University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

Publications, Agencies and Staff of the U.S. Department of Commerce

U.S. Department of Commerce

2005

Population histories of right whales (Cetacea: *Eubalaena*) Inferred from Mitochondrial Sequence Diversities and Divergences of Their Whale Lice (Amphipoda: *Cyamus*)

Zofia A. Kaliszewska
Harvard University

Jon Seger
University of Utah

Victoria J. Rowntree
Ocean Alliance/Whale Conservation Institute, Lincoln, Massachusetts & Instituto de Conservación de Ballenas, Miñones 1986, Buenos Aires , Argentina

Amy R. Knowlton
Ocean Alliance/Whale Conservation Institute, Lincoln, Massachusetts & Instituto de Conservación de Ballenas, Miñones 1986, Buenos Aires, Argentina

Kim Marshalltilas
Ocean Alliance/Whale Conservation Institute, Lincoln, Massachusetts, & Instituto de Conservación de Ballenas, Buenos Aires, Argentina

Follow this and additional works at: https://digitalcommons.unl.edu/usdeptcommercepub

Part of the Environmental Sciences Commons

See next page for additional authors

Kaliszewska, Zofia A.; Seger, Jon; Rowntree, Victoria J.; Knowlton, Amy R.; Marshalltilas, Kim; Patenaude, Nathalie J.; Rivarola, Mariana; Schaeff, Catherine M.; Sironi, Mariano; Smith, Wendy A.; Yamada, Tadasu K.; Barco, Susan G.; Benegas, Rafael; Best, Peter B.; Brown, Moira W.; Brownell, Robert L. Jr.; Harcourt, Robert; and Carribero, Alejandro, "Population histories of right whales (Cetacea: *Eubalaena*) Inferred from Mitochondrial Sequence Diversities and Divergences of Their Whale Lice (Amphipoda: *Cyamus*)" (2005). *Publications, Agencies and Staff of the U.S. Department of Commerce*. 88.
https://digitalcommons.unl.edu/usdeptcommercepub/88

Authors

Zofia A. Kaliszewska, Jon Seger, Victoria J. Rowntree, Amy R. Knowlton, Kim Marshalltilas, Nathalie J. Patenaude, Mariana Rivarola, Catherine M. Schaeff, Mariano Sironi, Wendy A. Smith, Tadasu K. Yamada, Susan G. Barco, Rafael Benegas, Peter B. Best, Moira W. Brown, Robert L. Brownell Jr., Robert Harcourt, and Alejandro Carribero

Population histories of right whales (Cetacea: *Eubalaena*) inferred from mitochondrial sequence diversities and divergences of their whale lice (Amphipoda: *Cyamus*)

ZOFIA A. KALISZEWSKA,* JON SEGER,† VICTORIA J. ROWNTREE,†‡§ SUSAN G. BARCO,¶ RAFAEL BENEGAS,** PETER B. BEST,†† MOIRA W. BROWN,‡‡ ROBERT L. BROWNELL JR,§§ ALEJANDRO CARRIBERO,¶¶ ROBERT HARCOURT,*** AMY R. KNOWLTON,‡‡ KIM MARSHALL-TILAS,‡ NATHALIE J. PATENAUDE,*** MARIANA RIVAROLA,††† CATHERINE M. SCHAEFF,‡‡‡ MARIANO SIRONI,§ WENDY A. SMITH† and TADASU K. YAMADA§§§

**Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, Massachusetts 02138, USA,* †*Department of Biology, University of Utah, 257 South 1400 East, Salt Lake City, Utah 84112, USA,* ‡*Ocean Alliance/Whale Conservation Institute, 191 Weston Rd, Lincoln, Massachusetts 01773, USA,* §*Instituto de Conservación de Ballenas, Miñones 1986, Buenos Aires C1428ATD, Argentina,* ¶*Virginia Marine Science Museum, 717 General Booth Boulevard, Virgina Beach, Virginia 23451, USA,* ***Hydro Sport, Av. de las Ballenas S/N, Puerto Pirámides, Chubut, Argentina,* ††*Mammal Research Institute, University of Pretoria, Pretoria 0002, South Africa,* ‡‡*New England Aquarium, Central Wharf, Boston, Massachusetts 02110, USA,* §§*Southwest Fisheries Science Center, National Marine Fisheries Service, Pacific Grove, California 93950, USA,* ¶¶*Ecocentro, Julio Verne 3784, U9129OJA, Puerto Madryn, Chubut, Argentina,* ****Marine Mammal Research Group, Graduate School of Environment, Macquarie University, Sydney, NSW 2109, Australia,* †††*ALUAR, Ruta Provincial N° 4, U9120OIA Puerto Madryn, Chubut, Argentina,* ‡‡‡*Department of Biology, American University, 4400 Massachusetts Avenue, N.W., Washington, DC 20016, USA,* §§§*National Science Museum, 3-23-1 Hyakunin-cho, Shinjuku-ku, Tokyo 169-0073, Japan*

Abstract

Right whales carry large populations of three 'whale lice' (*Cyamus ovalis*, *Cyamus gracilis*, *Cyamus erraticus*) that have no other hosts. We used sequence variation in the mitochondrial *COI* gene to ask (i) whether cyamid population structures might reveal associations among right whale individuals and subpopulations, (ii) whether the divergences of the three nominally conspecific cyamid species on North Atlantic, North Pacific, and southern right whales (*Eubalaena glacialis*, *Eubalaena japonica*, *Eubalaena australis*) might indicate their times of separation, and (iii) whether the shapes of cyamid gene trees might contain information about changes in the population sizes of right whales. We found high levels of nucleotide diversity but almost no population structure within oceans, indicating large effective population sizes and high rates of transfer between whales and subpopulations. North Atlantic and Southern Ocean populations of all three species are reciprocally monophyletic, and North Pacific *C. erraticus* is well separated from North Atlantic and southern *C. erraticus*. Mitochondrial clock calibrations suggest that these divergences occurred around 6 million years ago (Ma), and that the *Eubalaena* mitochondrial clock is very slow. North Pacific *C. ovalis* forms a clade inside the southern *C. ovalis* gene tree, implying that at least one right whale has crossed the equator in the Pacific Ocean within the last 1–2 million years (Myr). Low-frequency polymorphisms are more common than expected under neutrality for populations of constant size, but there is no obvious signal of rapid, interspecifically congruent expansion of the kind that would be expected if North Atlantic or southern right whales had experienced a prolonged population bottleneck within the last 0.5 Myr.

Keywords: *Cyamus*, cytochrome *c* oxidase I (*COI*), *Eubalaena*, migration, mitochondrial introgression, population structure, speciation

Received 1 January 2005; revision accepted 2 June 2005

Introduction

Most cetaceans carry populations of benign ectoparasites called cyamids or 'whale lice', which are amphipod crustaceans (Caprellidea; Cyamidae). Cyamids spend all life stages on whales, feeding on the outer surface of their host's epidermis (Rowntree 1983, 1996). They roam widely on their hosts, but because they have no free-swimming stage they can migrate only between whales that contact each other. The ecological universe of a cyamid population is therefore well defined and sharply bounded, consisting of many virtually identical and constantly moving habitat islands (whales) that occasionally replace themselves through a simple birth-and-death process.

As a consequence of this symbiotic relationship, the sizes, structures and histories of cyamid populations are coupled directly to those of their hosts. Cyamid populations can be vastly larger than whale populations, however, so they are expected to maintain higher levels of genetic variation. They also have an additional level of structure, because each individual whale supports a self-sustaining sub-population of cyamids. A calf's initial infestations arrive by vertical transmission from its mother. If subsequent horizontal transfer between unrelated whales were to occur infrequently, then the genetic relationships of the cyamid subpopulations on conspecific whales might reflect patterns of behavioural interaction among those whales. Here we ask whether information about the local interactions, large-scale movements and population histories of right whales (*Eubalaena* spp.) can be extracted from the distribution of mitochondrial DNA sequence variation in cyamids from right whales in the western North Atlantic, the western North Pacific, and three locations in the Southern Hemisphere.

Cyamid diversity and ecology

Around 35 species of cyamids in seven genera have been described to date, but systematic revisions using new morphological and molecular data are expected to increase this number substantially (Todd Haney, personal communication). Some cyamid species occur on more than one cetacean species, but few cetaceans normally host more than one species of cyamid (Leung 1967; Rowntree 1983, 1996).

The slow-swimming right whales and grey whale (*Eschrichtius robustus*) are unusual in that each carries large populations of three cyamid species that do not occur regularly on any other cetaceans. The grey whale cyamids (*Cyamus ceti*, *Cyamus kessleri*, *Cyamus scammoni*) are only distantly related to each other and to the right whale cyamids, and vice versa (Z.A.K., W.A.S., J.S. and T. Haney, unpublished genetic and morphological data). The right whale cyamids are traditionally referred to as *Cyamus ovalis*, *Cyamus gracilis* and *Cyamus erraticus*. However, genetic

evidence described below strongly implies that the North Atlantic and Southern Hemisphere populations of each of these three 'species' have been separated long enough to be considered distinct biological and phylogenetic species, and that at least two (and probably all three) of the cyamids on North Pacific right whales are also genetically isolated from their nominal conspecifics in the other two oceans. A systematic revision of *Cyamus* is in preparation (T. Haney, personal communication). Here we combine the traditional species names of right whale cyamids with the ocean-system names to distinguish the (apparently) nine geographical sibling species; for example, we recognize 'North Atlantic *C. ovalis*', 'North Pacific *C. ovalis*', and 'Southern Ocean *C. ovalis*'.

The three traditional (nominal) species occupy different parts of the whale that correspond to ecological niches (Roussel de Vauzème 1834) (Fig. 1). *Cyamus ovalis* blankets the callosities (raised patches of roughened skin on the head); their white bodies make the callosities stand out brightly against the black skin of the whale's head and back (Fig. 2). Each right whale has an individually distinctive callosity pattern that can be identified in clear photographs (Payne *et al.* 1983; Kraus *et al.* 1986; Best 1990) owing to the visual contrast and sharp delineation of callosity boundaries created by a living blanket of *C. ovalis*. *C. gracilis* occupies pits and grooves between the elevated patches of callosity tissue; adults of this species are smaller and thinner than adult *C. ovalis*. *C. erraticus* occupies smooth skin in the genital and mammary slits; it is highly mobile and opportunistically colonizes wounds and other areas of reduced water flow not occupied by the other two species; it also 'blooms' on the heads of young calves, but these concentrations disappear within a few months. *C. ovalis* and *C. gracilis* appear to be better than *C. erraticus* at clinging to
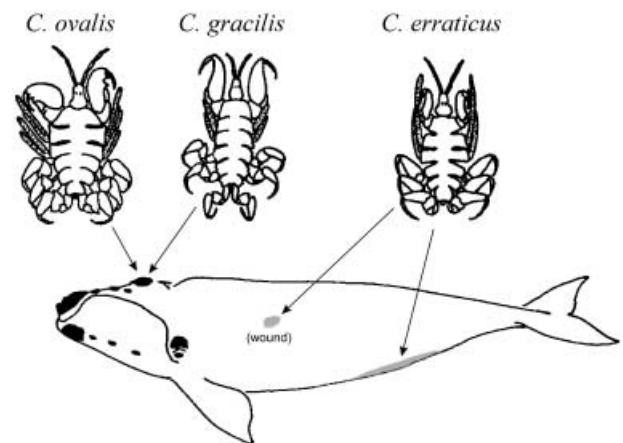

