MOLECULAR SYSTEMATICS AND POPULATION GENETICS OF WHALE LICE
(AMPHIPODA: CYAMIDAE) LIVING ON GRAY WHALE ISLANDS

by

Christopher Michael Callahan

A Thesis

Presented to
The Faculty of Humboldt State University

In Partial Fulfillment
of the Requirements for the Degree

Master of Arts

In Biological Sciences

May 2008
MOLECULAR SYSTEMATICS AND POPULATION GENETICS OF WHALE LICE
(AMPHIPODA: CYAMIDAE) LIVING ON GRAY WHALE ISLANDS

By

Christopher Michael Callahan

We certify that we have read this study and that it conforms to acceptable standards of scholarly presentation and is fully acceptable, in scope and quality, as a thesis for the degree of Master of Arts.

Brian S. Arbogast, Major Professor

P. Dawn Goley, Committee Member

Edward C. Metz, Committee Member

Jeffrey W. White, Committee Member

Michael R. Mesler, Graduate Coordinator

Chris A. Hopper, Interim Dean
Research, Graduate Studies
ABSTRACT

MOLECULAR SYSTEMATICS AND POPULATION GENETICS OF WHALE LICE (AMPHIPODA: CYAMIDAE) LIVING ON GRAY-WHALE ISLANDS

by

Christopher Michael Callahan

Master of Arts in Biological Sciences

Gray whales (Eschrichtius robustus) are ‘living islands’ to a diverse assemblage of crustacean ectoparasites that includes at least three species of highly specialized whale lice (Cyamus spp.). These lice are obligate parasites that undergo direct development on the gray-whale host, and are dependent on direct physical contact to colonize a new host whale. Given their high degree of morphological specialization and obligate relationship with whales, whale lice might be expected to have a close, long-term evolutionary association with gray whales. Such a relationship can lead to an evolutionary history of the parasite that closely mirrors that of its host and a highly correlated demographic history between the host and its parasites. Here I use a 738 base pair fragment of the mitochondrial DNA (mtDNA) cytochrome c oxidase subunit I gene (COI) gene to: (i) examine the phylogenetic relationships among, and genetic diversity within, gray-whale lice (C. scammoni, C. kessleri, and C. ceti); and (ii) to infer historical demographic patterns within each species of whale louse. Whale lice samples were collected from five different gray-whale hosts. Gray-whale lice exhibited relatively high levels of genetic diversity suggesting large effective population sizes and gene flow among different gray-whale hosts. Each species of louse was phylogenetically distinct and reciprocally monophyletic, indicating congruence between morphological and mtDNA phylogenies. The phylogeny also suggests that collectively these whale lice do not form a monophyletic group, supporting the hypothesis of independent, historical colonizations.
onto the gray-whale host. All louse species exhibited relatively high levels of genetic
diversity. Mismatch distributions for all three gray-whale lice are consistent with long-
term, stable historical population sizes and all three lineages coalesce to approximately
800,000 years before present. The polyphyletic relationships of gray-whale lice provide
three independent replicates for indirectly examining the demographic history of their
host, the gray whale.
ACKNOWLEDGMENTS

First, I would like to thank my advisor Brian Arbogast. Brian has and continues
to be a good friend and mentor, I am grateful for his guidance, patience, and faith in my
abilities. There are numerous others who have helped me along the way. The members
of the Arbogast Lab and Vertebrate Museum have been my surrogate family during my
stay at Humboldt State. Leslie Avalos, Jessica Blois, and Anthony Baker were pivotal in
my formative stages in the molecular lab; Bryan Jennings later helped me to master those
skills. My cohorts, Nick Kerhoulas and Erin Atkin, provided companionship and good
conversation. Tim and Barbara Lawlor provided useful advice along the way; Tim also
graciously edited several iterations of this draft. My committee members, Dawn Goley,
Ed Metz, Jeffrey White, and Michael Mesler, provided useful comments and advice
throughout this project.

In the field I was supported by many people, a few in particular, Michael King,
Aicha Ougzin, and Jeff Jacobsen, were always willing to go out and collect dead whales
with me on a moments notice. Dawn Goley and the numerous members of the Marine
Mammal Education and Research Program provided a community for me to stay
connected with. Andrew Tappy was instrumental in sorting through hundreds of whale
lice samples and spent many hours helping me in lab. Justin Fulkerson and Matthew
Lowe provided assistance in the lab as well as companionship throughout this project.

Whale lice samples were collected from many sources and donated to this study.
John Ososky, Charlie Potter, and Jim Mead at the Smithsonian Institution, National
Museum of Natural History gave me access to their collection and allowed me to collect
whale lice for this study. Jim Dines and John Heyning at the Natural History Museum of
Los Angeles County provided whale lice from their research collection. Laurie Hall at
the Moss Landing Marine Lab collected whale lice and donated them to this project. Jeff
Jacobsen and Jared Wolfe generously donated whale lice from their personal collections.
Jon Seger provided primer sequences and unpublished data that were instrumental in this project.

I would also like to thank the following organizations for funding: the American Society of Mammalogists, grant-in-aid of research; HSU Department of Biological Sciences, master’s grant; HSU Office of Research and Graduate Studies, research and creative activities grant; and the California State University Program for Education and Research in Biotechnology, faculty research grant.

I would like to thank Michael Callahan, Michael Mourer, Justin Fulkerson, and Matthew Lowe; their love, loyalty, and friendship have provided me with the stability needed to survive graduate school. I am forever indebted to them. Finally, I thank my family for accepting me as an individual and allowing me to pursue my childhood dreams of becoming a whale biologist. Their love and support have helped me to realize my potential, and for that I am grateful.
Conclusions and future analyses ................................................................. 47
LITERATURE CITED .......................................................................................... 49
LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Summary of host specificity and dispersal mechanisms for crustacean ectoparasites that live on gray whales</td>
<td>10</td>
</tr>
<tr>
<td>2</td>
<td>Samples used in this study and their corresponding gray-whale host</td>
<td>19</td>
</tr>
<tr>
<td>3</td>
<td>Summary of haplotype distributions for gray-whale lice</td>
<td>20</td>
</tr>
<tr>
<td>4</td>
<td>Likelihood ratio test of molecular clock behavior for each nominal species of gray-whale louse</td>
<td>25</td>
</tr>
<tr>
<td>5</td>
<td>Summary of diversity measures for a 740 bp fragment of the mtDNA COI gene from gray-whale lice sampled in this study</td>
<td>26</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Host specificity of gray-whale lice</td>
<td>11</td>
</tr>
<tr>
<td>2</td>
<td>Morphological phylogeny of cyamid whale lice</td>
<td>12</td>
</tr>
<tr>
<td>3</td>
<td>Neighbor-joining analysis of gray-whale lice monophyly</td>
<td>27</td>
</tr>
<tr>
<td>4</td>
<td>Maximum-parsimony analysis of gray-whale lice monophyly</td>
<td>28</td>
</tr>
<tr>
<td>5</td>
<td>Maximum-likelihood phylogeny of gray-whale lice and relatives</td>
<td>30</td>
</tr>
<tr>
<td>6</td>
<td>Ultrametric gene tree of <em>C. scammoni</em> haplotypes</td>
<td>31</td>
</tr>
<tr>
<td>7</td>
<td>Ultrametric gene tree of <em>C. ceti</em> haplotypes</td>
<td>32</td>
</tr>
<tr>
<td>8</td>
<td>Ultrametric gene tree of <em>C. kessleri</em> haplotypes</td>
<td>33</td>
</tr>
<tr>
<td>9</td>
<td>Mismatch distribution of observed vs. expected pairwise differences between <em>C. scammoni</em> haplotypes</td>
<td>34</td>
</tr>
<tr>
<td>10</td>
<td>Mismatch distribution of observed vs. expected pairwise differences between <em>C. ceti</em> haplotypes</td>
<td>35</td>
</tr>
<tr>
<td>11</td>
<td>Mismatch distribution of observed vs. expected pairwise differences between <em>C. kessleri</em> haplotypes</td>
<td>36</td>
</tr>
</tbody>
</table>
INTRODUCTION

Islands, especially oceanic islands, have played a fundamental role in shaping biogeographic theory. Relative to mainland landmasses, islands are isolated, often relatively small, and have simplified ecosystems. These features combine to make key biological processes, such as immigration, extinction, natural selection, and dispersal, much more tractable on islands than on more complex mainland settings.

