Research

Complete mitochondrial genome phylogeographic analysis of killer whales (*Orcinus orca*) indicates multiple species

Phillip A. Morin,[1,2,8] Frederick I. Archer,[1] Andrew D. Foote,[3,4] Julia Vilstrup,[3] Eric E. Allen,[2] Paul Wade,[5] John Durban,[5] Kim Parsons,[5] Robert Pitman,[1] Lewyn Li,[6] Pascal Bouffard,[6] Sandra C. Abel Nielsen,[3] Morten Rasmussen,[3] Eske Willerslev,[3] M. Thomas P. Gilbert,[3] and Timothy Harkins[7]

[1]National Marine Fisheries Service, Southwest Fisheries Science Center, La Jolla, California 92037, USA; [2]Scripps Institution of Oceanography, University of California, San Diego, La Jolla, California 92037, USA; [3]Centre for GeoGenetics, Natural History Museum of Denmark, University of Copenhagen, 1350 Copenhagen, Denmark; [4]University of Aberdeen, Aberdeen IV11 8YJ, United Kingdom; [5]Alaska Fisheries Science Center, NOAA Fisheries, Seattle, Washington 98115, USA; [6]454 Life Sciences [Roche], Branford, Connecticut 06405, USA; [7]Roche Applied Science, Indianapolis, Indiana 46250, USA

Killer whales (*Orcinus orca*) currently comprise a single, cosmopolitan species with a diverse diet. However, studies over the last 30 yr have revealed populations of sympatric "ecotypes" with discrete prey preferences, morphology, and behaviors. Although these ecotypes avoid social interactions and are not known to interbreed, genetic studies to date have found extremely low levels of diversity in the mitochondrial control region, and few clear phylogeographic patterns worldwide. This low level of diversity is likely due to low mitochondrial mutation rates that are common to cetaceans. Using killer whales as a case study, we have developed a method to readily sequence, assemble, and analyze complete mitochondrial genomes from large numbers of samples to more accurately assess phylogeography and estimate divergence times. This represents an important tool for wildlife management, not only for killer whales but for many marine taxa. We used high-throughput sequencing to survey whole mitochondrial genome variation of 139 samples from the North Pacific, North Atlantic, and southern oceans. Phylogenetic analysis indicated that each of the known ecotypes represents a strongly supported clade with divergence times ranging from ~150,000 to 700,000 yr ago. We recommend that three named ecotypes be elevated to full species, and that the remaining types be recognized as subspecies pending additional data. Establishing appropriate taxonomic designations will greatly aid in understanding the ecological impacts and conservation needs of these important marine predators. We predict that phylogeographic mitogenomics will become an important tool for improved statistical phylogeography and more precise estimates of divergence times.

[Supplemental material is available online at http://www.genome.org. The sequence data from this study have been submitted to GenBank (http://www.ncbi.nlm.nih.gov/genbank) under accession nos. GU187153–GU187164, GU187166–GU187219, and HM060332–HM060334.]

Theory and empirical studies have shown ecology to be a driving force in speciation (Schluter 2009). Ancestral populations can subdivide and radiate into novel ecological niches and are then subject to divergent selection and subsequent adaptive divergence, which can lead to reproductive isolation and speciation (Gavrilets and Losos 2009; Schluter 2009). This process of radiation into novel and divergent ecological niches is often characterized by a rapid burst of phenotypic diversification, which then slows as disparate ecological niches become filled (Gavrilets and Losos 2009), and is consistent with phylogenies showing the greatest ecological differences early in a clade's history (Grant and Grant 2008; Losos 2009).

Maintaining high levels of biodiversity by conserving both this process and the resultant genetic and phenotypic variation is an important goal of management bodies (Moritz 2002). Determining units on divergent evolutionary trajectories can facilitate

this, and a number of criteria and concepts have been suggested for defining species, subspecies, or management units. Some concepts are based purely upon genetic criteria such as reciprocal monophyly of mitochondrial DNA (mtDNA) haplotypes and significant divergence of allele frequencies at nuclear loci (e.g., Moritz 1994); others incorporate ecological and phenotypic data to assess "exchangeability" between putative species (e.g., Crandall et al. 2000; De Queiroz 2007). De Queiroz (2007) argued that the many concepts all agree in the basic description of species as independently evolving metapopulations, and that the criteria for defining these species all boil down to different types of supporting evidence.

Phylogenetic analysis using mtDNA is a widely used tool for the genetic component of delineating species (Moritz 1994). The generally rapid rate of mtDNA sequence evolution and lineage sorting (relative to the nuclear genome) facilitate inference of evolutionary patterns (Brown et al. 1982; Avise 1989; Moore 1995), especially for social species with a matrilineal group structure, which is common among terrestrial and oceanic mammals (e.g., Lyrholm and Gyllensten 1998; Nyakaana and Arctander 1999; Okello et al. 2008). Despite the general assumptions and its wide use, however, mtDNA

sequence can in some cases be an uninformative marker for phylogenetics and species delineation if only portions of the genome are used (Galtier et al. 2009b). Indeed, the mitochondrial "molecular clock" varies widely and has been shown to be especially slow in some taxa, e.g., cetaceans and sharks (Nabholz et al. 2008a,b, 2009; Galtier et al. 2009a). Short sections of the mtDNA genome can therefore be uninformative phylogeographic markers in these taxa (e.g., Hoelzel et al. 2002). A further limitation of traditional mtDNA sequencing that focused on either the control region or cytochrome *b* sequences has been the inability to resolve relationships when a radiation was very rapid. Greater resolution of phylogenies can be achieved by increasing the amount of sequence data (Saitou and Nei 1986; Ruvolo et al. 1991; Cummings et al. 1995; DeFilippis and Moore 2000; Rokas and Carroll 2005).

The advent of highly parallel long-read pyrosequencing technologies with targeted resequencing of large genetic regions from genetically tagged and pooled samples now makes it possible to rapidly and efficiently obtain orders-of-magnitude more sequence data than was previously possible with Sanger sequencing (Meyer et al. 2008). Although whole mitochondrial genomes (mitogenomes) are now available for a number of species, they have typically been generated for deep phylogenetic analysis, and to allow more precise estimates of long divergence times (Arnason et al. 2004, 2008; Xiong et al. 2009). To date, only a few studies have made use of phylogeographic mitogenomics to investigate patterns within a genus or species, and most have involved the use of few genomes, typically from species of medical or economic interest (e.g., Zarowiecki et al. 2007; Carr and Marshall 2008; Torriani et al. 2008). The ability to cost-effectively sequence the entire mitochondrial genome from larger numbers of samples for phylogeographic studies should help to resolve previously intractable polytomies resulting from low levels of sequence divergence or rapid radiations of many more species.

