Global phylogeography and genetic diversity of the long-finned pilot whale *Globicephala melas*, with new data from the southeastern Pacific

Sebastián Kraft1, MJosé Pérez-Álvarez1,2,3*, Carlos Olavarría3,4 & Elie Poulin1

The matrilineal long-finned pilot whale presents an antitropical distribution and is divided into two subspecies, one in the temperate seas of the Southern Hemisphere and the other restricted to the North Atlantic and Mediterranean. Until now, population genetic and phylogeographic studies have included localities of most of its Northern Hemisphere distribution, while only the southwestern Pacific has been sampled in the Southern Hemisphere. We add new genetic data from the southeastern Pacific to the published sequences. Low mitochondrial and nuclear diversity was encountered in this new area, as previously reported for other localities. Four haplotypes were found with only one new for the species. Fifteen haplotypes were detected in the global dataset, underlining the species’ low diversity. As previously reported, the subspecies shared two haplotypes and presented a strong phylogeographic structure. The extant distribution of this species has been related to dispersal events during the Last Glacial Maximum. Using the genetic data and Approximate Bayesian Calculations, this study supports this historical biogeographic scenario. From a taxonomic perspective, even if genetic analyses do not support the subspecies category, this study endorses the incipient divergence process between hemispheres, thus maintaining their status and addressing them as Demographically Independent Populations is recommended.

Cetaceans have a diverse -and many times contrasting- array of geographic distributions. Some of the widest ranging species include the cosmopolitan sperm whale *Physeter macrocephalus*1 and the orca *Orcinus orca*2. Other species, especially small-sized odontocetes, generally present coastal and more restricted distributions, like the extreme case of the vaquita, *Phocoena sinus*1.

Some cetaceans have a particular distribution pattern known as disjunct or antitropical, in which the taxon is present at high latitudes in both hemispheres while being absent from lower latitudes4. Examples include the mysticete genus *Eubalaena*, with *E. borealis* inhabiting the North Pacific, *E. japonica* the North Atlantic and their austral equivalent in the Southern Hemisphere, *E. australis*. The phocoenid species pair *Phocoena phocoena* and *P. spinipinnis* is another example, found in the Northern Hemisphere and the coasts of South America, respectively4. Among delphinids, the two *Lissodelphis* species also present this antitropical distribution pattern; *L. peronii* is present around the Southern Hemisphere, while *L. borealis* is found in the North Pacific3. A similar antitropical distribution is presented by the two subspecies of long-finned pilot whales *Globicephala melas*, where *G. m. edwardii* inhabits the temperate to subpolar waters of the Southern Hemisphere, while *G. m. melas* is restricted to the North Atlantic3. Extinct populations of long-finned pilot whales have been reported in the North Pacific, from Japan4 and Alaska10, dating back 8 000–12 000 years and 2 500–3 500 years, respectively4,10.

Pilot whales are a highly social matrilineal odontocete, thought to be among the most gregarious cetaceans. The species is known to form groups with a mean size around tens of individuals11 that can increase to hundreds.

---

1Instituto de Ecología y Biodiversidad, Laboratorio de Ecología Molecular, Departamento de Ciencias Ecológicas, Facultad de Ciencias, Universidad de Chile, Santiago, Chile. 2Centro de Investigación Eutropia, Santiago, Chile. 3Escuela de Medicina Veterinaria, Facultad de Ciencias, Universidad Mayor, Santiago, Chile. 4Centro de Estudios Avanzados en Zonas Áridas (CEAZA), La Serena, Chile. *email: maria.perez@umayor.cl
Table 1. Summary of number of haplotypes and genetic indices for each included locality and per subspecies (abbreviations are detailed in the methods section). The first row indicates the names of each haplotype according to the original authors. Numbers 60 and 62 represent the last two digits of the GenBank codes provided by Miralles et al., (2016) for identification. Number of haplotypes (h), haplotype diversity (Hd), pairwise differences between sequences (Π) and nucleotide diversity (π) are also detailed.

|                | S + R | P + U | Q + Y | R2 | T | O | V | O2 | W | Z | E + G | 60 | 62 | X | D | Total | h  | S   | Hd  | %  | Π | Total |
|----------------|-------|-------|-------|----|---|---|---|----|---|---|-------|----|----|---|---|-------|----|-----|-----|---|----|-------|
| TAS            | 51    | 32    | 72    |    |   |   |   |    |   |   |       | 215| 6  | 4  | 0.783| 0.394| 1.356|
| NZ             | 6     | 333   | 3     | 1  | 14| 1  |    |    |   |   |       | 335| 6  | 5  | 0.133| 0.040| 0.136|
| CL             | 31    | 45    | 12    | 2  |   |   |   |    |   |   |       | 90 | 4  | 4  | 0.620| 0.230| 0.791|
| NWA            | 73    | 1     | 4     |    |   |   |   |    |   |   |       | 78 | 3  | 2  | 0.123| 0.064| 0.220|
| FI             | 55    | 17    |       |    |   |   |   |    |   |   |       | 72 | 2  | 1  | 0.365| 0.100| 0.365|
| UK             | 34    | 1     | 4     |    |   |   |   |    |   |   |       | 39 | 3  | 2  | 0.235| 0.070| 0.240|
| IB             | 39    | 1     | 1     | 1  |   |   |   |    |   |   |       | 42 | 3  | 1  | 0.139| 0.055| 0.188|
| NEA            | 20    |       | 1     | 1  |   |   |   |    |   |   |       | 21 | 2  | 1  | 0.095| 0.027| 0.095|
| GIB            | 44    |       | 20    | 64 | 2 | 1  |   |    |   |   |       | 64 | 10 | 9  | 0.580| 0.233| 0.802|
| MED            | 12    |       | 21    | 33 | 2 | 1  |   |    |   |   |       | 47 | 10 | 9  | 0.477| 0.138| 0.477|
| G. m. edwardii | 88    | 410   | 87    | 2  | 14| 1  | 14| 15 | 31| 6 |       | 663| 10 | 9  | 0.580| 0.233| 0.802|
| G. m. melas    | 277   | 19    | 1     | 1  | 4 | 41 | 349| 8  | 5 |   |       | 353| 8  | 5  | 0.353| 0.117| 0.404|
| Total          | 365   | 429   | 87    | 2  | 14| 1  | 14| 15 | 31| 6 |       | 1012| 15| 12 | 0.680| 0.264| 0.909|

