MOLECULAR SYSTEMATICS AND PHYLOGEOGRAPHY OF *LAGENORHYNCHUS OBSCURUS* DERIVED FROM NUCLEAR AND MITOCHONDRIAL LOCI

A Dissertation

by

APRIL DAWN HARLIN

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

December 2004

Major Subject: Wildlife and Fisheries Sciences

MOLECULAR SYSTEMATICS AND PHYLOGEOGRAPHY OF *LAGENORHYNCHUS OBSCURUS* DERIVED FROM NUCLEAR AND MITOCHONDRIAL LOCI

A Dissertation

by

APRIL DAWN HARLIN

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Approved as to style and content by:

Rodney L. Honeycutt (Co-Chair of Committee)	Jim Woolley (Member)
Bernd Würsig (Co-Chair of Committee)	Lee Fitzgerald (Member)
	Robert Brown (Head of Department)

December 2004

Major Subject: Wildlife and Fisheries Sciences

ABSTRACT

Molecular Systematics and Phylogeography of *Lagenorhynchus obscurus* Derived from Nuclear and Mitochondrial Loci. (December 2004)

April Dawn Harlin, B.S., University of California at Davis;

M.S., Texas A&M University

Co-Chairs of Advisory Committee: Dr. Rodney L. Honeycutt
Dr. Bernd Würsig

This study presents evidence from mitochondrial and nuclear loci that there is genetic divergence among and within geographic populations of *Lagenorhynchus obscurus*. The effect of seasonal variation on the genetic structure within New Zealand was examined with mitochondrial DNA control region sequences from 4 localities. Analysis of nested haplotype clades indicated genetic fragmentation and at least 1 historical population expansion within New Zealand. AMOVA and Fst values from nuclear and mitochondrial DNA sequences suggested significant divergence between New Zealand, South Africa, Argentina, and Peru. Dispersal via the west-wind drift was not supported by patterns of population structure among regions. Alternatively, these data support reciprocal exchange among all four regions with 100% posterior probability for a root of origin in the Indian/Atlantic Oceans. The degree of divergence between Peru and other regions indicates the isolation of Peruvian stock is temporally correlated with the constriction of Drake's passage in the Plio-Pleistocene. There is evidence that the Plio-Pliestocene paleoceanography of the Indian and Southern Atlantic Oceans

influenced phylogeography with shifts of temperate sea surface temperatures northward ~5º of latitude, disrupting the dispersal corridor between New Zealand and Atlantic populations. A preference for temperate waters along continental shelves is proposed as an explanation for lack of contemporary genetic exchange among regions. This study supports the polyphyly of the genus *Lagenorhynchus*. North Atlantic species form a monophyletic *Lagenorhynchus*. In the Southern Hemisphere, *L. australis*/*L. cruciger* and *L. obliquidens*/*L. obscurus* do not form a monophyletic group. I discuss the taxonomic implications and propose taxonomic revision of the genus based on these results. Measures of character interaction indicate that combined evidence from nuclear and mitochondrial genes provide better phylogenetic resolution among delphinid lineages than any data partition independently, despite some indications of conflict among mitochondrial and nuclear data.

DEDICATION

To my grandmother, Betty Jean Ousley Wheetley, with love and admiration. Without you, grandma, I would not be here. With you, I have learned the meaning of endurance and self-sacrifice. From your example, I hope to become a better woman—to take the world by storm, carrying the burden of all, but to do so graciously, with the kindness and temperance of a loving heart.