As a case study to investigate the potential of using whole mitochondrial genomes for phylogeography, we have undertaken a study of one such "difficult" species. Killer whales (*Orcinus orca*) are apex predators found in all the world's oceans (Forney and Wade 2006). Currently considered a single species (Rice 1998), local variation in a number of characteristics, including body size, color patterning, social structure, vocalization pattern, and behavior, has led to the recognition of several named killer whale types (often referred to as "ecotypes") (Barrett-Lennard et al. 1996; Ford et al. 1998; Baird 2000; Pitman and Ensor 2003; Deecke et al. 2005; Pitman et al. 2007; Foote et al. 2009; Parsons et al. 2009). In particular, prey specialization appears to be a defining characteristic of these types, with partially or fully sympatric populations having specific, sometimes nonoverlapping prey preferences (e.g., fish vs. marine mammals) (Ford et al. 1998; Saulitis et al. 2000; Pitman and Ensor 2003; Herman et al. 2005; Krahn et al. 2007b). Although ecological specialization is not uncommon (Gavrilets and Losos 2009; Schluter 2009), the fact that killer whales exhibit specialization within an ecosystem that is largely based on social mechanisms is of great interest, suggesting that speciation may have occurred in the absence of physical barriers to gene flow. Many killer whale populations are being negatively impacted by human activities, such as over-fishing and pollution, and such threats are likely to vary substantially between types (e.g., Ross et al. 2000; Ylitalo et al. 2001). Effective management therefore requires the delineation of conservation units (Moritz 1994) within the genus *Orcinus* to facilitate different management strategies.

Despite a worldwide distribution and phenotypic differences among killer whale types, genetic diversity of mtDNA is low, and

the control region and other mtDNA loci have been used with limited success to determine population structure and phylogeography. In a survey of ~1000 bp of the control region from over 100 samples from various locations around the world, Hoelzel et al. (2002) found only 13 haplotypes and no clear pattern of genetic association with ocean basin or type. They concluded that killer whales had gone through a worldwide bottleneck ~145,000–210,000 yr ago (i.e., during the Pleistocene), and that the genetic patterns reflected stochastic distribution of mitochondrial haplotypes following the post-bottleneck expansion, rather than phylogenetic lineages reflecting the evolution of ecotypes. Analysis of an expanded set of mtDNA control region sequences by LeDuc et al. (2008), including 80 samples from three described types in the Southern Ocean around Antarctica, found similar patterns, but also found that two Antarctic types associated with the ice edge were each monophyletic, albeit with very low levels of differentiation. Indeed, levels of differentiation among types worldwide have been marked by only one to six fixed differences and total genetic distances of <0.3% for the most divergent control region lineages. This low level of mtDNA diversity has resulted in only weak inference of phylogeographic patterns and divergence times in killer whales, limiting our ability to understand their evolution and taxonomy. Studies of nuclear microsatellites have begun to clarify population structure within ecotypes, and propose even more recent divergence of regional ecotypes within the last 20,000–40,000 yr (Hoelzel et al. 2007; Pilot et al. 2010).

Killer whales are therefore an ideal candidate species for applying new high-throughput sequencing techniques to allow the production of a highly corroborated mitogenome tree and the testing of hypotheses of the timing of coalescence of killer whale populations (e.g., Hoelzel et al. 2002), with a precision of temporal discrimination not previously possible. Specifically, we test the hypotheses that killer whale ecotypes radiated toward the end of the Pleistocene, that ecotypes diversified regionally within ocean basins, and that mitochondrial haplotypes are stochastically distributed among ecotypes.

Results

Full-length mitochondrial genomes of ~16,390 bp (16,386–16,392 bp) were sequenced and assembled for 143 killer whale samples (Fig. 1) and three other cetacean species (false killer whale *Pseudorca crassidens* and long- and short-finned pilot whales *Globicephala macrorhynchus*, *G. melas*) for use as outgroups. An additional five partial mitochondrial sequences were generated, with one or more gaps in the sequence ranging from 35 bp to ~9 kb. Analysis of eight full mitochondrial sequences that were replicated yielded only two differences (not including polynucleotide repeats), for a sequence error rate of ~0.00076%. These included one sample sequenced in separate U.S. and Danish sequencing facilities as well as intra-lab replication. Mutation rate estimates for the whole mitogenome were 2.6×10^{-3} (1.50–3.83×10^{-3}) substitutions per site per million years for *Orcinus*. This rate is lower than the mean for mammals (3.3×10^{-2}), and similar to rates estimated for other cetaceans (2.3×10^{-3}, extrapolated from third positions only) (Nabholz et al. 2008a).

Previously published mitogenome sequences were combined with sequences generated in this study to estimate the time of divergence for the genus *Orcinus* (Fig. 2A). Within *Orcinus*, 139 mitogenomes (66 distinct haplotypes) were used for further analysis, after removal of duplicates and one poor-quality sequence. Bayesian analysis of the 66 unique haplotypes produced the phylogenetic tree
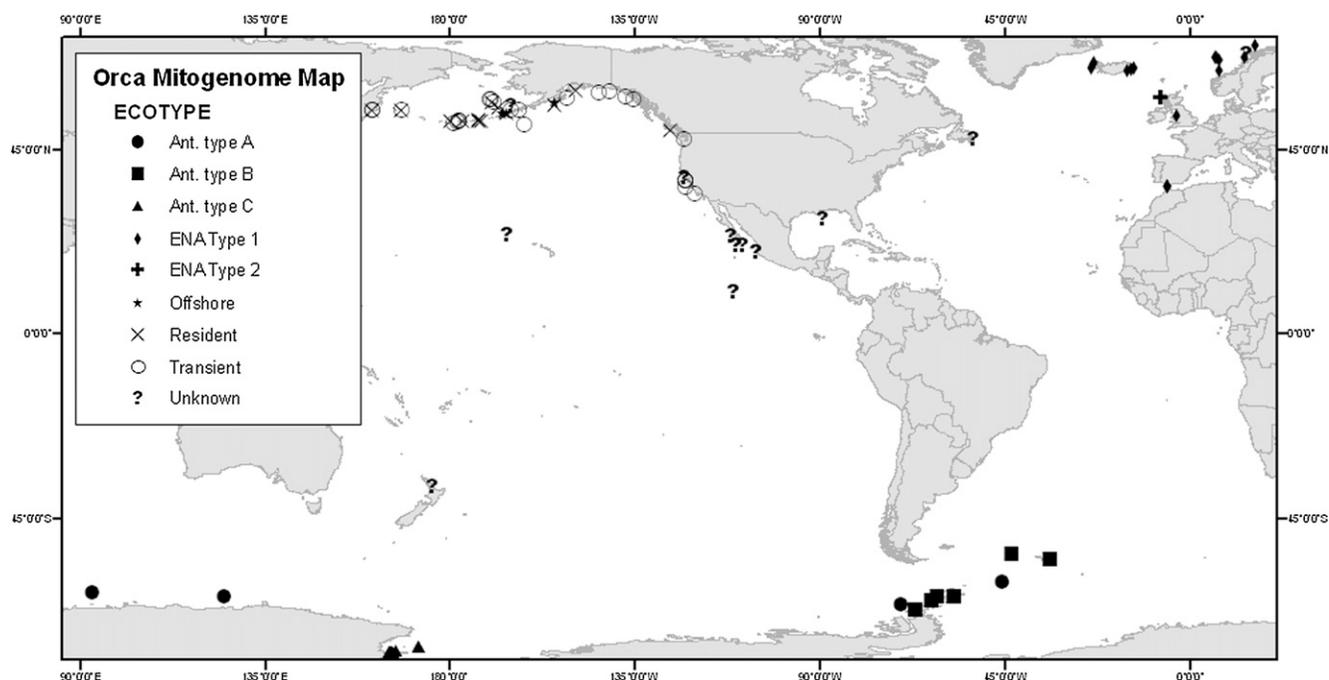
Figure 1. Sample collection locations with indication of type when known.