Results

Southeastern pacific sampling and genetic diversity. In total, 90 mtDNA sequences were obtained, defining four haplotypes. One haplotype was previously unreported (hereon referred to as haplotype R2). Haplotype and nucleotide diversities were low: Hd = 0.62 and π = 0.23% (Table 1).

Microsatellites were successfully genotyped in n = 44 samples (32 from Isla Clemente and 12 from Isla Navarino), across fourteen polymorphic loci. DirFCB6, and GTS1 presented an excess of homozygotes, but were not used in the comparisons among hemispheres. Regional diversity values of usable loci were Hr = 0.655; Hr = 0.700 and nA = 7.5 (Table 2).

Global diversity analysis. After adding previously published sequences from other oceanic basins20,23–26, a total of 15 haplotypes were obtained after the elimination of the site that was generating phylogeographic noise in the total 1012 sequences. This led to the combination of four previously described haplotypes in the following pairs: S with R, P with U, Q with Y, and E with G (Table 1). Nine of the twelve haplotypes reported by Miralles et al., (2016), originally of a consensus fragment size of 703 bp, were also merged into haplotype S + R.

Global and local haplotype networks. Adding previously published sequences from other oceanic basins, we detected that (1) two haplotypes were shared between G. m. edwardii (SP) and G. m. melas (NA and MED), and (2) one haplotype was shared by all SP localities but was absent from NA and MED (Table 1, Fig. 1).
Haplotypes $P + U$ and $S + R$ were the only ones detected within the distribution range of both $G. melas$ subspecies (Table 1, Fig. 2a): haplotype $P + U$ is present in the SP and NA, while $S + R$ was found in all three basins (SP, NA and MED), and was pivotal in the development of local diversity. Haplotype $Q + Y$ was shared by all SP localities, but absent from NA and MED. The remaining 12 haplotypes represent in situ diversification within each corresponding basin: eight for SP, four for NA and one for MED. The haplotypes found in the Mediterranean area are represented either by $S + R$ or by $D$, which derives directly from it. The same occurs with half (2) of haplotypes exclusive to the SP.

In the haplotype network by locality, haplotype $S + R$ is present in all ten studied areas and originates much of the local diversity within them (Table 1, Fig. 2b). Haplotype $P + U$ plays a similar, but downscaled role, in the South Pacific; it originates half (4) of the diversity that is exclusive to this area. This haplotype is also much more abundant in the localities corresponding to $G. m. edwardii$ than in the range of $G. m. melas$, where it was detected.

Table 2. Comparisons among shared microsatellite loci from three studies including Tasmania and New Zealand$^{27}$, Chile (this study) and the North Atlantic$^{28}$. Dashes indicate either unreported information or no locus amplification.
in lower numbers in three localities (17 samples from FI, one from NWA and one from UK). One of the haplotypes that originates from P+U, haplotype Q+Y, is exclusive to the South Pacific and is present in all three localities. Ten of the total fifteen haplotypes (67%) were exclusively present in one of the following five localities: TAS (3), NZ (3), CL (1), IB (2) and NWA (1). In the South Pacific, 70% of haplotypes were private to one of the three localities; it was the only basin with at least one private haplotype in each. Three private haplotypes were encountered in the North Atlantic: one in the NWA and two in the IB.

**Global and local genetic structure.** Significant genetic and phylogeographic structure was also detected between the subspecies, i.e. South Pacific with North Atlantic and Mediterranean (\(F_{ST} = 0.439, \varphi_{ST} = 0.464\), both \(P < 0.000004\)). The Snn test for genetic differentiation was statistically significant among the two disjoint distribution areas of the two subspecies (i.e. \(G. m. edwardii\) and \(G. m. melas\); \(Snn = 0.830, P < 0.001\)). Subsequently, samples were organized in this same way for the AMOVA. Similar structure results were obtained grouping the samples by the 10 worldwide localities. Average \(\varphi_{ST}\) values were highest within the SP (avg. \(\varphi_{ST} = 0.283\), lowest in the NA (avg. \(\varphi_{ST} = 0.06\)), and intermediate among the two Mediterranean localities (avg. \(\varphi_{ST} = 0.174\) (Table 3). All \(\varphi_{ST}\) comparisons were statistically significant within the South Pacific. In the North Atlantic, only the Faroe Islands showed a statistically significant phylogeographic structure with all other localities, while NWA did with FI and UK. All comparisons among localities from different basins showed statistically significant differences. The AMOVA showed significant differentiation among the two distribution ranges, with greater variation among groups than within them (Table 4).

In the Correspondence Analysis, the first axis separated the localities according to their respective hemisphere of origin. The three South Pacific localities (TAS, NZ and CL) were clustered separately from those of the Northern Hemisphere. The second axis divided the Strait of Gibraltar and the Mediterranean Sea from the remaining localities in the Northern Hemisphere (Fig. 3). The calculation of Nei’s (1987) net nucleotide divergence (\(d_{st}\)) among the subspecies \(G. m. edwardii\) and \(G. m. melas\) resulted in \(d_{st} = 0.00158\) (SD = 0.00019).

