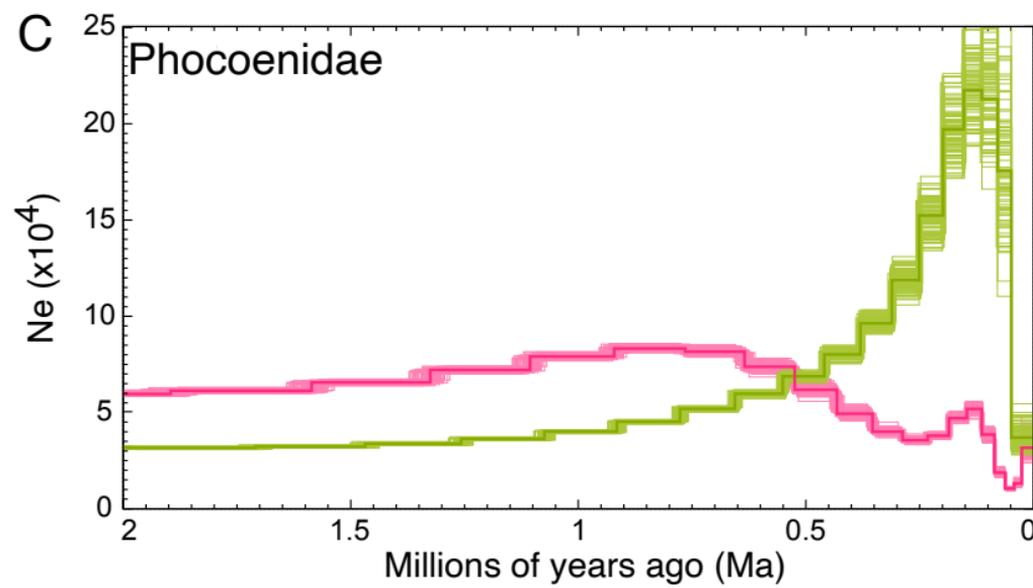
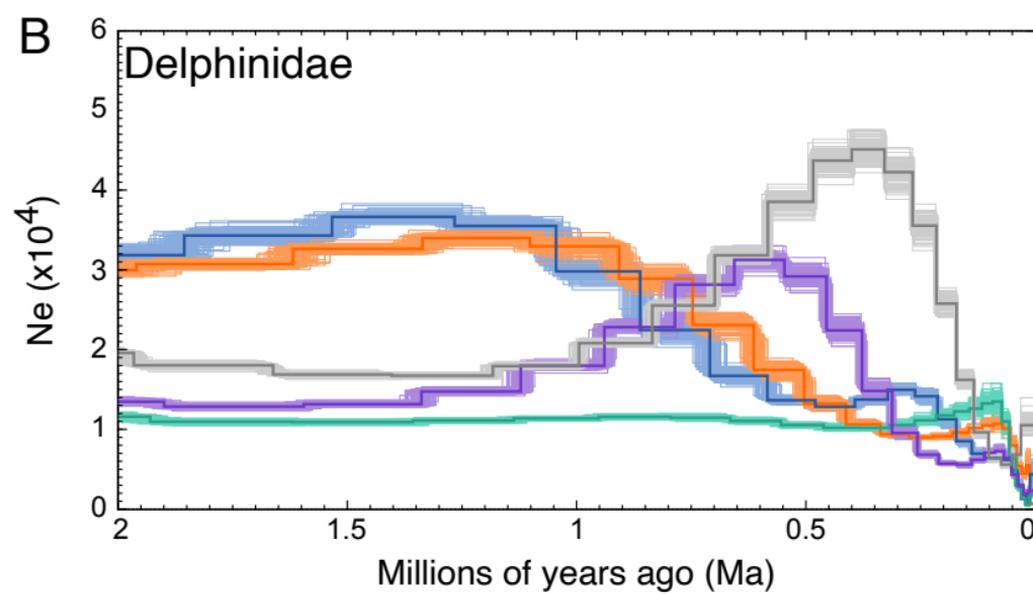
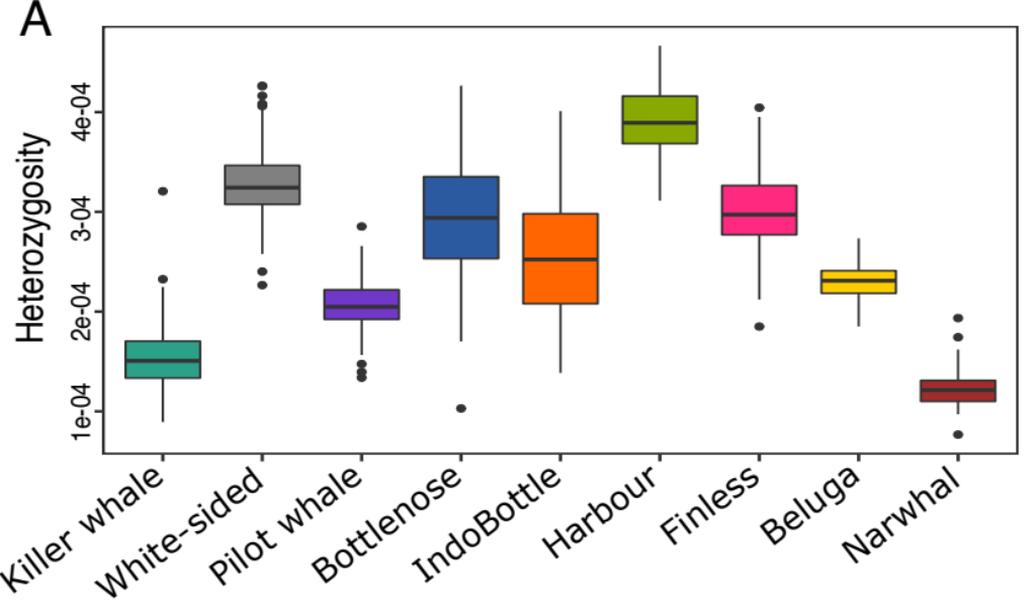


### B Between families



### C





## Supplementary information

**Supplementary table S1:** Proportions of the most frequent five topologies based on window sizes. NA - not in the five most frequent for that window size. Whitesided - Pacific white-sided dolphin, Pilotwhale - long-finned pilot whale, IndoBottlenose - Indo-Pacific bottlenose dolphin, Bottlenose - bottlenose dolphin, Killerwhale - killer whale, Beluga - beluga, Narwhal - narwhal, Harbour - harbour porpoise, Finless - finless porpoise, Baiji - Baiji (outgroup).

| 50kb | 100kb | 500kb | 1Mb  | Topology  |
|------|-------|-------|------|---|
| 0.24 | 0.32  | 0.64  | 0.79 | ((((Whitesided,(Pilotwhale,(IndoBottlenose,Bottlenose))),Killerwhale),(Beluga,Narwhal),(Harbour,Finless))),Baiji);  |
| 0.14 | 0.14  | 0.09  | 0.05 | ((((Pilotwhale,(IndoBottlenose,Bottlenose)),(Whitesided,Killerwhale),(Beluga,Narwhal),(Harbour,Finless))),Baiji);   |
| 0.13 | 0.14  | 0.14  | 0.10 | ((((Pilotwhale,(Whitesided,(IndoBottlenose,Bottlenose))),Killerwhale),(Beluga,Narwhal),(Harbour,Finless))),Baiji);  |
| 0.09 | 0.08  | 0.04  | 0.02 | ((((Pilotwhale,Whitesided),(IndoBottlenose,Bottlenose)),Killerwhale),(Beluga,Narwhal),(Harbour,Finless))),Baiji);   |
| 0.08 | NA    | NA    | NA   | ((((Killerwhale,(Pilotwhale,(IndoBottlenose,Bottlenose))),Whitesided),(Beluga,Narwhal),(Harbour,Finless))),Baiji);  |
| NA   | 0.07  | 0.03  | 0.02 | ((((Whitesided,((Pilotwhale,(IndoBottlenose,Bottlenose))),Killerwhale),(Beluga,Narwhal),(Harbour,Finless))),Baiji); |
| 0.69 | 0.76  | 0.94  | 0.98 | Top 5 topologies combined   |