shown in Figure 2B. In contrast to results based on the control region sequences alone (Hoelzel et al. 2002; LeDuc et al. 2008), no haplotypes were shared between killer whale types or ocean basins. However, animals of unknown type from multiple oceanic regions were grouped in the phylogenetic tree with the Offshore, eastern North Atlantic (ENA) Type 1, Antarctic Type A, and Transient types (Fig. 2B), indicating possible common ancestry of widely separated populations, or, in the case of the one putative Antarctic A type that clusters with eastern North Pacific (ENP) Transient type (predominantly coastal mammal-eating specialist; Ford and Ellis 1999; Baird 2000), potentially an as yet undescribed new killer whale type in the Southern Ocean.

The most striking feature of the phylogenetic tree is the relatively deep divergence of the ENP Transient clade from all others, including two sympatric groups, the Offshore and Resident types. The estimated divergence time is 700,000 yr ago (95% highest posterior density interval [HPDI] = 488,000–960,000), and this clade has 40 fixed differences from all other samples when the single Type A sample is excluded (36 when included). This is 17% of all of the variable sites detected genus-wide. All other types differ from each other by three to 25 fixed differences (Table 1). The time to most recent common ancestor (TMRCA) within Transients is estimated at ~190,000 yr ago, and other types have mean TMRCAs ranging from 59,000 to 117,000 yr ago (Table 2). The Antarctic types (including the single sample from the Gulf of Mexico) together form a clade with a mean TMRCA of ~330,000 yr.

The "Resident" clade includes all fish-eating, coastal Resident types (Baird 2000). The "Offshore" clade includes all ENP Offshore types (a relatively little known type thought to specialize on bony fish and elasmobranchs; Baird 2000; Dahlheim et al. 2008) and other Pacific Ocean samples from off western Mexico and Clipperton Island that were not previously identified as part of the Offshore population. Interestingly, one sample from Newfoundland (western North Atlantic; WNA) also clusters with the ENP

Offshore haplotypes, indicating either an origin of the ENP Offshore and Resident groups from ancestral populations in the North Atlantic, or a remigration of animals in the ENP to the WNA via the Northwest Passage during periods of climate warming.

The "ENA 1-2" clade contains two haplotypes that diverged at approximately the same time as the ENP Offshore and Resident clades. These are from animals sampled near Iceland, Scotland, and England. The Icelandic and English samples were previously categorized as a generalist type (North Atlantic Type 1) that includes individuals specializing on fish and individuals that are thought to predate both fish and mammals (Foote et al. 2009). The Scottish sample was categorized as being from a poorly characterized North Atlantic specialist type (Type 2) and had previously been clustered with Antarctic killer whales based on control region data (Foote et al. 2009). ENA type 2 killer whales were represented by only a single sample, and the sequence contained a large gap (3848 bp). This type has been characterized primarily using museum specimens (Foote et al. 2009), not suitable for long-range PCR. Sequencing methods used to obtain ancient mitogenome sequences (see Ho and Gilbert 2010) may be more suitable for further investigating the phylogenetic relationship between types within the ENA.

The "ENA Type 1" clade clusters a sample from New Zealand with whales from Iceland, Norway, and the Strait of Gibraltar. ENA type 1 has recently been described based on diet and morphology (Foote et al. 2009). Together, the ENA samples cluster closely to the ENP Resident and Offshore types. Samples from both ocean basins appear to have similar levels of haplotypic diversity, so it is unclear whether the ancestral population was in the North Atlantic or the North Pacific, or if both arose from a broader generalist population in lower latitudes that is not yet adequately sampled.

There are three described Antarctic types that differ in diet and morphology (Pitman and Ensor 2003; Pitman et al. 2007). All
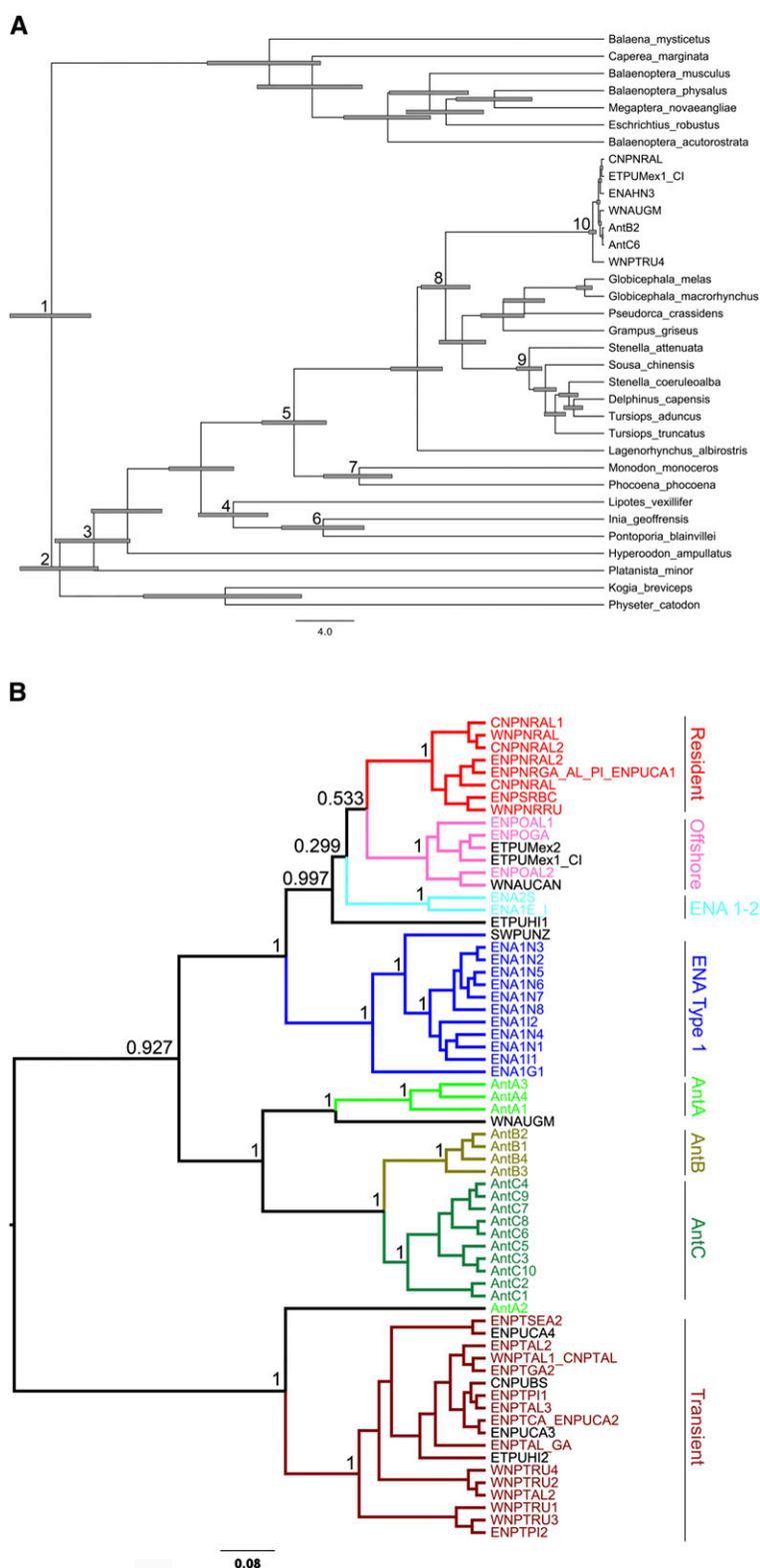
Figure 2. (*A*) Bayesian phylogenetic tree of cetacean phylogeny of whole mitochondrial genomes from the public databases and new species from this study, including 95% highest posterior density interval (HPDI) bars. Nodes with divergence priors are indicated by numbers corresponding to taxonomic groups described in Supplemental Table S5. (*B*) Whole mitochondrial genome phylogeny of 66 unique killer whale haplotypes. Posterior probabilities are indicated for nodes of interest. Whales of known type are indicated in color, and those of unknown type are in black type.