From a biogeographical perspective, parasite hosts represent discrete, island-like habitats, and parasites often develop intimate, and even obligate, relationships with a specific host species. Host-parasite systems provide an opportunity to investigate a wide range of evolutionary and biogeographic phenomena, including transmission dynamics of parasites and potential co-speciation events between parasites and their hosts (Kuris et al. 1980, Hafner and Nadler 1988, Page 1994). Recently, a number of researchers have examined the co-evolutionary and biogeographic relationships among parasitic lice (Arthropoda: Insecta) and their vertebrate hosts (Hafner and Nadler 1988, 1990; Demastes and Hafner 1993, Hafner et al. 1994, Paterson et al. 2000, Johnson et al. 2001). For example, based on their respective phylogenies, pocket gophers (Rodentia: Geomyidae) and their chewing lice (Mallophaga: Trichodectidae) appear to have experienced a series of shared speciation events, resulting in host and parasite phylogenies that closely mirror one another in topology (Hafner and Nadler 1988, 1990; Hafner et al. 1994, Page and Hafner 1996). Long-term evolutionary relationships such as these also provide an opportunity to independently infer parameters related to the historical biogeography and evolutionary history of hosts by examining relationships among their respective parasites. For example, Reed et al. (2004) used parasitic lice found on humans to infer the historical biogeography of early Homo spp. where there was an absence of suitable host data from the fossil record. Furthermore, the human mtDNA
genome appears to coalesce relatively recently, which obscures the deep evolutionary history of modern *Homo spp.* This shallow coalescence is likely the result of a population bottleneck event in the recent past (Rogers and Harpending 1992). Despite this bottleneck, Kittler *et al.* (2003) found that human lice, *Pediculus humanus*, have two ancient lineages that precede the coalescence of modern *Homo spp.* DNA by an order of magnitude (Reed *et al.* 2004). In the absence of suitable markers for recovering human history, Reed *et al.* (2004) used *P. humanus* as a proxy to indirectly infer the evolutionary history of modern humans and found that the two lineages of lice co-diverged with ancient *Homo spp.* around 1.8 million years ago.

Host-parasite systems also provide an excellent comparative framework for examining evolutionary phenomena operating at the molecular level, especially when the parasite has an obligate relationship with its host. Rates of molecular evolution in parasites have been shown to be substantially faster than in their hosts, despite a long history of co-speciation between the two groups (Hafner *et al.* 1994, Page and Hafner 1996, Kaliszewska *et al.* 2005). For example, Hafner *et al.* (1994) found that chewing lice living on pocket gophers had a non-synonymous nucleotide substitution rate approximately three times higher than that of their hosts with synonymous substitutions about an order of magnitude higher in lice. Similarly, Kaliszewska *et al.* (2005) found synonymous sequence divergences in parasitic whale lice (Cyamidae) to be 10 times faster than in their whale hosts for homologous markers used in their study. The differing rates in sequence divergence between these hosts and parasites are likely owing to two factors: disparate rates in generation time and larger effective population sizes of the parasites. In many cases, parasites likely complete multiple generations in the lifespan of one host. A recent study of bowhead whales (*Balaena mysticetus*) showed that some individuals lived well past 100 years, with one individual estimated to be over 200 years
old (George et al. 1999). Although detailed information on generation times for whale lice is lacking, they likely complete one or more generations per year (Kaliszewska et al. 2005). Leung (1976) reported that one species of whale louse (Cyamus scammoni) completed its life cycle in eight to nine months. These parasites also have larger effective populations than their hosts; for example adult right whales (Eubalaena spp.) can have over 5,000 whale lice per individual host (Kaliszewska et al. 2005). The above factors make these parasites ideal candidates for genetic studies, because whale lice can accumulate genetic variation at a faster rate than their gray-whale hosts, and larger effective populations are less likely to lose genetic diversity to stochastic events such as genetic drift.

Although intimate ecological and evolutionary relationships have been described between many cetaceans and their parasites (Leung 1967, 1976; Arvy 1982, Rowntree 1983, 1996; Haney 1999, Kaliszewska et al. 2005), details of these relationships remain poorly understood, especially from a biogeographic perspective. As a group, cetaceans host a diverse assemblage of ectoparasites, most notably a variety of crustaceans including barnacles, amphipods, and copepods (Arvy 1982). Gray whales (Eschrichtius robustus) have a diverse ectoparasitic fauna, serving as host for at least three species of whale lice (Amphipoda: Cyamidae) and one species of barnacle, Cryptolepas rhachianecti (Table 1; Hurley and Mohr 1957, Leung 1976, Rice and Wolman 1971, Samaras and Durham 1985). Rice and Wolman (1971) examined 316 adult gray whales and found that over 95% harbored all four species. The only other cetaceans that consistently harbor high diversities of whale lice are right whales (Eubalaena spp.; Leung 1967, Rowntree 1983, 1996; Haney 1999, Margolis et al. 2000). Rowntree (1996) characterized the ecology of three whale-louse species living on northern right whales.
(Eubalaena glacialis), and found that most individuals harbor all three species of whale lice. In contrast to gray and right whales, most other baleen whales typically harbor only a single species of whale louse (Leung 1967, Rowntree 1983, Haney 1999, Margolis et al. 2000).

Cyamid diversity and ecology

All whale lice (cyamids) are thought to be obligate cetacean ectoparasites (i.e., they spend their entire lives living on whales and cannot survive elsewhere). Like other peracarids, cyamids have no free-living aquatic stage in their development, but instead develop within a marsupium and emerge fully developed (Leung 1976, Rowntree 1983, Pfeiffer and Viers 1998). The diet of cyamids also ties them intimately to their whale host(s). Cyamids have specialized mouthparts and digestive systems that allow them to feed on the epidermal skin layer of whales (Keith 1974, Schell et al. 2000). These characteristics of cyamid lice suggest that they must move from one whale host to another primarily by direct whale-to-whale contact.

In gray whales, the two most likely mechanisms of louse transmission are from mother to offspring during nursing and close contact between adults (e.g., mating events; Balbuena and Raga 1991). Both of these events occur annually in the lagoons of Baja California, Mexico, during the boreal winter. Following transmission, whale lice are capable of moving about the body of the host whale and may selectively occupy different regions of the body (Rowntree 1996). Typically, lice aggregate in areas of the body that provide them protection from water currents, such as skin folds (Kasuya and Rice 1970, Briggs and Morejohn 1972, Berzin and Vlasova 1982).

Although data are limited, the distribution of different species of lice appears non-random across the gray-whale host, with each species selectively occupying a different region of the host’s body (Rice and Wolman 1971, Leung 1976, Berzin and Vlasova
1982, Balbuena and Raga 1991, Rowntree 1996). For example, Rowntree (1996) found that three species of whale lice exhibited clear partitioning of habitat (space) on their host, the northern right whale, with each louse species occupying different regions of the body. Rice and Wolman (1971) also found similar differences in the distribution of whale lice on gray whales. The degree to which these observations represent a generalized spatial pattern is not known.

Host-specificity is another important aspect of whale-louse biogeography that is poorly understood. Occasionally a particular species of whale louse can co-occur on more than one host species (Figure 1; Table 1). For example, of the three described species of gray-whale lice, only two (*Cyamus scammoni* and *C. kessleri*) are thought to be host-specific to gray whales. The third louse species, *C. ceti*, occurs also on bowhead whales (Rowntree 1983, Haney 1999). The presence of *C. ceti* on these two host species poses an interesting biogeographic question because the above-mentioned host species are not considered to be each other’s closest relatives. The monotypic gray whale, family Eschrichtiidae, appears to be more closely related to rorqual whales (family Balaenopteridae; *e.g.*, blue whales, *Balaenoptera musculus*, and humpback whales, *Megaptera novaeangliae*) than to the family Balaenidae (bowhead and right whales; Milinkovitch *et al.* 1994, Hasegawa *et al.* 1997, Geisler and Sanders 2003, Rychel *et al.* 2004). The presence of *C. ceti* on these two host species suggests that horizontal transmission (host-switching), rather than strict association by descent (Hafner and Nadler 1990, Brooks and McClellan 1991), has played an important role in shaping the evolutionary and biogeographic histories of at least some lineages of whale lice.

Morphological data from whale lice also suggest that host switching may have been common in the evolutionary history of whale-whale louse relationships. In Haney’s (1999) phylogeny for the family Cyamidae (Figure 2), based on 80 morphological
characters, the monophyly of the genus *Cyamus* was supported; however, the three nominal species of gray-whale lice did not collectively form a monophyletic group. This relationship is similar to that seen in right whales (*Eubalaena spp.*) wherein the three nominal species of right-whale lice do not form a monophyletic relationship to each other (Haney 1999). The polyphyletic relationship of these taxa suggests that colonization of whale lice onto gray whales occurred via three separate evolutionary events.

*Population histories of gray whales*

Historically gray whales occurred in the north Atlantic (NA), western north Pacific (WNP), and eastern north Pacific (ENP). The NA gray whale went extinct in modern times (Rice and Wolman 1971, Mead and Mitchell 1984), and the WNP gray whale population has been reduced to approximately 100 individuals (Bradford *et al.* 2006). The ENP gray whale, once listed as an endangered species, has recovered with current census estimates in the range of 18,000 to 29,000 individuals (Rugh *et al.* 2005). Recent genetic analysis revealed little gene flow between the WNP and ENP gray whale populations (LeDuc *et al.* 2002).

Despite the lack of gene flow between ENP and WNP gray whale populations, there appear to be no restrictions to gene flow and little population structure within ENP gray whales (Steeves *et al.* 2001). Gray whales form loose feeding aggregations from northern California to the Arctic waters around Alaska, and some whales showed strong site fidelity to feeding grounds (Steeves *et al.* 2001, Calambokidis *et al.* 2002). In a ‘resident’ feeding group in Clayoquot Sound, British Columbia, 35 to 50 animals showed strong fidelity to this site; however, analysis of mtDNA showed lack of segregation between resident and non-resident gray whale haplotypes (Steeves *et al.* 2001).
multi-locus study using microsatellite and DNA sequence data of ENP gray whales, Alter et al. (2007) found no evidence of population sub-structure within the ENP, suggesting one large breeding population, which is also in agreement with mtDNA data (Swartz et al. 2006).