**Supplementary table S2:** Proportions of the most frequent five topologies based on GC content and a window size of 50kb. NA - not in the five most frequent for that window size. Whitesided - Pacific white-sided dolphin, Pilotwhale - long-finned pilot whale, IndoBottlenose - Indo-Pacific bottlenose dolphin, Bottlenose - bottlenose dolphin, Killerwhale - killer whale, Beluga - beluga, Narwhal - narwhal, Harbour - harbour porpoise, Finless - finless porpoise, Baiji - Baiji (outgroup).

| Low GC | Medium GC | High GC | Topology   |
|--------|-----------|---------|--|
| 2814   | 3395      | 4227    | ((((Killerwhale,(Whitesided,((IndoBottlenose,Bottlenose),Pilotwhale))),((Beluga,Narwhal),(Harbour,Finless))),Baiji); |
| 2023   | 2107      | 2085    | ((((Pilotwhale,(IndoBottlenose,Bottlenose)),(Whitesided,Killerwhale),(Beluga,Narwhal),(Harbour,Finless))),Baiji);    |
| 1740   | 1898      | 1976    | ((((Pilotwhale,(Whitesided,(IndoBottlenose,Bottlenose))),Killerwhale),(Beluga,Narwhal),(Harbour,Finless))),Baiji);   |
| 1287   | 1289      | 1317    | ((((Pilotwhale,Whitesided),(IndoBottlenose,Bottlenose)),Killerwhale),(Beluga,Narwhal),(Harbour,Finless))),Baiji);    |
| 1152   | NA        | NA      | ((((Whitesided,(IndoBottlenose,Bottlenose)),(Pilotwhale,Killerwhale),(Beluga,Narwhal),(Harbour,Finless))),Baiji);    |
| NA     | 1190      | 1149    | ((((Whitesided,((Pilotwhale,(IndoBottlenose,Bottlenose))),Killerwhale),(Beluga,Narwhal),(Harbour,Finless))),Baiji);  |

**Supplementary table S3:** QuIBL results when using every twentieth tree from the 50kb sliding window analysis - attached as spreadsheet. QuIBL analyses all triplet combinations ((A, B), C) in a given set of phylogenetic trees. Here we only present the alternative topologies within Delphinidae, that are in disagreement with the species tree, and may have arisen due to ILS or gene flow. The gene flow pair shows individuals A and B and outgroup is C. Two BIC scores are presented - one for ILS alone and one for ILS and gene flow. A BIC difference >10 suggests ILS and gene flow both as factors giving rise to the discordance topologies. % of total trees shows the percentage of all trees in the dataset having said triplet topology, whereas % of trees supporting topology explained by gene flow shows the percentage of the trees supporting said triplet topology that likely arose due to gene flow (based on branch length) instead of ILS. - attached as spreadsheet

**Supplementary table S4:** QuIBL results from trees constructed using 20kb windows with a 1Mb slide - attached as spreadsheet. QuIBL analyses all triplet combinations ((A, B), C) in a given set of phylogenetic trees. Here we only present the alternative topologies within Delphinidae, that are in disagreement with the species tree, and may have arisen due to ILS or gene flow. The gene flow pair shows individuals A and B and outgroup is C. Two BIC scores are presented - one for ILS alone and one for ILS and gene flow. A BIC difference >10 suggests ILS and gene flow both as factors giving rise to the discordance topologies. ‘% of total trees’ shows the percentage of all trees in the dataset having said triplet topology. ‘% of trees supporting topology explained by gene flow’ shows the percentage of the trees supporting said triplet topology that likely arose due to gene flow (based on branch length) instead of ILS. - attached as spreadsheet

**Supplementary table S5:** D-statistics results for all triplet combinations phylogenetically concurrent with our results shown in Figure 1. Baiji was used as the outgroup/ancestral sequence. A non-significant result ( $|Z| < 3$ ) is indicated in bold. Colours indicate the family of the given individual. Red = Delphinidae, yellow = Phocoenidae, blue = Monodontidae.

| H1         | H2             | H3           | nABBA     | nBABA     | D-score | Z-score |
|------------|----------------|--------------|-----------|-----------|---------|---------|
| Bottlenose | IndoBottlenose | Killer whale | 597,251   | 554,780   | 0.037   | 23.26   |
| Bottlenose | IndoBottlenose | Pilotwhale   | 748,948   | 691,844   | 0.040   | 24.13   |
| Bottlenose | IndoBottlenose | Whitesided   | 721,498   | 665,420   | 0.040   | 25.20   |
| Pilotwhale | Whitesided     | Killer whale | 2,224,888 | 2,119,068 | 0.024   | 11.77   |
| Pilotwhale | Bottlenose     | Killer whale | 1,998,297 | 1,795,444 | 0.053   | 26.15   |
| Pilotwhale | IndoBottlenose | Killer whale | 2,004,478 | 1,757,429 | 0.066   | 31.95   |
| Pilotwhale | Bottlenose     | Whitesided   | 2,490,189 | 2,051,579 | 0.097   | 42.67   |
| Pilotwhale | IndoBottlenose | Whitesided   | 2,508,755 | 2,007,966 | 0.111   | 48.64   |
| Whitesided | Bottlenose     | Killer whale | 2,111,742 | 2,014,525 | 0.024   | 11.88   |
| Whitesided | IndoBottlenose | Killer whale | 2,117,925 | 1,975,800 | 0.035   | 17.25   |

|              |                |                  |           |         |       |        |
|--------------|----------------|------------------|-----------|---------|-------|--------|
| Killer whale | Pilotwhale     | Finless          | 928,942   | 840,273 | 0.050 | 51.99  |
| Killer whale | Whitesided     | Finless          | 924,323   | 829,525 | 0.054 | 56.12  |
| Killer whale | Pilotwhale     | Harbour porpoise | 959,748   | 851,885 | 0.060 | 60.74  |
| Killer whale | Whitesided     | Harbour porpoise | 956,686   | 840,318 | 0.065 | 65.46  |
| Killer whale | Bottlenose     | Finless          | 942,684   | 757,495 | 0.109 | 107.12 |
| Killer whale | Bottlenose     | Harbour porpoise | 974,032   | 767,636 | 0.119 | 116.98 |
| Killer whale | IndoBottlenose | Finless          | 943,526   | 728,185 | 0.129 | 120.99 |
| Killer whale | IndoBottlenose | Harbour porpoise | 974,967   | 739,024 | 0.138 | 130.60 |
| Pilotwhale   | Whitesided     | Finless          | 861,276   | 855,083 | 0.004 | 4.41   |
| Pilotwhale   | Whitesided     | Harbour porpoise | 892,930   | 884,620 | 0.005 | 5.64   |
| Pilotwhale   | Bottlenose     | Finless          | 828,193   | 724,397 | 0.067 | 73.75  |
| Pilotwhale   | Bottlenose     | Harbour porpoise | 857,823   | 749,827 | 0.067 | 76.38  |
| Pilotwhale   | IndoBottlenose | Finless          | 829,393   | 692,413 | 0.090 | 97.23  |
| Pilotwhale   | IndoBottlenose | Harbour porpoise | 859,146   | 718,044 | 0.089 | 98.69  |
| Whitesided   | Bottlenose     | Harbour porpoise | 887,876   | 787,914 | 0.060 | 68.88  |
| Whitesided   | Bottlenose     | Finless          | 857,483   | 760,224 | 0.060 | 69.75  |
| Whitesided   | IndoBottlenose | Harbour porpoise | 888,872   | 755,955 | 0.081 | 92.25  |
| Whitesided   | IndoBottlenose | Finless          | 858,523   | 727,924 | 0.082 | 92.84  |
| Bottlenose   | IndoBottlenose | Narwhal          | 414,272   | 380,995 | 0.042 | 33.84  |
| Bottlenose   | IndoBottlenose | Beluga           | 434,366   | 396,566 | 0.045 | 37.67  |
| Killer whale | Pilotwhale     | Narwhal          | 955,756   | 837,598 | 0.066 | 61.58  |
| Killer whale | Pilotwhale     | Beluga           | 984,462   | 854,528 | 0.071 | 65.67  |
| Killer whale | Whitesided     | Narwhal          | 953,496   | 826,881 | 0.071 | 66.17  |
| Killer whale | Whitesided     | Beluga           | 982,162   | 844,661 | 0.075 | 67.95  |
| Killer whale | Bottlenose     | Narwhal          | 971,164   | 751,458 | 0.128 | 111.86 |
| Killer whale | Bottlenose     | Beluga           | 1,001,546 | 767,422 | 0.132 | 113.69 |
| Killer whale | IndoBottlenose | Narwhal          | 974,507   | 722,249 | 0.149 | 126.51 |
| Killer whale | IndoBottlenose | Beluga           | 1,007,582 | 736,424 | 0.155 | 128.87 |
| Pilotwhale   | Whitesided     | Beluga           | 918,941   | 911,423 | 0.004 | 4.93   |
| Pilotwhale   | Whitesided     | Narwhal          | 891,298   | 883,114 | 0.005 | 5.61   |
| Pilotwhale   | Bottlenose     | Narwhal          | 859,652   | 743,735 | 0.072 | 78.60  |
| Pilotwhale   | Bottlenose     | Beluga           | 887,196   | 766,562 | 0.073 | 81.55  |
| Pilotwhale   | IndoBottlenose | Narwhal          | 863,608   | 710,777 | 0.097 | 103.83 |