Fig. 1 Schematic drawing of a right whale and its cyamids, showing where the cyamid species commonly occur. Adult right whales are 12–16 m in length. Adult right whale cyamids are 5–12 mm (female *Cyamus gracilis* at the lower end of this range, male *Cyamus erraticus* at the higher end).
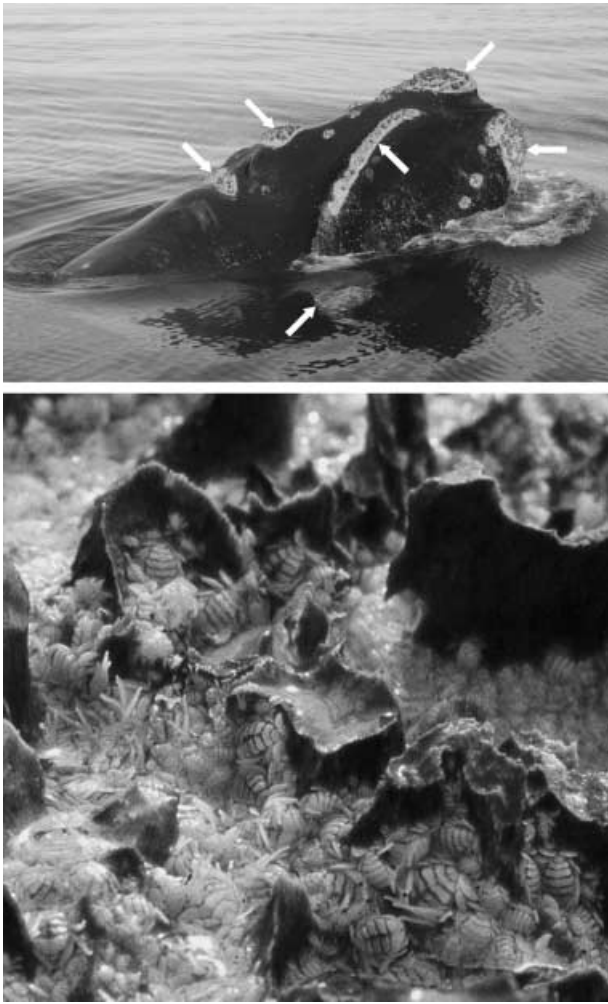
Fig. 2 Right whale callosities whitened by dense populations of cyamids. The upper panel shows the head of a southern right whale (*Eubalaena australis*) at Península Valdés, Argentina. Arrows point to callosities behind and in front of the blowholes, above the eye (under water), along the top edge of the lower lip, at the tip of the rostrum ('bonnet'), and on the right side of the chin. Smaller round callosities also occur along the lower jaw between the chin callosity and the eyes, and along the rostrum between the bonnet and the anterior blowhole callosity ('coaming'). The large bonnet, chin and eyebrow callosities are typically around 50 cm in diameter. (Photo by Mariano Sironi.) The lower panel shows *Cyamus ovalis* (and a few *Cyamus gracilis*) on a small portion of the bonnet of a different individual. The field of view is roughly twice the width of the white arrows in the upper panel. (Photo by Iain Kerr.)

mature callosity tissue, which is dense and hard; thus *C. erraticus* may be forced to retreat to softer, more protected substrates as the calf's callosities develop (Rowntree 1996).

Every adult right whale carries populations of all three nominal cyamid species. On a typical adult, the callosities provide about 0.5 m^2 of substrate for *C. ovalis* which is by far the most abundant of the three species with about 5000 adults per adult right whale, based on an estimated density of roughly one adult cyamid per cm^2 (from photographic and defined-area census data, not shown; see Fig. 2). The genital and mammary slits provide a smaller substrate for *C. erraticus*, and the pits and grooves between callosities appear to provide an even more limited substrate for *C. gracilis*. These two species are less abundant than *C. ovalis*, but their average population sizes cannot yet be estimated as accurately. Based on their relative abundances in collections (which derive mainly from opportunistic sampling on stranded whales), *C. erraticus* and *C. gracilis* would appear to have typical numbers on the order of 2000 and 500 per whale, respectively, although *C. gracilis* is probably underrepresented in collections owing to its small size, its relatively concealed location in recesses, and the fact that few collectors realize they should look for two species of cyamids on the whale's head.

Sizes and structures of right whale populations

Right whales occur in all the major oceans of the world. North Atlantic right whales (*Eubalaena glacialis*) have been hunted for about 1000 years and have long been effectively extinct in the eastern North Atlantic; the remnant western North Atlantic population consists of roughly 300 individuals and is growing slowly, if at all (Knowlton *et al.* 1994; Caswell *et al.* 1999; Kraus *et al.* 2001). Southern right whales (*Eubalaena australis*) were exploited more recently and are recovering rapidly; in the late 1990s their population was estimated to number around 7500 and to be growing at a rate of roughly 7% per year (International Whaling Commission 2001). Pre-exploitation populations of *E. glacialis* and *E. australis* are thought to have been around 40 000–150 000 (International Whaling Commission 2001; Baker & Clapham 2004). Right whales also occur in the North Pacific, and phylogeographic analyses of mitochondrial control region and nuclear gene sequences imply that this population represents a third distinct species, *Eubalaena japonica* (Rosenbaum *et al.* 2000; Gaines *et al.* 2005); the remnant population is small and poorly known (Brownell *et al.* 2001). If the three right whale species have had long-term average population sizes on the order of 4×10^4, then the effective population sizes of right whale cyamids would appear to be on the order of 2×10^7 for *C. gracilis* ($500 \times 4 \times 10^4$) to 2×10^8 for *C. ovalis* ($5000 \times 4 \times 10^4$).

Southern right whales gather in coastal aggregations to calve and nurse their young during the austral winter and spring (June–December). The largest aggregations occur in three areas centred on the coastal waters of Argentina, South Africa and Australia. In spring the whales swim south (and perhaps in some cases north) to feeding grounds that remain poorly defined (International Whaling Commission 2001). The wintering aggregations have long been thought to represent subpopulations or 'stocks' between

which relatively little migration occurs, and recent analyses of mitochondrial control region sequences support this view. The aggregations show substantial differentiation, with F_{ST} values ranging from 0.15 to 0.5 (Portway 1998; Patenaude 2002; Patenaude *et al*. in review). On the assumption that these levels of differentiation represent equilibria between migration and drift, the long-term average number of female migrants per generation can be estimated as $N_f m = (1 - F_{ST})/2F_{ST}$ (Wright 1951; Takahata & Palumbi 1985). For the three major aggregations, these estimates range from 1.8 to 2.8 females per right whale generation. Southern right whale females typically begin breeding at 9 years of age (International Whaling Commission 2001). Average and maximum lifespans have not yet been determined, but a female longevity greater than 60 years has been documented, as have several reproductive spans approaching 30 years (International Whaling Commission 2001), so the average generation time seems likely to exceed 20 (and possibly 30) years, implying average migration rates of one or fewer females per decade between the major Southern Ocean aggregations.

Right whale cyamids as 'replicated evolutionary experiments'

Because each right whale species supports three species of cyamids, the cyamids' population histories can be viewed as replicated ecological and evolutionary experiments of two kinds. The first and more obvious kind of replication derives from the speciation of right whales. The North Atlantic, North Pacific, and Southern Ocean populations of each nominal cyamid species form sibling-species trios that are still very similar in all obvious respects; this pattern demonstrates that the three nominal cyamid species had established their current ways of life before their hosts stopped crossing the equator several million years ago. The three right whale species also have changed very little since separation; no nonmolecular character differences have been found that reliably distinguish them, and some systematists have argued, even recently, that all right whales should be classified as members of the same species (Rice 1998). Thus the environments of the North Atlantic, North Pacific and Southern Ocean populations representing each of the three nominal cyamid species seem likely to be very similar, and their independent histories since the time when right whales stopped crossing the equator can be interpreted as replicates of substantially the same scenario, beginning from the same initial genetic conditions.

The second kind of replication derives from the more ancient separation of the three lineages that define the nominal cyamid species. Within each ocean, all three nominal species occur on the same whales, in predictable but very different numbers. They should be affected similarly by any environmental changes in their ocean, including changes in the size or structure of the right whale population on which they live. Their responses might differ quantitatively, however, owing to their different population sizes.

We used sequence variation in the mitochondrial gene encoding cytochrome *c* oxidase subunit 1 (*COI, cox1*) to ask three questions. First, do the nominal cyamid species show patterns of differentiation among subpopulations that appear to reflect patterns of association among right whale individuals and subpopulations within oceans? Second, do patterns of molecular divergence among Northern and Southern Hemisphere cyamids of the same nominal species support the view that North Atlantic, North Pacific and Southern Ocean right whales are reproductively isolated, and if so, do levels of divergence between Northern and Southern Hemisphere cyamid sibling species suggest plausible times of separation for the three right whale species? Third, do patterns of coalescence in cyamid gene trees suggest histories of cyamid population-size change that imply corresponding changes in the populations of their hosts?

Materials and methods

DNA extraction, PCR, and sequencing

Cyamids were collected from stranded (beach-cast) whales and stored in 70–95% ethanol. Collection information is summarized in Table 1. Individual cyamids were dried at 50 °C, frozen on dry ice, and ground in 1.5 mL centrifuge tubes using disposable pestles. Following digestion in proteinase K and SDS at 50 °C for 12–24 h with continual mixing, and extraction with chloroform, DNA was adsorbed to diatomaceous silica (Carter & Milton 1993), washed, and eluted into 10 mM Tris buffer. Alternatively, some digests were extracted with phenol, phenol–chloroform and chloroform, followed by ethanol precipitation of nucleic acids (Sambrook *et al*. 1989).