**Historical biogeography.** The origin of the current disjunct distribution of *Globicephala melas* was explored with DIYABC, using two historical biogeographic models associated to the Last Glacial Maximum (LGM). The first model, scenario 1, represented a colonization event from the Southern Hemisphere to the Northern Hemisphere, followed by isolation and population growth in the Northern Hemisphere. The second axis divided the Strait of Gibraltar and the Mediterranean Sea from the remaining localities in the Northern Hemisphere (Fig. 3). The calculation of Nei’s (1987) net nucleotide divergence (\(d_{st}\)) among the subspecies \(G. m. edwardii\) and \(G. m. melas\) resulted in \(d_{st} = 0.00158\) (SD = 0.00019).

**Microsatellite diversity.** Previous data on microsatellite diversity was available for the southwestern Pacific27 and the North Atlantic23,25,28. The diversity by locus reported27,28 is shown in Table 2, with five loci in common to both studies. The number of alleles for each shared locus were, respectively: 409/470 (10 vs. 9), 415/416 (9 vs. 5), 464/465 (9 vs. 6), EV37 (10 vs. 6) and EV94 (7 vs. 7). Two of these loci were also present in
our database of 44 analysed individuals from the southeastern Pacific (464/465 and EV37), with 9 and 15 alleles in total. The upper-tail exact test among the five shared loci of *G. m. edwardii* and *G. m. melas* showed that South Pacific samples had significantly more alleles per locus (p = 0.028). By contrast, the upper-tail exact test

| Source of variation | d.f. | Sum of squares | Variance components | Percentage of variation |
|---------------------|------|---------------|---------------------|------------------------|
| Among groups        | 1    | 124.096       | 0.23802 Va          | 39.19                  |
| Among populations within groups | 8     | 78.725       | 0.11293 Vb          | 18.60                  |
| Within populations  | 1002 | 256.860      | 0.25635 Vc          | 48.33                  |
| Total               | 1011 | 459.681      | 0.6073              |                        |

Table 4. Results of the AMOVA among both subspecies of long-finned pilot whales in the South Pacific (i.e. *G. m. edwardii*) and North Atlantic with the Mediterranean (i.e. *G. m. melas*).

Figure 3. (a) Correspondence Analysis scatter plot of presence/absence of haplotypes per locality: Tasmania (TAS), New Zealand (NZ), Chile (CL), northwestern Atlantic (NWA), Faroe Islands (FI), United Kingdom (UK), Iberian Peninsula (IB), northeastern Atlantic (NEA), Strait of Gibraltar (GIB) and the Mediterranean Sea (MED). Localities within the distribution of *G. m. melas* are encircled in yellow and the South Pacific localities of *G. m. edwardii* in green.
performed with the 10 loci shared between *G. m. edwardii* from the southwestern27 and southeastern Pacific (this study) did not detect a statistically significant difference (*p* = 0.166).

**Discussion**

Sampling cetacean species for population genetic studies has proven to be logistically challenging. A large number of individuals need to be sampled for this kind of studies, but they are generally widespread geographically and low in density. Mass strandings of cetaceans represent unique opportunities to collect large numbers of samples, as in the case of the franciscana *Pontoporia blainvillei*29, bottlenose dolphin *Tursiops truncatus*30, Cuvier’s beaked whale *Ziphius cavirostris*31 and the short-beaked common dolphin *Delphinus delphis*32. The long-finned pilot whale is a species for which almost all genetic data has been obtained from sampling mass strandings. A large part of its distribution range has been already sampled in the northwestern and northeastern Atlantic23–26, but within their broader distribution in the Southern Hemisphere, only the southwestern Pacific has been covered20. This study presents new data from two recent mass strandings in southern Chile, in order to produce the first genetic information for *Globicephala melas* in the southeastern Pacific and to integrate this new information into the global phylogeography of the species.

Contrasting tissue quality was found between both strandings, which we attribute to the time of sampling of each event. Our sampling of the Isla Clemente stranding took place an estimated two to three months after occurring33, while the Isla Navarino individuals were sampled five days after stranding34, precluding tissue the deterioration encountered in the former event. This highlights the importance of swift responses to stranding events, in order to secure the best tissue quality possible.

Prior to conducting mtDNA analyses, we first decided to remove one polymorphic site that exhibited a strong homoplasy signal. This problem in mtDNA sequences of *G. melas* was already detected20, but not taken into account for the genetic analyses. Homoplasic sites have been described to hinder the resolution of mtDNA gene trees35 and have also been pointed out as potential confounders of evolutionary analysis in the mtDNA control region of humpback whales36 and human mtDNA coding regions37, among others.

Despite the addition of 90 new sequences from the southeastern Pacific to the set of available sequences in the literature - over 9000 km from New Zealand, the nearest sampled locality - the global haplotype network was expanded by just one private low-frequency haplotype, further confirming the worldwide low mitochondrial diversity of *G. melas*. Such low mtDNA diversity has been reported for other matrilineal odontocetes, including orcas (*π* = 0.52%)38, false killer whales (*π* = 0.01–0.30%)39 and sperm whales (*π* = 0.131–0.407%)40. In contrast, cetaceans with more labile social cohesion such as mysticetes and non-matrilineal odontocetes, generally present much higher nucleotide diversity, a feature that appears to be common among these species41. Cultural hitch-hiking has been regarded as a driver of their low mitochondrial diversity41,42. An analogous effect of behaviour on mitochondrial diversity was described in a resident coastal bottlenose dolphin population in central Chile43. That study found that the genetic diversity of this matrilineal population (Hd = 0.63; *π* = 0.8%) was lower than that of the non-matrilineal transient-pelagic group adjacent to them (Hd = 0.95;
π = 1.4%), suggesting the importance of social structure in shaping the pattern of genetic diversity in cetaceans (at least in odontocetes).