|                  |                         |                  |                |                |              |             |
|------------------|-------------------------|------------------|----------------|----------------|--------------|-------------|
| Pilotwhale       | IndoBottlenose          | Beluga           | 895,023        | 731,826        | 0.100        | 105.92      |
| Whitesided       | Bottlenose              | Narwhal          | 888,390        | 780,573        | 0.065        | 74.77       |
| Whitesided       | Bottlenose              | Beluga           | 917,400        | 804,237        | 0.066        | 76.44       |
| Whitesided       | IndoBottlenose          | Narwhal          | 892,496        | 747,539        | 0.088        | 97.69       |
| Whitesided       | IndoBottlenose          | Beluga           | 925,091        | 769,228        | 0.092        | 102.86      |
| <b>Finless</b>   | <b>Harbour porpoise</b> | <b>Narwhal</b>   | <b>452,411</b> | <b>450,657</b> | <b>0.002</b> | <b>1.59</b> |
| Harbour porpoise | Finless                 | Beluga           | 570,767        | 552,830        | 0.016        | 13.47       |
| Narwhal          | Beluga                  | Harbour porpoise | 532,605        | 502,660        | 0.029        | 25.72       |
| Narwhal          | Beluga                  | Finless          | 514,273        | 466,273        | 0.049        | 41.75       |
| Finless          | Narwhal                 | Killer whale     | 973,140        | 885,678        | 0.047        | 47.30       |
| Finless          | Narwhal                 | Bottlenose       | 1,077,206      | 966,370        | 0.054        | 55.93       |
| Finless          | Narwhal                 | IndoBottlenose   | 1,080,812      | 970,600        | 0.054        | 56.63       |
| Finless          | Narwhal                 | Pilotwhale       | 1,059,846      | 950,178        | 0.055        | 57.27       |
| Finless          | Beluga                  | Killer whale     | 989,901        | 875,364        | 0.061        | 57.51       |
| Finless          | Narwhal                 | Whitesided       | 1,062,632      | 951,040        | 0.055        | 57.94       |
| Finless          | Beluga                  | Bottlenose       | 1,103,352      | 951,967        | 0.074        | 68.54       |
| Finless          | Beluga                  | Pilotwhale       | 1,084,679      | 936,511        | 0.073        | 68.84       |
| Finless          | Beluga                  | IndoBottlenose   | 1,109,158      | 955,589        | 0.074        | 69.72       |
| Finless          | Beluga                  | Whitesided       | 1,087,277      | 938,148        | 0.074        | 69.88       |
| Harbour porpoise | Narwhal                 | Killer whale     | 1,004,793      | 891,909        | 0.060        | 59.43       |
| Harbour porpoise | Beluga                  | Killer whale     | 1,028,676      | 885,849        | 0.075        | 69.85       |
| Harbour porpoise | Narwhal                 | Pilotwhale       | 1,124,641      | 974,232        | 0.072        | 75.43       |
| Harbour porpoise | Narwhal                 | Bottlenose       | 1,145,470      | 990,640        | 0.072        | 75.66       |
| Harbour porpoise | Narwhal                 | Whitesided       | 1,127,578      | 976,951        | 0.072        | 75.84       |
| Harbour porpoise | Narwhal                 | IndoBottlenose   | 1,153,263      | 994,022        | 0.074        | 78.93       |
| Harbour porpoise | Beluga                  | Pilotwhale       | 1,163,136      | 965,266        | 0.093        | 88.73       |
| Harbour porpoise | Beluga                  | Whitesided       | 1,165,862      | 968,086        | 0.093        | 89.42       |
| Harbour porpoise | Beluga                  | Bottlenose       | 1,185,612      | 981,030        | 0.094        | 89.66       |
| Harbour porpoise | Beluga                  | IndoBottlenose   | 1,197,547      | 984,311        | 0.098        | 93.10       |

**Supplementary table S6:** 100kb non-overlapping sliding window D-foil results for all quadruplet combinations [[H1,H2][H3,H4]] phylogenetically concurrent with our consensus topology shown in figure 1. Baiji was used as the outgroup/ancestral sequence. - attached as a spreadsheet. NA indicates not enough data in the window. None indicates no gene flow. As we implemented many different combinations, the species designation to H1 - H4 is indicated

at the top of the table. Numbers within the table show the number of windows that show evidence to the gene flow event depicted. - attached as spreadsheet

**Supplementary table S7:** The pre-divergence  $N_e$ , divergence time intervals, and the increments specified for each of the species pair used for the simulations to compare against the hPSMC results.

| <b>Species pair</b>                                | <b>Pre-divergence<br/><math>N_e</math></b> | <b>Range (Ma)</b> | <b>Increments<br/>(years)</b> |
|--|--|-------------------|-------------------------------|
| Beluga whale + Narwhal                             | 30,000                                     | 1-2               | 100,000                       |
| Beluga whale + Finless porpoise                    | 60,000                                     | 3-7               | 200,000                       |
| Beluga whale + Harbour porpoise                    | 60,000                                     | 3-7               | 200,000                       |
| Narwhal + Finless porpoise                         | 60,000                                     | 3-7               | 200,000                       |
| Narwhal + Harbour porpoise                         | 60,000                                     | 3-7               | 200,000                       |
| Beluga whale + Bottlenose dolphin                  | 105,000                                    | 3.9-8.5           | 200,000                       |
| Beluga whale + Indo-Pacific bottlenose dolphin     | 105,000                                    | 3.9-8.5           | 200,000                       |
| Narwhal + Bottlenose dolphin                       | 105,000                                    | 3.9-8.5           | 200,000                       |
| Narwhal + Indo-Pacific bottlenose dolphin          | 105,000                                    | 3.9-8.5           | 200,000                       |
| Narwhal + Killer whale                             | 105,000                                    | 3.9-8.5           | 200,000                       |
| Narwhal + Long-finned pilot whale                  | 105,000                                    | 3.9-8.5           | 200,000                       |
| Narwhal + Pacific white-sided dolphin              | 105,000                                    | 3.9-8.5           | 200,000                       |
| Beluga whale + Killer whale                        | 105,000                                    | 3.9-8.5           | 200,000                       |
| Beluga whale + Long-finned pilot whale             | 105,000                                    | 3.9-8.5           | 200,000                       |
| Beluga whale + Pacific white-sided dolphin         | 105,000                                    | 3.9-8.5           | 200,000                       |
| Harbour porpoise + Bottlenose dolphin              | 105,000                                    | 3.9-8.5           | 200,000                       |
| Harbour porpoise + Indo-Pacific bottlenose dolphin | 105,000                                    | 3.9-8.5           | 200,000                       |
| Finless porpoise + Bottlenose dolphin              | 105,000                                    | 3.9-8.5           | 200,000                       |