ACKNOWLEDGEMENTS

This research was supported with a Doctoral Dissertation Improvement Grant 0300480 from the National Science Foundation; an Interdisciplinary Research Initiative (IRI 97-24) from the Research Enhancement Program, Office of the Vice President for Research and Associate Provost for Graduate Studies, Texas A&M University; from the Earthwatch Center for Field Research; and from the Committee for Research and Exploration of the National Geographic Society. Many special thanks to Tim Markowitz for his assistance in the field, and most importantly, for the love and moral support that made this study possible. My Ph.D. experience was greatly enhanced by the friendship of H. López-Fernández ("El Chamo") and A. Cognato—movies, dinners, good bottles of wine! To Anthony I owe the deepest gratitude for his loving support and patience—I could not have survived these last years without you! To my mother and grandmother whose faith in me never wavered and who gave me the tenacity to endure. To Rodney Honeycutt, my friend and advisor, I owe the deepest gratitude for 5 years of friendship and guidance, for enduring my quirks, honing my skills; for patience with my countless interruptions, for always having an open door, for reading hundreds of drafts of the same papers, four humorous East Texan words of wisdom, and for setting the standard by which I measure my own success. To Bernd Würsig, for taking on a floundering young scientist and for having faith in me when all others failed. To Jim Woolley, who taught me all I know about quantitative phylogenetics, "pseudo-statistical mumbo-jumbo", and a good Zin ($, $$, $$$). To Lee Fitzgerald, for his mentorship and consistent words of

encouragement. To C. Scott Baker for taking me under his wing, teaching me how to collect data, and for moral and intellectual support during all phases of my academic development.

I acknowledge the NMFS Southwest Fisheries Sciences Center, La Jolla, California, for genetic material. I thank M. Meyers and D. Koetze, Sea Fisheries Research Institute, South Africa; S. Dans and K. Crespo, Laboratorio de Mamíferos Marinos Centro Nacional Patagónico, CONICET, Argentina; and J. Gibbons, Universidad de Magallanes, Chile for their invaluable support in the field. F. Pichler graciously provided access to specimens with permission from N. Goodall, C. Olavarria, and from Te Papa Tongarewa, National Museum of New Zealand. Permits for collection of tissue samples were issued by the New Zealand Department of Conservation, and approved by Texas A&M University and University of Auckland Animal Care and Use Committees. Tissue samples from beach-cast dolphins were provided courtesy of F. Cipriano. Fieldwork was made possible with help from A. Acevedo-Gutiérrez, K. Mazzarella, S. Stanley, L. McOmber, N. Brown, H. Petersen, and Earthwatch team members. Many thanks to C. Ingram, D. Rowe, L. Frabotta, and G. Proudfoot for being good friends and lab mates. I thank Shirley, Diann, Vicki, Janice, Jennifer, and others of our departmental staff for their help in countless ways. F. Pichler, J. Murrell, G. Lento, K. Dunn, J. Bielawski, helped with analyses and useful discussions. Thanks to my mates at Dolphin Encounter, Kaikoura, and the Edward Percival Field Station, University of Canterbury, for various elements of friendship and support of the dusky dolphin project in Kaikoura.

TABLE OF CONTENTS

Page

ABSTRACT iii

DEDICATION v

ACKNOWLEDGEMENTS vi

TABLE OF CONTENTS viii

LIST OF FIGURES x

LIST OF TABLES xii

CHAPTER

I INTRODUCTION 1

Objective 1: Patterns of variation within New Zealand 3
Objective 2: Molecular systematics of the genus *Lagenorhynchus* 5
Objective 3: Phylogeography of *L. obscurus* 8

II GENETIC STRUCTURE, DIVERSITY, AND HISTORICAL DEMOGRAPHY OF NEW ZEALAND'S DUSKY DOLPHIN (*LAGENORHYNCHUS OBSCURUS*) 12

Introduction 12
Materials and methods 17
Results 25
Discussion 34

III ASSESSMENT OF CHARACTER INTERACTION AND PHYLOGENETIC UTILITY OF MULTIPLE DATA PARTITIONS: RELATIONSHIPS AMONG DOLPHINS OF THE SUBFAMILY LISSODELPHININAE 41

Introduction 41
Materials and methods 47
Results 57

CHAPTER Page

Discussion 70

IV CATCHING THE DRIFT: PHYLOGEOGRAPHY OF THE DUSKY DOLPHIN (*LAGENORHYNCHUS OBSCURUS*) AND THE WEST-WIND DRIFT HYPOTHESIS 77