As previously reported for other areas20,23,24,26, our results show that southeastern Pacific long-finned pilot whales (n = 90) have low genetic diversity, in particular for haplotype richness (h = 4) and nucleotide diversity (π = 0.23%, Table 1). Within the South Pacific, the diversity indices of southeastern Pacific (i.e., Chilean) samples were similar to the values obtained for Tasmania. Nevertheless, genetic diversity in these two localities was very different from that of New Zealand. Despite accounting for 54% of the samples in this basin, the genetic and nucleotide diversity of pilot whales from New Zealand were the lowest (Table 1), particularly because of the high predominance of one haplotype (93%, haplotype P + U). Such overrepresentation of a single haplotype in a matrilineal species may derive from sampling bias, for example, by sampling a single, large mass stranding event. However, in that study, the sampling period spanned from 1993 to 2007 and included numerous single and mass strandings that took place in various localities20. Therefore, we can assume that this diversity accurately reflects what is present in this area. The distinctiveness of New Zealand from Tasmania and Chile is further supported by high ϕST values (ϕST = 0.421 and 0.342 respectively) (Table 3), which are four and five times higher than between Chile and Tasmania (ϕST = 0.086). The absence of obvious geographic or oceanographic barriers between New Zealand and the other localities in the South Pacific does not allow a simple interpretation of the pattern of genetic structure found here. This genetic differentiation could have been attained through ecological specialization, as previously pointed out20. Similarly, sea surface temperature and its influence on prey distribution has been regarded as a possible ecological factor underlying genetic differentiation in extant G. melas in North Atlantic waters28 and similar trends have been observed on the Scotian shelf44.

Differentiation over relatively short distances without any conspicuous geographical barriers has also been detected in other odontocetes. For example, genetic differentiation was detected in the Chilean dolphin Cephalorhynchus eutropia between two differing coastal habitats along the uninterrupted Chilean coastline, attributed to habitat adaptation and specific hunting strategies45. Also, the Eastern Pacific Barrier has been proposed as a driver behind the genetic differentiation of the short-finned pilot whale Globicephala macrorhynchus into two subspecies46.

The haplotype network of the specimens from New Zealand presents a typical star-like shape (Fig. 5a), suggesting that long-finned pilot whales around New Zealand represent a young population, perhaps tracing back to the Last Glacial Maximum, as suggested by DIYABC analyses. During this period, a 6–10 °C cooling occurred in superficial waters of southeast New Zealand, the strongest temperature drop reported in this area of the southwestern Pacific47. Such changes in environmental conditions, probably associated to a shift in the distribution
of marine biota, may have provoked a typical population contraction-expansion in the long-finned pilot whale population in this area, as described for cold-temperate and polar marine species34, including cetaceans45-50. A similar case of strong genetic differentiation among populations of long-finned pilot whales is observed between the North Atlantic and Mediterranean populations, where the latter exhibits high phylogeographic and genetic differences with the North Atlantic localities33. In this case, geographic and oceanographic discontinuities between the Mediterranean Sea and the Atlantic Ocean provide a robust explanation for the observed genetic separation. The population of Mediterranean populations from North Atlantic ones has been previously reported in various marine species, such as shallow water crustaceans31, sea stars32, white sharks33 and other odontocetes, like in sperm whales34, striped dolphins Stenella coeruleoalba35 as well as Cuvier’s beaked whales Ziphius cavirostris31 and Risso’s dolphins Grampus griseus56.

Widespread cetacean taxa occurring in both the Northern and Southern Hemispheres generally exhibit strong phylogeographic structure and genetic divergence between regions. Such genetic differentiation has been generally exemplified by fixed substitutions in mtDNA control region sequences. This is the case of the fin whale Balaenoptera physalus, with one fixed difference between South and North Atlantic samples, thus presenting no shared haplotypes37, the harbour porpoise Phocoena phocoena, with high divergence among ocean basins38 and the more closely related false killer whale Pseudorca crassidens39. In the latter species, although the study had a low sample size in the North Atlantic, no shared haplotypes were found and at least 10 substitutions separated the Atlantic populations from those in the Indo-Pacific. In the case of G. melas, despite its antitropical distribution and the large geographic discontinuity between northern and southern distribution areas, the subspecies shared their two most abundant haplotypes. This genetic pattern may reflect (1) contemporary gene flow between hemispheres, or alternatively (2) an ancestral polymorphism resulting from an incipient divergence process. The strong phylogeographic structure detected among the subspecies supports the second hypothesis.

The South Pacific network holds much of the species’ genetic diversity and is therefore very similar to the global network in overall shape (Fig. 5b). In contrast, the star-like haplotype network of North Atlantic and Mediterranean samples is typically presented by recently expanding populations. A biogeographic scenario of a dispersal event over the equator during previous glacial period has been proposed22, with a split in distribution occurred after the Last Glacial Maximum, 10 000–15 000 years ago. This hypothesis was further expanded26, mentioning that a trans-equatorial dispersal event rather than vicariance might have taken place, on account of the lower diversity in the Northern Hemisphere subspecies. Thus, the hypothesis is supported by (1) the contrasting haplotype networks (Fig. 5c, d), (2) the higher mitochondrial and microsatellite diversity indices of South Pacific long-finned pilot whales compared to their counterparts in the North Atlantic and Mediterranean (Tables 1 and 2) and (3) the clear separation of areas by hemisphere in the haplotype presence/absence Correspondence Analysis (Fig. 3). The validity of the previously proposed scenario of population history was further explored here with DIYABC population history simulations, comparing a founder effect scenario with a vicariance scenario. The patterns of population genetic diversity and structure observed in G. melas were consistent with the data simulated under the former scenario, a trans-equatorial dispersal event followed by divergence (Fig. 4). Posterior estimation of scenario parameters allowed estimating the time at which the different events occurred, setting the dispersal from the Southern Hemisphere to the North Atlantic at around 13 000 years ago, followed by a population demographic expansion around 9 380 years ago. However, even considering this scenario of divergence, the absence of reciprocal monophyly does not qualify northern and southern G. melas as different Evolutionary Significant Units (ESU) following the criterion of Moritz (1994)21. It is likely that not enough time has passed to sort lineages, or some level of gene flow still occurs46, as they still share haplotypes, but in differing frequencies29. Additionally, net nucleotide divergence dN and Percent Diagnosability (PD) had been previously calculated for the long-finned pilot whale subspecies40, obtaining values of dN = 0.00128 and PD = 0.84286. Our calculation of net nucleotide divergence, which included new sequences from the southeastern Pacific and integrated other published sequences from the Northern Hemisphere, resulted in a slightly higher value of dN = 0.00158, which still is below the proposed subspecies interval41. Therefore, addressing them as Demographically Independent Populations (DIP), a transitional state between a panmictic population and separate ESUs59 might be more accurate.