|   |         |                 |         |
|---|---------|-----------------|---------|
| Finless porpoise + Indo-Pacific bottlenose dolphin            | 105,000 | 3.9-8.5         | 200,000 |
| Finless porpoise + Killer whale                               | 105,000 | 3.9-8.5         | 200,000 |
| Finless porpoise + Long-finned pilot whale                    | 105,000 | 3.9-8.5         | 200,000 |
| Finless porpoise + Pacific white-sided dolphin                | 105,000 | 3.9-8.5         | 200,000 |
| Harbour porpoise + Killer whale                               | 105,000 | 3.9-8.5         | 200,000 |
| Harbour porpoise + Long-finned pilot whale                    | 105,000 | 3.9-8.5         | 200,000 |
| Harbour porpoise + Pacific white-sided dolphin                | 105,000 | 3.9-8.5         | 200,000 |
| Harbour porpoise + Finless porpoise                           | 40,000  | 0.3-1.4         | 100,000 |
| Indo-Pacific Bottlenose dolphin + Bottlenose dolphin          | 20,000  | 0.2-1.2         | 100,000 |
| Indo-Pacific bottlenose dolphin + Killer whale                | 50,000  | 0.9-2.1 & 3.4-7 | 200,000 |
| Indo-Pacific bottlenose dolphin + Long-finned pilot whale     | 50,000  | 0.9-2.1 & 3.4-7 | 200,000 |
| Indo-Pacific bottlenose dolphin + Pacific white-sided dolphin | 50,000  | 0.9-2.1 & 3.4-7 | 200,000 |
| Bottlenose dolphin + Killer whale                             | 50,000  | 0.9-2.1 & 3.4-7 | 200,000 |
| Bottlenose dolphin + Long-finned pilot whale                  | 50,000  | 0.9-2.1 & 3.4-7 | 200,000 |
| Bottlenose dolphin + Pacific white-sided dolphin              | 50,000  | 0.9-2.1 & 3.4-7 | 200,000 |
| Long-finned pilot whale + Killer whale                        | 60,000  | 0.9-2.1 & 3.4-7 | 200,000 |
| Pacific white-sided dolphin + Killer whale                    | 50,000  | 0.9-2.1 & 3.4-7 | 200,000 |
| Pacific white-sided dolphin + Long-finned pilot whale         | 50,000  | 0.9-2.1 & 3.4-7 | 200,000 |

**Supplementary table S8:** Mapping statistics of each Delphinoidea species used in this study when specifying the reference genome as the baiji assembly.

| <b>Common name</b>              | <b>Raw read pairs</b> | <b>Mapped reads</b> | <b>Coverage</b> | <b>Bp-mapped</b> |
|---------------------------------|-----------------------|---------------------|-----------------|------------------|
| Beluga                          | 466,374,135           | 476,814,543         | 31.44           | 69,807,010,359   |
| Bottlenose dolphin              | 578,690,171           | 732,418,659         | 47.61           | 105,524,983,813  |
| Harbour porpoise                | 289,063,910           | 418,431,029         | 23.17           | 50,830,083,145   |
| Indo-Pacific bottlenose dolphin | 466,306,082           | 551,837,703         | 35.62           | 78,749,625,267   |
| Indo-Pacific finless porpoise   | 523,612,238           | 557,766,873         | 24.96           | 54,450,935,944   |
| Killer whale                    | 1,467,089,287         | 1,047,260,000       | 39.53           | 88,692,400,000   |
| Long-finned pilot whale         | 428,064,233           | 504,482,080         | 28.61           | 63,276,638,573   |
| Narwhal                         | 384,563,392           | 468,429,237         | 31.09           | 68,247,058,370   |
| Pacific white-sided dolphin     | 453,348,710           | 499,704,592         | 28.83           | 63,800,396,300   |

**Supplementary table S9:** Mapping statistics of each Delphinoidea species used in this study when specifying the reference genome as a conspecific assembly.

| <b>Common name</b>              | <b>Raw read pairs</b> | <b>Mapped reads</b> | <b>Coverage</b> | <b>Bp-mapped</b> |
|---------------------------------|-----------------------|---------------------|-----------------|------------------|
| Beluga                          | 466,374,135           | 531,535,936         | 34.47           | 79,218,898,913   |
| Bottlenose dolphin              | 578,690,171           | 779,210,277         | 54.03           | 114,530,169,747  |
| Harbour porpoise                | 289,063,910           | 431,762,883         | 23.74           | 52,067,455,809   |
| Indo-Pacific bottlenose dolphin | 466,306,082           | 587,440,922         | 37.88           | 85,032,333,848   |
| Indo-Pacific finless porpoise   | 523,612,238           | 620,580,505         | 27.33           | 61,286,732,910   |
| Killer whale                    | 1,467,089,287         | 1,213,221,913       | 44.93           | 100,903,316,971  |
| Long-finned pilot whale         | 428,064,233           | 598,612,204         | 32.79           | 75,639,560,432   |
| Narwhal                         | 384,563,392           | 529,082,769         | 33.85           | 78,238,763,386   |

|                             |             |             |       |                |
|-----------------------------|-------------|-------------|-------|----------------|
| Pacific white-sided dolphin | 453,348,710 | 592,814,373 | 33.02 | 76,299,243,217 |
|-----------------------------|-------------|-------------|-------|----------------|

**Supplementary table S10:** Genome-wide pairwise distance matrix of the nine Delphinoidea included in this study. Bottlenose = bottlenose dolphin, Finless = finless porpoise, Harbour = harbour porpoise, Indobottle = Indo-Pacific bottlenose dolphin, Killer = killer whale, Pilot = pilot whale, White = Pacific white-sided dolphin.

|            |        |        |        |        |        |        |        |        |        |
|------------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| Beluga     | 0.0000 | 0.0211 | 0.0151 | 0.0153 | 0.0211 | 0.0205 | 0.0056 | 0.0210 | 0.0209 |
| Bottlenose | 0.0211 | 0.0000 | 0.0230 | 0.0231 | 0.0040 | 0.0113 | 0.0210 | 0.0102 | 0.0107 |
| Finless    | 0.0151 | 0.0230 | 0.0000 | 0.0056 | 0.0230 | 0.0224 | 0.0151 | 0.0229 | 0.0228 |
| Harbour    | 0.0153 | 0.0231 | 0.0056 | 0.0000 | 0.0231 | 0.0225 | 0.0152 | 0.0231 | 0.0230 |
| Indobottle | 0.0211 | 0.0040 | 0.0230 | 0.0231 | 0.0000 | 0.0113 | 0.0210 | 0.0102 | 0.0107 |
| Killer     | 0.0205 | 0.0113 | 0.0224 | 0.0225 | 0.0113 | 0.0000 | 0.0204 | 0.0113 | 0.0112 |
| Narwhal    | 0.0056 | 0.0210 | 0.0151 | 0.0152 | 0.0210 | 0.0204 | 0.0000 | 0.0209 | 0.0208 |
| Pilot      | 0.0210 | 0.0102 | 0.0229 | 0.0231 | 0.0102 | 0.0113 | 0.0209 | 0.0000 | 0.0109 |
| White      | 0.0209 | 0.0107 | 0.0228 | 0.0230 | 0.0107 | 0.0112 | 0.0208 | 0.0109 | 0.0000 |

**Supplementary table S11:** Metrics used to calculate the mutation rate per year with the equation mutation rate = divergence time / 2x genetic distance. Mean divergences were taken from the full dataset 10-partition AR from McGowen et al 2020 (McGowen et al., 2020) and average genetic distances were calculated from the results shown in supplementary table S5.