Antarctic samples except one (AntA2) cluster in a monophyletic group that also includes one sample from the Gulf of Mexico in the WNA, and this group is more closely related to all other types in both the ENA and ENP than to the ENP Transients. Two of the three types (B and C) in the Antarctic are also monophyletic, with three fixed differences between types B and C and 24–25 fixed differences between type A (excluding AntA2) and types B and C, respectively. The single sample from the Southern Ocean that does not cluster with the Antarctic types clusters basal to the Transient-type samples, indicating that it may in fact represent a separate population of Transient-like whales in the southern hemisphere. Further sampling in this region is warranted, especially in light of recent observations of a possible fourth pelagic killer whale type in the southern oceans (Jefferson et al. 2007).

Mitochondrial DNA, though useful in phylogeographic studies, has the limitation of being a single, maternally inherited locus. Nuclear markers are needed to obtain data from multiple loci and from both male and female genetic components. Although microsatellites are a poor marker for taxonomic questions, they are the marker of choice for many intraspecific studies, and have been used to study killer whale types. In the North Pacific, previous analysis of 16 nuclear microsatellites has shown that types are genetically very distinct (Hoelzel et al. 2007). Our data from 26 microsatellites genotyped from samples in the North Pacific and Antarctic indicate similarly high levels of differentiation among all killer whale types (Table 3), when divergence is calculated as Hedrick's G'_{ST} to control for intrapopulation levels of diversity (Hedrick 2005). These data are somewhat preliminary due to small sample sizes for three of the killer whale types and no data from ENA types, and though they do not rule out historical or even ongoing low levels of gene flow, they do indicate substantial genetic divergence among types, greater than reported among the majority of comparisons in a recent review of conspecific divergence (Heller and Siegismund 2009). In that review, comparisons that exhibited G'_{ST} levels similar to *sympatric* killer whale types typically involved species where extrinsic barriers to gene flow existed and/or populations exhibited local adaptation (e.g., European wild boars separated by the Alps [Scandura et al. 2008] and locally adapted Atlantic salmon [Dionne et al. 2008]).

Table 1. Fixed differences between killer whale types and oceanic regions

Ecotype comparison	Fixed differences	Variable sites	Haplotypes
North Pacific Transient vs. North Pacific Resident	57	110	27
North Pacific Transient vs. North Pacific Offshore[a]	57	113	25
North Pacific Resident vs. North Pacific Offshore[a]	10	30	14
Antarctic A[b] vs. Antarctic B	24	41	7
Antarctic A[b] vs. Antarctic C	25	55	13
Antarctic B vs. Antarctic C	3	25	14
All Antarctic[c]	5	236	64
Transient[d] vs. non-Transient	40	236	64

Two haplotypes with large sequence gaps (ENASU and AntC7) were not included in calculations.
[a]Offshores include unknown ecotype sample from western Atlantic, Newfoundland (WNAUCAN), and unknown type samples from Mexico and Clipperton Island.
[b]All Antarctic A except sample AntA2, which clusters with the North Pacific Transients (see text).
[c]All Antarctic except sample AntA2, compared with all other complete haplotypes.
[d]All samples in Transient clade (not including AntA2), compared with all other complete haplotypes.

Discussion

Killer whale phylogenetics has been troublesome because of the extremely low levels of diversity found in the mitochondrial control region, so phylogenetic inference was weak or nonexistent (Hoelzel et al. 2002; LeDuc et al. 2008). With highly parallel sequencing technologies, we have developed methods to sequence and assemble whole mitochondrial genomes from representative geographic and ecotype samples to provide strong inference of killer whale evolutionary patterns for the first time. The percent divergence among killer whale types was typically ≥50% higher in the control region than over the whole mitogenome (data not shown), suggesting that simply adding other short segments of mtDNA to the analysis would not significantly improve phylogenetic resolution. Additionally, Bayesian analysis in two phases allowed us to provide a much more accurate and precise date for the most recent common ancestor of all killer whales, and to use that date to estimate divergence times for each of the killer whale types.

Our estimated dates of divergence based on the entire mitogenome are much older than those inferred from short mitochondrial and nuclear loci. Previous studies using mitochondrial control region sequences and microsatellites have inferred that there was a Pleistocene bottleneck ~145,000–200,000 yr ago that reduced variation in killer whale populations globally, followed by recent divergence among the known ecotypes in high-latitude coastal regions. The divergence times were estimated at 20,000–40,000 yr, with wide confidence intervals (Hoelzel et al. 2002, 2007). Using these inferences as our hypotheses, we used whole mitogenomes to infer killer whale evolutionary patterns, and our results indicate much deeper initial separation (either geographic or ecological) between the mammal-eating Transient clade in the North Pacific and a second clade in the Atlantic or lower latitudes ~700,000 yr ago, followed by ecological and/or geographical diversification of the second clade into the present day types at high latitudes, including secondary contact with Transients. These splits between types date from ~150,000–700,000 yr ago rather than 20,000–40,000 yr ago, consistent with species or subspecies level designations. Given the clear lack of phylogenetic information in mitochondrial control region sequences, and the high mutation rates that could cause microsatellites to reach saturation over the time periods that we have inferred from mitogenomes, we believe the mitogenome data provide much stronger support for inference of divergence times.

Recent reviews on subspecies definitions in general have recognized the difficulty in coming up with criteria that will work for defining subspecies and species in all taxa, but they generally agree that data should support discreteness of the subspecies in relation to the remainder of the species and biological significance of the subspecies (Haig et al. 2006). Species concepts are no closer to being universally accepted, but De Queiroz (2007) pointed out that all such concepts agree that species are separately evolving metapopulation lineages, and their delineation is primarily done by accumulation of lines of evidence. Using these criteria, we argue that the combined genetic, morphological, and behavioral evidence of divergence among sympatric types in the high-latitude regions of the North Pacific and Southern Ocean support the recognition of these types as separately evolving metapopulation lineages, and the elevation of three types to species, and several others to subspecies status.