Introduction 77
Materials and methods 81
Results 88
Discussion 98

V SUMMARY 108

Objective 1: Patterns of variation within New Zealand 108
Objective 2: Molecular systematics of the genus *Lagenorhynchus*.....109
Objective 3: Phylogeography of *L. obscurus*.....110

REFERENCES.....112

VITA 136

LIST OF FIGURES

FIGURE		Page
1	Sampling localities in New Zealand	19
2	Unrooted neighbor-joining phylogram of 169 mtDNA control region haplotypes from dusky dolphins.	27
3	Nested clade network of haplotypes.	31
4	Continuation of nested clade network, Levels 1-5..	32
5	Summary of demographic events supported with nested clade analysis	33
6	Graph depicting results of lineages-through-time analysis.	35
7	Distribution of the genus *Lagenorhynchus* and the phylogeny of Delphinidae derived from cytochrome *b* sequences.	46
8	An example of SDI calculation with a simulated data set	58
9	Cladogram derived from simultaneous analysis (SA) of 4 data partitions	60
10	Cladograms derived from separate analyses of mitochondrial and nuclear genes, respectively.	63
11	Saturation plots.	64
12	Distribution of branch support (BS) by character partition	66
13	Distribution and relative frequency of perfectly consistent characters (PC) in the simultaneous analysis (SA) cladogram	67
14	Distribution of the dusky dolphin (*Lagenorhynchus obscurus*).	80
15	Statistical parsimony network of *Lagenorhynchus obscurus* cytochrome *b* haplotypes	90
16	Statistical parsimony network of control region haplotypes	91

FIGURE Page

17 Parsimony network of actin haplotypes.....92

18 World-wide distribution of dusky dolphins in relation to oceanographic conditions.....100

19 Changes in ocean current patterns and temperatures in the last 2 million years.....101

LIST OF TABLES

TABLE Page

1 Pairwise distances between 4 regions along South Island coast of New Zealand.. 18

2 Summary of inferences regarding demographic events deduced from clades with significant nested clade values.....29

3 List of taxa and sample size for each data partition. 49

4 Summary of results from parsimony analysis of data partitions.....59

5 Partition branch support (PBS) for each node of the simultaneous analysis (SA) cladogram.....65

6 Summary of support indices for branches in the simultaneous analysis (SA) and independent analyses (IA) of data partitions.....69

7 Summary of diversity indices for cytochrome b (Cytb), the control region (Dloop) from the mitochondrial genome, and actin intron 1 (actin) from the nucleus among geographic localities.....83

8 Genetic divergence among *L. obscurus* populations. 94

9 Absolute frequency of haplotypes among regions for *L. obscurus* mitochondrial cytochrome b (Cytb), and control region (Dloop). 95

10 Frequency and geographic distribution of actin intron haplotypes.....98

CHAPTER I

INTRODUCTION

The use of genetic markers for reconstructing the demographic and evolutionary history of a species has several basic requirements. First, multiple molecular markers, with different rates of coalescence, are necessary for an accurate evaluation of population-level processes that are associated with species diversification (e.g., barriers to gene flow, population expansion). Second, patterns of allelic or haplotypic divergence are best observed through the phylogenetic reconstruction of gene trees that consider overall genetic diversity within and between populations. Finally, genetic diversity within a species should be interpreted from a broader phylogenetic and geographic perspective, whereby the phylogeographic history of the species is incorporated into the broader evolutionary history of the genus. More recent methods to examine intra- and inter-specific evolutionary history apply population genetic and phylogenetic methods to multiple genetic loci (i.e., mitochondrial and nuclear) to corroborate demographic events across temporal and geographical scales (Bernardi et al. 1993; Palumbi & Baker 1994; Templeton 2002; Antunes et al. 2002).

A rapid rate of coalescence, uniparental inheritance, and lack of recombination (Brown 1979; Giles et al. 1980) enhance the usefulness of mitochondrial DNA (mtDNA) as the molecule of choice for studies of intraspecific phylogeography (Avise 1987).