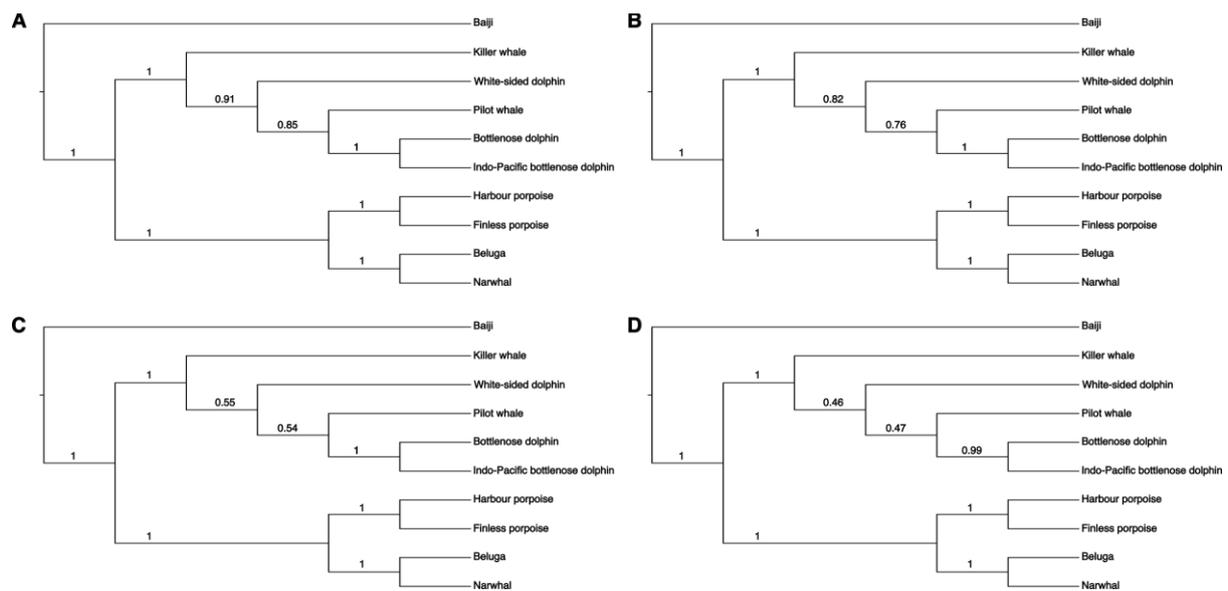
| Species            | Closest relative                | Divergence (Ma) | Distance | Mutation rate per year |
|--------------------|---------------------------------|-----------------|----------|------------------------|
| Beluga             | Narwhal                         | 7.72            | 0.0056   | $3.63 \times 10^{-10}$ |
| Killer whale       | Delphinidae                     | 10.16           | 0.0113   | $5.56 \times 10^{-10}$ |
| Bottlenose dolphin | Indo-Pacific bottlenose dolphin | 2.69            | 0.0040   | $7.51 \times 10^{-10}$ |
| Harbour porpoise   | Finless porpoise                | 5.36            | 0.0056   | $5.25 \times 10^{-10}$ |
| Long-finned pilot  | <i>Tursiops</i> spp.            | 7.46            | 0.0102   | $6.83 \times 10^{-10}$ |

|                             |                                |      |        |                        |
|-----------------------------|--------------------------------|------|--------|------------------------|
| whale                       |                                |      |        |                        |
| Pacific white-sided dolphin | <i>Tursiops + Globicephala</i> | 9.48 | 0.0108 | $5.69 \times 10^{-10}$ |

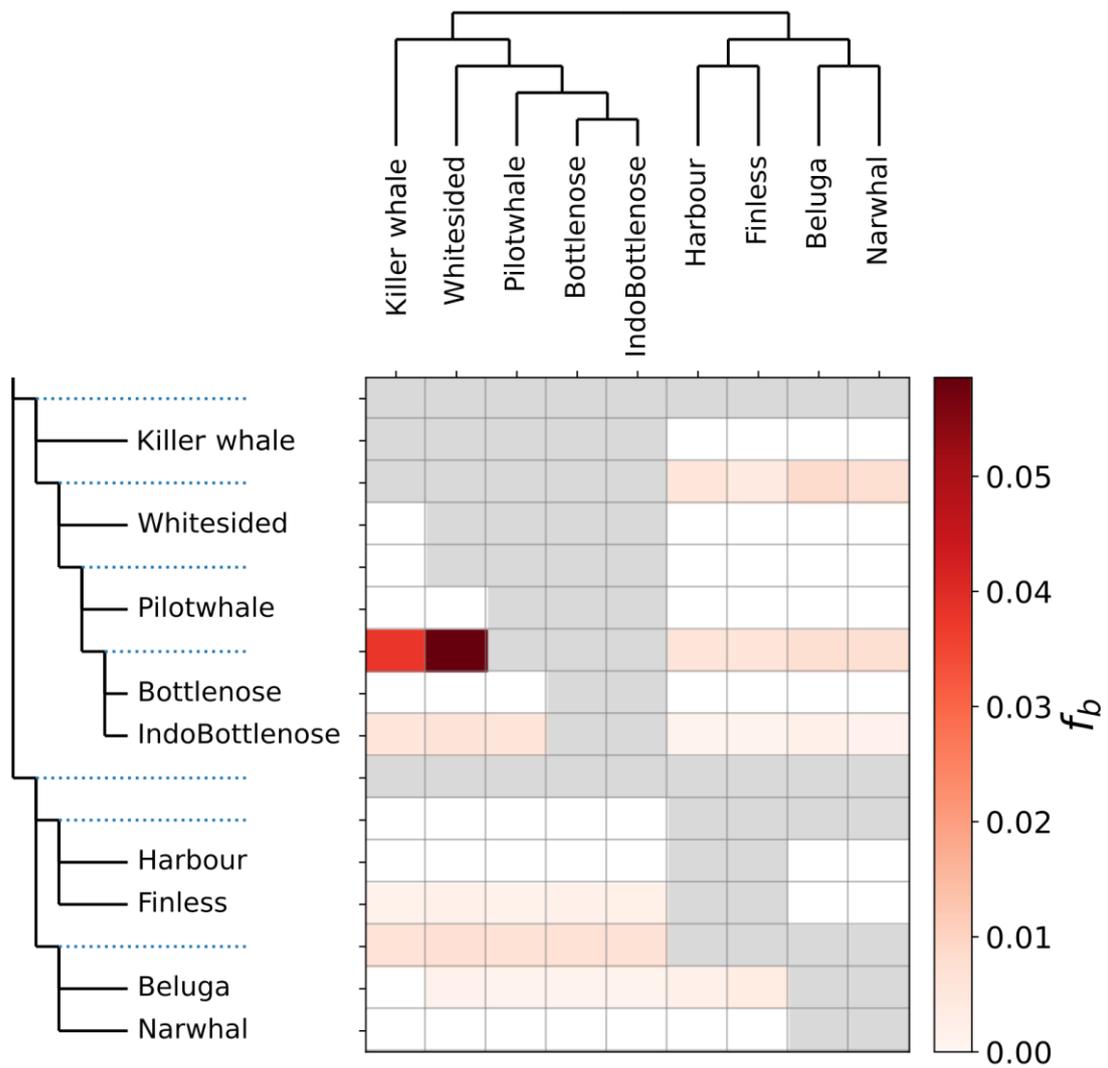
**Supplementary table S12:** Generation times, generational mutation rates and references for the generation times for each of the nine Delphinoidea species used in this study.

| <b>Common name</b>              | <b>Generation time</b> | <b>Generational mutation rate</b> | <b>Generation time reference</b> | <b>Bp-mapped</b> |
|---------------------------------|------------------------|-----------------------------------|----------------------------------|------------------|
| Beluga                          | 32                     | $1.16 \times 10^{-8}$             | (Garde et al., 2015)             | 79,218,898,913   |
| Bottlenose dolphin              | 21                     | $1.58 \times 10^{-8}$             | (Taylor et al., 2007)            | 114,530,169,747  |
| Harbour porpoise                | 10                     | $5.25 \times 10^{-9}$             | (Birkun and Frantzis, 2008)      | 52,067,455,809   |
| Indo-Pacific bottlenose dolphin | 21                     | $1.58 \times 10^{-8}$             | (Taylor et al., 2007)            | 85,032,333,848   |
| Indo-Pacific finless porpoise   | 8                      | $4.20 \times 10^{-9}$             | (Zhou et al., 2018)              | 61,286,732,910   |
| Killer whale                    | 26                     | $1.43 \times 10^{-8}$             | (Foote et al., 2016)             | 100,903,316,971  |
| Long-finned pilot whale         | 24                     | $1.64 \times 10^{-8}$             | (Taylor et al., 2007)            | 75,639,560,432   |
| Narwhal                         | 30                     | $1.09 \times 10^{-8}$             | (Garde et al., 2015)             | 78,238,763,386   |
| Pacific white-sided dolphin     | 21                     | $1.21 \times 10^{-8}$             | (Taylor et al., 2007)            | 76,299,243,217   |