ENP gray whales migrate annually to breeding grounds in Baja California from sub-arctic waters of the northern seas (where they feed from late spring through early fall). The breeding season is the only time of year that gray whales aggregate in large social groups (Rice and Wolman 1971). Social aggregations facilitate ectoparasite dispersal, especially for gray-whale barnacles, which are endemic to the gray whale. These barnacles release larvae seasonally to coincide with the breeding events of the gray whale, which facilitates settlement of larvae onto a suitable host. The first infestations of whale lice onto a neonate-whale comes from intimate contact with its mother, and presumably young whales will not accumulate further infestations until they mix with the breeding population.

Coalescent approaches to estimating evolutionary history

Analysis of genetic data within a coalescent or genealogical framework can yield reliable estimates of demographic parameters. This type of approach traces alleles (haplotypes) within a population to their most recent common ancestor and can be used to date speciation events, estimate historical and extant effective population sizes, measure genetic diversity within and among populations, and determine migration rates between populations (Slatkin 1989, Beerli and Felsenstein 2001, Roman and Palumbi 2003, Beerli 2006). These data can provide insight into the population structure, spatial ecology, and long-term evolutionary potential of gray-whale lice.

The disparity between generation times and effective population sizes of whale hosts and whale lice make past evolutionary events more tractable in whale lice and thus
ideal for inferring the historical demographies of the host (Kaliszewska et al. 2005). For example, Kaliszewska et al. (2005) used whale lice to infer the population histories and speciational events of right whales in the northern and southern hemispheres. Coalescent estimates for gray-whale lice can also be used to infer the historical demography of its host, the gray whale, because of their long-term association. Using a multi-locus (nuclear and mitochondrial DNA) coalescent approach, Alter et al. (2007) estimated historical abundances of gray whales to be three to five times greater than estimates of present-day numbers. Gray-whale lice provide three independent replicates, essentially a multi-marker approach for measuring gray whale history and can be compared to the results of Alter et al. (2007).

**Objectives**

The specific goals of this study were to use molecular phylogenetic and coalescent approaches to describe the evolutionary relationships of whale lice living on gray whales and to investigate levels of genetic diversity within each nominal species of whale louse. Here I use sequence variation from the mitochondrial DNA (mtDNA) cytochrome c oxidase subunit I gene (COI) to address the following questions: (i) Does each nominal species of gray-whale louse represent a distinct evolutionary lineage? (ii) If gray-whale lice form distinct evolutionary lineages, do they collectively form a monophyletic group? If so, this would suggest that these lineages diversified following a single colonization of the gray whale. If not (i.e., if they are polyphyletic), this would support the morphological data (i.e., Haney 1999) in suggesting that the extant cyamid fauna of gray whales is the result of multiple independent colonization events. (iii) Do distinct evolutionary lineages of whale lice found on the gray whale have population histories that are similar to one another and to other congeneric lice living on other host
whales as reported in the literature (e.g., Kaliszewska et al. 2005)? (iv) Do gray-whale lice have similar demographic histories to their host, the gray whale? For example, coalescent estimates of gray-whale lice diversity can be used to infer historical abundance estimates for gray whales, which are currently estimated to be higher than present-day numbers (Alter et al. 2007).
Table 1. Summary of host specificity and dispersal mechanisms for crustacean ectoparasites that live on gray whales.

<table>
<thead>
<tr>
<th>Ectoparasites</th>
<th>Cetacean Hosts</th>
<th>Free Living Stage?</th>
<th>Dispersal Mechanism</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Barnacles</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Cryptolepas rhachianecti</em></td>
<td>Gray Whale</td>
<td>Yes</td>
<td>Planktonic Larvae</td>
</tr>
<tr>
<td><strong>Whale Lice</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Cyamus scammoni</em></td>
<td>Gray Whale</td>
<td>No</td>
<td>Direct Contact Between Hosts</td>
</tr>
<tr>
<td><em>Cyamus kessleri</em></td>
<td>Gray Whale</td>
<td>No</td>
<td>Direct Contact Between Hosts</td>
</tr>
<tr>
<td><em>Cyamus ceti</em></td>
<td>Gray Whale</td>
<td>No</td>
<td>Direct Contact Between Hosts</td>
</tr>
<tr>
<td><em>Cyamus ceti</em></td>
<td>Bowhead Whale?</td>
<td>No</td>
<td>Direct Contact Between Hosts</td>
</tr>
</tbody>
</table>
Figure 1. (A) Aggregation of *Cyamus scammoni* and *C. ceti* whale lice on the anterior dorsal region of a gray whale. Lice are clustered around gray-whale barnacles. (B) Host specificity of gray-whale lice. *C. kessleri* and *C. scammoni* are host specific to gray whales, whereas *C. ceti* co-occurs on bowhead whales, *Balaena mysticetus*. (C) Ventral view of female (left) and male (right) *C. ceti* (note that adult females are characterized by the presence of a marsupial brood pouch shown by white arrow). Whale louse images adapted from Rowntree (1983); whale images adapted from the American Cetacean Society.
Figure 2. Phylogenetic hypothesis for whale lice (*Cyamus*) based on 80 morphological characteristics (Haney 1999). Gray-whale lice (indicated by asterisks) form a polyphyletic relationship.
MATERIALS AND METHODS

Specimen sampling and DNA extractions

Whale lice were sampled directly from stranded, dead gray-whale carcasses and from museum collections. A total of five gray-whale hosts were examined in this study (Table 2). The adult female (VM 2534) was collected prior to the design of this project and thus was not sampled systematically to collect all nominal species of gray-whale lice. Similarly, the neonate calf (MLML 127802) was collected by staff at Moss Landing Marine Lab and was also not systematically sampled. The remaining three host whales were sampled in this study and attempts were made to collect every louse specimen observed. *Cyamus scammoni* was collected from all five host animals; *C. ceti* was collected from four host whales; and *C. kessleri* was only collected from two host whales (Table 2). Specimens were either frozen (-20°C) or stored in 70% ethanol. Lice were inventoried and sorted into lots based on the location from which they were collected on the gray-whale host. Each individual selected for the study was assigned a voucher number and was stored separately in ethanol.

I attempted to extract amplifiable DNA from 15 individual louse specimens of each nominal species from each host whale using the QIAGEN DNeasy tissue extraction kit (Qiagen Inc., Valencia, California) following the protocol for animal tissue. Some host whales lacked all three nominal species; others harbored fewer than 15 animals (Table 2). In these instances I extracted DNA from as many lice as were available.
Some individuals did not produce amplifiable DNA and museum specimens preserved in formalin were unusable in this study. The net number of usable extractions included 53 for *C. scammoni*, 26 for *C. ceti*, and 29 for *C. kessleri* (Table 2).

Whale lice from other whale species were also obtained to evaluate the interspecific relationships of gray-whale lice within a phylogenetic framework. Congeneric whale lice were sampled from right and humpback whales. Outgroup taxa included whale lice from the genera *Isocyamus* and *Neocyamus* and a caprellid skeleton shrimp, *Caprella sp.* (caprellid amphipods are thought to be the sister group of whale lice; Haney 1999). These sequences were obtained from J. Seger, except for humpback and pilot whale (*Globicephala macrorhynchus*) lice, which were collected specifically for this study.

**DNA amplification and sequencing**

Polymerase chain reaction (PCR) was used to amplify a 750 bp fragment of the COI coding gene using the primers Patcy (5’-ACT AGC ACA TTT ATC TGT CAC ATT A-3’; reverse) and Jerky (5’-TAC CAA CAT TTA TTC TGR TTT TTY GG-3’; forward) as described by Kaliszewska *et al.* (2005) and PCR Master Mix (Promega Corp., Madison, WI). PCR products were amplified using the following profile: an initial DNA denaturation of 180-s at 94°C, followed by 5 repeating cycles of 30-s denaturations at 94°C, a 90-s annealing step at 45°C, and a 60-s extension step at 72°C. The annealing temperature was then raised to 51°C for an additional 35 cycles followed by a final extension of 72°C for 5-m. This profile generally worked well for all amplifications, including samples with low-quantity template DNA. All amplifications were run on a PTC-100 thermocycler (MJ Research, Waltham, MA) in 25 µl reactions.
Amplified PCR products were visualized using gel-electrophoresis with a standard 1% agarose gel and 1% Tris-acetate-EDTA (TAE) buffer solution (Murphy et al. 1996). PCR products were purified using the QIAquick® PCR Purification Kit (Qiagen Inc., Valencia, California) and sequenced on an ABI Prism 3100 capillary sequencer at the Microchemical Core Facility at San Diego State University (San Diego, California). Most samples were sequenced with the reverse primer; and samples with ambiguities were re-sequenced with the forward primer to verify correct nucleotide sequence. Mitochondrial DNA sequences from each sample were visually examined for errors and edited using SEQUENCHER (software package ver. 4.2; Gene Codes Corp. Ann Arbor, Michigan). Sequences were aligned with CLUSTALX (Thompson et al. 1997). Once edited, most haplotypes were reduced to 740 bp. Final data matrices were further reduced to 738 bp to incorporate in-frame coding sequences (i.e., no partial codons).