Finally, from a conservation perspective, even if genetic analyses do not support the subspecies category, we recommend maintaining their current taxonomic status, since these DIP might be undergoing a recent divergence process which has yet to mature into fully sorted lineages.

Concluding Remarks
The results here presented should be considered as preliminary evidence, as the use of a single mtDNA marker for phylogeographic and demographic inference has been deemed problematic before50. As previously stated, such molecular studies should include nuclear markers together with mitochondrial DNA, even more when delimiting species and subspecies55. However, to date, the mtDNA control region is the only molecular marker for which sequences are available with sufficient sample size from different ocean basins to perform a worldwide phylogeographic study in G. melas56. To incorporate new genetic markers to a global phylogeographic study would require tremendous sampling effort and expenses, and may also strongly depend on the occurrence of mass strandings in all these regions of the world.

Finally, we believe that collaborative studies surveying uncharted areas, especially within the greater distribution range of long-finned pilot whales in the Southern Hemisphere, are fundamental to obtain complete data on the worldwide phylogeography and taxonomy of G. melas and to lay the groundwork for future research on these topics. Modern genetic tools, such as complete mitogenome and SNPs, have been already used to revise the taxonomic status of short-finned pilot whales46, which could be replicated in long-finned pilot whales.
Methods

Southeastern pacific sample collection. Tissue samples were collected from twelve individuals in a stranding event that occurred in Isla Navarino in 2006 (55°15'S; 67°30'W (Fig. 1). In 2016, 124 Globicephala melas were sampled from the mass stranding event that occurred in Islas Centleante 45°35'57.50'S; 74°34'30.32"W. All samples were preserved in 90–95% ethanol.

DNA extraction, mitochondrial control region sequencing and microsatellite genotyping. DNA extractions were performed following a modified salt-extraction protocol34, adding a second step of digestion with proteinase K one hour after the first one.

Mitochondrial control region data. The mtDNA control region was amplified using the primers M13 Dlp1.5 5′-TGTAACACGACACGCTACGTCACCCAAAAGCTG RARTTCT-3′ (forward) and 8 G 5′-GGAGTACTATGTCCTGTAACCA-3′ (reverse)31. The amplification protocol was as follows: 25.6 µL reaction volume for each PCR reaction consisted of 12.7 µL water, 5 µL 10X Buffer (Invitrogen), 2 µL 50 mM MgCl2 (Invitrogen), 2 µL 10 pM dNTPs (Invitrogen), 1 µL 10 pM of each primer (2 µL total), 0.5 µL Taq polymerase (Invitrogen) and 70–150 ng of DNA. A Thermo Hybaid PSE 0.5 thermal cycler was used for all amplifications, with the following profile: Preliminary denaturation of 2 minutes at 94°C; followed by 30 cycles of: denaturation for 30 s at 94°C, annealing for 40 s at 56°C, polymerase extension for 40 s at 72°C; and a final polymerase extension for 10 minutes at 72°C and an infinite hold temperature of 4°C. Each PCR run included positive and negative controls. Fragments were run in a 1% agarose gel, each well containing 3 µL of PCR product mixed with an equal volume of loading dye with 0.3% Gel Red and visualized in a gel documentation system (Maestrogen SMU-01).

PCR product purification and sequencing in both directions were done at Macrogen Inc., Seoul, South Korea with a 3730XL DNA Analyzer (Applied Biosystems). All sequences obtained were aligned manually in ProSeq 3.5a4 and polymorphic sites were visually checked. Prior to molecular analyses, the species for each sample was corroborated with two platforms of comparative analysis of sequences: BLAST (Basic Local Alignment Search Tool, www.blast.ncbi.nlm.nih.gov) and DNA Surveillance65.

An additional 922 control region sequences were obtained from five other sources: (1) Oremus et al., (2009)30 (n = 573, Tasmania and New Zealand, GenBank access codes: FJ513342-54); (2) Siemann (1994)30 (n = 59; western North Atlantic, Cape Cod, Newfoundland, Nova Scotia, Scotland and England; GenBank access codes: U20926-28); (3) Monteiro (2013)34 (n = 116, eastern coast of the United States, Faroe Islands, United Kingdom and Iberian Peninsula, GenBank access codes: KC934932-34); (4) Verborgh (2015)25 (n = 117, eastern coast of the United States, Faroe Islands, United Kingdom, Euskadi, northeastern Atlantic, Iberian Peninsula, Strait of Gibraltar and the Mediterranean Sea; haplotypes reconstructed by hand) and (5) Miralles et al., (2016)43 (n = 57, Faroe Islands and Iberian Peninsula, GenBank access codes: KJ740360-71) (Table 1).