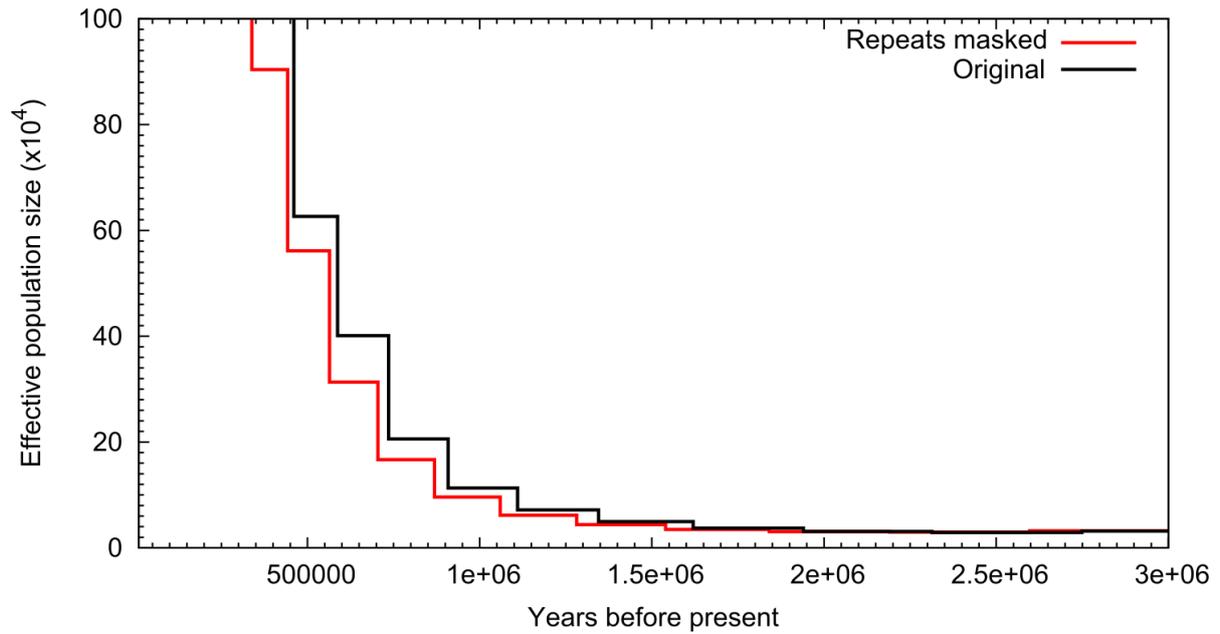
## Supplementary figures



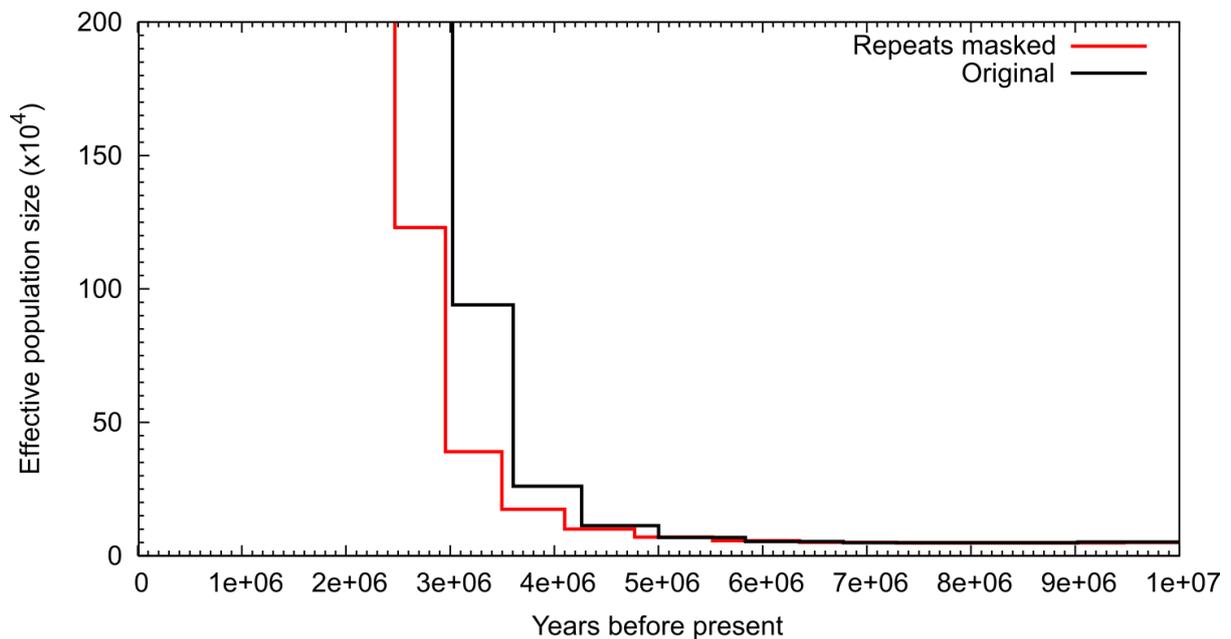
**Supplementary figure S1:** Consensus trees of independent Maximum-Likelihood trees constructed from non-overlapping sliding windows of (A) 1Mb, (B) 500kb, (C) 100kb, or (D) 50kb in length. Branch numbers represent the number of independent trees supporting each node.



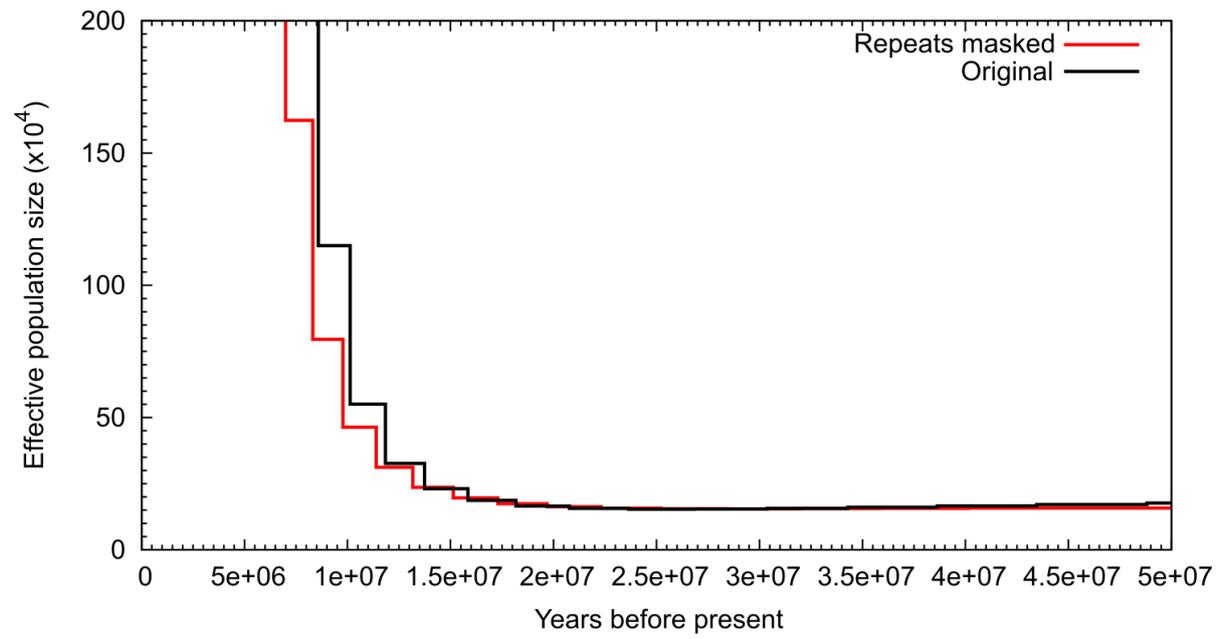
**Supplementary figure S2:** X chromosome Fbranch results. The species tree is displayed above while the trees to the left and right of the matrix are an expanded form, including internal branches as dotted lines. The values in the matrix refer to excess allele sharing between the expanded tree branch (relative to its sister branch) and the species on the  $x$ -axis.