Partial sequences of *COI*, *COII* and the intervening leucine tRNA gene were determined for an individual of North Atlantic *Cyamus ovalis* from PCR products amplified with the 'universal' arthropod primers Jerry (C1-J-2183), Pat (L2-N-3014), Marilyn (C2-N-3389) and Barbara (C2-N-3661) (reviewed by Simon *et al*. 1994). Partial *ND2* sequences, the 5′ end of *COI*, and the intervening tryptophan and glycine tRNAs were determined for an individual of North Atlantic *Cyamus gracilis* from PCR products amplified with a forward primer designed by us from an alignment of publicly available crustacean *ND2* sequences (Ind2cy: 5′-TAGGGGGTTTACCTCCTTTTACGGG-3′; subsequent primer sequences are also written 5′ to 3′) and a reverse primer designed from our initial partial *COI* sequences for several cyamid species (C1L1: CTGTCCCTACTCCTCTT-TCTAC). From the resulting sequences flanking *COI* we designed a forward primer in the glycine tRNA gene
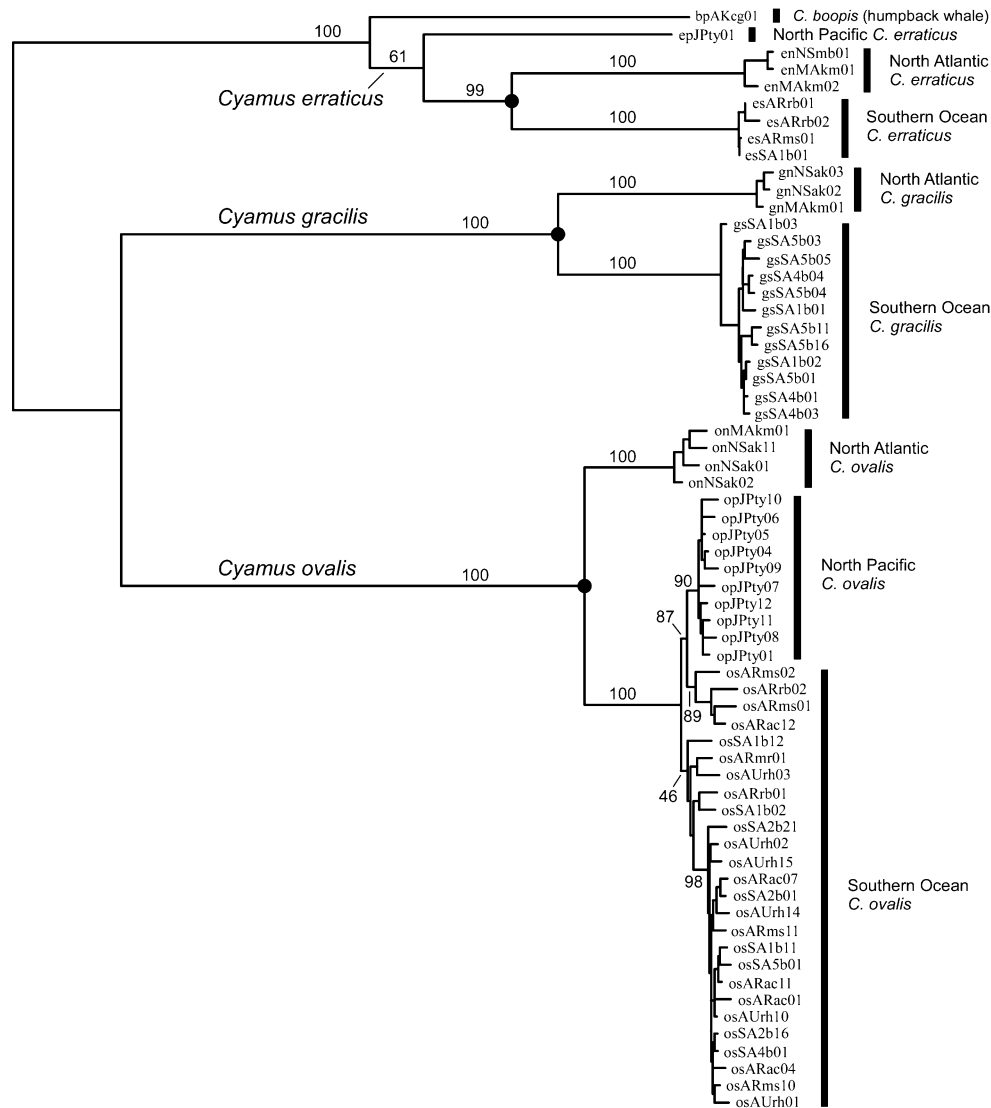
Table 1 Collection information

Stranding location	Lat., Long.	Date	Collector	Whale
Nova Scotia (Freeport)	44.3°N, 66.4°W	20 August 1997	A. Knowlton	ak
Massachusetts (Wellfleet)	41.9°N, 70.0°W	21 June 1999	K. Marshall-Tilas	km
Bay of Fundy	44.5°N, 66.5°W	13 September 1999	M. Brown	mb
Maryland (Ocean City)	38.4°N, 74.6°W	22 August 2002	T. Litwiler	tl
Virginia (Virginia Beach)	36.8°N, 75.8°W	11 February 2004	S. Barco	sb
Argentina (Península Valdés)	42.5°S, 64.3°W	5 May 1995	M. Rivarola	mr
Argentina (Pta. Tehuelche, GSJ)	42.4°S, 64.3°W	17 November 1995	R. Benegas	rb
Argentina (Pto. Pirámides, G. Nuevo)	42.5°S, 64.3°W	3 November 1999	M. Sironi	ms
Argentina (Fracaso, Golfo San Jose)	42.4°S, 64.1°W	15 October 2000	A. Carribero	ac
South Africa (Cape Province)	34.5°S, 19.3°E	16 September 1999	P. Best	1b
South Africa (Dwarskersbos)	32.6°S, 18.3°E	18 September 2000	P. Best	2b
South Africa (Elands Bay)	32.3°S, 18.3°E	29 July 2000	P. Best	3b
South Africa (Cape Pt. Nat. Res.)	34.3°S, 18.4°E	13 October 2000	P. Best	4b
South Africa (Soetwater, Cape Pen.)	34.2°S, 18.3°E	28 June 2001	P. Best	5b
Australia (Cape Jervis)	35.3°S, 137.3°E	21 July 2001	R. Harcourt	rh
Japan (Hitachi City)	36.6°N, 140.8°E	3 April 2004	S. Nishida	ty

Identities of whales: ak, Eg#2450; km, Eg#1014 'Staccato'; mb, Eg#2030; sb, Eg#1004 'Stumpy'; 1b, 99/05; 2b, 00/12; 3b, 00/10; 4b, 00/14; 5b, 01/04.

immediately 5′ to *COI* (tGcy: AGGCTTGAAAAAAGT-TAATCTTAGG) and a reverse primer, derived from Pat, in the leucine tRNA gene immediately 3′ to *COI* (Patcy: ACTAGCACATTTATCTGTCACATTA). These primers have worked well in all cyamids tested to date.

The full-length (1.5 kb) *COI* sequences described here were determined from PCR products amplified with tGcy and Patcy, and the half-length (0.8 kb) sequences were generated with a cyamid-optimized version of Jerry (Jercy: TACCAACATTTATTCTGRTTTTTYGG) and Patcy. Typical 20-μL polymerase chain reactions contained 0.5 unit of Jumpstart *Taq* DNA polymerase (Sigma), 2 μL 10× buffer, 2 mM $MgCl_2$, 0.03% BSA, 200 μM each dNTP, 200 nM each primer, and approximately 1 ng genomic DNA. Typical reaction conditions were 2 min at 93 °C followed by 36 cycles of 20 s at 94 °C, 30 s at 50 °C, and 60 s at 70 °C. Blank (no template) controls were included in every experiment.

Following cleanup with exonuclease I and shrimp alkaline phosphatase, PCR products were directly sequenced on ABI instruments using dye terminator chemistry at the University of Utah Health Sciences DNA Sequencing Core Facility. PCR products were sequenced on both strands, and outward-pointing reverse (CIA: AAAATAAAGGG-TACCGATATCTTTRTG; and K5cy: CCTACTGTAAATA-TATGGTGBGCTCA) and forward (C1RF2: CAYGACAC-TTACTATGRAGTYGCTCA) primers were used as needed to obtain unambiguous reads at positions near the ends of PCR products. Sequences were assembled and edited using SEQUENCHER 3.0–4.1 (Gene Codes Corporation) and aligned with CLUSTAL_X 1.81 (Thompson *et al.* 1997) to produce in-frame 1536- or 819-bp alignments. After all sequencing was completed, the nonsynonymous and singleton polymorphisms were carefully checked for accuracy.

Complete mitochondrial genome sequences for a southern *C. ovalis* individual and a *Cyamus* nr. *ceti* individual from a bowhead whale (*Balaena mysticetus*) were subsequently determined by primer-walking on long PCR products spanning the intervals between *COI* and *16S* (*lrRNA*) in both directions. These sequences will be described else-where. Primers developed in the course of this project were used to extend one or a few of the *COI* sequences from each right whale cyamid species through the *COII*, *ATP8*, *ATP6*, *COIII*, and intervening tRNA genes, and 306 bp into the 5′ end of *ND3*. These multigene arcs include 4.1 kb of coding sequence, which was used to estimate divergences of the Northern and Southern Hemisphere members of each sibling species pair; the concatenated 6-gene coding sequences were also included in the alignment used to infer the species tree shown in Fig. 3. The cyamid mitochondrial sequences have been deposited in GenBank under Acces-sion numbers DQ094899–DQ095150.

Full-length *COI* and cyt *b* sequences were determined for two southern right whales (Ea123 and Ea604) and for one northern right whale (Eg#1014, 'Staccato'), using primers designed from publicly available mitochondrial genome sequences for whales and other cetartiodactyls. *COI* was amplified with L5749 (CTCAACCTCTGTCTT-TAGATTTAC) and H7357 (GGTTATGATATTGGCTTGA-AACC), which sit in flanking tRNA genes; PCR products were sequenced in addition with internal primers CO1X (GGCGTATCCTCTATCCTYGGRGC), CO1Y (TGGTATT-GGGTTATGGCYGGRGG), and CO1Z (TGATCTCCT-GCCCTAATRTGRGC). Cyt *b* was amplified with Pal1m

Fig. 3 Multispecies mitochondrial gene tree for right whale cyamids. Nine species are represented (all right whale cyamids except North Pacific *Cyamus gracilis*, plus *Cyamus boopis* which is closely related to *Cyamus erraticus* but regularly occurs only on the humpback whale, *Megaptera novaeangliae*). Sequences are full-length *COI* or 4.1-kb-multigene arcs, except that North Pacific *Cyamus ovalis* is represented by a mixture of full and half-length *COI* sequences; each species other than North Pacific *C. ovalis* is represented by at least one multigene arc. Sequence names refer to individual cyamids and are constructed to denote (in order) the nominal species, the ocean, the collecting locality, the whale (see Table 1), and the specimen's ID number in the series on that whale. The tree is a maximum-likelihood estimate inferred by PHYML from all coding sites under a GTR + G model with six rate categories. Bootstrap values (out of 100 replicates) are shown next to critical internal branches. Large filled dots mark the speciation events that appear to have been caused by the separation of North Atlantic and Southern Ocean right whales; these nodes correspond to synonymous divergences on the order of 0.5 (Table 6) and to a time around 6 Ma (as discussed in the text), implying that the two deepest nodes representing the separations of the three nominal species correspond to synonymous divergences much greater than 1.0 and times greater than 20 Ma. The surprisingly deep divergence of North Pacific *C. erraticus* might be explained by a period of occasional genetic exchange with the population ancestral to modern *C. boopis*, early in the history of their separation into species highly specialized for life on right whales (ancestral *C. erraticus*) and on humpback whales (ancestral *C. boopis*). Subsequently, the *C. erraticus* populations in all three ocean systems appear to have become fully isolated from each other and from *C. boopis*.

(TGACATGAAACATCATCGTTG) and Thr1 (GGTATTT-TATTATACTACAAAGAC), which also sit in flanking tRNA genes; both are modified versions of primers described by LeDuc *et al.* (1999). Cyt *b* PCR products were sequenced in addition with CB1 (CB-J-10933) and CB2-H (CB-N-10920), as described by Simon *et al.* (1994). The right whale *COI* and cyt *b* sequences have been deposited in GenBank under Accession numbers DQ095151–DQ095156.

Polymorphism within and differentiation among populations

Standard polymorphism statistics including θ_S (the neutral mutation parameter $2N_f\mu$ estimated from numbers of segregating sites) and π (the average nucleotide heterozygosity, which is also an estimate of the neutral parameter) were calculated by DNASP 4.00.5 (Rozas *et al.* 2003), ARLEQUIN 2.0 (Schneider *et al.* 2000), and POPALN (written by J. S. and available upon request). Upper and lower confidence limits for θ_S were estimated by the method of Kreitman & Hudson (1991), which involves directly calculating the expected distribution of the number of segregating sites (S) for a given sample size and θ under the infinite-sites standard neutral model (Watterson 1975; Hudson 1990). Synonymous and nonsynonymous substitutions between species were estimated by the maximum-likelihood algorithm implemented in CODEML (Yang 1997, 2002) under the F3 × 4 substitution model and the arthropod mitochondrial code. Inclusive amino acid translations (profiles) were generated for all population samples by TRANALN (J. S.). Hierarchical *F*-statistics for intraspecific population subdivisions were estimated by analysis of molecular variance (AMOVA) (Weir & Cockerham 1984; Excoffier *et al.* 1992; Weir 1996) as implemented in ARLEQUIN 2.0, using pairwise synonymous divergences estimated by CODEML under the F3 × 4 model.