Sample grouping. Samples were grouped in two ways: (1) first in ten groups, according to their respective sampling locality: Tasmania (TAS), New Zealand (NZ), Chile (CL), northwestern Atlantic (NWA), Faroe Islands (FI), United Kingdom (UK), Iberian Peninsula (IB), northeastern Atlantic (NEA), Strait of Gibraltar (GIB) and the Mediterranean Sea (MED). NEA included samples ranging from UK to IB, and was combined with the locality Euskadi from the original study, since no significant FST structure was detected among them in that study25 (Fig. 1, Table 1). (2) The second way of grouping samples was according to the subspecies categories, i.e. G. m. edwardii from the South Pacific (SP) including the localities TAS, NZ and CL and G. m. melas from the North Atlantic (NA), including NWA, FI, UK, IB, NEA and the Mediterranean (GIB and MED).

Sequence editing. After alignment and trimming, a haplotype network was constructed in Network 5.066. With an exploratory examination of the global haplotype network, it was noted that site 156 of the alignment generated three loops in the network. This hypervariable site was considered to be interfering with the phylogeographic signal of the data and was consequently removed, in order to eliminate a potential homoplasmy signal. Additionally, a repeated TA motif starting at position 90 was identified as a possible microsatellite. We modified the sequences by deleting one of the nucleotide positions within each repeat, so each motif was considered as a single mutational step, instead of each nucleotide separately. Thus the final fragment length was of 345 bp.

Genetic diversity and structure. The genetic diversity indices number of haplotypes (h), number of polymorphic sites (S), haplotype diversity (Hd), nucleotide diversity (π) and pairwise differences between sequences (D) were estimated in Arlequin v3.5.27. Analyses of genetic structure (FST), phylogenographic structure (ϕST) and analysis of molecular variance (AMOVA) were conducted in Arlequin v3.5.2 with 1000 permutations and a significance level of 0.05. Phylogeographic structure was also explored with Snn tests of genetic differentiation46, performed in DnaSP 5.10.0166. For the AMOVA, the ten localities were grouped according to the distribution of each subspecies. A Correspondence Analysis (CA) was performed on all localities with the software Past 3.1970, using the matrix of Table 1 in the form of presence/absence of haplotypes.

Additionally, as suggested by the guidelines for the delimitation of cetacean subspecies using genetic data of Taylor et al., Nei's (1987) net nucleotide divergence (dA, equation 10.21)41 was calculated among the two putative subspecies in DnaSP. According to these guidelines, the net nucleotide divergence among two subspecies should be within the range of dA = 0.004–0.04.

Historical biogeography. The population history of the species was tested on the program DIYABC v2.1.027. This software evaluates population histories using Approximate Bayesian Computation (ABC) with genetic data, by testing scenarios built through the combination of population divergence, admixture and population size changes. Two models were evaluated. The first model was defined based on a scenario previously proposed24, together with evidence from the genetic results provided in the present study. The model considers
a trans-equatorial, Last Glacial Maximum (LGM)-associated dispersal event from the Southern Hemisphere to the Northern Hemisphere, followed by a split in distribution and instantaneous population growth (Fig. 6). The alternative model differs from the previous one in that it considers a vicariance event, rather than a founder effect. The program was used to evaluate the accordance of these two scenarios with our genetic data. Priors were set as follows: Effective population size ($N_e$) of ancient population $= 1000–100 000$; Effective population size of founder effect ($N_f$) $= 10–1 000$; time of dispersal event $t_2 = 10 000–35 000$; time of instantaneous population growth $t_1 = 2 000–15 000$ (with $t_2 > t_1$) and mutation rate $u = 1.5 e^{-7}–1.5 e^{-8}$. In accordance with the recommendations of the authors of the software, we performed 6 000 000 simulations.

Microsatellite data. A total of 19 loci were amplified: 409/470, 464/46573, DlrFCB1, DlrFCB674, Ev1, Ev14, EV3775, GATA5376, GT6, GT5177, GT21, GT509, GT57578, MK5, MK979 and PPHO13158. PCR reactions were done with a Multiplex PCR kit (Qiagen), each reaction containing: 12.5 µL of water, 5 µL of MM2x Buffer and 1 µL of each primer at 10 pM. Between two and four loci with different fluorescent dyes were combined in each reaction. Allele scoring was done with the software GeneMarker v2.6.0 (www.softgenetics.com) with a 500 liz standard. The dataset was tested for scoring errors, allele dropout and null alleles in Micro-Checker v.2.2.380. Observed heterozygosity ($H_o$), expected heterozygosity ($H_e$), average number of alleles per locus (nA) and genetic structure ($F_{ST}$) were estimated in Genetix v.4.05.281. Published microsatellite data on this species was available from Tasmania and New Zealand27, NW A, UK28, Faroe Islands23,28, Iberian Peninsula23, northeastern Atlantic, Strait of Gibraltar and the Mediterranean25. However, because of the differences in the loci used in these studies, only a partial comparison of genetic diversity could be performed between populations of $G. m. melas$ and $G. m. edwardii$. Only comparisons of allelic richness could be done, which were carried out in Random Pro 1.182 with 10 000 randomizations. Comparisons were intra-subspecies among southwestern and southeastern Pacific samples, and inter-subspecies among southwestern Pacific and North Atlantic samples.

Approval. We confirm that all methods were carried out in accordance with relevant guidelines and regulations. Samples were taken from stranded, deceased animals with permission from the National Fisheries Service (SERNAPESCA, document ID 2016-11-13). All experimental protocols were approved by the Postgraduate Evaluation Committee at the Faculty of Science of the Universidad de Chile.

Accession codes. Haplotype R2 (GenBank accession number pending. Submission number #2305260).