It has previously been suggested that the Southern Ocean B and C types warrant species status on morphological grounds under the biological species concept (BSC), pending confirmation from genetic studies (Pitman and Ensor 2003; Pitman et al. 2007). The genetic data presented here provide such confirmation, demonstrating that the two pagophilic (ice-associated) forms (B and C) are reciprocally monophyletic and form sister taxa substantially divergent from both open-water type A and all other killer whales. Therefore, we recommend that they be designated as distinct species that have diverged from one another for ~150,000 yr and suggest that further analysis of nuclear sequence data should be performed for confirmation.

Our mitogenome data also indicate that the North Pacific Transients should be considered an independent species. Not only are they ecologically and morphologically distinct from other high-latitude killer whales, but genetically they are the most divergent type, diverging from all other killer whale types ~700,000 yr ago. Taxonomic status for the Antarctic A type, North Atlantic types, and North Pacific Resident and Offshore types is less clear due to limited morphological divergence or limited morphological and ecological information, and/or small genetic sample sizes. As such, lines of evidence are not strong for species designation, and

Table 2. Median time to most recent common ancestor (TMRCA) and 95% highest posterior density interval (HPDI) for killer whale types

Clade	TMRCA (million yr)	95% HPDI
Antarctic A	0.117	0.044–0.212
Antarctic B	0.059	0.013–0.136
Antarctic C	0.116	0.051–0.214
Antartic B and C	0.151	0.067–0.275
All Antarctic[a]	0.331	0.182–0.555
North Pacific Resident	0.080	0.031–0.157
North Pacific Offshore	0.087	0.036–0.166
North Pacific Transient	0.188	0.097–0.312
North Pacific Resident and Offshore	0.177	0.085–0.305
Eastern North Atlantic Type 1[b]	0.168	0.078–0.303
All killer whales	0.702	0.489–0.956

[a]Including single sample from the Gulf of Mexico.
[b]Including single sample from New Zealand.

Table 3. The harmonic mean of standardized differentiation (G′$_{ST}$) below diagonal, nonstandardized differentiation (G$_{ST}$) above, based on 26 microsatellite loci

	Resident	Offshore	Transient	Antarctic A	Antarctic B	Antarctic C
Resident	—	0.05	0.07	0.07	0.17	0.15
Offshore	0.19	—	0.04	0.03	0.14	0.12
Transient	0.28	0.15	—	0.02	0.07	0.06
Antarctic A	0.32	0.12	0.10	—	0.06	0.05
Antarctic B	0.56	0.52	0.32	0.30	—	0.02
Antarctic C	0.58	0.53	0.32	0.33	0.11	—

we recommend subspecies status pending additional nuclear sequence and morphological data.

Low levels of mtDNA variation have limited our ability to resolve evolutionary patterns in killer whales and some other cetacean species. The ability to use whole mitogenome sequences has allowed resolution of phylogeographic patterns. In killer whales these patterns are consistent with historical ecological specialization of small populations (or even single maternal groups) in each region reinforced by either temporary allopatry or in sympatry with social and behavioral isolating mechanisms. Given the typically very small population sizes of killer whale populations, numbering in the low hundreds to thousands, with higher densities in high-latitude coastal regions (Barrett-Lennard and Ellis 2001; Forney and Wade 2006), monophyly might arise quickly in divergent types with little or no female dispersal (Parsons et al. 2009), or dispersal limited to groups with similar vocal and behavioral patterns (Baird et al. 1992; Baird 2000).

The limited sampling in lower latitudes, where diversity is relatively high but density is typically low (Forney and Wade 2006), may mean that we have missed sharing of haplotypes across ocean basins in subtropical and tropical waters, but the patterns for the high latitudes are strongly supported by the more extensive sampling presented here. Most recognized types (except ENA types) have fixed differences in the mitogenomes, indicating independent evolution of each type. The pattern of relatedness among clades is consistent with the independent evolution of feeding specialization in different ocean basins, and apparent diversification of most types within each ocean basin, with the exception of the early separation of Transients from all others.

Type-specific prey specialization largely defines the ecological roles of killer whales and also determines their exposure to human impacts such as fisheries depletion of prey and bioaccumulation of pollutants. Each of these potential species, subspecies, or ecotypes represents a top predator in its ecosystem (or multiple top predators in areas where they are sympatric). As such, and because they are globally distributed, killer whales are critical components of the ocean ecosystems, and represent substantial biological and ecological diversity. Human impacts including over-fishing, persistent organic pollutants, and climate change are already affecting some killer whale populations (e.g., Hickie et al. 2007; Krahn et al. 2007a). Establishing appropriate taxonomic designations is critical for understanding the ecological impacts and conservation needs of these important marine predators, and for maintaining biological diversity and ecosystem health.

As indicated in this killer whale case study, previously reported limitations of using short DNA sequences can be overcome by using whole mitogenomes. High-throughput mitogenomics provides a new tool for intra- and interspecific phylogeography that addresses many of the problems of limited diversity and variable mutation rates and patterns found in short segments of mitochondrial loci.

In addition, the long sequences provide both greater power for phylogenetic inference, and greater precision in estimation of divergence times. We expect that, as sequencing technologies continue to allow more samples, more sequence, and lower cost, the application of mitogenomics will become the default approach to phylogeography, as was previously the use of control region and cytochrome *b* sequence analysis.

Methods

DNA extraction and long-range PCR amplification

Skin biopsy samples were obtained from free-ranging killer whales by dart biopsy (e.g., Barrett-Lennard et al. 1996), or from stranded animals. Samples were selected to cover the broadest geographic range as well as genetic and killer whale type diversity. For the mitogenome analysis, most samples were selected from separate collection dates and, when known, identified groups to minimize chances of collecting close relatives or replicate individuals, except in the Southern Ocean, where all samples that had been assigned to one of the three types were included in the sample list (though not all were successfully sequenced). The sequenced sample set included: five (four, after removal of duplicate samples) Antarctic type A (Ant_A), 18 (15) Antarctic Type B (Ant_B), 39 (36) Antarctic type C (Ant_C), five ENP offshore, 11 ENP resident, 17 ENP transient, 12 ENP unknown, 12 eastern tropical Pacific (ETP) unknown, one Gulf of Mexico (unknown), one Newfoundland western Atlantic (unknown), one western South Pacific (New Zealand, unknown), 20 ENA Type 1, one ENA Type 2. In the Antarctic sample sets, we used all available samples, including one to four individuals each from 11 social groups of type C, and one to seven individuals from six social groups of type B whales. In the ETP, multiple (two to four) samples from four social groups were included in the analysis. All other samples are thought to be single samples from a social group. Sample collection locations and types are shown in Figure 1, and additional sample information is shown in Supplemental Table S1.