Gene tree estimates

The relationships of haplotypes were estimated by maximum likelihood under a variety of substitution models using DNAML and DNAMLK from PHYLIP *3.6* (Felsenstein 2002), PAUP* 4.0b10 (Swofford 1998), and PHYML 2.3 (Guindon & Gascuel 2003). Duplicate haplotypes were removed from the alignments used to estimate gene trees. The multispecies tree shown here (Fig. 3) is a bootstrapped maximum-likelihood estimate under the GTR + G model, using full-length *COI* and longer multigene sequences. The single-species trees (Fig. 4) are maximum-likelihood estimates
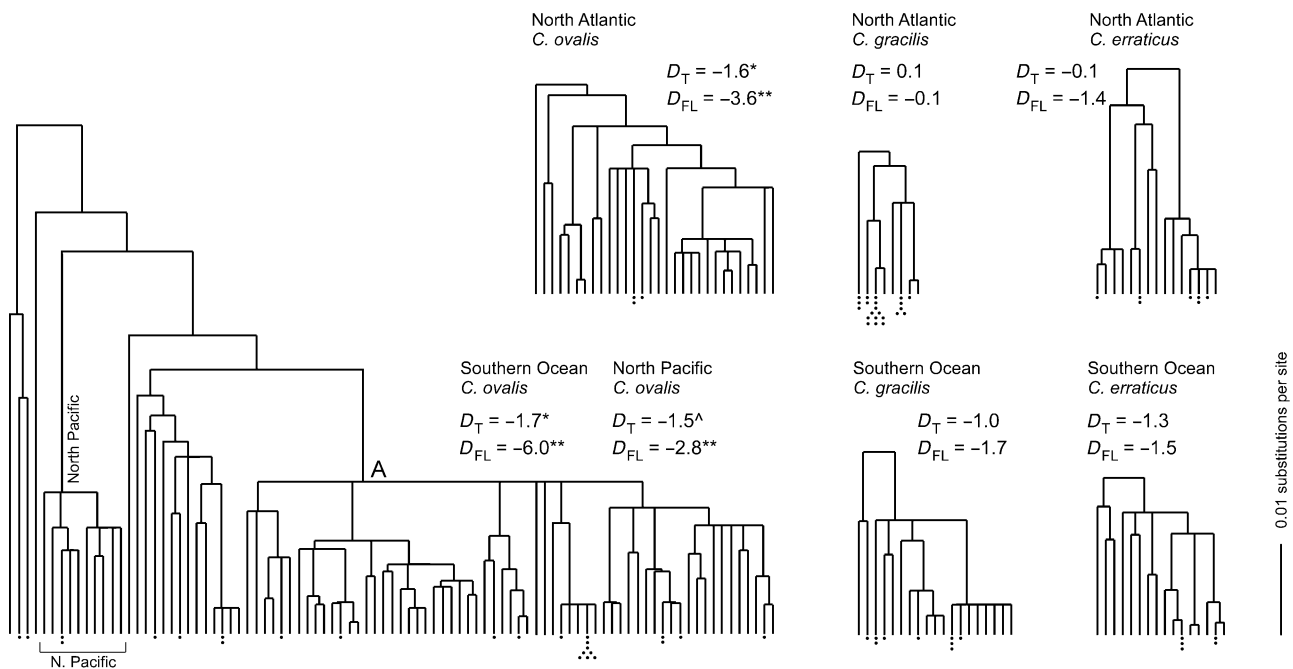
Fig. 4 Single-species *COI* gene trees for right whale cyamids. Seven species are represented, with North Pacific *Cyamus ovalis* inside southern *C. ovalis*. The trees are ML estimates inferred by PAUP* from the synonymous data partition, under the molecular clock and various substitution models as recommended by MODELTEST; the recommended model is fully parameterized (GTR + G) for the large southern *C. ovalis* data set, but less so for the others, roughly in proportion to sample sizes. The clade marked 'A' on the southern *C. ovalis* tree appears to have expanded rapidly around 400 000 years ago, as discussed in the text. The tree-shape statistics and statistical significance symbols are taken from Table 7. In all of these trees the deeper branches tend to be proportionally shorter than expected for neutral coalescents at equilibrium in populations of constant size; for example, the last coalescent interval (where two lines of descent lead to the root) is expected to account for roughly half of the tree's total depth, on average, and the last four intervals together (five and fewer lines of descent) are expected to account for roughly 80% of the total depth (Hudson 1990). The positive D_T for North Atlantic *Cyamus gracilis* reflects the fact that only eight distinct haplotypes occur in the sample of 30 sequences (Table 3), as discussed in the text. Dots below the tips of terminal branches represent duplicate haplotypes (Tables 2–4 give sample sizes, haplotype numbers, and other statistics).

under the molecular clock assumption and various substitution models as recommended by MODELTEST 3.5 (Posada & Crandall 1998), using half-length (819 bp) *COI* sequences.

Especially within species, almost all substitutions occur in third codon positions and in the first positions of leucine codons. The nucleotide compositions of these sites differ from those of second and typical first positions, so we estimated the single-species trees from alignments of synonymous positions, augmented by the few first and / or second position sites with nonsynonymous substitutions, so as to use all available phylogenetic information. The equilibrium nucleotide frequencies and substitution rate parameters estimated from synonymous sites more accurately model the evolution of those sites than do parameters estimated from all sites. This procedure does not select only those sites that are observed to vary (which would seriously bias the resulting parameter estimates); instead, it selects a functionally defined data partition within which almost all of the variation occurs, although most of those sites are fixed even in the largest samples. Adding other variable sites violates this principle only slightly, because there are few such sites. An additional advantage of using the synonymous data partition is that relatively simple models of nucleotide substitution (e.g. ones without invariant sites) work well and are preferred over more parameter-rich models by likelihood-ratio criteria such as the one implemented in MODELTEST.

Tests of equilibrium under neutrality and constant population size

Pairwise difference and site-frequency distributions were calculated by POPALN. Tajima's (1989a) *D* statistic and Fu's (1997) F_S statistic were calculated by ARLEQUIN 2.0 and their significance was estimated by coalescent simulations. Fu & Li's (1993) *D* statistic was calculated from numbers of substitutions inferred by MACCLADE 4.0 (Maddison & Maddison 2000) under the program's default assumptions, on maximum-likelihood gene trees estimated by PAUP* (see Fig. 4). MACCLADE's 'Tree Length' statistic was used as the total number of substitutions on a given tree (η); the number of substitutions on external branches (η_e) was then counted by hand from graphic displays resulting from application of the 'trace all (unambiguous) changes' command. Branches leading to duplicate haplotypes were not counted as external, even though they appear as such on the trees.

Results and discussion

Consistently high levels of synonymous polymorphism

Most haplotypes are unique (even in southern *Cyamus ovalis* with a sample size of over 100), and levels of nucleotide polymorphism are also high (Tables 2–4). Almost all polymorphisms are synonymous, and there are few amino acid

Table 2 Polymorphism summary for *Cyamus ovalis*

Sample (subsample)	*n*	*h*	*S*	θ_S	(SD)	Π	π	(SD)
North Atlantic	33	30	71	0.021	(0.0068)	10.0	0.012	(0.0057)
(whale ak)	20	18	48	0.017	(0.0060)	10.1	0.012	(0.0059)
(whale km)	1	1	—	—	—	—	—	—
(whale sb)	12	12	41	0.017	(0.0068)	10.0	0.012	(0.0060)
North Pacific	12	10	33	0.013	(0.0055)	7.3	0.009	(0.0045)
(whale ty)	12	10	33	0.013	(0.0055)	7.3	0.009	(0.0045)
Southern Ocean	104	81	129	0.030	(0.0076)	12.4	0.015	(0.0069)
(Argentina)	42	31	79	0.022	(0.0068)	13.7	0.017	(0.0077)
(whale ac)	13	12	52	0.021	(0.0081)	13.7	0.017	(0.0080)
(whale mr)	8	5	22	0.010	(0.0049)	8.7	0.011	(0.0055)
(whale ms)	16	12	52	0.019	(0.0072)	13.9	0.017	(0.0081)
(whale rb)	5	5	41	0.024	(0.0124)	18.7	0.023	(0.0123)
(South Africa)	42	36	88	0.025	(0.0075)	10.8	0.013	(0.0061)
(whale 1b)	12	11	50	0.020	(0.0082)	14.1	0.017	(0.0083)
(whale 2b)	21	18	53	0.018	(0.0064)	9.8	0.012	(0.0057)
(whale 3b)	7	7	19	0.010	(0.0047)	7.0	0.009	(0.0045)
(whale 4b)	1	1	—	—	—	—	—	—
(whale 5b)	1	1	—	—	—	—	—	—
(Australia)	20	17	59	0.020	(0.0072)	12.0	0.015	(0.0069)
(whale rh)	20	17	59	0.020	(0.0072)	12.0	0.015	(0.0069)

n, number of sequences in sample (or subsample); *h*, number of distinct haplotypes; *S*, number of segregating (polymorphic) sites; θ_S, population mutation rate per site, estimated from *S* (Watterson's θ/L); Π, mean number of pairwise differences (per sequence); π, mean number of pairwise differences (per site); SD, standard deviation (sampling and stochastic variation).

Table 3 Polymorphism summary for *Cyamus gracilis*

Sample (subsample)	n	h	S	θ_S	(SD)	Π	π	(SD)
North Atlantic	30	8	22	0.007	(0.0025)	5.6	0.007	(0.0034)
(whale ak)	2	2	5	0.006	(0.0047)	5.0	0.006	(0.0047)
(whale km)	1	1	—	—	—	—	—	—
(whale sb)	27	6	18	0.006	(0.0022)	5.3	0.007	(0.0032)
Southern Ocean	28	19	33	0.010	(0.0036)	6.3	0.008	(0.0038)
(whale 1b)	3	3	11	0.009	(0.0058)	7.3	0.009	(0.0058)
(whale 4b)	6	5	9	0.005	(0.0027)	4.5	0.006	(0.0032)
(whale 5b)	19	13	27	0.009	(0.0036)	6.5	0.008	(0.0039)

Column headings as in Table 2.