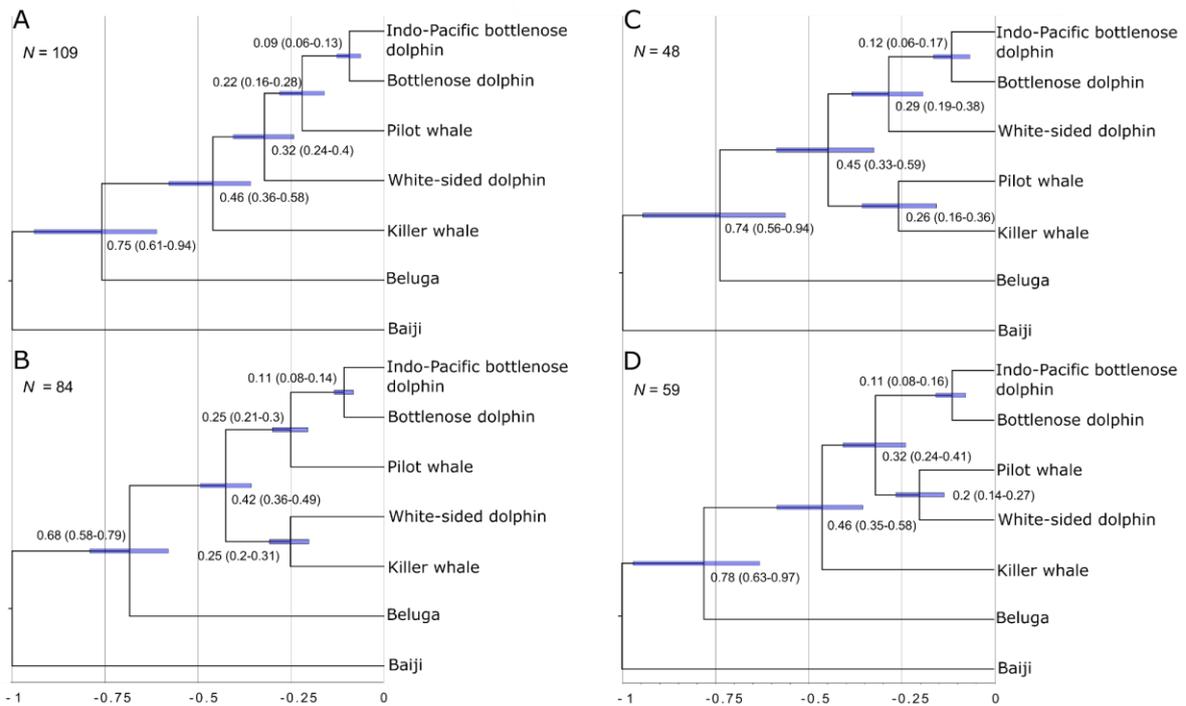
**Supplementary figure S3:** Comparison of hPSMC results using a pseudodiploid sequence from the bottlenose and Indo-Pacific bottlenose dolphins (shallow divergence) with either repeat regions masked or not.



**Supplementary figure S4:** Comparison of hPSMC results using a pseudodiploid sequence from the beluga and narwhal (medium divergence) with either repeat regions masked or not.



**Supplementary figure S5:** Comparison of hPSMC results using a pseudodiploid sequence from the bottlenose dolphin and beluga (deep divergence) with either repeat regions masked or not.



**Supplementary figure S6:** Relative divergence times of alternative topologies assumed to arise due to incomplete lineage sorting (ILS) or gene flow. N represents the number of independent loci supporting said topology. A) Consensus species topology. B) ILS/gene flow between the killer whale and Pacific white-sided dolphin. C) ILS/gene flow between killer whale and long-finned pilot whale. D) ILS/gene flow between Pacific white-sided dolphin and the long-finned pilot whale. Blue bars and numbers in parentheses show 95% credibility intervals.

### Supplementary results - hPSMC

Additional plots of the hPSMC empirical and simulated data can be found under the following link: [https://sid.erda.dk/cgi-sid/lis.py?share\\_id=ewvczfS2hH](https://sid.erda.dk/cgi-sid/lis.py?share_id=ewvczfS2hH) on the University of Copenhagen's electronic research data archive (ERDA). Bold lines show the hPSMC empirical data, faded lines show the simulated data, and the black lines show the simulated data that most closely match the empirical data without overlapping it between 1.5x and 10x the pre-divergence  $N_e$ .

## Supplementary references

- Birkun AA Jr, Frantzi A. 2008. *Phocoena phocoena* ssp. relict. The IUCN Red List of Threatened Species : e.T17027A50369903. <https://dx.doi.org/10.2305/IUCN.UK.2020-2.RLTS.T17027A50369903.en>.
- Foote AD, Vijay N, Ávila-Arcos MC, Baird RW, Durban JW, Fumagalli M, Gibbs RA, Hanson MB, Korneliussen TS, Martin MD, Robertson KM, Sousa VC, Vieira FG, Vinař T, Wade P, Worley KC, Excoffier L, Morin PA, Gilbert MTP, Wolf JBW. 2016. Genome-culture coevolution promotes rapid divergence of killer whale ecotypes. *Nat Commun* **7**:11693.
- Garde E, Hansen SH, Ditlevsen S, Tvermosegaard KB, Hansen J, Harding KC, Heide-Jørgensen MP. 2015. Life history parameters of narwhals (*Monodon monoceros*) from Greenland. *J Mammal* **96**:866–879.
- McGowen MR, Tsagkogeorga G, Álvarez-Carretero S, Dos Reis M, Struebig M, Deaville R, Jepson PD, Jarman S, Polanowski A, Morin PA, Rossiter SJ. 2020. Phylogenomic Resolution of the Cetacean Tree of Life Using Target Sequence Capture. *Syst Biol* **69**:479–501.
- Taylor BL, Chivers SJ, Larese J, Perrin WF. 2007. Generation length and percent mature estimates for IUCN assessments of cetaceans (No. Administrative Report LJ-07-01 ). National Marine Fisheries Service, Southwest Fisheries Science Center.
- Zhou X, Guang X, Sun D, Xu S, Li M, Seim I, Jie W, Yang L, Zhu Q, Xu J, Gao Q, Kaya A, Dou Q, Chen B, Ren W, Li S, Zhou K, Gladyshev VN, Nielsen R, Fang X, Yang G. 2018. Population genomics of finless porpoises reveal an incipient cetacean species adapted to freshwater. *Nat Commun* **9**:1276.