This thesis follows the format and style of *Molecular Ecology*.

Despite these genetic attributes, mtDNA provides a single gene perspective depicting the genetic history of the female lineage only (Moritz 1994). Clearly, a decipherment of the phylogeographic history of a species, especially as it relates to population processes associated with demography, dispersal patterns, population structure, etc., requires the use of multiple markers inherited both uniparentally and biparentally (Sugg et al. 1996; Hoelzer et al. 1998; Ross 2001). Recently, multiple genetic loci (both mitochondrial and nuclear) have been used to address questions related to historical patterns of evolutionary divergence within and between species (Bernardi et al. 1993; Palumbi & Baker 1994; Templeton 2002; Antunes et al. 2002). In terms of phylogeography (Avise et al. 1987) the addition of neutral nuclear markers provides an opportunity to distinguish between alternative explanations for inferred population history (Hare 2001; Zhang & Hewitt 2003). In theory, neutral nuclear loci coalesce 3 times more slowly on average than mtDNA genes (Palumbi et al. 2001), thus providing more temporal depth for studying the phylogeography of recently diverged intra- or inter-specific lineages (Hare 2001; Zhang & Hewitt 2003).

Here I utilize a combination of nuclear and mitochondrial DNA sequences to infer the phylogeographic history of the dusky dolphin (*Lagenorhynchus obscurus*), and the molecular systematics of the genus *Lagenorhynchus*, a member of the subfamily Lissodelphininae (LeDuc et al. 1999). Members of the genus *Lagenorhynchus* are generally meso-pelagic, small-bodied species occurring in cool, temperate waters of the North Pacific, North Atlantic, and Southern Ocean. The antitropical distribution of *Lagenorhynchus*, and the mix of endemic and widespread species within the genus,

provides a unique opportunity to simultaneously examine inter- and intra-specific processes that have shaped the mode and tempo of speciation within aquatic mammal lineages.

Objective 1: Patterns of variation within New Zealand

In New Zealand, dusky dolphins occur around the North and South Islands from Stewart Island in the south, Hawke Bay in the north, and the Chatham Islands in the east (Würsig et al. 1997). Historical sightings suggest that groups are very rare around the North Island but are common and abundant on the east coast of the South Island where deeper, cooler waters approach close to shore. This is especially true off the coasts of the Kaikoura and Otago Peninsulas, where groups of hundreds are seasonally common. Although abundant in these two regions, dusky dolphins are rarely seen along the shallow shelf waters of the Chatham Rise that lies between the Kaikoura and Otago regions and extends eastward 467 nautical miles to the Chatham Islands (Würsig et al. 1997). In contrast to the east coast, anecdotal evidence and casual observations during short coastal surveys (Bräger & Schneider 1998) suggest dusky dolphin groups on the west coast of the South Island are comparatively small and rare. The northern half of the west coast, like the Chatham Rise, has a wide, shallow continental shelf. The pattern of increased abundance of dusky dolphins in the deep, coastal waters of New Zealand also is seen in Chile and Peru, where a narrow continental shelf and deep near-shore waters appear to be correlated with large aggregations of dusky dolphins (Würsig et al. 1997). Stomach contents from dolphins incidentally caught in fishing nets suggest that their

presence in these areas is related to the diurnal abundance of meso-pelagic fishes and squid (Cipriano 1992; Würsig et al. 1997).

There is evidence for behavioral segregation of dolphins within New Zealand that may represent one or more genetically distinct populations. The typical group pattern in Kaikoura is a large, "main" pod (50 to >1000 individuals) with small "satellite" groups of 5-15 individuals concentrated within 2 to 3 kilometers of the main pod. In the Marlborough Sounds, approximately 250 kilometers north of Kaikoura, dusky dolphins have been observed to forage cooperatively on schooling fishes during the daytime (Harlin et al. 2001; Markowitz 2004;), similar to the cooperative feeding behavior observed in Argentina. In contrast, the feeding behavior of dolphins in Kaikoura is dominated by diurnal vertical migration of fishes and squid in the water column that regularly support groups of thousands of dusky dolphins (Würsig et al. 1997). Photo-identification and molecular gender determination suggest that at least some individuals migrate from Kaikoura to the Marlborough Sounds, but that these groups are comprised of 10-15 adult males (Harlin et al. 2001).