Table 4 Polymorphism summary for *Cyamus erraticus*

Sample (subsample)	n	h	S	θ_S	(SD)	Π	π	(SD)
North Atlantic	22	15	39	0.013	(0.0047)	10.3	0.013	(0.0060)
(whale km)	10	9	33	0.014	(0.0062)	11.2	0.014	(0.0068)
(whale mb)	1	1	—	—	—	—	—	—
(whale tl)	11	7	25	0.010	(0.0045)	10.0	0.012	(0.0061)
Southern Ocean	21	16	38	0.013	(0.0047)	7.1	0.009	(0.0042)
(Argentina)	10	10	25	0.011	(0.0048)	7.2	0.009	(0.0045)
(whale ms)	5	5	15	0.009	(0.0048)	7.4	0.009	(0.0051)
(whale rb)	5	5	16	0.009	(0.0051)	6.6	0.008	(0.0046)
(South Africa)	11	8	26	0.011	(0.0047)	7.0	0.009	(0.0043)
(whale 1b)	6	6	22	0.012	(0.0059)	8.6	0.011	(0.0057)
(whale 2b)	5	4	10	0.006	(0.0033)	5.0	0.006	(0.0036)

Column headings as in Table 2.

differences even among the three nominal species, despite synonymous divergences (d_S) well above 1.0 (see Fig. 3). Neutral mutation parameters ($\theta = 2N_f\mu$) estimated from numbers of segregating sites (θ_S) range from 0.007 (in North Atlantic *Cyamus gracilis*) to 0.030 (in southern *C. ovalis*); mean heterozygosities per site (π) range from 0.007 to 0.015; and mean synonymous divergences (d_S) range from 0.027 to 0.067. These levels of mitochondrial synonymous polymorphism are high relative to those seen in other metazoan populations without apparent geographical structure (Avise 2000). Assuming two generations per year and typical arthropod mitochondrial mutation rates (3–5 $\times 10^{-8}$/year, from the snapping-shrimp calibration discussed below), these diversities imply effective population sizes in the millions. Human exploitation of right whale populations has reduced whale and cyamid populations simultaneously, but these changes are too recent to have had a large effect on levels of nucleotide variation in cyamid populations, which are hundreds to thousands of times larger than those of the whales.

North Atlantic right whales are themselves less genetically variable than southern right whales, both for mitochondrial and nuclear markers (Malik *et al*. 2000; Waldick

et al. 2002). Right whale cyamids show this pattern only weakly, if at all (Tables 2–4). Averaging over the three nominal species, $\theta_S = 0.014$ in the North Atlantic and 0.018 in the Southern Ocean; the average value of π is 0.011 in both hemispheres. If levels of nuclear variation in cyamids also turn out to be similar in the Northern and Southern Hemispheres (as suggested by preliminary data for one nuclear locus, EF-1α, discussed briefly below), then the close relationship between cyamid and right whale population histories will imply that genetic diversity in North Atlantic right whales was reduced in their recent evolutionary history, and that in the more distant past northern and southern right whales had similar average population sizes.

Levels of variation differ among the three nominal cyamid species in a pattern of the kind expected, given their very different population sizes: *C. ovalis* is most variable, *C. gracilis* is least variable, and *Cyamus erraticus* is intermediate (Tables 2–4). This pattern holds without exception, both within and between ocean systems, for θ_S. However, θ_S is estimated on the assumption that the number of polymorphic sites is the same as the number of mutations in the gene tree (i.e. no site has mutated more than once). This

assumption is violated in every sample, most severely for southern *C. ovalis* where the number of polymorphic sites ($S = 129$) is less than half the phylogenetically estimated number of mutations ($\eta = 272$) (Table 7). For this reason, and because confidence intervals for θ_S can be strongly asymmetric, we used the method of Kreitman & Hudson (1991) to estimate 95% confidence intervals around estimates of θ_S derived from phylogenetically estimated numbers of mutations rather than from numbers of segregating sites. The resulting lower bound for southern *C. ovalis* ($\theta_L = 0.039$) exceeds the upper bounds for North Atlantic and Southern Ocean *C. gracilis* and *C. erraticus* ($\theta_U = 0.015–0.033$); similarly, the lower bound for North Atlantic *C. ovalis* ($\theta_L = 0.015$) equals the upper bound for North Atlantic *C. gracilis* ($\theta_U = 0.015$). All other comparisons result in overlapping confidence intervals and thus (by this criterion) nonsignificant differences between estimates of θ_S. Estimates of π show a qualitatively similar pattern among the species, except that North Atlantic *C. erraticus* ($\pi = 0.012$) is equal to North Atlantic *C. ovalis* and higher than North Pacific *C. ovalis* ($\pi = 0.009$). However, none of the pairwise differences in π approach formal significance, owing to the large stochastic variance of π (Tables 2–4).

Interestingly, the relative magnitudes of the differences among the species (in θ_S as well as in π) are consistently smaller than expected from the apparently large differences among their census population sizes, and all of the observed levels of variation are absolutely much lower than expected on the basis of conservative direct estimates of the species' long-term effective population sizes, as discussed below.

Almost no population structure within oceans

The hierarchical analyses of molecular variance (AMOVA) summarized in Table 5 estimate differentiation among the cyamid populations on individual whales within local breeding aggregations and, for the Southern Hemisphere, differentiation among the geographically defined aggrega-

tions. For southern *C. ovalis*, we assigned the eight samples from individual whales to three groups representing the major aggregations (four whales in Argentina, three in South Africa, and one in Australia). For southern *C. gracilis* there are two samples in one group (two whales in South Africa), and for southern *C. erraticus* there are four samples in two groups (two whales in Argentina and two in South Africa). For all three North Atlantic species there are two whales in one group (though not always the same two whales).

In five of the six cases, almost all of the variation (95–100%) is estimated to occur on individual whales within groups, and in the two cases with more than one group, none of the variation is estimated to occur among groups (i.e. $F_{ST} = 0$ among the Southern Ocean breeding aggregations). In southern *C. ovalis*, slightly less than 5% of the variation is estimated to occur among whales within groups ($F_{WS} = 0.048$), and this is significantly greater than zero ($P < 0.05$). Southern *C. gracilis* shows a similar level of differentiation among whales within groups ($F_{WS} = 0.052$, not significant). Even these modest levels of genetic differentiation among conspecific cyamid populations on different whales within breeding aggregations may be overestimates, because most of our samples come from calves and juveniles whose cyamid populations may represent only a portion of the diversity present on their mothers, with relatively little horizontal input (so early in their lives) from unrelated whales. But even if our estimates of F_{WS} accurately represent the values typical of cyamids on adult right whales, they imply that several to many female cyamids transfer onto a typical whale per cyamid generation.

There is a striking lack of haplotypic diversity in the sample of North Atlantic *C. gracilis*, with just eight distinct haplotypes among 30 sequences ($h/n = 8/30 = 0.27$ haplotypes/sequence; see Table 3). By contrast, southern *C. gracilis* shows 19 haplotypes in a sample of 28 sequences ($h/n = 0.68$), and the other species show h/n ratios of 0.68 (North Atlantic *C. erraticus*) to 0.91 (North Atlantic *C. ovalis*). Most of the North Atlantic *C. gracilis* sample comes from

Table 5 Fixation indices estimated by AMOVA

Species		No. of whales sampled in subpopulation				F_{WS}	F_{ST}
		NA	AR	SA	AU		
C. ovalis	N. Atlantic	2				−0.005	
	S. Ocean		4	3	1	0.048*	−0.0004
C. gracilis	N. Atlantic	(2)†				(0.214)†	
	S. Ocean			2		0.052	
C. erraticus	N. Atlantic	2				−0.040	
	S. Ocean		2	2		0.016	−0.009

NA, North Atlantic subpopulation (stock, calving aggregation); AR, Argentina; SA, South Africa; AU, Australia; F_{WS}, variation among whales within subpopulations (stocks); F_{ST}, subpopulations within the total (Southern Hemisphere); *$P < 0.05$; ()† 2 individuals from whale ak, 27 from whale sb, $P \approx 0.06$.

one whale (Eg#1004, 'Stumpy', referred to as whale 'sb' in the tables), and these 27 sequences represent just six haplotypes ($h/n = 0.22$) (see Table 3). The other three North Atlantic *C. gracilis* sequences come from two whales, and all three are different from each other ($h/n = 1.0$). Despite its low haplotypic diversity, the North Atlantic *C. gracilis* sample shows levels of nucleotide diversity (π and θ_S) nearly equal to those of southern *C. gracilis* (Table 3). Taken together, these patterns imply that the *C. gracilis* population on Stumpy is atypical and that, as a consequence, little meaning should be attached to the high estimate of F_{WS} (0.214) for North Atlantic *C. gracilis*.

We do not see any obvious explanation for the low haplotypic diversity of the *C. gracilis* sample from Stumpy; she was a mature female, and her *C. ovalis* population is highly diverse ($h/n = 12/12 = 1.0$; see Table 2). However, we recently detected the intracellular parasitic bacterium *Wolbachia* in several *C. gracilis* individuals from Stumpy. In many other arthropods, *Wolbachia* infections cause cytoplasmic incompatibility and other reproductive anomalies that may transiently and locally boost the frequencies of associated mitochondrial haplotypes (Hoffmann & Turelli 1997; Charlat *et al*. 2001). We are currently surveying all of our right whale cyamid DNAs for the possible presence of *Wolbachia* and functionally analogous microsporidia (Terry *et al*. 2004); these investigations will be reported elsewhere.

Cyamid migration rates among the Southern Hemisphere breeding aggregations cannot be estimated accurately because the estimates of F_{ST} cannot be distinguished from zero. If the true values of F_{ST} were as high as 0.01, then the minimum number of migrating female cyamids would be on the order of 100 per cyamid generation. Cyamid generation times are not known but seem likely to be less than a year and more than 2 months. If the effective generation time were around 6 months, then roughly 200 cyamid females would need to migrate each year between the major Southern Ocean populations. At least for *C. ovalis*, with populations on the order of 10^3 to 10^4 per whale, such migration rates might easily be achieved by moving just one female right whale between aggregations every whale generation (20–30 years), consistent with the whales' substantial levels of mitochondrial differentiation among aggregations (as reviewed in the Introduction). But cyamids travel on male right whales as well as females, so if males migrate at rates as high or higher than females, as seems likely, then even *C. gracilis*, with populations on the order of 10^2 to 10^3 per whale, could remain effectively undifferentiated throughout the Southern Hemisphere. Cyamid migration rates might be further increased by casual contacts between whales from different subpopulations on Antarctic summer feeding grounds, or by peripatetic males that briefly visit other breeding aggregations but then return to their natal aggregation.

Reciprocal monophyly of North Atlantic and Southern Ocean cyamids, but not of North Pacific and Southern Ocean C. ovalis

Nominally conspecific cyamids of all three species from the North Atlantic and the Southern Ocean form reciprocally monophyletic clades with high bootstrap support (Fig. 3). This pattern implies that North Atlantic right whales (*Eubalaena glacialis*) have long been isolated from other right whales.

North Pacific right whale cyamids are seldom collected, and to date we have obtained only one small sample consisting almost entirely of *C. ovalis*. The one *C. erraticus* individual in this sample carries a mitochondrial haplotype that appears to be at least as distant from North Atlantic and Southern Ocean *C. erraticus* as they are from each other (Fig. 3).

By contrast, 12 individuals (10 haplotypes) from North Pacific *C. ovalis* form a clade nested within the southern *C. ovalis* gene tree (Figs 3 and 4). This pattern cannot be explained without some movement of right whales across the equator in the Pacific much more recently than in the Atlantic. The simplest model would be one in which a single southern right whale migrated to the North Pacific carrying a *C. ovalis* mitochondrial genome that subsequently swept through the North Pacific *C. ovalis* population, whereas all of its *C. erraticus* mitochondrial genomes failed to introgress into the North Pacific *C. erraticus* population. Ballard & Whitlock (2004) review evidence of such introgression in other animal taxa. Alternatively, the migrating whale's *C. ovalis* population might simply have replaced the existing North Pacific *C. ovalis* species, without interbreeding. These alternative models can be distinguished by analysing patterns of differentiation at nuclear loci, and work to this end is in progress; initial results for EF-1α show similar levels of differentiation among *C. ovalis* populations in all three ocean systems, consistent with mitochondrial introgression from the Southern Ocean, but not with a recent complete replacement (W.A.S. & J.S., unpublished).