Phylogenetic analyses and tests of monophyly

Aligned sequences were used to examine the phylogenetic relationships among gray-whale lice and relatives using a combination of maximum parsimony (MP), neighbor-joining (NJ) distance, and model-based methods. Maximum-likelihood (ML), MP, and NJ phylogenies were inferred using PAUP* (ver. 4.0b10; Swofford 2003), and a Bayesian (BA) phylogeny was inferred using MrBAYES (ver. 3.2-cvs; Huelsenbeck and Ronquist 2001, Huelsenbeck et al. 2001). MACCLADE (ver. 4; Maddison and Maddison 2000) was used to construct a constrained topology to test for gray-whale lice monophyly.

A NJ analysis with 1000 bootstrap replicates (Felsenstein 1985) under maximum-likelihood distances and a MP analysis with 1000 fast, stepwise-additions with 50%
majority consensus were performed for gray-whale lice to test for reciprocal monophyly among morphospecies. These analyses also were used to identify unique haplotypes. Shared haplotypes were re-examined using SEQUENCHER to visually confirm congruence among like-strands and were collapsed for ML and BA analyses to reduce computation time.

A reduced data matrix with four unique haplotypes from each gray-whale louse species along with homologous sequences from congeneric whale lice and outgroup taxa was then used to construct a multi-species gene tree. The ML data matrix was analyzed under the best-fit substitution model as selected by MODELTEST (ver. 3.5; Posada and Crandall 1998) using the Akaike information criterion (AIC). A heuristic tree search with random addition of taxa and tree-bissection-reconnection (TBR) branch swapping estimated the best –log likelihood (−ln L) tree score. Nodal support for the ML tree was estimated by generating 100 maximum-likelihood non-parametric bootstrap replicates; scores of 70% or greater were considered strong support (Felsenstein 1985, Hillis and Bull 1993).

A likelihood ratio test (LRT; Huelsenbeck and Rannala 1997) was used to test for a molecular clock. The outgroup taxa were removed from the best tree and re-scored in PAUP* with and without a molecular clock enforced to produce −ln L tree scores. In a LRT, P-values <0.05, using a $\chi^2$ contingency table, indicate that the tree topology is significantly worse than the best −ln L tree score.

The best tree generated by the ML analysis was edited in MACCLADE to constrain gray-whale lice to form a monophyletic clade. All other ingroup congeners were collapsed into a sister polytomy with outgroup taxa basal to the ingroup. The constrained tree was then evaluated in PAUP* with the original data matrix (from the best tree) under maximum-likelihood criteria to generate a −ln L tree score. A LRT was used to compare both tree scores.
The BA phylogeny was evaluated under the best-fit model as selected by MrMODELTEST (ver. 2.2; Nylander 2004), also using AIC. The BA consisted of two runs of $1 \times 10^6$ Markov chain Monte Carlo generations. Trees were sampled every 100 generations for a total of 10,000 trees. Inspection of likelihood scores indicated that both runs had converged by 1,000,000 generations and the first 2,500 trees were discarded as burn-in. Posterior probabilities were estimated after maximum-likelihood scores had stabilized.

**Demographic histories and diversity measures**

Sequence variations among the COI gene haplotypes were also used to estimate the intraspecific genetic diversity within each gray-whale louse species. Collapsed haplotype data matrices were generated for each nominal species. Unrooted single-species gene trees were constructed for each gray-whale louse species under maximum-likelihood criteria and various models as selected by MODELTEST in PAUP*. Individual louse samples were color coded to indicate from which whale host they were collected (Table 2). Each terminal branch represents a unique haplotype (allele) and each tree was visually inspected for reciprocal monophyly of haplotypes collected from each gray-whale host (i.e., does each whale host harbor a distinct population of whale lice?). The best ML tree for each species was then rescored under a molecular clock and evaluated with a LRT. Clock trees were converted to ultrametric phylograms and calibrated to estimate divergence time.

Standard diversity measures such as haplotype diversity ($H$), number of polymorphic sites ($S$), pairwise differences per site ($\pi$) diversity (Tajima 1983, Nei 1987), and divergence of haplotype sequences ($\Pi$) were estimated using the software program ARLEQUIN (ver. 2.00; Schneider et al. 2000). Pairwise divergences were then
evaluated between each nominal louse species from values estimated by ARLEQUIN. Mismatch distributions of observed vs. expected pairwise differences between haplotypes were inspected for signs of recent population expansions. A unimodal distribution is consistent with a recent population expansion, whereas a multi-modal distribution is more consistent with a stable population in equilibrium (Rogers and Harpending 1992, Slatkin and Hudson 1991). The effective population size ($N_e$) from each nominal species of whale louse was estimated with the mutation rate ($\mu$) of 0.0175 substitutions per site per million years and theta ($\theta$) values generated by the software program MIGRATE (ver. 2.4.3; Beerli 1997, Beerli and Felsenstein 1999; 2001) under Bayesian criteria. Parameter values were sampled $1 \times 10^8$ steps with updates stored every $1 \times 10^6$ steps for a total of 100 recorded updates. Nucleotide base frequencies and the transition/transversion (Ti/Tv) ratio were estimated in MODELTEST under AIC and hierarchical LRT settings; other parameters were left at default settings.
Table 2. Samples used in this study and their corresponding gray-whale host.

<table>
<thead>
<tr>
<th>Field Number</th>
<th>Age Class</th>
<th>Collection Site</th>
<th>Cyamus scammoni</th>
<th>Cyamus ceti</th>
<th>Cyamus kessleri</th>
</tr>
</thead>
<tbody>
<tr>
<td>VM 2534</td>
<td>adult female</td>
<td>Humboldt Co., CA</td>
<td>10</td>
<td>8</td>
<td>2</td>
</tr>
<tr>
<td>VM 2584</td>
<td>neonate female</td>
<td>Humboldt Co., CA</td>
<td>57</td>
<td>14</td>
<td>58</td>
</tr>
<tr>
<td>VM 2645</td>
<td>sub-adult male</td>
<td>Humboldt Co., CA</td>
<td>57</td>
<td>14</td>
<td>35</td>
</tr>
<tr>
<td>MLML 127802</td>
<td>neonate unknown</td>
<td>Monterey Co., CA</td>
<td>15</td>
<td>12</td>
<td>--</td>
</tr>
<tr>
<td>VM 2664</td>
<td>neonate female</td>
<td>Humboldt Co., CA</td>
<td>14</td>
<td>5</td>
<td>90</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td>153</td>
<td>53</td>
<td>185</td>
</tr>
</tbody>
</table>

Cyamus scammoni: 8
Cyamus ceti: 1
Cyamus kessleri:
---

Cyamus scammoni: 58
Cyamus ceti: 11
Cyamus kessleri:
---

Cyamus scammoni: 35
Cyamus ceti: 8
Cyamus kessleri:
---

Cyamus scammoni: 12
Cyamus ceti: --
Cyamus kessleri:
---

Cyamus scammoni: 90
Cyamus ceti: 6
Cyamus kessleri:
---

Cyamus scammoni: 135
Cyamus ceti: 29
Cyamus kessleri:
---

Total:
Cyamus scammoni: 185
Cyamus ceti: 26
Cyamus kessleri: 29
Table 3. Summary of haplotype distributions for each nominal species of gray-whale louse. Each louse sampled in this study was given a unique voucher number: *C. scammoni*, SCA-#; *C. ceti*, CET-#; and *C. kessleri*, KES-#. Host whales correspond to

Note that *C. scammoni* haplotypes 2, 4, 5, and 8 are shared among more than one host whale; all other haplotypes were private to a specific whale host.

<table>
<thead>
<tr>
<th>Louse Species</th>
<th>Haplotype #</th>
<th>Lice Shared Among Haplotype</th>
<th>Host Whale(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Cyamus scammoni</em></td>
<td>H1</td>
<td>SCA-2, SCA-6, SCA-31</td>
<td>VM 2584</td>
</tr>
<tr>
<td></td>
<td>H2</td>
<td>SCA-8, SCA-23</td>
<td>VM 2584</td>
</tr>
<tr>
<td></td>
<td>H3</td>
<td>SCA-1, SCA-25</td>
<td>VM 2534, VM 2645</td>
</tr>
<tr>
<td></td>
<td>H4</td>
<td>SCA-15, SCA-29</td>
<td>VM 2584, VM 2645</td>
</tr>
<tr>
<td></td>
<td>H5</td>
<td>SCA-14, SCA-24</td>
<td>VM 2645</td>
</tr>
<tr>
<td></td>
<td>H6</td>
<td>SCA-36, SCA-37</td>
<td>VM 2645</td>
</tr>
<tr>
<td></td>
<td>H7</td>
<td>SCA-12, SCA-39</td>
<td>VM 2645, MLML 127802</td>
</tr>
<tr>
<td></td>
<td>H8</td>
<td>SCA-42, SCA-49</td>
<td>MLML 127802</td>
</tr>
<tr>
<td></td>
<td>H9</td>
<td>SCA-43, SCA-52</td>
<td>MLML 127802</td>
</tr>
<tr>
<td></td>
<td>H10</td>
<td>SCA-40, SCA-53</td>
<td>MLML 127802</td>
</tr>
<tr>
<td></td>
<td>H11</td>
<td>SCA-59, SCA-60</td>
<td>VM 2664</td>
</tr>
<tr>
<td><em>Cyamus ceti</em></td>
<td>H1</td>
<td>CET-22, CET-24, CET-34, CET-36</td>
<td>VM 2645</td>
</tr>
<tr>
<td></td>
<td>H2</td>
<td>CET-17, CET-32</td>
<td>VM 2645</td>
</tr>
<tr>
<td></td>
<td>H3</td>
<td>CET-3, CET-8</td>
<td>VM 2584</td>
</tr>
<tr>
<td><em>Cyamus kessleri</em></td>
<td>H1</td>
<td>KES-5, KES-11, KES-15</td>
<td>VM 2584</td>
</tr>
<tr>
<td></td>
<td>H2</td>
<td>KES-6, KES-9, KES-10, KES-14, KES-17</td>
<td>VM 2584</td>
</tr>
<tr>
<td></td>
<td>H3</td>
<td>KES-4, KES-13, KES-16</td>
<td>VM 2584</td>
</tr>
<tr>
<td></td>
<td>H4</td>
<td>KES-1, KES-2, KES-3</td>
<td>VM 2584</td>
</tr>
<tr>
<td></td>
<td>H5</td>
<td>KES-19, KES-22, KES-24, KES-26, KES-29</td>
<td>VM 2664</td>
</tr>
<tr>
<td></td>
<td>H6</td>
<td>KES-27, KES-30, KES-31, KES-32, KES-33</td>
<td>VM 2664</td>
</tr>
<tr>
<td></td>
<td>H7</td>
<td>KES-28, KES-34</td>
<td>VM 2664</td>
</tr>
</tbody>
</table>
RESULTS