DNA was extracted using a variety of common extraction methods, including silica-based filter membranes (Qiagen), standard phenol/chloroform extraction (modified from Sambrook et al. 1989), and lithium chloride (Gemmell and Akiyama 1996). Outgroup sequences were obtained from the NCBI GenBank or generated for this study (Supplemental Table S2).

PCR primers were designed from alignment of published whole mitochondrial genomes of other cetacean species, and partial mitochondrial genome sequences of killer whales to amplify the whole mitogenome in two to five overlapping fragments (Supplemental Table S3). PCR conditions for long-range amplification are given in Supplemental Table S3.

454 Life Sciences (Roche) sequencing

PCR products were quantified using Nanodrop (Thermo Scientific) or QuantIt Pico Green (Invitrogen) and pooled in equimolar concentrations for each sample. Samples were made into shotgun sequencing libraries following the manufacturer's instructions (454 Life Sciences [Roche]). Sample pools were grouped into sets, and within each sample set individual libraries were made to contain a different multiplexing identifier (MID) allowing for the combining of the libraries prior to emulsion PCR. Libraries were tagged for multiplexing according to the manufacturer's instructions (454 Life Sciences [Roche]) or Meyer et al. (2008). Sequencing libraries

were quantified by qPCR (Meyer et al. 2007) and pooled at equimolar concentrations. Library pools were divided among regions on GS FLX and Titanium sequencing runs. Sequences for 66 unique killer whale mitogenome haplotypes were deposited in GenBank (accession nos. GU187153–GU187164, GU187166–GU187219), as well as new mitogenome sequences for three outgroup species (accession nos. HM060332–HM060334). All accession numbers used for analysis are listed in Supplemental Table S2.

Microsatellites

A set of 26 microsatellite loci was used to genotype all Antarctic samples and samples collected across the northern North Pacific for population analysis and identification of samples representing duplicate biopsy sampling of the same individual. Forty-four samples were intentionally genotyped in duplicate to estimate error rates, and an additional 15 samples were found to be duplicate samplings of the same individual. One of each pair of unintentional duplicates was removed prior to statistical analysis. Sample sizes (after removal of duplicates) were: Transient, 126; Resident, 245; Offshore, six; Antarctic A, eight; Antarctic B, 15; Antarctic , 42. To test for genotyping errors, we compared replicated genotypes across all loci for replicated samples, and found a per-allele error rate of 0.2%, which is in the low range for published studies (Morin et al. 2010). Population differentiation was calculated using G_{ST} and Hedrick's G'_{ST} to control for the effect of heterozygosity, using the program SMOGD (Crawford 2010). The approximate harmonic mean (H) was calculated from the mean and variance across loci using the equation

$$H \sim 1/[(1/A) + \text{var}(D)(1/A)^3]$$

where A = average divergence across loci and var(D) = variance of divergence across loci (SMOGD website, http://www.ngcrawford.com/django/jost/).

Sequence assembly and phylogenetic analysis

Sequence reads for each sample were sorted by tag sequences, and a single sample was assembled de novo into a single 16,388-bp contig using 454 de novo Assembler software (Roche Applied Science). The consensus sequence and assembly reads were exported as an ACE file and edited with Consed (Gordon et al. 1998), and used as the reference sequence for all subsequent assemblies using the 454 Reference Mapper software (Roche Applied Science). Consensus sequences were aligned in Sequencher (v.4.7, Gene Codes Corporation) or GENEIOUS (Biomatters Ltd), and ambiguities in polynucleotide repeats were individually checked in the 454 Reference Mapper assembly viewer and edited in Sequencher or GENEIOUS. For a region of between nine and 14 Cs in a row (positions 1130–1144 in the original alignment), and another region of seven to eight As in a row (positions 5210–5217), the assembly was unreliable, so the regions were shortened to a fixed set of nine Cs and seven As, respectively, for phylogenetic analysis to avoid introducing potentially erroneous variation. Eight samples were sequenced twice and analyzed for differences between replicates. Sequence alignments of other cetacean sequences and killer whale sequence were performed using Clustal v2.0.4 (Larkin et al. 2007). A figure showing variable sites is shown in Supplemental material S4.

Neighbor-joining trees (MEGA4; Tamura et al. 2007) were constructed initially to select a subset of samples that represented the diversity in the killer whale clades. Bayesian phylogenetic trees and estimates of time since divergence of clades were conducted using BEAST v1.4.8 (Drummond and Rambaut 2007). The HKY

nucleotide substitution model was used, with relaxed clock and a Yule speciation process. We performed two sequential analyses to first estimate divergence times for the genus *Orcinus*, then for types within killer whales. In the first analysis, posterior distributions for divergence times of other cetaceans were used to estimate the divergence time for killer whales, using a set of seven samples representing each of the killer whale clades (Supplemental Table S5). The posterior distribution of divergence time for *Orcinus* from this analysis was then used as a prior distribution with all unique haplotypes to generate the killer whale phylogenetic tree and divergence time estimates for types. The phase I analysis used a burn-in period of 100,000 Markov chain Monte Carlo (MCMC) steps, 100 million total MCMC steps, and samples taken every 1000 steps. Phase II analysis was identical, except that the burn-in period was 80,000 and 80 million MCMC steps were used. Acceptable mixing and convergence to the stationary distribution were checked by visual inspection of posterior samples. Effective sample sizes were 1701 for *Orcinus* in the phase I analysis, 2700–5500 for type clades in the phase II analysis.

References

Arnason U, Gullberg A, Janke A. 2004. Mitogenomic analyses provide new insights into cetacean origin and evolution. *Gene* **333:** 27–34.

Arnason U, Adegoke JA, Gullberg A, Harley EH, Janke A, Kullberg M. 2008. Mitogenomic relationships of placental mammals and molecular estimates of their divergences. *Gene* **421:** 37–51.

Avise JC. 1989. Gene trees and organismal histories—a phylogenetic approach to population biology. *Evolution* **43:** 1192–1208.

Baird RW. 2000. The killer whale—foraging specializations and group hunting. In *Cetacean societies: Field studies of dolphins and whales* (ed. J Mann et al.), pp. 127–153. University of Chicago Press, Chicago, IL.

Baird RW, Abrams PA, Dill LM. 1992. Possible indirect interactions between transient and resident killer whales: Implications for the evolution of foraging specializations in the genus *Orcinus*. *Oecologia* **89:** 125–132.

Barrett-Lennard LG, Ellis GM. 2001. *Population structure and genetic variability in northeastern Pacific killer whales: Towards an assessment of population viability.* Canadian Science Advisory Secretariat, Ottawa, Canada.

Barrett-Lennard LG, Ford JKB, Heise K. 1996. The mixed blessing of echolocation: Differences in sonar use by fish-eating and mammal-eating killer whales. *Anim Behav* **51:** 553–565.

Brown WM, Prager EM, Wang A, Wilson AC. 1982. Mitochondrial DNA sequences of primates: Tempo and mode of evolution. *J Mol Evol* **18:** 225–239.