Cetaceans generally have high dispersal potential (e.g., Baker et al. 1993; Rosel et al. 1995, Bakke et al. 1996;). Therefore, one might not expect genetic differentiation between populations within small regions. Nevertheless, photo-identification (Würsig & Jefferson 1990), behavioral data, and studies of mitochondrial DNA (mtDNA) lineages indicate that genetic divergence can occur over small geographic scales with gene flow limited by either (a) geographic barriers (e.g., Dowling & Brown 1993), (b) resource heterogeneity (e.g., Hoelzel & Dover 1991; Hoelzel 1998; Hoelzel et al. 1998a,b), (c)

habitat specialization (e.g., Pichler et al. 1998), or (d) feeding or breeding site fidelity (e.g., Shane et al. 1986; Baker et al. 1990, 1993; Duffield & Wells 1991). The **first objective** of this dissertation is to determine the effects that habitat heterogeneity (e.g., deep canyons vs. shallow glacial bays) and seasonal movement patterns have had on the genetic structure of dusky dolphin populations in New Zealand. If there were behavioral barriers to dispersal, one would predict a seasonal division of dusky dolphins into genetically diagnosable populations within New Zealand.

Objective 2: Molecular systematics of the genus *Lagenorhynchus*.

In terms of a global perspective of evolutionary history, there are to date no studies of the dusky dolphin, or the genus *Lagenorhynchus*, that have included samples from all regions of the dusky dolphin's current range. Van Warebeek (1993) demonstrated that the skulls of dusky dolphins from Peru are significantly larger and have fewer teeth on average than those from either New Zealand or South Africa (Van Warebeek 1993; Miyazaki & Shikano 1997), and are more similar in size to those of *L. obliquidens*, the sister taxon to the dusky dolphin (Cipriano 1992; Miyazaki & Shikano 1997; LeDuc et al. 1999). These morphological data, in addition to more recent genetic analyses of mtDNA sequence variation in Peru and New Zealand (Cipriano 1992; LeDuc 1999), indicate that the New Zealand/South African dusky dolphin populations have shared a more recent common ancestry than either region has to Peru. According to Van Warebeek (1993), the similarity in skull size between Peruvian dusky dolphins and the Northern Hemisphere *L. obliquidens* suggests that the Peruvian “form” of the dusky dolphin is ancestral to the smaller South African and New Zealand “forms”.

Nevertheless, neither the morphological nor molecular studies included samples of *L. obscurus* from all three major biogeographic regions (e.g., South America, South Africa, and New Zealand), nor were the patterns suggested by the statistical analysis of the morphological data (Van Warebeek 1993) interpretable in a phylogenetic framework.

Members of the genus *Lagenorhynchus* have been described as a monophyletic group on the basis of shared similarities in color pattern, and skull and beak morphology (Gray 1866; Fraser 1966; Mitchell 1970). Based on these criteria, *L. albirostris* and *L. acutus* were considered as North Atlantic members of the genus. More recently, a molecular phylogeny derived from mitochondrial cytochrome *b* sequences (LeDuc et al. 1999) suggested that with the inclusion of the North Atlantic species, the genus *Lagenorhynchus* is polyphyletic. Furthermore, this study suggested that *L. acutus* remain in the genus *Lagenorhyncus*, *L. albirostris* be placed in the genus *Leucopleurus*, and the genus *Sagmatius* be resurrected for *L. australis* and *L. cruciger* to represent the actual phylogenetic relationships of these taxa. Cipriano (1997) used a molecular phylogeny to infer a Plio-Pleistocene antitropical speciation event in the genus *Lagenorhynchus* that occurred as a result of transgression across cool equatorial waters (CLIMAP 1976) followed by restricted gene flow with an increase in post-glacial water temperatures (Davies 1963; White 1986; Lindberg 1991). Without the North Atlantic species, the genus *Lagenorhynchus* contains taxa that occur only in the Southern Hemisphere and North Pacific; therefore, the antitropical speciation event in the genus is more likely to have occurred in the Pacific Ocean basin. Recently, Hare et al. (2002) used mitochondrial and nuclear genes to examine the demographic history of *L.*