Phylogenetic analyses and tests of monophyly

Phylogenetic relationships among gray-whale lice show a high degree of consistency between the different methods used in this study. Both tests of monophyly (NJ and MP trees) exhibited identical topologies with each morphospecies of whale louse forming reciprocally monophyletic groupings (Figures 3 - 4). Both trees were constructed with a data matrix consisting of only unique gray-whale louse haplotypes (n=64; Table 3) and excluded outgroups. The NJ tree was constructed using maximum-likelihood distances estimated using the TrN + I + Γ model of substitution, as selected by MODELTEST. The proportion of invariable sites (I) was 0.636 and alpha, the shape of the gamma distribution, was 0.931. MP and NJ tree topologies were congruent, with the monophyly of each nominal gray-whale louse species receiving 100% bootstrap support (Figures 3 - 4).

The maximum-likelihood and Bayesian multi-species gene trees, which included multiple outgroup taxa, were also congruent and supported the monophyly of each gray-whale louse species; however, collectively, the three species did not form a monophyletic group (i.e., they were not each others closest relatives). The model selected for this broader phylogenetic analysis was HKY85 + I + Γ (transition/transversion ratio, Ti/Tv=4.239; I=0.553; alpha=1.259). The best likelihood tree (-ln L 5550.194) was bootstrapped with 100 ML replicates and nodal support ≥ 70% is shown below the major clades in Figure 5; branch lengths > 0.01 for critical nodes are shown above branches. The Bayesian analysis recovered the same topology and sister-clade relationships as the ML tree, so posterior probability scores were added to Figure 5; scores of >0.9 are considered strong support and are shown below critical branches next
to the bootstrap scores. The best ML tree was then constrained in MACCLADE to force the three nominal species of gray-whale lice to form a monophyletic relationship with each other. The constrained topology was then rescored in PAUP*. A LRT (\(-\ln L 556.295, \Delta \ln L 12.202, \chi^2=36.415, df=24, P=0.9776\)) indicated that the constrained tree was not a significantly worse fit to the data than the best ML tree. The best ML tree was also tested for clock-like behavior. The best likelihood tree, constrained to ingroup taxa only (i.e., *Cyamus* spp.), was rescored with and without a molecular clock enforced (\(-\ln L 3848.012 \text{ and } –\ln L 3827.127\), respectively). A LRT (\(\Delta \ln L 41.770, \chi^2=30.14, df=19, P=0.0019\)) rejected clock-like behavior for these data.

**Demographic histories and diversity measures**

Single-species gene trees (Figures 6–8) failed to show significant population structure within each nominal species of gray-whale louse (i.e., haplotypes from each host-whale did not form reciprocally monophyletic groups). Each single-species gene tree was estimated under the best-fit substitution model as selected by MODELTEST (*C. scammoni*, HKY85 + I + \(\Gamma\); *C. ceti*, HKY85 + I; and *C. kessleri*, TrN + I). The best likelihood tree for each species was then rescored with a molecular clock enforced; the LRT’s failed to reject clock-like behavior for each nominal gray-whale louse species (Table 4).

All species of gray-whale lice had high levels of haplotype diversity (Table 5). The number of haplotypes (\(h\)) varied among gray-whale lice species, with *C. ceti*, \(n=26\), having the most haplotypes (\(h=21\)) and *C. kessleri*, \(n=29\), having the fewest (\(h=10\)). The most intensively sampled species, *C. scammoni*, \(n=53\), had intermediate haplotype diversity (\(h=33\)). Haplotypes of the three species of whale lice also exhibited differences
in how they were distributed among host whales. The species with the most haplotypes, *C. scammoni*, had 12 haplotypes that were shared among multiple whale lice. Furthermore, four of those shared haplotypes were detected on more than one host whale (Table 3). The most abundant haplotype, SCA-H2, was detected on three different host whales. All other haplotypes for *C. scammoni* were restricted to a single host whale. *C. ceti* had three haplotypes found on multiple host whales; all haplotypes for *C. kessleri* were private to individual host whales (Table 3).

All species had high levels of nucleotide diversity (*C. scammoni*, $H=0.986$; *C. ceti*, $H=0.975$; and *C. kessleri*, $H=0.901$). The number of segregating sites ($S$) within species was similar for *C. scammoni* and *C. ceti* ($S=43$ and $S=44$, respectively), whereas *C. kessleri* had half that number ($S=22$). The population mutation rate per site ($\theta_s$) and mean number of pairwise differences per sequence ($\Pi$) and per site ($\pi$) were similar for all three species (Table 5).

Corrected values of sequence divergence for each nominal species of gray-whale louse were estimated by summing the number of substitutions per site between haplotypes with the longest branch lengths to the node where they coalesced. Values for each species were estimated independently based on their best ML gene tree. *C. scammoni* was estimated to have 0.011-0.013 substitutions per site (equivalent to 1.1-1.3% corrected sequence divergence). Corrected sequence divergences for *C. ceti* (0.014-0.017 substitutions per site) and *C. kessleri* (0.014-0.015 substitutions per site) were 1.4-1.7% and 1.4-1.5%, respectively. Estimated rates of evolution for mtDNA coding regions in arthropods typically range from 1.5-2.0% per million years (Knowlton *et al.* 1993). The mean average rate of evolution (1.75%) was used to approximate coalescence of haplotypes for each single-species gene tree. All three trees coalesce to
approximately 0.668-0.954 million years before present (\textit{C. scammoni}, 0.668-0.748\textit{my}; \textit{C. ceti}, 0.806-0.954\textit{my}; and \textit{C. kessleri}, 0.834-0.874\textit{my}). Single-species trees constrained to follow clock-like behavior (Figures 6 - 8) were scaled with the mean value of these estimates in order to date divergence time within each species of gray-whale louse.

Mismatch distributions for all three nominal species of gray-whale louse were ragged and multi-modal, a pattern consistent with long-term, stable population sizes (Figures 9 - 11). However, statistically, these were not significantly different than the distribution predicted under an exponential growth model, that is, a unimodal distribution (\textit{C. scammoni}, \(P=0.639\); \textit{C. ceti}, \(P=0.630\); and \textit{C. kessleri}, \(P=0.300\)). Substitution models used to calculated theta were the same as those in the molecular-clock analyses. The TrN model selected for \textit{C. kessleri} did not specify a Ti/Tv ratio; therefore, the HKY85 + I model was used instead. Estimated values (\(\theta=4N_{e}\mu\)) for all three species were: \textit{C. scammoni}, \(\theta=0.514, N_{e}=7.3\times10^6\); \textit{C. ceti}, \(\theta=0.356, N_{e}=5.1\times10^6\); \textit{C. kessleri}, \(\theta=0.078, N_{e}=1.1\times10^6\).
Table 4. Likelihood ratio test of molecular clock behavior for each nominal species of gray-whale louse. Trees were midpoint rooted and scored under maximum-likelihood criteria in PAUP* under various substitution models as selected by MODELTEST.

<table>
<thead>
<tr>
<th>Species</th>
<th>no clock -ln L</th>
<th>-ln L</th>
<th>Δ in L</th>
<th>$\chi^2$</th>
<th>d.f.</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>C. scammoni</td>
<td>1373.67</td>
<td>1393.93</td>
<td>40.51</td>
<td>44.98</td>
<td>31</td>
<td>0.118</td>
</tr>
<tr>
<td>C. ceti</td>
<td>1345.34</td>
<td>1354.45</td>
<td>18.23</td>
<td>30.14</td>
<td>19</td>
<td>0.507</td>
</tr>
<tr>
<td>C. kessleri</td>
<td>1145.54</td>
<td>1152.85</td>
<td>14.62</td>
<td>15.51</td>
<td>8</td>
<td>0.067</td>
</tr>
</tbody>
</table>
Table 5. Summary of diversity measures for a 740 bp fragment of the mtDNA COI gene from gray-whale lice sampled in this study.