Consistent divergence times of sibling species pairs

To estimate the divergence times of North Atlantic and Southern Ocean cyamids we used a COI molecular clock calibration for another group of small marine crustaceans. Many shallow-water species were separated into Carribean and Pacific sibling species pairs by the rise of the Isthmus of Panama, a gradual process that culminated 3 Ma with the emergence of a continuous land bridge. Knowlton & Weigt (1998) sequenced 564 bp of the *COI* gene from 15 sibling species pairs of snapping shrimp (Crustacea; Decapoda: *Alpheus*). The two most similar pairs (23 and 33 substitutions) occupy extreme inshore (mangrove) habitats.

The authors argue that the pair with the smaller difference probably diverged most recently (at 3 Ma), while the pair with the larger difference diverged earlier (as did many deeper-water *Alpheus* which show even more substitutions in *COI*). However, large sampling and stochastic errors are associated with such small numbers of substitutions, so it seems more prudent to view them as suggesting a range within which typical substitution rates are likely to fall. In any case, these two pairs of *Alpheus* sequences have diverged at overall rates similar to estimates for the *COI* genes of insects and other Crustacea (1.5–2.0% per Myr) (reviewed by Brower 1994 and Quek *et al*. 2004).

We used CODEML from PAML 3.13a (Yang 1997, 2002) to estimate synonymous and nonsynonymous substitutions between the sibling species pairs of mangrove-inhabiting *Alpheus*, and between North Atlantic and Southern Ocean pairs of right whale *Cyamus*. The estimated numbers of synonymous substitutions per synonymous site (d_S) for the two sibling species pairs of snapping shrimp are 0.20 and 0.33. The half-length (0.8 kb) *COI* sequences of North Atlantic and southern *C. ovalis* and *C. gracilis* are separated by identical mean pairwise synonymous divergences (d_S = 0.51), but the *COI* sequences of northern and southern *C. erraticus* are nearly twice as diverged (d_S = 0.98) (Table 6).

The longer (4.1 kb) 6-gene arcs extending from *COI* through *ND3* show a somewhat different pattern: the divergence of northern and southern *C. ovalis* declines to 0.40, *C. erraticus* declines to 0.64, and *C. gracilis* increases to 0.66 (Table 6). These differences suggest that synonymous substitution rates have varied among genes and species in a complicated way, for reasons that seem likely to remain obscure until the patterns can be interpreted in the context of an independent phylogeny derived from multiple nuclear genes. If all three sibling species pairs became reproductively isolated at the same time, then *C. ovalis* would appear to have a lower mitochondrial substitution rate than the other two pairs. Alternatively, *C. gracilis* and *C. erraticus* could have become resistant to mitochondrial introgression before *C. ovalis* did (i.e. earlier in the speciation of North Atlantic and southern right whales), as suggested by the seemingly large difference in times of separation for *C. erraticus* and *C. ovalis* in the North Pacific (Fig. 3).

Table 6 Divergences of North Atlantic and Southern Ocean *Cyamus* and *Eubalaena* species

Species	Gene(s)	NA	SO	L	D	P	d_N	d_S	(SE)
C. ovalis	COI	30	81	819	64.0	0.078	0.004	0.506	(0.085)
	COI-ND3	2	3	4131	273.3	0.066	0.010	0.396	(0.030)
C. gracilis	COI	8	19	819	64.8	0.079	0.003	0.506	(0.084)
	COI-ND3	1	2	4131	381.0	0.092	0.011	0.662	(0.050)
C. erraticus	COI	15	16	819	100.3	0.122	0.004	0.976	(0.163)
	COI-ND3	1	1	4131	438.0	0.106	0.021	0.643	(0.043)
E. glacialis, E. australis	COI + *cytb*	1	2	2676	26.0	0.010	0.000	0.040	(0.005)

NA, numbers of unique sequences in North Atlantic samples; SO, numbers of unique sequences in Southern Ocean samples; L, length of in-frame coding sequence (no stops or partial codons); D, mean number of nucleotide differences between NA and SO sequences; P, uncorrected mean proportion of differences (D/L); d_N, estimated mean nonsynonymous substitutions per nonsynonymous site (NA–SO); d_S, estimated mean synonymous substitutions per synonymous site (NA–SO); SE, mean pairwise standard error of the estimate of d_S.

Table 7 Tests of equilibrium under neutrality

Species	n	h	S	Π	Θ_W	η	η_e	D_T	D_{FL}	F_S
C. ovalis NA	33	30	71	10.0	17.5	91	58	−1.6*	−3.6**	−18.9**
C. ovalis NP	12	10	33	7.3	10.9	35	28	−1.5***	−2.8**	−2.3
C. ovalis SO	104	81	129	12.4	24.7	272	158	−1.7*	−6.0**	−24.2**
clade A	80	63	88	8.0	17.8	149	100	−1.8*	−6.3**	−24.8**
not clade A	24	18	81	17.7	21.7	118	59	−0.7	−1.9***	−1.5
C. gracilis NA	30	8	22	5.6	5.6	23	6	0.1	−0.1	2.5
C. gracilis SO	28	19	33	6.3	8.5	42	20	−1.0	−1.8	−7.0**
C. erraticus NA	22	15	39	10.3	10.7	45	21	−0.1	−1.4	−1.9
C. erraticus SO	21	16	38	7.1	10.6	44	21	−1.3	−1.5	−5.3*

n, number of sequences in sample; h, number of unique haplotypes; S, number of segregating sites; Π, mean number of pairwise differences (per sequence); Θ_W, population mutation rate per sequence, estimated from S (Watterson's Θ); η, number of mutations in sample, estimated phylogenetically; η_e, number of mutations on external branches, estimated phylogenetically; D_T, Tajima's (1989a) D statistic; D_{FL}, Fu & Li's (1993) D statistic; F_S, Fu's (1997) F_S statistic; ***$P < 0.06$; *$P < 0.05$; **$P < 0.01$.

The average 6-gene (4.1 kb) synonymous divergence for all three species of North Atlantic and Southern Ocean cyamids is 0.57, and the average partial *COI* synonymous divergence for the two mangrove-inhabiting *Alpheus* is 0.27. On the assumption that *Alpheus* and *Cyamus* mitochondrial clocks run at the same average rates, and on the assumption that the *Alpheus* pairs speciated 3 Ma, then the right whale cyamids are estimated to have speciated (3) $(0.57/0.27) = 6.3$ Ma. This falls within the range (3–12 Ma) estimated from North Atlantic and southern right whale control region sequences by Malik *et al.* (2000) using a calibration from other cetaceans (Hoelzel *et al.* 1991). The most extreme estimates implied by the 6-gene divergences in Table 6 and the two *Alpheus* calibrations also fall within this range; the smallest is 3.6 Ma [= $(3)(0.40/0.33)$, for *C. ovalis* and the faster pair of *Alpheus* species], and the largest is 9.9 Ma [= $(3)(0.66/0.20)$, for *C. gracilis* and the slower *Alpheus*].

More accurate estimates of the cyamid divergences would account for coalescence within the populations ancestral to modern *Alpheus* and *Cyamus* sibling species pairs, but we lack (i) population samples that are needed to estimate coalescence times for *Alpheus* species and (ii) long sequences that are needed to reduce the variances of the estimates of their average divergences. In addition, decapod (*Alpheus*) and amphipod (*Cyamus*) Crustacea are distantly related. A more direct calibration of cyamid molecular clocks could be made using trans-isthmian sibling species pairs within their sister family, the Caprellidae.

North Atlantic and Southern Ocean right whales (*Eubalaena glacialis* and *Eubalaena australis*) show mitochondrial synonymous divergences one-tenth as large as the smallest seen in their cyamids: $d_S = 0.04$ for *COI* and cyt *b* in right whales, while $d_S = 0.40$ for *COI* and five other genes in *C. ovalis* (Table 6). If synonymous substitutions in right whale mitochondria are effectively neutral, and if North Atlantic and southern right whales diverged around 6 Ma, then the mitochondrial mutation rate in right whales is roughly 0.040 subs/site/12×10^6 year $= 3.3 \times 10^{-9}$ subs/site/ year, about one-tenth as large as typical rates estimated for other mammals (Brown *et al.* 1982; Parsons *et al.* 1997) and two-thirds to one-sixth as large as rates estimated for the hypervariable segment of the control region in cetaceans ($5–20 \times 10^{-9}$ subs/site/ year) (Hoelzel *et al.* 1991; Pesole *et al.* 1999; Rooney *et al.* 2001).

A phylogenetic analysis of baleen whale mitochondrial genome sequences was recently published by Sasaki *et al.* (2005). These authors estimated the ages of internal nodes in their tree using the 'relaxed clock' model of Thorne *et al.* (1998) and Kishino *et al.* (2001), with calibration provided by two fossil dates. This procedure gives a date of 4.4 ± 2.5 Ma for the divergence of North Pacific and southern right whales. Sasaki *et al.* (2005) did not include the North Atlantic right whale in their study, but other analyses of mitochondrial

control region and nuclear gene sequences suggest that North Pacific and southern right whales became isolated from each other shortly after separation from the North Atlantic population (Rosenbaum *et al.* 2000; Gaines *et al.* 2005). Thus the date estimated by Sasaki *et al.* (2005) appears to be fully consistent with our estimates derived from cyamid and snapping shrimp divergences.

Nonequilibrium levels and configurations of sequence variation

The levels of variation reported here are comparatively high but far lower than expected under neutral-theory assumptions. For a haploid population of size N_f at mutation–drift equilibrium, the expected nucleotide diversity at neutral sites is $\pi \approx \theta = 2N_f\mu$, where μ is the neutral mutation rate per generation. If we assume, to be very conservative, that the long-term average number of right whales in an ocean is 20 000 and the effective number of breeding female *C. ovalis* per right whale is 1000, then $N_f = 2 \times 10^7$. A conservative estimate of μ is 1×10^{-8} (from the clock calibrations derived above, assuming speciation 5 Ma and four generations per year, or speciation 10 Ma and two generations per year). Combining these conservative assumptions, we expect a synonymous-site nucleotide diversity of at least $\pi = (2)(2 \times 10^7)(1 \times 10^{-8}) = 0.4$ (after correcting for multiple hits). The highest actual synonymous polymorphism (mean pairwise d_S) is 0.067, for southern *C. ovalis*. Thus the largest of seven estimates falls nearly an order of magnitude below an unrealistically low expectation; the actual shortfall seems likely to approach two orders of magnitude, at least for *C. ovalis*. For example, if there are 50 000 whales and 2000 female cyamids per whale, and if $\mu \approx 2 \times 10^{-8}$, then $\theta \approx 4$ and every synonymous site should be highly polymorphic. This striking lack of agreement between theory and observation implies that one or more of the assumptions used to derive the expectations must be false. The question is, which one(s)?