obscurus and *L. obliquidens*, and concluded that the dusky dolphin likely arose in the Southern Hemisphere from a propagule of >10 *L. obliquidens* approximately 0.74 million years ago. However, their analyses assumed that *L. obscurus* from New Zealand and Peru, and *L. obliquidens* from the coast of Northern California, adequately represent the genetic variation of each species, and that an equatorial transgression occurred only once. They interpreted the lack of lineage sorting (Avise et al. 1984) within nuclear gene trees to be due to a large effective population size in both species. Yet an alternative explanation for this pattern could be that an incomplete sample of lineage diversity has lead to an under-representation of the number of nodes at the distal portion of the phylogenies, which can mimic a slowdown in lineage coalescence (Nee 1994).

The pattern of mtDNA variation presented by Harlin et al. (2003) reveals a more complex pattern of dusky dolphin evolutionary history than that presented in other studies. This pattern may very well reflect historical and more recent exchanges among the various biogeographic regions of the dusky dolphin's distribution, and perhaps even among species in the genus *Lagenorhynchus*. Clearly, a more adequate test of the hypotheses presented in previous studies requires a more extensive assessment of variation across the range of the dusky dolphin. LeDuc et al. (1999) and Pichler et al. (2001) failed to find any resolution among the relationships of Southern Hemisphere *Lagenorhynchus*. This suggests that the inclusion of other members of the genus from the Southern Hemisphere also is important for understanding the evolutionary history of the dusky dolphin and the genus as a whole.

The **second objective** of this dissertation is to place patterns of diversification within the dusky dolphin in a broader evolutionary context through a detailed examination of phylogenetic relationships of the dusky dolphin and other members of the genus *Lagenorhynchus*, distributed in the North Atlantic (*L. acutus* and *L. albirostris*), the North Pacific (*L. obliquidens*), and the Southern Hemisphere (*L. australis, L. cruciger*).

Objective 3: Phylogeography of *L. obscurus*

Recent advances in analytical methods have increased the potential for nuclear DNA (nDNA) to be incorporated into phylogeographic studies. The interpretation of nuclear gene phylogenies are complicated by the potential for recombination (Schierup & Hein 2000), gene tree polyphyly, and the blurred species boundaries in recently diverged taxa (Schierup & Hein 2000; Carbone & Kohn 2001; Hare 2001). However, polyphyly of gene trees can be resolved potentially by increasing the number of characters analyzed, and methods such as Templeton and Sing's (1993) statistical parsimony approach can be used to identify regions of recombination and modify phylogenetic analysis accordingly (Templeton et al. 2000). The development of these more-sophisticated approaches that incorporate nDNA into studies of inter- and intra-specific evolution increases the breadth of phylogeographic inference (Sugg et al. 1996; Hare 2001; Ross 2001). Even if limitations associated with nDNA gene-tree polyphyly and recombination are not completely resolved, the resulting phylogeny provides a means of developing and testing alternative hypotheses. In addition, demographic history, including changes in effective population size and the rate of lineage

diversification over time can be estimated to explore alternative historical processes that have shaped evolutionary history.