<table>
<thead>
<tr>
<th>Louse Species (Whale Host)</th>
<th>n</th>
<th>h</th>
<th>S</th>
<th>(\theta_s) (SD)</th>
<th>H (SD)</th>
<th>(\Pi) (SD)</th>
<th>(\pi) (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Cyamus scammoni</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(VM 2534)</td>
<td>53</td>
<td>33</td>
<td>43</td>
<td>0.013</td>
<td>0.004</td>
<td>0.986</td>
<td>0.006</td>
</tr>
<tr>
<td>(VM 2584)</td>
<td>8</td>
<td>6</td>
<td>11</td>
<td>0.006</td>
<td>0.003</td>
<td>0.892</td>
<td>0.111</td>
</tr>
<tr>
<td>(VM 2645)</td>
<td>14</td>
<td>9</td>
<td>16</td>
<td>0.007</td>
<td>0.003</td>
<td>0.923</td>
<td>0.500</td>
</tr>
<tr>
<td>(MLML 127802)</td>
<td>12</td>
<td>9</td>
<td>23</td>
<td>0.010</td>
<td>0.004</td>
<td>0.954</td>
<td>0.046</td>
</tr>
<tr>
<td>(VM 2664)</td>
<td>5</td>
<td>4</td>
<td>15</td>
<td>0.010</td>
<td>0.005</td>
<td>0.900</td>
<td>0.161</td>
</tr>
<tr>
<td><em>Cyamus ceti</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(VM 2534)</td>
<td>26</td>
<td>21</td>
<td>44</td>
<td>0.016</td>
<td>0.005</td>
<td>0.975</td>
<td>0.020</td>
</tr>
<tr>
<td>(VM 2584)</td>
<td>11</td>
<td>10</td>
<td>33</td>
<td>0.015</td>
<td>0.006</td>
<td>0.981</td>
<td>0.046</td>
</tr>
<tr>
<td>(VM 2645)</td>
<td>8</td>
<td>4</td>
<td>9</td>
<td>0.005</td>
<td>0.002</td>
<td>0.750</td>
<td>0.139</td>
</tr>
<tr>
<td>(VM 2664)</td>
<td>6</td>
<td>6</td>
<td>16</td>
<td>0.009</td>
<td>0.005</td>
<td>1.000</td>
<td>0.096</td>
</tr>
<tr>
<td><em>Cyamus kessleri</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(VM 2584)</td>
<td>29</td>
<td>10</td>
<td>22</td>
<td>0.008</td>
<td>0.003</td>
<td>0.901</td>
<td>0.024</td>
</tr>
<tr>
<td>(VM 2664)</td>
<td>13</td>
<td>4</td>
<td>13</td>
<td>0.006</td>
<td>0.003</td>
<td>0.730</td>
<td>0.078</td>
</tr>
</tbody>
</table>

\(n\), number of sequences per species (or from individual host); \(h\), number of unique haplotypes; \(S\), number of polymorphic sites; \(\theta_s\), population mutation rate per site; \(H\), haplotype diversity; \(\Pi\), mean number of pairwise differences (per sequence); \(\pi\), mean number of pairwise differences (per site); SD, standard deviation.
Figure 3. Multi-species gene tree of gray-whale lice under a neighbor-joining (NJ) analysis using maximum-likelihood corrected genetic distances. Best-fit model TrN + I + Γ estimated from MODELTEST. Bootstrap scores were estimated from 1000 NJ replicates in PAUP* (ver. 4.0b10); scores for major clades are shown above branches. Each terminal branch represents a distinct haplotype and all three louse species coalesce into reciprocally monophyletic clades (black terminal branches, *C. scammoni*; blue terminal branches, *C. ceti*; and red terminal branches *C. kessleri*).
Figure 4. Multi-species gene tree of gray-whale lice under a parsimony analysis. Bootstrap scores estimated under 1000 fast stepwise-additions with 50% majority consensus enforced in PAUP* (ver. 4.0b10); scores are shown above major clades. Each terminal branch represents a distinct haplotype and all three louse species coalesce into reciprocally monophyletic clades (black terminal branches, *C. scammoni*; blue terminal branches, *C. ceti*; and red terminal branches *C. kessleri*).
Figure 5. Multi-species gene tree of gray whale-lice and relatives. Gray-whale lice are shown in blue and are represented by four samples from each nominal species. All other whale lice are represented by one individual each. Cyamus erraticus, C. gracilis, and C. ovalis sequences were obtained from Genbank (Kaliszewska et al. 2005) and represent individuals from the North Atlantic, NA; North Pacific, NP; and Southern Ocean, SO. Outgroup lice (Neocyamus sp. and Isocyamus sp.) and caprellid shrimp sequences were obtained from J. Seger (unpublished) except for Isocyamus globicipitis and C. boopis, which were collected for this study. Sequences are 738 bp fragments of the mitochondrial DNA, cytochrome c oxidase sub unit I gene. The maximum-likelihood tree was inferred in PAUP* (ver. 4.0b10) under the best fit model (HKY85 + I + Γ) as selected by MODELTEST (ver. 3.7). Branch lengths for critical nodes greater than 0.01 substitutions per site are shown above branches. Bayesian posterior probabilities estimated under the best fit model (GTR + I + Γ) as selected by MrMODELTEST (ver. 2.2) are shown immediately below branches. Maximum-likelihood bootstrap values (n=100 replicates) greater than 70% are shown next to Bayesian posterior probabilities.
Figure 6. The single-species gene tree for *C. scammoni* estimated with a molecular clock enforced under maximum-likelihood criteria in PAUP* (ver. 4.0b10). Best likelihood tree (-ln L 1373.6747) was constrained to test for a molecular clock. The clock tree (-ln L 1393.9285) was not significantly worse than the best tree based on a likelihood ratio test (Table 4). The tree was calibrated using a divergence rate of 1.75% per million years estimated from the average divergence rate of 1.5-2.0% in arthropods (Knowlton *et al.* 1993); all lineages coalesce around 0.708 (range 0.543-0.873) million years ago. Each terminal branch represents a distinct haplotype (*n*=33) and its color corresponds to the gray-whale host from which it was sampled: red, adult female gray whale (VM2534); purple, neonate female (VM2584); blue, sub-adult male (VM2645); green, neonate unknown sex (MLML127802); orange, neonate female (VM2664); black terminal branches correspond to haplotypes that were observed on multiple host whales and the numbers next to branches are color-coded to correspond with host whale.
Figure 7. The single-species gene tree for *C. ceti* estimated with a molecular clock enforced under maximum-likelihood criteria in PAUP* (ver. 4.0b10). Best likelihood tree (-ln L 1345.3383) was constrained to test for a molecular clock. The clock tree (-ln L 1345.3383) was not significantly worse than the best tree based on a likelihood ratio test (Table 5). All lineages coalesce to 0.880 (range 0.806-0.954) million years ago. Tree calibration and colored branches are the same as in Figure 6; however, no shared haplotypes were observed on more than one host whale (Table 3). No lice were collected from the neonate gray whale MLML127802.
Figure 8. The single-species gene tree, for *C. kessleri* estimated with a molecular clock enforced under maximum-likelihood criteria in PAUP* (ver. 4.0b10). Best likelihood tree (-ln L 1145.5449) was constrained to test for a molecular clock. The clock tree (-ln L 1152.8524) was not significantly worse than the best tree based on a likelihood ratio test (Table 5). All lineages coalesce to 0.850 (range 0.830-0.850) million years ago. Tree calibration and colored branches are the same as in Figure 6; however, no shared haplotypes were observed on more than one host whale (Table 3). *C. kessleri* samples were collected from only two host animals, VM 2584 and VM 2664.
Figure 9. Mismatch distribution of observed vs. expected pairwise differences between haplotypes for *C. scammoni* (*n*=53). Although not significantly different than predicted under an exponential growth model (*P*=0.639), the multi-modal distribution is consistent with a relatively stable historic population size.
Figure 10. Mismatch distribution of observed vs. expected pairwise differences between haplotypes for *C. ceti* (*n*=26). Although not significantly different than predicted under an exponential growth model (*P*=0.630), the multi-modal distribution is consistent with a relatively stable historic population size.
Figure 11. Mismatch distribution of observed vs. expected pairwise differences between haplotypes for *C. kessleri* (*n*=29). Although not significantly different than predicted under an exponential growth model (*P*=0.300), the multi-modal distribution is consistent with a relatively stable historic population size.
DISCUSSION

*Phylogenetic relationships among gray-whale lice*

The high degree of consistency between molecular phylogenies were in congruence with Haney’s (1999) morphological phylogeny; however, several sister-species relationships differed. The incongruence among morphological and molecular phylogenies could be due to a number of factors. First, relationships inferred from morphological characters could be misleading due to homoplasy (*i.e.*, retained ancestral character traits). Second, the COI mtDNA data used in this study may simply have lacked sufficient power to recover some parts of the phylogeny. Finally, molecular resolution of the whale-louse phylogeny may have been hampered due to incomplete sampling of all congeneric whale lice. Only seven of the twelve *Cyamus spp.* described by Haney (1999) were available for this study. Attempts were made to extract DNA from missing taxa from archived museum specimens; however, most samples were fixed in formalin and failed to yield usable DNA. Furthermore, it appears that the genus *Cyamus* is more speciose than current morphological estimates suggest (Leung 1967, Haney 1999, Margolis *et al.* 2000).