Carr SM, Marshall HD. 2008. Intraspecific phylogeographic genomics from multiple complete mtDNA genomes in Atlantic cod (*Gadus morhua*): Origins of the "codmother," transatlantic vicariance and midglacial population expansion. *Genetics* **180:** 381–389.

Crandall KA, Bininda-Emonds OR, Mace GM, Wayne RK. 2000. Considering evolutionary processes in conservation biology. *Trends Ecol Evol* **15:** 290–295.

Crawford NG. 2010. SMOGD: Software for the measurement of genetic diversity. *Mol Ecol Resources* **10:** 556–557.

Cummings MP, Otto SP, Wakeley J. 1995. Sampling properties of DNA sequence data in phylogenetic analysis. *Mol Biol Evol* **12:** 814–822.

Dahlheim ME, Schulman-Janiger A, Black N, Ternullo R, Ellifrit D, Balcomb KC III. 2008. Eastern temperate North Pacific offshore killer whales (*Orcinus orca*): Occurrence, movements, and insights into feeding ecology. *Mar Mamm Sci* **24:** 719–729.

De Queiroz K. 2007. Species concepts and species delimitation. *Syst Biol* **56:** 879–886.

Deecke VB, Ford JKB, Slater PJB. 2005. The vocal behaviour of mammal-eating killer whales: Communicating with costly calls. *Anim Behav* **69:** 395–405.

DeFilippis VR, Moore WS. 2000. Resolution of phylogenetic relationships among recently evolved species as a function of amount of DNA sequence: An empirical study based on woodpeckers (Aves: Picidae). *Mol Phylogenet Evol* **16:** 143–160.

Dionne M, Caron F, Dodson JJ, Bernatchez L. 2008. Landscape genetics and hierarchical genetic structure in Atlantic salmon: The interaction of gene flow and local adaptation. *Mol Ecol* **17:** 2382–2396.

Drummond AJ, Rambaut A. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol Biol* **7:** 214. doi: 10.1186/1471-2148-7-214.

Foote AD, Newton J, Piertney SB, Willerslev E, Gilbert MTP. 2009. Ecological, morphological and genetic divergence of sympatric North Atlantic killer whale populations. *Mol Ecol* **18:** 5207–5217.

Ford JKB, Ellis GM. 1999. *Transients: Mammal-hunting killer whales of British Columbia, Washington, and Southeastern Alaska* . UBC Press, Vancouver.

Ford JKB, Ellis GM, Barrett-Lennard LG, Morton AB, Palm RS, Balcomb KC. 1998. Dietary specialization in two sympatric populations of killer whales (*Orcinus orca*) in coastal British Columbia and adjacent waters. *Can J Zool* **76:** 1456–1471.

Forney KA, Wade PR. 2006. Worldwide distribution and abundance of killer whales. In *Whales, whaling and ocean ecosystems* (ed. JA Estes et al.), pp. 145–162. University of California Press, Berkeley.

Galtier N, Jobson RW, Nabholz B, Glemin S, Blier PU. 2009a. Mitochondrial whims: Metabolic rate, longevity and the rate of molecular evolution. *Biol Lett* **5:** 413–416.

Galtier N, Nabholz B, Glemin S, Hurst GD. 2009b. Mitochondrial DNA as a marker of molecular diversity: A reappraisal. *Mol Ecol* **18:** 4541–4550.

Gavrilets S, Losos JB. 2009. Adaptive radiation: Contrasting theory with data. *Science* **323:** 732–737.

Gemmell NJ, Akiyama S. 1996. An efficient method for the extraction of DNA from vertebrate tissues. *Trends Genet* **12:** 338–339.

Gordon D, Abajian C, Green P. 1998. *Consed*: A graphical tool for sequence finishing. *Genome Res* **8:** 195–202.

Grant PR, Grant BR. 2008. *How and why species multiply: The radiation of Darwin's finches*. Princeton University Press, Princeton, NJ.

Haig SM, Beever EA, Chambers SM, Draheim HM, Dugger BD, Dunham S, Elliott-Smith E, Fontaine JB, Kesler DC, Knaus BJ, et al. 2006. Taxonomic considerations in listing subspecies under the US Endangered Species Act. *Conserv Biol* **20:** 1584–1594.

Hedrick PW. 2005. A standardized genetic differentiation measure. *Evolution* **59:** 1633–1638.

Heller R, Siegismund HR. 2009. Relationship between three measures of genetic differentiation G_{ST}, D_{EST} and G'_{ST}: How wrong have we been? *Mol Ecol* **18:** 2080–2083.

Herman DP, Burrows DG, Wade PR, Durban JW, Matkin CO, LeDuc RG, Barrett-Lennard LG, Krahn MM. 2005. Feeding ecology of eastern North Pacific killer whales *Orcinus orca* from fatty acid, stable isotope, and organochlorine analyses of blubber biopsies. *Mar Ecol Prog Ser* **302:** 275–291.

Hickie BE, Ross PS, MacDonald RW, Ford JKB. 2007. Killer whales (*Orcinus orca*) face protracted health risks associated with lifetime exposure to PCBs. *Environ Sci Technol* **41:** 6613–6619.

Ho SY, Gilbert MT. 2010. Ancient mitogenomics. *Mitochondrion* **10:** 1–11.

Hoelzel AR, Natoli A, Dahlheim ME, Olavarria C, Baird RW, Black NA. 2002. Low worldwide genetic diversity in the killer whale (*Orcinus orca*): Implications for demographic history. *Proc R Soc Lond B Biol Sci* **269:** 1467–1473.

Hoelzel AR, Hey J, Dahlheim ME, Nicholson C, Burkanov V, Black N. 2007. Evolution of population structure in a highly social top predator, the killer whale. *Mol Biol Evol* **24:** 1407–1415.

Jefferson TA, Webber MA, Pitman RL. 2007. *Marine mammals of the world*. Academic Press, San Diego, CA.

Krahn MM, Hanson MB, Baird RW, Boyer RH, Burrows DG, Emmons CK, Ford JK, Jones LL, Noren DP, Ross PS, et al. 2007a. Persistent organic pollutants and stable isotopes in biopsy samples (2004/2006) from Southern Resident killer whales. *Mar Pollut Bull* **54:** 1903–1911.

Krahn MM, Herman DP, Matkin CO, Durban JW, Barrett-Lennard L, Burrows DG, Dahlheim ME, Black N, LeDuc RG, Wade PR. 2007b. Use of chemical tracers in assessing the diet and foraging regions of eastern North Pacific killer whales. *Mar Environ Res* **63:** 91–114.

Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, et al. 2007. Clustal W and Clustal X version 2.0. *Bioinformatics* **23:** 2947–2948.

LeDuc RG, Robertson KM, Pitman RL. 2008. Mitochondrial sequence divergence among Antarctic killer whale ecotypes is consistent with multiple species. *Biol Lett* **4:** 426–429.

Losos JB. 2009. *Lizards in an evolutionary tree: Ecology and adaptive radiation of anoles*. University of California Press, Berkeley, CA.

Lyrholm T, Gyllensten U. 1998. Global matrilineal population structure in sperm whales as indicated by mitochondrial DNA sequences. *Proc Biol Sci* **265:** 1679–1684.