For example, effective female population sizes could be far less than census sizes if only small fractions of the adults on a given whale ever succeed in reproducing. Competition for space on the whale's callosities is obviously severe (Fig. 2), and females can produce more than 100 offspring in a single brood, so conditions permitting high levels of female reproductive skew are clearly present. 'Recruitment sweepstakes' syndromes have been proposed to explain why other marine invertebrates also show much lower levels of genetic polymorphism than would be expected under neutral-theory assumptions, given their apparently huge population sizes (Hedgecock 1994; Flowers *et al.* 2002; Árnason 2004). But unlike the species for which sweepstakes models have been proposed, cyamids do not have planktonic larvae that drift through unpredictable environments; instead, cyamids remain firmly attached

to very stable, predictable, and firmly bounded environments, so some important aspects of the sweepstakes scenario are missing. In addition, our data contradict a prediction of the sweepstakes model: if only a few females reproduce on a given whale at a given time, then their mitochondrial haplotypes should be very common at least among the juvenile cohorts, and there should be considerable genetic differentiation (F_{WS}) among the cyamid subpopulations on whales within local breeding aggregations. Instead, we find that most mitochondrial haplotypes are rare, even on a given whale (Tables 2–4), and that there is very little differentiation among the subpopulations on whales within aggregations (Table 5). The only way to rescue the reproductive-skew hypothesis (given high haplotype diversity and low differentiation) is to assume what would appear to be implausibly high rates of adult cyamid migration between whales.

Alternatively, right whale populations might have been much smaller in the past than they were when whaling began. In that case the gene trees of right whales and their cyamids should all show the signatures of simultaneous and substantial population expansion. Sudden expansion from a prolonged bottleneck creates star- or comb-shaped gene trees, strongly 'humped' (unimodal) distributions of pairwise differences, excessive numbers of rare polymorphisms and (as a consequence) negative values of tree-shape statistics such as Tajima's (1989a) D, which compares the mean pairwise difference to the number of polymorphic sites (Tajima 1989b; Slatkin & Hudson 1991; Rogers & Harpending 1992; Fu 1997). The observed tree-shape statistics are indeed mostly negative — often significantly so (Table 7 and Fig. 4). There is also one seemingly dramatic (but partial) case of a comb-shaped gene tree (clade 'A' in southern *C. ovalis*; Fig. 4), and it gives rise to a strongly and smoothly humped pairwise-difference distribution (not shown). However, these findings do not make a strong case for recent population expansion, because such an event should leave correlated signals in the gene trees of all three cyamid species associated with the right whale species whose population expanded. Instead, the pattern appears to be one in which the deep branches in all seven gene trees tend to be shorter than expected under neutrality, regardless of the absolute time depth.

The mitochondrial control region sequences of southern right whales also imply large long-term effective population sizes for the whales, and they show no obvious signal of population growth. The *E. australis* mitochondrial gene tree is very deep, with a root that appears to be at least half the age of the species (Portway 1998; Baker *et al.* 1999; Rosenbaum *et al.* 2000), and levels of control region polymorphism are comparable to those in humpback whales which have a large, globally distributed population (Baker *et al.* 1993, 1994, 1998). Patenaude (2002) and Patenaude *et al.* (in review) analyse 136 *E. australis* control region

sequences representing all of the major Southern Ocean breeding aggregations. There are 39 polymorphic sites in an alignment of 275 bp, giving $\theta_S = 0.026$, which is nearly identical to the mean pairwise difference, $\pi = 0.027$. Tajima's $D = 0.14$, which is statistically indistinguishable from zero. Thus the depth and shape of the *E. australis* gene tree does not suggest that the species passed through a significant bottleneck in its recent evolutionary history, prior to human exploitation.

The *E. australis* population size implied by its own mitochondrial polymorphism is large. If the mean female generation time is 25 years and $\mu = 1-2 \times 10^{-8}$ subs/site/year (Pesole *et al.* 1999; Rooney *et al.* 2001), then the per-generation mutation rate is $2.5-5 \times 10^{-7}$, implying long-term female effective population sizes of 30 000–60 000, consistent with or larger than estimates based on historic catch records (Baker & Clapham 2004). Population-size estimates derived from genetic polymorphisms are subject to many potential sources of error (Roman & Palumbi 2003; Baker & Clapham 2004), especially when the data consist entirely of short mitochondrial sequences, so this crude estimate of the *E. australis* population size should be viewed with caution. (Among other problems, it does not take account of the mitochondrial differentiation shown by *E. australis* subpopulations.) We suggest only that the data available for southern right whales do not support an explanation of the cyamid 'variation shortfall' based on evolutionarily recent expansion of right whale populations. Credible estimates of long-term right whale and cyamid population histories will require multilocus nuclear data sets and coalescent-based analyses.

Selection provides another possible explanation for the low levels of nucleotide diversity seen in cyamid mitochondria, and for the consistently negative values of Tajima's D and other tree-shape statistics. Neutral variation is temporarily 'swept away' when the chromosome linked to an adaptive mutation 'hitches a ride' to fixation (Maynard Smith & Haigh 1974). Neutral polymorphism then slowly recovers, and during this recovery process gene trees in the genetic neighbourhood of the adaptive substitution tend to be comb-shaped, just as if the population had expanded following a bottleneck (Kaplan *et al.* 1989; Braverman *et al.* 1995). Likewise, selection against unconditionally deleterious mutations ('background' selection) can reduce the genetic effective population size by creating a large class of 'loaded' chromosomes that rarely contribute descendants to the distant future (Charlesworth *et al.* 1993, 1995).

Mitochondrial genomes have high mutation rates, no recombination, and gene products vital to fitness, so both adaptive hitch-hiking and background selection could affect their levels and patterns of neutral variation, especially in very large populations where selection can discriminate between tiny fitness differences (Gillespie 2001). We have

studied a model, similar to that of Tachida (2000), which combines elements of both processes and easily reproduces the qualitative features of right whale cyamid mitochondrial variation including (i) chronically but weakly comb-shaped gene trees and (ii) large reductions in neutral variation that are relatively greater in species with larger population sizes, thereby giving rise to compressed ranges of θ_S and π among species. This model will be described elsewhere. Such reductions of variation (and skews in the distribution of variation) are expected to be more pronounced for the mitochondrion than for nuclear loci, especially those with high local rates of recombination; by contrast, population history affects all loci in similar ways. Work is in progress to test the alternative hypotheses (selection vs. population history) by comparing patterns of nuclear and mitochondrial variation in all of the sampled populations and species of right whale cyamids.

Conclusions

Cyamid subpopulations on individual right whales exhibit high levels of haplotype diversity and low levels of genetic differentiation, contrary to our initial expectation. This pattern indicates that cyamids of all three nominal species transfer between unrelated whales so often that it would be difficult or impossible to infer differential rates of behavioural interaction from patterns of genetic similarity and difference among the cyamid subpopulations on individual whales.

There is no detectable differentiation of cyamids on right whales from the major Southern Ocean breeding aggregations that calve off South America, southern Africa, and Australia. At this hemispheric scale, each of the three southern cyamid species appears to form a single effectively panmictic population in which, for example, two mitochondria sampled off Argentina are as different from each other, on average, as each is from mitochondria sampled off South Africa or Australia. This finding could have been predicted, because drift should be extremely slow within the enormous cyamid populations that occur at the scale of right-whale breeding aggregations; low rates of interaggregation migration (consistent with the observed differentiation of female southern right whales) would then be enough to prevent detectable differentiation of right whale cyamids. Actual rates of migration could be higher than this minimum, of course, but our present data lack power to estimate these rates with any precision; larger multilocus data sets may do better.

By contrast, cyamid populations on opposite sides of the equator appear to have been fully (or almost fully) isolated for several million years. This finding strongly supports the view that North Atlantic, North Pacific, and southern right whales also have been isolated for several million years and therefore should be considered distinct species (Rosenbaum *et al*. 2000; Gaines *et al*. 2005). At least one right whale has crossed the equator in the Pacific during this time, but such exchanges appear to be rare. Multilocus nuclear sequence data for cyamids (calibrated by trans-isthmian caprellids) should allow the times of separation between northern and southern populations to be estimated with some precision.

Cyamid mitochondria consistently show relatively too many low-frequency mutations and absolutely less nucleotide diversity than expected under neutral-theory assumptions for populations at equilibrium. These patterns do not appear to have been caused by dramatic, long-lasting changes in the population sizes of right whales, at least not during the last few hundred thousand years. Long-lasting bottlenecks would have left strong, correlated signatures in the gene trees of all three cyamids associated with a given species of right whale (reflecting the threefold 'replication' of the parasite–host relationship). No such patterns are apparent at time depths where all of the species still have many surviving mitochondrial lineages. However, all of the mitochondrial gene trees except that of southern *C. ovalis* appear to coalesce fully within the last half million years, so multilocus nuclear sequence data will be needed to estimate the population histories of right whale cyamids and their hosts farther into the past, and with better resolution. Work to this end is in progress, and we hypothesize that it will support a model in which levels of mitochondrial synonymous polymorphism in right whale cyamids have been greatly depressed by background selection, owing to the very large effective population sizes of these species.

Acknowledgements

References

Árnason E (2004) Mitochondrial cytochrome *b* DNA variation in the high-fecundity Atlantic cod: trans-Atlantic clines and shallow gene genealogy. *Genetics*, **166**, 1871–1885.

Avise J (2000) *Phylogeography: The History and Formation of Species.* Harvard University Press, Cambridge, Massachusetts.

Baker CS, Clapham PJ (2004) Modelling the past and future of whales and whaling. *Trends in Ecology & Evolution*, **19**, 365–371.

Baker CS, Perry A, Bannister JL *et al.* (1993) Abundant mitochondrial DNA variation and world-wide population structure in humpback whales. *Proceedings of the National Academy of Sciences, USA*, **90**, 8239–8243.

Baker CS, Slade RW, Bannister JL *et al.* (1994) Hierarchical structure of mitochondrial DNA gene flow among humpback whales *Megaptera novaeangliae*, world-wide. *Molecular Ecology*, **3**, 313–327.

Baker CS, Medrando-Gonzalez L, Calambokidis J *et al.* (1998) Population structure of nuclear and mitochondrial DNA variation among humpback whales in the North Pacific. *Molecular Ecology*, **7**, 695–707.

Baker CS, Patenaude NJ, Bannister JL, Robins J, Kato H (1999) Distribution and diversity of mtDNA lineages among southern right whales (*Eubalaena australis*) from Australia and New Zealand. *Marine Biology*, **134**, 1–7.

Ballard JWO, Whitlock MC (2004) The incomplete natural history of mitochondria. *Molecular Ecology*, **13**, 729–744.

Best P (1990) Natural markings and their use in determining calving intervals in right whales off South Africa. *South African Journal of Zoology*, **25**, 114–123.

Braverman JM, Hudson RR, Kaplan NL, Langley CH, Stephan W (1995) The hitchhiking effect on the site frequency spectrum of DNA polymorphisms. *Genetics*, **140**, 783–796.

Brower AVZ (1994) Rapid morphological radiation and convergence among races of the butterfly *Heliconius erato* inferred from patterns of mitochondrial DNA evolution. *Proceedings of the National Academy of Sciences, USA*, **91**, 6491–6495.

Brown WM, Prager EM, Wang A, Wilson AC (1982) Mitochondrial sequences of primates: tempo and mode of evolution. *Journal of Molecular Evolution*, **18**, 225–239.

Brownell RL, Clapham PJ, Miyashita T, Kasuya T (2001) Conservation status of North Pacific right whales. *Journal of Cetacean Research and Management (Special Issue)*, **2**, 269–286.

Carter MJ, Milton ID (1993) An inexpensive and simple method for DNA purification on silica particles. *Nucleic Acids Research*, **21**, 1044.

Caswell H, Fujiwara M, Brault S (1999) Declining survival probability threatens the North Atlantic right whale. *Proceedings of the National Academy of Sciences, USA*, **96**, 3308–3313.