Received: 14 May 2019; Accepted: 6 January 2020;
Published online: 04 February 2020

References
1. Whitehead, H. Sperm whale Physeter macrocephalus. In Würsig, B. G., Thewissen, J. G. M. & Kovacs, K. M. (eds.) Encyclopedia of marine mammals, 919–925, 3rd edition (Academic Press, an imprint of Elsevier, London, United Kingdom; San Diego, CA, United States, 2018).
2. Ford, J. K. Killer whale Orcinus orca. In Würsig, B. G., Thewissen, J. G. M. & Kovacs, K. M. (eds.) Encyclopedia of marine mammals, 531–537, 3rd edition. (Academic Press, an imprint of Elsevier, London, United Kingdom; San Diego, CA, United States, 2018).
3. Rojas-Bracho, L., Jaramillo-Legorreta, A. M., Nieto-García, E. & Cádiz-Hinojosa, G. Vaquita, Phocoena sinus. In Würsig, B. G., Thewissen, J. G. M. & Kovacs, K. M. (eds.) Encyclopedia of marine mammals, 1031–1035, 3rd edition. (Academic Press, an imprint of Elsevier, London, United Kingdom; San Diego, CA, United States, 2018).
4. LeDuc, R. Biogeography. In Würsig, B. G., Thewissen, J. G. M. & Kovacs, K. M. (eds.) Encyclopedia of marine mammals, 99–103, 3rd edition. (Academic Press, an imprint of Elsevier, London, United Kingdom; San Diego, CA, United States, 2018).

5. Kenney, R. D. Right Whales. In Würsig, B. G., Thewissen, J. G. M. & Kovacs, K. M. (eds.) Encyclopedia of marine mammals, 817–822, 3rd edition. (Academic Press, an imprint of Elsevier, London, United Kingdom; San Diego, CA, United States, 2018).

6. Jefferson, T. A., Leatherwood, S. & Webber, M. A. Marine mammals of the world (FAO, Rome, 1996).

7. Lipsky, J. D. Right Whale Dolphins (Lissodelphis peronii). In Perrin, W. F., Würsig, B. G. & Thewissen, J. G. M. (eds.) Encyclopedia of marine mammals, 2nd edition. (Academic Press, San Diego, 2002).

8. Olson, P. & Reilly, S. Pilot Whales. In Perrin, W., Würsig, B. & Thewissen, H. (eds.) Encyclopedia of marine mammals, 898–903, 2nd edition. (Academic Press, San Diego, 2002).

9. Kassuya, T. Past occurrence of Globicephala melaena in the western North Pacific. Scientific report of the Whales Research Institute 27, 95–110 (1975).

10. Frey, A., Crockford, S. J., Meyer, M. & O’Corry-Crowe, G. M. Genetic analysis of prehistoric marine mammal bones from an ancient atoll village in the southeastern Bering sea. Biennial Conference on the biology of marine mammals. San Diego, California, 12–16 December (2005).

11. Ottensmeyer, C. A. & Whitehead, B. Behavioural evidence for social units in long-finned pilot whales. Canadian Journal of Zoology 81, 1327–1338, https://doi.org/10.1139/z03-127 (2003).

12. Bernard, H. & Reilly, S. Pilot whales Globicephala Lesson, 1828. In Ridgway, S. & Harrison, R. (eds.) Handbook of marine mammals (Academic Press, San Diego, CA, 1994).

13. Sergeant, D. Mass stranding of toothed whales (Odontoceti) as a population phenomenon. The scientific reports of the Whales Research Institute 34, 1–47 (1982).

14. Amos, R., Barrett, J. & Dover, G. A. Breeding behaviour of pilot whales revealed by DNA fingerprinting. Heredity 67, 49–55 (1991).

15. Rayner, G. Globicephala leucosagmaphora, a new species of the genus Globicephala. The Annals & Magazine of Natural History. 11, 543–544 (1939).

16. Davies, J. L. The Southern Form of the Pilot Whale. The Annals & Magazine of Natural History. 13, 23–43 (1876).

17. Sergeant, D. On the External Characters of the Blackfish or Pilot Whales (Genus Globicephala). In Würsig, B. G., Thewissen, J. G. M. & Kovacs, K. M. (eds.) Encyclopedia of marine mammals, 933–935 (2002).

18. Kenney, R. D. Right Whales. In Würsig, B. G., Thewissen, J. G. M. & Kovacs, K. M. (eds.) Encyclopedia of marine mammals, 910–913 (2002).

19. LeDuc, R. Biogepgraphy. In Würsig, B. G., Thewissen, J. G. M. & Kovacs, K. M. (eds.) Encyclopedia of marine mammals, 812–817 (2002).

20. Oremus, M. Physical and biological traits. In Würsig, B. G., Thewissen, J. G. M. & Kovacs, K. M. (eds.) Encyclopedia of marine mammals, 817–822 (2002).

21. Moritz, C. D. ‘Evolutionarily Significant Units’ for conservation. Trends in Ecology & Evolution 6, 173–185, https://doi.org/10.1016/S0169-5348(94)90057-4 (1994).

22. Davies, J. L. The Antitropical Factor in Cetacean Speciation. Polar Biology 4, 949–958 (2000).

23. Crockford, S. J., Meyer, M. & Hoelzel, A. R. Genetic patterns of bottlenose dolphins (Tursiops truncatus) from the Leeuwin continental shelf. Marine Mammal Science 17, 660–676, https://doi.org/10.1111/j.1365-2907.2001.00193.x (2005).

24. Verborgh, P. In Würsig, B. G., Thewissen, J. G. M. & Kovacs, K. M. (eds.) Encyclopedia of marine mammals, 910–913 (2002).

25. Verborgh, P. In Würsig, B. G., Thewissen, J. G. M. & Kovacs, K. M. (eds.) Encyclopedia of marine mammals, 910–913 (2002).

26. Fullard, K. J. et al. Population structure of long-finned pilot whales in the North Atlantic: a correlation with sea surface temperature? Molecular Ecology 9, 949–958 (2000).

27. Oremus, M., Gales, R., Kettles, H. & Baker, C. S. Genetic evidence of multiple matrilines and spatial disruption of kinship bonds in Mass strandings of long-finned pilot whales (Globicephala melas). Journal of Heredity 104, 301–311, https://doi.org/10.1093/jhered/est007 (2013).