Meyer M, Stenzel U, Myles S, Prufer K, Hofreiter M. 2007. Targeted high-throughput sequencing of tagged nucleic acid samples. *Nucleic Acids Res* **35:** e97. doi: 10.1093/nar/gkm566.

Meyer M, Stenzel U, Hofreiter M. 2008. Parallel tagged sequencing on the 454 platform. *Nat Protoc* **3:** 267–278.

Moore WS. 1995. Inferring phylogenies from mtDNA variation—mitochondrial-gene trees versus nuclear-gene trees. *Evolution* **49:** 718–726.

Morin PA, Martien KK, Archer FI, Cipriano F, Steel D, Jackson J, Taylor BL. 2010. Applied conservation genetics and the need for quality control and reporting of genetic data used in fisheries and wildlife management. *J Hered* **10:** 1–10.

Moritz C. 1994. Defining 'evolutionarily significant units' for conservation. *Trends Ecol Evol* **9:** 373–375.

Moritz C. 2002. Strategies to protect biological diversity and the evolutionary processes that sustain it. *Syst Biol* **51:** 238–254.

Nabholz B, Glemin S, Galtier N. 2008a. Strong variations of mitochondrial mutation rate across mammals—the longevity hypothesis. *Mol Biol Evol* **25:** 120–130.

Nabholz B, Mauffrey JF, Bazin E, Galtier N, Glemin S. 2008b. Determination of mitochondrial genetic diversity in mammals. *Genetics* **178:** 351–361.

Nabholz B, Glemin S, Galtier N. 2009. The erratic mitochondrial clock: Variations of mutation rate, not population size, affect mtDNA diversity across birds and mammals. *BMC Evol Biol* **9:** 54. doi: 10.1186/1471-2148-9-54.

Nyakaana S, Arctander P. 1999. Population genetic structure of the African elephant in Uganda based on variation at mitochondrial and nuclear loci: Evidence for male-biased gene flow. *Mol Ecol* **8:** 1105–1115.

Okello JB, Masembe C, Rasmussen HB, Wittemyer G, Omondi P, Kahindi O, Muwanika VB, Arctander P, Douglas-Hamilton I, Nyakaana S, et al. 2008. Population genetic structure of savannah elephants in Kenya: Conservation and management implications. *J Hered* **99:** 443–452.

Parsons K, Balcomb KC, Ford JKB, Durban JW. 2009. The social dynamics of southern resident killer whales and conservation implications for this endangered population. *Anim Behav* **77:** 963–971.

Pilot M, Dahlheim ME, Hoelzel AR. 2010. Social cohesion among kin, gene flow without dispersal and the evolution of population genetic structure in the killer whale (*Orcinus orca*). *J Evol Biol* **23:** 20–31.

Pitman RL, Ensor P. 2003. Three forms of killer whales (*Orcinus orca*) in Antarctic waters. *J Cetacean Res Manag* **5:** 131–139.

Pitman RL, Perryman WL, LeRoi D, Eilers E. 2007. A dwarf form of killer whale in the Antarctic. *J Mammal* **88:** 43–48.

Rice DW. 1998. *Marine mammals of the world. Systematics and distribution*. Special Publication No. 4. The Society for Marine Mammalogy, Lawrence, KS.

Rokas A, Carroll SB. 2005. More genes or more taxa? The relative contribution of gene number and taxon number to phylogenetic accuracy. *Mol Biol Evol* **22:** 1337–1344.

Ross PS, Ellis GM, Ikonomou MG, Barrett-Lennard LG, Addison RF. 2000. High PCB concentrations in free-ranging Pacific killer whales, *Orcinus orca*: Effects of age, sex and dietary preference. *Mar Pollut Bull* **40:** 504–515.

Ruvolo M, Disotell TR, Allard MW, Brown WM, Honeycutt RL. 1991. Resolution of the African hominoid trichotomy by use of a mitochondrial gene sequence. *Proc Natl Acad Sci* **88:** 1570–1574.

Saitou N, Nei M. 1986. The number of nucleotides required to determine the branching order of three species, with special reference to the human-chimpanzee-gorilla divergence. *J Mol Evol* **24:** 189–204.

Sambrook J, Fritsch EF, Maniatis T. 1989. *Molecular cloning: A laboratory manual, 2nd ed*. Cold Spring Harbor Laboratory Press, New York.

Saulitis E, Matkin C, Barrett-Lennard LG, Heise K, Ellis GM. 2000. Foraging strategies of sympatric killer whale (*Orcinus orca*) populations in Prince William Sound, Alaska. *Mar Mamm Sci* **16:** 94–109.

Scandura M, Iacolina L, Crestanello B, Pecchioli E, Di Benedetto MF, Russo V, Davoli R, Apollonio M, Bertorelle G. 2008. Ancient vs. recent processes as factors shaping the genetic variation of the European wild boar: Are the effects of the last glaciation still detectable? *Mol Ecol* **17:** 1745–1762.

Schluter D. 2009. Evidence for ecological speciation and its alternative. *Science* **323:** 737–741.

Tamura K, Dudley J, Nei M, Kumar S. 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol* **24:** 1596–1599.

Torriani SF, Goodwin SB, Kema GH, Pangilinan JL, McDonald BA. 2008. Intraspecific comparison and annotation of two complete mitochondrial genome sequences from the plant pathogenic fungus *Mycosphaerella graminicola*. *Fungal Genet Biol* **45:** 628–637.

Xiong Y, Brandley MC, Xu S, Zhou K, Yang G. 2009. Seven new dolphin mitochondrial genomes and a time-calibrated phylogeny of whales. *BMC Evol Biol* **9:** 20. doi: 10.1186/1471-2148-9-20.

Ylitalo GM, Matkin CO, Buzitis J, Krahn MM, Jones LL, Rowles T, Stein JE. 2001. Influence of life-history parameters on organochlorine concentrations in free-ranging killer whales (*Orcinus orca*) from Prince William Sound, AK. *Sci Total Environ* **281:** 183–203.

Zarowiecki MZ, Huyse T, Littlewood DT. 2007. Making the most of mitochondrial genomes—markers for phylogeny, molecular ecology and barcodes in (Platyhelminthes: Digenea). *Int J Parasitol* **37:** 1401–1418.

Complete mitochondrial genome phylogeographic analysis of killer whales (*Orcinus orca*) indicates multiple species

Phillip A. Morin, Frederick I. Archer, Andrew D. Foote, et al.

Supplemental Material	http://genome.cshlp.org/content/suppl/2010/04/08/gr.102954.109.DC1.html
References	This article cites 63 articles, 19 of which can be accessed free at: http://genome.cshlp.org/content/20/7/908.full.html#ref-list-1
Creative Commons License	This article is distributed exclusively by Cold Spring Harbor Laboratory Press for the first six months after the full-issue publication date (see http://genome.cshlp.org/site/misc/terms.xhtml). After six months, it is available under a Creative Commons License (Attribution-NonCommercial 3.0 Unported License), as described at http://creativecommons.org/licenses/by-nc/3.0/.
Email Alerting Service	Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or **click here.**