Understanding this potential cryptic diversity would help to provide a clearer picture of whale to whale louse relationships. For example, Kaliszewska *et al.* (2005) studied cyamid diversities living on right whales in the northern and southern hemispheres. They found that each right whale species (north Atlantic, north Pacific, and Southern Ocean) harbored the same three whale lice species (*C. gracilis*, *C. erraticus*, and *C. ovalis*), creating three sets of sibling species, one for each species of right whale (*e.g.*, three distinct lineages of *C. gracilis*, one for each right-whale species). Although
morphologically similar, each of the right whale species, as well as their whale lice, has been separated geographically for several million years (Rosenbaum et al. 2000, Kaliszewska et al. 2005). Kaliszewska et al. (2005) estimated the divergence time between Southern Ocean and northern hemisphere *C. erraticus* to be around $6 \times 10^6$ million years ago. Given the long separation in these three sibling trios, it appears that each could be recognized as a separate species, thus yielding nine species in total (Figure 5). Furthermore, recent studies have described two new species of whale lice in the genus *Cyamus*, *C.* novel species A, occurring on bowhead whales (Haney 1999), and *C. eschrichtii*, occurring on gray whales (Margolis et al. 2000). The description of these two novel species is based on external morphology, and both are similar in appearance to *C. ceti*. Inspection of external morphology and nucleotide sequences of specimens used in this study did not detect the presence of *C. eschrichtii*. Molecular analysis of a paratype specimen, as described by Margolis et al. (2000), would help verify the existence of this species on gray whales.

Another example of possible cryptic diversity is the presence of *C. ceti* on bowhead and gray whales. Individuals of *C. ceti* sampled from gray whales formed a reciprocally monophyletic clade (Figure 5). This clade is sister to a monophyletic group of *C. ceti* samples from bowhead whales (data not shown; Z. Kaliszewska and J. Seger pers. comm.). The distance between these two sister clades is much greater than the average pairwise differences observed within the gray-whale *C. ceti* clade ($\Pi=6.55$; Table 5). Although the sample size of *C. ceti* from bowhead whales was small ($n=2$), the nucleotide divergence between clades and the range of pairwise nucleotide differences observed between other species of gray-whale lice are large enough to suggest that enough time has passed for lineage sorting to have occurred. The sister-clade
relationship of these two populations differs from Haney’s (1999) morphological analysis. Lice collected from bowhead whales were first described as *Cyamus mysticeti* by Lütken in 1870, but Margolis (1955) considered *C. ceti* and *C. mysticeti* to be synonymous and *C. ceti* was given priority (Haney 1999). Haney (1999) reported some morphological differences in overall body size and number of mandibular incisors of *C. ceti* from gray and bowhead whales, but he considered these differences to be insufficient to resurrect separate species-level status. Given the sister relationship between these two groups, it is likely that the shared morphology represents symplesiomorphic character traits retained from a recent common ancestor. Because gray and bowhead whales do not themselves share a recent common ancestor (Milinkovitch *et al.* 1994, Hasegawa *et al.* 1997, Geisler and Sanders 2003, Rychel *et al.* 2004), the presence of these two lice on their respective hosts appears to reflect a host-switching event in the distant past. Considering the level of mtDNA divergence and apparent morphological differences described by Haney (1999), *C. ceti* on gray and bowhead whales appear to represent distinct evolutionary lineages. This hypothesis warrants further examination.

The mtDNA marker used in this study was a 738 bp region of the COI coding gene, which is an important gene in cellular respiration. Like most coding regions of the genome, the COI is more conserved than non-coding regions. That is, the COI gene is under functional constraints that are reflected in the distribution of mutations among codon positions. Most substitutions occur at third nucleotide positions of a codon; substitutions at this position typically do not alter the amino acid being coded for because of the degenerate nature of the genetic code. In contrast, mutations in the first or second base (especially the latter) often change the amino acid sequence. Although third base mutations are phylogenetically informative at the population level, they typically become less so for deeper time scales because they become saturated (*i.e.*, multiple substitutions...
occur at a single nucleotide position; Arbogast et al. 2002). In this study, substitutions at the more constrained first and second bases were minimal. The combined result is that the fragment of the COI gene used in this study appears to lack sufficient power to resolve sister-species relationships among cyamid whale lice with a high level of confidence. Typically, in a situation such as this, the next step would be to examine longer sequences and/or more genes in an effort to better elucidate the complex intrageneric relationships. However, in the case of cyamids, substantially longer sequences do not appear to further clarify the phylogeny; Kaliszewska and Seger (pers. comm.) have used 4.1 kilo bases (kb) of mtDNA from multiple genes to examine the relationships of Cyamus spp. and have still been unable to obtain strong nodal support for sister-species relationships. In addition, each gray-whale and right-whale louse species appears to fully coalesce to less than $1 \times 10^6$ years ago (Kaliszewska et al. 2005). It is likely that the speciation within the genus Cyamus was relatively rapid and fairly recent. As such, it may be very difficult to recover the sequence of speciation events in this group with high levels of confidence.

Levels of genetic diversity within and among gray-whale lice

All three louse species showed high levels of genetic diversity (Table 5), which is consistent with COI data for cyamid whale lice found on other mysticete whales (e.g., right whales; Kaliszewska et al. 2005). The number of private haplotypes ($h/n$), per louse species, was highest in C. ceti, with 80% of haplotypes being unique. Sixty percent of haplotypes were unique in C. scammoni and 34% were unique in C. kessleri. Kaliszewska et al. (2005) consistently found relatively high levels of haplotype diversity in right-whale lice (26-90% unique haplotypes). The relatively lower level of haplotype
diversity observed in *C. kessleri* in this study may be explained by the fact that they were collected from only two host whales, both of which were neonates.

Samaras and Durham (1985) suggested that dispersal of whale lice typically involved younger, smaller lice. If younger lice are dispersing, then it might be reasonable to expect that newly dispersing lice may be more closely related. Leung (1976) reported that female whale lice can rear hundreds of young in one generation, thus one cohort of whale lice would all have the same haplotype as the mother. Also, neonate whales have not had the opportunity to accumulate diverse populations of whale lice (*i.e.*, they have not mixed with the breeding population), potentially limiting the source pool of whale lice considerably. However, samples of *C. scammoni* and *C. ceti* collected from the same two neonate calves did not show similarly low levels of haplotype diversity. Rather, they exhibited levels of haplotype diversity consistent with populations of whale lice collected from older whale hosts (*C. scammoni*, 64-80%; and *C. ceti*, 90-100% unique haplotypes).

The discrepancy between levels of haplotype diversity in *C. kessleri* and the other two gray-whale louse species could be influenced, at least in part, by the overall number and densities of each species of louse present on the mother gray whale. *C. kessleri* is typically restricted to the urogenital region and occurs in lower numbers and densities than the other two louse species, each of which occur on the head region of the host whale (Rice and Wolman 1971, Sumaras and Durham 1985). The number of individuals of *C. kessleri* collected from the neonate whales is comparable to those collected for the other two species of whale lice (Table 2). This suggests that the lower level of haplotype diversity observed in the former may be due to a lower level of diversity in the source populations (*i.e.*, those of the mother whales), rather than a sample-size effect. For example, sample size did not appear to influence haplotype diversity in right-whale lice;
of 104 individual *C. ovalis* sampled from Southern Ocean right whales, there were a total of 81 unique haplotypes.

Kaliszewska *et al.* (2005) did not age-class their host whales, so it is unclear if young right-whale calves harbor less genetically diverse populations of whale lice compared to older individuals. However, *C. erraticus* is analogous to *C. kessleri* in that it also occupies the urogenital regions of its host. Individuals of *C. erraticus* sampled by Kaliszewska *et al.* (2005) did not show low levels of haplotypic diversity (68-100%), but again, it is unclear whether whales were adults, juveniles, or neonate calves. Only one right-whale louse population, *C. gracilis*, had low haplotype diversity (22%) in their study. The authors suggested that this population was atypical, apparently having suffered from an internal bacterial infection that may have reduced the genetic diversity of this population (Kaliszewska *et al.* 2005). Other diversity measures, such as the mean number of pairwise differences per sequence and per site and population mutation rates, were similar for all gray-whale lice and were comparable to those observed in right whale lice (Table 5; Kaliszewska *et al.* 2005).

*Divergence times in gray-whale lice*

I identified three distinct lineages of gray-whale lice, which is consistent with the morphological data. Each lineage coalesced independently at approximately 6.6-9.5x10^5 years ago (Figures 6 - 8). These estimates are older than values reported by Kaliszewska *et al.* (2005) for right-whale lice; however, their estimates were calculated from synonymous substitutions of the mtDNA data and divergence times of trans-isthmus snapping shrimp (*Alpheus spp.*). Their estimates for all lineages, except for one, coalesced to approximately 5x10^5 years ago. Differences in coalescent estimates could
reflect real differences in the ages of the two groups of lice or be due to differences in the methods used to estimate the coalescent dates. Direct comparisons of synonymous substitution rates in gray-whale lice and a multi-species gene tree, with a molecular clock assumption, may yield more consistent estimates of divergence times among cyamid whale lice. The multi-species gene tree in this study failed the molecular clock assumption, so divergence times were not calculated using that particular tree. Additional sampling of multiple nuclear markers may give better resolution of sister-species relationships, which may also help the DNA data meet the molecular clock assumption, thus giving better estimates of divergence times for cyamid whale lice.