| Triplet analysed   | Gene flow pair | Control taxon                   | BIC2Dist (IBS + Gene flow) | BIC1Dist (IBS alone) | BIC difference | Significant for gene flow (BIC difference >10) | Number of trees | Percentage of total trees (2161) from triplet | Percentage of trees supporting topology explained by gene flow |
|--|----------------|---------------------------------|----------------------------|----------------------|----------------|--|-----------------|---|--|
| Pilot whale_Bottlenose dolphin_Killer whale                                  | Bot-Orca       | Pilot whale                     | -4176.75                   | -4015.52             | -161.23        | Yes  | 363             | 16.80   | 44.13  |
| White-sided dolphin_Bottlenose dolphin_Killer whale                          | Bot-Orca       | White-sided dolphin             | -5203                      | -5001.75             | -201.25        | Yes  | 451             | 20.87   | 51.55  |
| Pilot whale_Indo-Pacific Bottlenose dolphin_Killer whale                     | Indo-Orca      | Pilot whale                     | -4163.39                   | -4003.35             | -160.04        | Yes  | 362             | 16.75   | 44.27  |
| White-sided dolphin_Indo-Pacific Bottlenose dolphin_Killer whale             | Indo-Orca      | White-sided dolphin             | -5157.77                   | -4961.79             | -195.98        | Yes  | 448             | 20.73   | 91.82  |
| Pilot whale_Bottlenose dolphin_Killer whale                                  | Pilot-Orca     | Bottlenose dolphin              | -4149.09                   | -3995.26             | -153.83        | Yes  | 353             | 16.34   | 26.63  |
| Pilot whale_Indo-Pacific Bottlenose dolphin_Killer whale                     | Pilot-Orca     | Indo-Pacific Bottlenose dolphin | -4145.01                   | -3991.4              | -153.61        | Yes  | 353             | 16.34   | 24.46  |
| White-sided dolphin_Pilot whale_Killer whale                                 | Pilot-Orca     | White-sided dolphin             | -5551.99                   | -5354.47             | -197.52        | Yes  | 479             | 22.17   | 30.52  |
| White-sided dolphin_Pilot whale_Bottlenose dolphin_Killer whale              | Pilot-White    | Bottlenose dolphin              | -5329.17                   | -5126.07             | -203.10        | Yes  | 459             | 21.24   | 44.05  |
| White-sided dolphin_Pilot whale_Indo-Pacific Bottlenose dolphin_Killer whale | Pilot-White    | Indo-Pacific Bottlenose dolphin | -5332.08                   | -5127.41             | -204.67        | Yes  | 459             | 21.24   | 37.09  |
| White-sided dolphin_Pilot whale_Bottlenose dolphin_Killer whale              | White-Bot      | Pilot whale                     | -7160.67                   | -6929.73             | -230.94        | Yes  | 629             | 29.11   | 86.33  |
| White-sided dolphin_Pilot whale_Indo-Pacific Bottlenose dolphin_Killer whale | White-Indo     | Pilot whale                     | -7154.12                   | -6919.18             | -234.94        | Yes  | 628             | 29.06   | 49.33  |
| White-sided dolphin_Bottlenose dolphin_Killer whale                          | White-Orca     | Bottlenose dolphin              | -5679.95                   | -5365.25             | -314.70        | Yes  | 478             | 22.12   | 29.40  |
| White-sided dolphin_Indo-Pacific Bottlenose dolphin_Killer whale             | White-Orca     | Indo-Pacific Bottlenose dolphin | -5687.27                   | -5373.09             | -314.18        | Yes  | 479             | 22.17   | 31.43  |
| White-sided dolphin_Pilot whale_Killer whale                                 | White-Orca     | Pilot whale                     | -6205.88                   | -5910.93             | -294.95        | Yes  | 529             | 24.48   | 50.04  |
| Indo-Pacific Bottlenose dolphin_Bottlenose dolphin_Killer whale              | Bot-Orca       | Indo-Pacific Bottlenose dolphin | -47.1718                   | -40.833              | -6.34          | No   | 4               | 0.19  | 1.09   |
| Indo-Pacific Bottlenose dolphin_Bottlenose dolphin_Killer whale              | Indo-Orca      | Bottlenose dolphin              | -35.0559                   | -32.055              | -3.00          | No   | 3               | 0.14  | 0.37   |
| Pilot whale_Indo-Pacific Bottlenose dolphin_Bottlenose dolphin_Killer whale  | Pilot-Bot      | Indo-Pacific Bottlenose dolphin | -56.1656                   | -53.3674             | -2.80          | No   | 5               | 0.23  | 1.09   |
| Pilot whale_Indo-Pacific Bottlenose dolphin_Bottlenose dolphin_Killer whale  | Pilot-Indo     | Bottlenose dolphin              | -43.6088                   | -44.5198             | 0.91           | No   | 4               | 0.19  | 0.15   |
| White-sided dolphin_Indo-Pacific Bottlenose dolphin_Killer whale             | White-Bot      | Indo-Pacific Bottlenose dolphin | -53.2849                   | -53.8868             | 0.60           | No   | 5               | 0.23  | 0.46   |
| White-sided dolphin_Indo-Pacific Bottlenose dolphin_Killer whale             | White-Indo     | Bottlenose dolphin              | -41.6525                   | -42.3186             | 0.67           | No   | 4               | 0.19  | 0.31   |

| Triplet analysed   | Gene flow pair | Control taxon                   | BIC2Dist (IBS + Gene flow) | BIC1Dist (IBS alone) | BIC difference | Significant for gene flow | Number of trees | % of total trees | % of trees supporting topology explained by gene flow |
|--|----------------|---------------------------------|----------------------------|----------------------|----------------|---------------------------|-----------------|------------------|---|
| Pilot whale_Bottlenose dolphin_Killer whale                      | Bot-Orca       | Pilot whale                     | -5877.09                   | -5828.01             | -49.08         | Yes                       | 543             | 19.89            | 12.79   |
| White-sided dolphin_Bottlenose dolphin_Killer whale              | Bot-Orca       | White-sided dolphin             | -6493.50                   | -6410.93             | -82.57         | Yes                       | 589             | 21.58            | 14.76   |
| Pilot whale_Indo-Pacific Bottlenose dolphin_Killer whale         | Indo-Orca      | Pilot whale                     | -5836.61                   | -5777.56             | -59.05         | Yes                       | 539             | 19.74            | 13.24   |
| White-sided dolphin_Indo-Pacific Bottlenose dolphin_Killer whale | Indo-Orca      | White-sided dolphin             | -6501.26                   | -6417.36             | -83.90         | Yes                       | 590             | 21.61            | 14.82   |
| Pilot whale_White-sided dolphin_Killer whale                     | Pilot-Orca     | White-sided dolphin             | -6892.35                   | -6861.90             | -30.45         | Yes                       | 631             | 23.11            | 12.75   |
| Pilot whale_White-sided dolphin_Bottlenose dolphin               | Pilot-White    | Bottlenose dolphin              | -7033.39                   | -6989.18             | -44.21         | Yes                       | 648             | 23.74            | 14.00   |
| Pilot whale_White-sided dolphin_Indo-Pacific Bottlenose dolphin  | Pilot-White    | Indo-Pacific Bottlenose dolphin | -7073.33                   | -7026.60             | -46.73         | Yes                       | 651             | 23.85            | 14.15   |
| Pilot whale_White-sided dolphin_Bottlenose dolphin               | White-Bot      | Pilot whale                     | -9197.44                   | -9186.93             | -10.51         | Yes                       | 865             | 31.68            | 16.05   |
| Pilot whale_White-sided dolphin_Killer whale                     | White-Orca     | Pilot whale                     | -8498.20                   | -8408.06             | -90.14         | Yes                       | 784             | 28.72            | 19.25   |
| White-sided dolphin_Bottlenose dolphin_Killer whale              | White-Orca     | Bottlenose dolphin              | -7986.93                   | -7853.23             | -133.70        | Yes                       | 726             | 26.59            | 19.83   |
| White-sided dolphin_Indo-Pacific Bottlenose dolphin_Killer whale | White-Orca     | Indo-Pacific Bottlenose dolphin | -7983.67                   | -7846.07             | -137.60        | Yes                       | 726             | 26.59            | 20.03   |
| Indo-Pacific Bottlenose dolphin_Bottlenose dolphin_Killer whale  | Bot-Orca       | Indo-Pacific Bottlenose dolphin | -143.55                    | -144.83              | 1.28           | No                        | 13              | 0.48             | 0.39  |
| Indo-Pacific Bottlenose dolphin_Bottlenose dolphin_Killer whale  | Indo-Orca      | Bottlenose dolphin              | -82.72                     | -81.61               | -1.11          | No                        | 8               | 0.29             | 0.25  |
| Pilot whale_Indo-Pacific Bottlenose dolphin_Bottlenose dolphin   | Pilot-Bot      | Indo-Pacific Bottlenose dolphin | -306.79                    | -305.15              | -1.64          | No                        | 28              | 1.03             | 0.82  |
| Pilot whale_Indo-Pacific Bottlenose dolphin_Bottlenose dolphin   | Pilot-Indo     | Bottlenose dolphin              | -330.52                    | -336.87              | 6.35           | No                        | 31              | 1.14             | 0.52  |
| Pilot whale_Bottlenose dolphin_Killer whale                      | Pilot-Orca     | Bottlenose dolphin              | -5643.28                   | -5648.29             | 5.01           | No                        | 521             | 19.08            | 9.13  |
| Pilot whale_Indo-Pacific Bottlenose dolphin_Killer whale         | Pilot-Orca     | Indo-Pacific Bottlenose dolphin | -5701.86                   | -5699.31             | -2.55          | No                        | 525             | 19.23            | 9.77  |
| White-sided dolphin_Indo-Pacific Bottlenose dolphin_Bottlenose d | White-Bot      | Indo-Pacific Bottlenose dolphin | -257.04                    | -258.27              | 1.24           | No                        | 24              | 0.88             | 0.56  |
| Pilot whale_White-sided dolphin_Indo-Pacific Bottlenose dolphin  | White-Indo     | Pilot whale                     | -9117.94                   | -9115.68             | -2.26          | No                        | 858             | 31.43            | 15.41   |
| White-sided dolphin_Indo-Pacific Bottlenose dolphin_Bottlenose d | White-Indo     | Bottlenose dolphin              | -170.67                    | -176.49              | 5.81           | No                        | 16              | 0.59             | 0.23  |