28. Fullard, K. J. et al. Population structure of long-finned pilot whales in the North Atlantic: a correlation with sea surface temperature? Molecular Ecology 9, 949–958 (2000).

29. Cunha, H. A. et al. Population structure of the Endangered Franciscana Dolphin (Pontoporia blainvillei): Reassessing Management Units. PLOS ONE 9, e85633, https://doi.org/10.1371/journal.pone.0085633 (2014).

30. Parsons, K. M., Noble, L. R., Reid, R. J. & Thompson, P. M. Mitochondrial genetic diversity and population structuring of UK bottlenose dolphins (Tursiops truncatus): is the NE Scotland population demographically and geographically isolated? Biological Conservation 108, 175–182, https://doi.org/10.1016/S0006-3207(02)00103-9 (2002).

31. Duke, P. W. et al. Worldwide structure of mtDNA diversity among Cuvier’s beaked whales (Ziphius cavirostris): implications for threatened populations. Molecular Ecology 14, 3353–3371, https://doi.org/10.1111/j.1365-294X.2005.02676.x (2005).

32. Viricel, A., Strand, A. E., Rosel, P. E., Ridoux, V. & Garcia, P. Insights on common dolphin (Delphinus delphis) social organization from genetic analysis of a mass-stranded pod. Behavioral Ecology and Sociobiology 65, 173–185, https://doi.org/10.1007/s00265-008-0468-7 (2008).

33. Alvarado-Rybák, M. A. Mass stranding event of long-finned pilot whales (Globicephala melas) in southern Chile. Aquatic Mammals 45, 447–455, https://doi.org/10.1578/AM.45.4.2019.447 (2019).

34. Mansilla, L., Olavarria, C. & Vega, M. A. Stomach contents of long-finned pilot whales (Globicephala melas) from southern Chile. Polar Biology 33, 1929–1933, https://doi.org/10.1007/s00300-012-1222-3 (2010).

35. McCracken, K. G. & Sorenson, M. D. Is Homoplaspy or Lineage Sorting the Source of Incongruent mtDNA and Nuclear Gene Trees in the Stiff-Tailed Ducks (Nomonyx-Oxyura)? Systematic Biology 54, 35–55, https://doi.org/10.1006/sysb.2003.1029 (2005).

36. Jackson, J. A. et al. Global diversity and oceanic divergence of humpback whales (Megaptera novaengliae). Proceedings of the Royal Society B: Biological Sciences 281, https://doi.org/10.1098/rspb.2013.3222 (2014).

37. Herrnstadt, C. et al. Reduced-Median-Network Analysis of Complete Mitochondrial DNA Coding-Region Sequences for the Major African, Asian, and European Haplogroups. The American Journal of Human Genetics 70, 1152–1171, https://doi.org/10.1086/339933 (2002).

38. Hoezel, A. R. et al. Low worldwide genetic diversity in the killer whale (Orcinus orca): implications for demographic history. Proceedings of the Royal Society B: Biological Sciences 269, 1467–1473, https://doi.org/10.1098/rspb.2002.2033 (2002).

39. Martien, K. K. et al. Nuclear and Mitochondrial Patterns of Population Structure in North Pacific False Killer Whales (Pseudorca crassidens). Journal of Heredity 105, 611–626, https://doi.org/10.1093/jhered/esu209 (2014).

40. Mason, S. et al. What influences the worldwide genetic structure of sperm whales (Physeter macrocephalus)? Mol. Ecol. 25, 2754–2772, https://doi.org/10.1111/mec.13638 (2016).

41. Whitehead, H., Vachon, F. & Fraser, T. R. Cultural Hitchhiking in the Matrilineral Whales. Behavioral Genetics 47, 324–334, https://doi.org/10.1007/s10519-017-9840-8 (2017).
81. Belkhir, K., Borsa, P., Chikhi, L., Raufaste, N. & Bonhomme, F. Genetix 4.05, logiciel sous windows pour la génétique des populations. laboratoire génome, populations, interactions, cnrs unir 5000. Univ. de Montpellier II, Montpellier, France (2004).
82. Jadwiszczak, P. Random Projects: an application for randomization and bootstrap testing, version 1.1. Computer program available from the web site http://pjadw.tripod.com. (2003).

Acknowledgements
We would like to thank the Chilean Navy for providing transportation and personnel for the sampling in Isla Clemente and Beagle Channel, and the National Fisheries Service (SERNAPESCA) for helping and organizing the Isla Clemente expedition and storing and sending the samples to our laboratory in Santiago. We would also like to thank Miguel Íñiguez, Elsa Cabrera, Barbara Galetti (CCC Chile) for the help during the collection of samples in the Isla Navarino stranding and Fundación CEQUA for their support. Funding: Beca CONICYT 2017 de Magister Nacional 22171215, Instituto de Ecología y Biodiversidad (IEB, Proyecto CONICYT PIA APOYO CCTE AFB170008), Proyecto PIA CONICYT ACT172065, Proyecto FONDECYT iniciación 11170182, Proyecto Fortalecimiento de Centros Regionales del Programa Regional CONICYT R16A10003.

Author contributions
Sample collection: S.K. collected the Isla Clemente samples, C.O. collected the Isla Navarino samples. Laboratory work: S.K. extracted and amplified both mitochondrial and microsatellite markers for the Isla Clemente samples, C.O. extracted and amplified the mitochondrial marker for the Isla Navarino samples. S.K. amplified the microsatellite markers for the Isla Navarino samples. S.K. prepared and analyzed the data and prepared the figures and tables. S.K., M.J.P.A. and E.P. interpreted and discussed the results. All authors reviewed the manuscript.

Competing interests
The authors declare no competing interests.

Additional information
Correspondence and requests for materials should be addressed to M.P.-Á.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2020