Charlat S, Bourtzis K, Merçot H (2001) *Wolbachia* induced cytoplasmic incompatibility. In: *Symbiosis: Mechanisms and Model Systems* (ed. Seckbach J), pp. 623–642. Kluwer Academic Publishers, Dordrecht, The Netherlands.

Charlesworth B, Morgan MT, Charlesworth D (1993) The effect of deleterious mutations on neutral molecular variation. *Genetics*, **134**, 1289–1303.

Charlesworth D, Charlesworth B, Morgan MT (1995) The pattern of neutral molecular variation under the background selection model. *Genetics*, **141**, 1619–1632.

Excoffier L, Smouse P, Quattro J (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics*, **131**, 479–491.

Felsenstein J (2002) PHYLIP (*Phylogeny Inference Package) version 3.6a3*, Distributed by the author. Department of Genome Sciences, University of Washington. (http://evolution.genetics.washington.edu/phylip.html).

Flowers JM, Schroeter SC, Burton RS (2002) The recruitment sweepstakes has many winners: genetic evidence from the sea urchin *Strongylocentrotus purpuratus*. *Evolution*, **56**, 1445–1453.

Fu YX (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics*, **147**, 915–925.

Fu YX, Li W-H (1993) Statistical tests of neutrality of mutations. *Genetics*, **133**, 693–709.

Gaines CA, Hare MP, Beck SE, Rosenbaum HC (2005) Nuclear markers confirm taxonomic status and relationships among highly endangered and closely related right whale species. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, **272**, 533–542.

Gillespie JH (2001) Is the population size of a species relevant to its evolution? *Evolution*, **55**, 2161–2169.

Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Systematic Biology*, **52**, 696–704. (http://atgc.lirmm.fr/phyml/).

Hedgecock D (1994) Does variance in reproductive success limit effective population sizes of marine organisms?. In: *Genetics and Evolution of Aquatic Organisms* (ed. Beaumont A), pp. 122–134. Chapman & Hall, London.

Hoelzel AR, Hancock JM, Dover GA (1991) Evolution of the cetacean mitochondrial D-loop region. *Molecular Biology and Evolution*, **8**, 475–493.

Hoffmann AA, Turelli M (1997) Cytoplasmic incompatibility in insects. In: *Influential Passengers: Inherited Microorganisms and Arthropod Reproduction* (eds O'Neill SL, Werren JH, Hoffmann AA), pp. 42–80. Oxford University Press, Oxford.

Hudson RR (1990) Gene genealogies and the coalescent process. In: *Oxford Surveys in Evolutionary Biology*, Vol. 7 (eds Futuyma D, Antonovics J), pp. 1–44. Oxford University Press, Oxford.

International Whaling Commission (2001) Report of the workshop on the comprehensive assessment of right whales: a worldwide comparison, Cape Town, 19–25 March 1998. *Journal of Cetacean Research and Management (Special Issue)*, **2**, 1–60.

Kaplan N, Hudson RR, Langley CH (1989) The 'hitchhiking effect' revisited. *Genetics*, **123**, 887–899.

Kishino H, Thorne JL, Bruno WJ (2001) Performance of a divergence time estimation method under a probabilistic model of rate evolution. *Molecular Biology and Evolution*, **18**, 352–361.

Knowlton N, Weigt LA (1998) New dates and new rates for divergence across the Isthmus of Panama. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, **265**, 2257–2263.

Knowlton AR, Kraus SD, Kenney RD (1994) Reproduction in North Atlantic right whales (*Eubalaena glacialis*). *Canadian Journal of Zoology*, **72**, 1297–1305.

Kraus SD, Moore KE, Price CA *et al.* (1986) The use of photographs to identify individual North Atlantic right whales (*Eubalaena glacialis*). In: *Right Whales: Past and Present Status* (eds Brownell RL Jr, Best PB, Prescott JH), *Reports of the International Whaling Commission (Special Issue 10)*, pp. 145–151. International Whaling Commission, Cambridge.

Kraus SD, Hamilton PK, Kenney RD, Knowlton AR, Slay CK (2001) Reproductive parameters of the North Atlantic right whale. *Journal of Cetacean Research and Management (Special Issue)*, **2**, 231–236.

Kreitman M, Hudson RR (1991) Inferring the evolutionary histories of the *Adh* and *Adh-dup* loci in *Drosophila melanogaster* from patterns of polymorphism and divergence. *Genetics*, **127**, 565–582.

LeDuc RG, Perrin WF, Dizon AE (1999) Phylogenetic relationships among the delphinid cetaceans based on full cytochrome *b* sequences. *Marine Mammal Science*, **15**, 619–648.

Leung Y-M (1967) An illustrated key to the species of whale-lice (Amphipoda, Cyamidae), ectoparasites of Cetacea, with a guide to the literature. *Crustaceana*, **12**, 279–291.

Maddison D, Maddison W (2000) MACCLADE 4: *Analysis of Phylogeny and Character Evolution*. Sinauer Associates, Sunderland, Massachusetts.

Malik S, Brown MW, Kraus SD, White BN (2000) Analysis of mitochondrial DNA diversity within and between North and South Atlantic right whales. *Marine Mammal Science*, **16**, 545–558.

Maynard Smith J, Haigh J (1974) The hitch-hiking effect of a favourable gene. *Genetical Research*, **23**, 23–35.

Parsons TJ, Muniec DS, Sullivan K *et al.* (1997) A high observed substitution rate in the human mitochondrial DNA control region. *Nature Genetics*, **15**, 363–368.

Patenaude NJ (2002) *Demographic and genetic status of southern right whales at the Auckland Islands, New Zealand*. PhD Thesis, Biological Sciences, University of Auckland, New Zealand.

Patenaude NJ, Portway VA, Schaeff CM *et al.* (in review) Genetic differentiation and population structure among southern right whales.

Payne R, Brazier O, Dorsey EM, Perkins JS, Rowntree VJ, Titus A (1983) External features in southern right whales (*Eubalaena australis*) and their use in identifying individuals. In: *Communication and Behavior of Whales* (ed. Payne R), pp. 371–445. AAAS Selected Symposium 76. Westview Press, Boulder, Colorado.

Pesole G, Gissi C, De Chirico A, Saccone C (1999) Nucleotide substitution rate of mammalian mitochondrial genes. *Journal of Molecular Evolution*, **48**, 427–438.

Portway VA (1998) *Genetic population structure of South Atlantic right whales (*Eubalaena australis*)*. MSc Thesis, American University, Washington, D.C.

Posada D, Crandall KA (1998) MODELTEST: testing the model of DNA substitution. *Bioinformatics*, **14**, 817–818. (http://darwin.uvigo.es/software/modeltest.html).

Quek S-P, Davies SJ, Itino T, Pierce NE (2004) Codiversification in an ant–plant mutualism: stem texture and the evolution of host use in *Crematogaster* (Formicidae: Myrmicinae) inhabitants of *Macaranga* (Euphorbiaceae). *Evolution*, **58**, 554–570.

Rice DW (1998) *Marine Mammals of the World: Systematics and Distribution*. Society for Marine Mammalogy Special Publication Number 4. Allen Press, Lawrence, Kansas.

Rogers AR, Harpending HC (1992) Population growth makes waves in the distribution of pairwise genetic differences. *Molecular Biology and Evolution*, **9**, 552–569.

Roman J, Palumbi SR (2003) Whales before whaling in the North Atlantic. *Science*, **301**, 508–510.

Rooney AP, Honeycutt RL, Derr JN (2001) Historical population size change of bowhead whales inferred from DNA sequence polymorphism data. *Evolution*, **55**, 1678–1685.

Rosenbaum HC, Brownell RL Jr, Brown MW *et al.* (2000) World-wide genetic differentiation of *Eubalaena*: questioning the number of right whale species. *Molecular Ecology*, **9**, 1793–1802.

Roussel de Vauzème M (1834) Mémoire sur le *Cyamus ceti* (Latr.) de la classe des Crustacés. *Annales des Sciences Naturelles, Zoologie, 2e Série*, **1**, 239–265.

Rowntree VJ (1983) Cyamids: the louse that moored. *Whalewatcher: Journal of the American Cetacean Society*, **17**, 14–17.

Rowntree VJ (1996) Feeding, distribution, and reproductive behavior of cyamids (Crustacea: Amphipoda) living on humpback and right whales. *Canadian Journal of Zoology*, **74**, 103–109.

Rozas J, Sánchez-DelBarrio JC, Messeguer X, Rozas R (2003) DNASP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics*, **19**, 2496–2497. (http://www.ub.es/dnasp).

Sambrook J, Fritsch EF, Maniatis T (1989) *Molecular Cloning: A Laboratory Manual*, 2nd edn. Cold Spring Harbor Laboratory Press, Plainview, New York.

Sasaki T, Nikaido M, Hamilton H *et al.* (2005) Mitochondrial phylogenetics and evolution of mysticete whales. *Systematic Biology*, **54**, 77–90.

Schneider S, Roessli D, Excoffier L (2000) ARLEQUIN version 2.000: *A Software for Population Genetics Data Analysis*. Genetics and Biometry Laboratory, Department of Anthropology and Ecology, University of Geneva, Switzerland. (http://anthro.unige.ch/arlequin).

Simon C, Frati F, Beckenbach A, Crespi B, Liu H, Flook P (1994) Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. *Annals of the Entomological Society of America*, **87**, 651–701.

Slatkin M, Hudson RR (1991) Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. *Genetics*, **129**, 555–562.

Swofford DL (1998) PAUP*. *Phylogenetic Analysis Using Parsimony (*and Other Methods)*, Version 4. Sinauer Associates, Sunderland, Massachusetts.

Tachida H (2000) Molecular evolution in a multisite nearly neutral mutation model. *Journal of Molecular Evolution*, **50**, 69–81.

Tajima F (1989a) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, **123**, 585–595.

Tajima F (1989b) The effect of change in population size on DNA polymorphism. *Genetics*, **123**, 597–601.

Takahata N, Palumbi SR (1985) Extranuclear differentiation and gene flow in the finite island model. *Genetics*, **109**, 441–457.

Terry RS, Smith JE, Sharpe RG *et al.* (2004) Widespread vertical transmission and associated host sex-ratio distortion within the eukaryotic phylum Microspora. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, **271**, 1783–1789.

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The CLUSTAL_X Windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research*, **24**, 4876–4882.

Thorne JL, Kishino H, Painter IS (1998) Estimating the rate of evolution of the rate of molecular evolution. *Molecular Biology and Evolution*, **15**, 1647–1657.

Waldick RC, Kraus SD, Brown M, White BN (2002) Evaluating the effects of historic bottleneck events: an assessment of

microsatellite variability in the endangered, North Atlantic right whale. *Molecular Ecology*, **11**, 2241–2249.

Watterson GA (1975) On the number of segregating sites in genetical models without recombination. *Theoretical Population Biology*, **7**, 256–276.

Weir BS (1996) *Genetic Data Analysis II: Methods for Discrete Population Genetic Data*. Sinauer Associates, Sunderland, Massachusetts.

Weir BS, Cockerham CC (1984) Estimating *F*-statistics for the analysis of population structure. *Evolution*, **38**, 1358–1370.

Wright S (1951) The genetical structure of populations. *Annals of Eugenics*, **15**, 323–354.

Yang Z (1997) paml: a program package for phylogenetic analysis by maximum likelihood. *CABIOS*, 13, 555–556.

Yang Z (2002) *Phylogenetic Analysis by Maximum Likelihood (paml)*, *Version 3.13a*, September 2002. Distributed by the author. (http://abacus.gene.ucl.ac.uk/software/paml).