Population estimates of gray-whale lice

The high levels of genetic diversity observed in gray-whale lice likely reflect large effective populations and short generation times relative to whale hosts. All three species of lice are found in high densities among individual gray whales. Rice and Wolman (1971) examined 316 gray whales and found that 98% of all animals had all three species of whale lice present. There are no good data for abundance estimates; however, the literature reports that C. scammoni is the most abundant species followed by C. ceti, and then C. kessleri (Rice and Wolman 1971, Samaras and Durham 1985). That pattern was not observed in this study, but that could simply reflect the fact that several of the whales examined in this study were neonates. One account reports that over 100,000 C. scammoni were collected off of a single gray whale (Leung 1965). This high abundance is atypical and may have been from an animal that was sick or injured. Kaliszewska et al. (2005) reported that injured right whales often harbor inflated populations of whale lice. This may be because lice infest an open wound, or because the
animal may have altered behavior, which may allow aggregates of lice to disperse onto areas of the body not normally occupied (Rowntree 1996).

The distribution patterns, and likely abundances, of gray-whale lice are similar to right-whale lice. The highest densities of lice on right whales are *C. ovalis*, which occupies the outer callosity regions of the head. This is similar to *C. scammoni*, which occupies the outer barnacle regions on the anterior dorsal region of the host. *C. gracilis* lives in callosity grooves and on the fringes of *C. ovalis* aggregations. On gray whales, *C. ceti* lives on the fringes of *C. scammoni* aggregations on the head region. Both of the latter two species are smaller and are present in lower densities. Lastly, *C. kessleri* (gray whales) and *C. erraticus* (right whales) occupy the urogenital regions of their hosts and occur in the lowest densities. The parallel life histories of these whale lice and their associate whale hosts provides the opportunity to infer estimates of gray-whale lice abundances in the absence of actual counts. The population neutral mutation rates, calculated from segregating sites (θs), and mean rate of pairwise differences per site (π), were similar for all populations of gray-whale and right-whale lice (Table 5). Populations of gray-whale lice exhibited neutral mutations (θs) of 0.005-0.016, compared to right-whale lice, which had values of θs that ranged from 0.005-0.030. Gray-whale lice and right-whale lice had values of π ranging from 0.004-0.011 and from 0.006-0.023, respectively. Although the values for gray-whale lice were slightly lower, this may be an artifact of smaller sample sizes and shorter DNA sequences (i.e., right-whale lice sequences were approximately 60 bp longer; Kaliszewska et al. 2005). However, overall the data suggest that right-whale and gray-whale lice populations may be behaving similarly to each other.
Whale louse population estimates for each individual right whale were: C. ovalis, $n=5,000$; C. gracilis, $n=500$; and C. erraticus, $n=2,000$, although estimates for C. gracilis were reported to be under-estimated and attributed to incomplete sampling (Kaliszewska et al. 2005). The only complete sampling of gray-whale lice in this study involved two neonate calves and one sub-adult male (Table 2). The sub-adult male did not have any C. kessleri present, probably because the urogenital region was exposed to scavenging by birds and showed evidence of bite wounds. Therefore, abundance estimates are based on the two calves and can be approximated as a low baseline for gray-whale lice abundances on their hosts. The average densities on the two calves were C. scammoni, $n=71$; C. ceti, $n=148$; and C. kessleri, $n=135$. Assuming the current population estimates for gray whales ($n=18,000$ to 29,000; Rugh et al. 2005), then overall population estimates for each species of whale louse would be C. scammoni, $n=1.3-2.1\times10^6$; C. ceti, $n=2.6-4.3\times10^6$; and C. kessleri, $n=2.4-3.9\times10^6$. These estimates are likely an order of magnitude, or more, lower than actual numbers, given that the effective population estimates range from $1.1\times10^6$ to $7.3\times10^6$. This range is consistent with observed densities; that is, C. kessleri has the lowest $N_e$ and C. scammoni has the highest.

All three species of gray-whale lice failed to show population structure. This finding is consistent with right-whale lice (Kaliszewska et al. 2005). Finally, mismatch distributions of observed versus expected pairwise differences between haplotypes for all gray-whale louse species were ragged and multi-modal (Figures 9 - 11). The multi-modal distributions are consistent with stable long-term historical populations, which would be expected given the relatively high genetic diversity seen in gray-whale lice.
Historically gray whales occurred in the northern Atlantic and Pacific oceans (Rice and Wolman 1971, Mead and Mitchell 1984). The Atlantic population went extinct during modern times and two populations remain in the eastern and western north Pacific (ENP and WNP, respectively). North Pacific gray whales were intensely harvested in recent times until their protection in the last few decades. ENP gray whales have appeared to recover to pre-exploitation numbers ($n=22,000$); however, new multi-locus coalescent estimates suggest that historical numbers may have been much larger (i.e., 76,000-118,000 individuals; Alter et al. 2007). The WNP population remains critically endangered with fewer than 150 individuals (Rugh et al. 2005, Bradford et al. 2006). There appears to be little gene flow between the two populations (LeDuc et al. 2002) and mitochondrial and nuclear DNA markers revealed very little genetic sub-structure within the ENP population (Steeves et al. 2001, Alter et al. 2007).

Effectively, ENP gray whales represent one large population. These findings are consistent with the data from the gray-whale lice, which also showed no significant population structure among hosts (Figures 6 - 8). Similarly, right-whale lice also had little structure within ocean basins (Kaliszewska et al. 2005). Mismatch distributions (Figures 9 - 11), although not significantly different from a rapid population expansion, suggest that all three species of whale lice have maintained a long-term association with gray whales.

The estimated generation time in gray whales is approximately 15.5 years (Alter et al. 2007). Kaliszewska et al. (2005) suggested that whale lice on right whales may complete two generations per year, and Leung (1976) suggested that gray-whale lice complete their life cycles in 8-9 months. Using the conservative estimate of one
generation per year, gray-whale lice would have approximately a 15-fold faster
generation time than their gray-whale hosts. These faster generation times, coupled with
the large effective population sizes of the whale lice, mean that we should be able to track
historical demographic patterns in whale lice more easily than in their whale hosts.

It is presumed that WNP gray whales harbor the same three species of whale lice
as ENP gray whales. Given the low rate of gene flow between the WNP and ENP
populations, whale lice of the two groups might be expected to show a clear separation in
haplotypes, such as that seen in right-whale lice from different ocean basins. Attempts
were made to sample cyamid whale lice from WNP gray whales; however, none was
available for this study. Future analyses between whale lice from the two gray whale
populations should enhance our ability to accurately estimate the divergence time
between the ENP and WNP populations of gray whales and their whale lice.

Conclusions and future analyses

Gray whales host three distinct evolutionary lineages of cyamid whale lice, each
endemic to the gray whale. The polyphyly of this group suggests that each species
colonized gray whales independently, and that each species has had a long-stable
association with gray whales. Host-switching may be a common event that has occurred
several times in the history of cetacean-cyamid associations, especially for gray whales.
Although sister-species relationships were unresolved, they were consistent with
morphological data, which suggests gray-whale lice are polyphyletic. Further sampling
of congeneric whale lice and multi-locus nuclear markers should help to further resolve
the phylogenetic relationships of this group.

Additionally, each louse species exhibited high levels of genetic diversity with no
population sub-structure within its gray whale hosts. Preliminary comparisons of the
demographic histories of the three gray-whale lice (this study) and their gray-whale hosts
(Alter et al. 2007) suggest that they have had a long-shared history living together.
The close and relatively long-term evolutionary association of the lice with gray whales provides an exciting opportunity to use estimates of historical demographic patterns from the lice to make inferences about the historical demography of the gray whales themselves. This possibility is enhanced in gray whales relative to many other species of cetacean, because their louse fauna provides three independent replicates (i.e., three different species of whale louse) for indirectly inferring signatures of population structure in the gray-whale host.

Systematic sampling of future gray-whale strandings, including the urogenital region, which is typically under-sampled, may give better estimates of whale lice abundances. These data would be useful in estimating demographic parameters for the whale lice and their gray-whale host. Whale lice samples from the WNP gray whale population may also reveal divergence times between the two whale populations. Further sampling of cryptic species, such as, *C. ceti* from bowhead whales and *C. eschrichtii* paratypes, will help resolve host-parasite relationships between their whale hosts. Finally, multi-locus nuclear DNA markers are needed to help resolve the phylogenetic relationships of cyamid whale lice.
LITERATURE CITED


Microchemical Core Facility, San Diego State University, CSL 326 MC1030, 5500 Campanile Drive, San Diego, CA 92182.


MJ Research, Inc. 1997. PTC-100 thermocycler. Waltham, MA, USA.


Promega Corporation, 2800 Woods Hollow Rd., Madison, WI 53711.


Qiagen, Inc. 2002b. QIAquick® PCR Purification Kit. Valencia, California, USA.


Sequencher, version 4.2. Gene Codes Corporation, 775 Technology Drive, Suite 100A, Ann Arbor, MI 48108.