Although its range is large, the dusky dolphin's distribution is discontinuous and represented by large populations along the coasts of New Zealand, Peru, Argentina, and South Africa where deep, temperate waters approach close to shore. In New Zealand and along the west coast of South America a narrow continental shelf and deep near-shore waters appear to support large aggregations that feed on meso-pelagic fishes and squid (Cipriano 1992; Würsig et al. 1997). The disjunct, yet broad distribution coupled with patterns of morphological and genetic (exclusively mtDNA) divergence have resulted in the proposal of several hypotheses related to the phylogeogeographic history of dusky dolphin. One of these, the west-wind drift hypothesis (WWD), relates patterns of divergence in the dusky dolphin and the dolphin genus *Cephlorhynchus* to the eastward flowing, virtually uninterrupted current system in the Southern Ocean that carries cool water around the Antarctic. Pichler et al.'s (2000) phylogenetic study of mtDNA lineages suggested that *Cephalorhynchus* speciated along continental coastlines following dispersal with the WWD, beginning in the eastern Atlantic (*C. heavisidii*) and moving eastward to South America (*C. eutropia* and *C. commersoni*) and New Zealand (*C. hectori*). Their affinity for near-shore river effluences, small group size, and high levels of site fidelity (Bräger & Schneider 1998) are thought to have contributed to the diversification of endemic species following the colonization of each region. In contrast, the dusky dolphin is a wide-ranging, meso-pelagic species that forms large social aggregations that are capable of long-distance movements. Nevertheless, congruent

patterns of distribution of the dusky dolphin with *Cephalorhynchus* species suggest that the same dispersal mechanisms, namely the WWD, were responsible for dispersal and colonization events that resulted in the contemporary distribution of dusky dolphin populations (Pichler et al. 2001).

There is evidence that paleoceanographic conditions played a role in the phylogeography of the dusky dolphin. Studies of molecular and morphological variation support isolation of Peruvian populations from all other regions, as would be expected with isolation of the Atlantic and the Pacific during the last glacial maximum (LGM) (Cipriano 1997; Cassens et al. 2003). Significant differences in skull morphology and parasite loads of Peru have been used as evidence to support the origin of the dusky dolphin in the Pacific Ocean (Van Warebeek 1993). In contrast, analysis of complete cytochrome *b* sequences from South Africa, Peru, and Argentina suggested a more basal position of Atlantic lineages in the gene tree of the dusky dolphin (Cassens et al. 2003). There are, however, patterns of population structure that allude to a complex phylogeographic history of the dusky dolphin history that do not fit the predictions of the WWD hypothesis. Cassens et al. (2003) revealed significant genetic discontinuity between Southwest Africa and Argentina, yet found no obvious geographic partitioning of mtDNA variation. Although focused on New Zealand, Harlin et al. (2003) found two distinct clades, one of which appeared to represent a more recent lineage expansion with the other possibly representing more ancestral mitochondrial lineages.

Clearly, the WWD and alternative hypotheses cannot be addressed adequately until all four biogeographic areas (especially New Zealand) containing the dusky dolphin

have been adequately incorporated into a detailed analysis. Furthermore, an assessment of recent and historical patterns of gene flow requires the use of both nuclear and mitochondrial markers, and these markers should be analyzed using similar methods. This study provides a global phylogeographic perspective by incorporating a large number of New Zealand samples along with samples from the other three biogeographic areas. The **third objective** of this dissertation is to use both nuclear and mitochondrial markers to test hypotheses related to: 1) the origin of dusky dolphin populations; 2) the phylogeographic history of those populations; and 3) the degree to which recent and historical processes have shaped current patterns of genetic divergence within and between populations.

CHAPTER II

GENETIC STRUCTURE, DIVERSITY, AND HISTORICAL DEMOGRAPHY OF NEW ZEALAND'S DUSKY DOLPHIN (*LAGENORHYNCHUS OBSCURUS*)[*]

Introduction

Cetaceans generally have high dispersal potential (Baker et al. 1993; Rosel et al. 1995; Bakke et al. 1996), thus decreasing the likelihood of genetic differentiation between populations. Nevertheless, geographic distance is not necessarily correlated with genetic divergence in cetaceans (Hoelzel 1998). Photo-identification (Würsig and Jefferson 1990), behavioral data, and studies of mitochondrial DNA (mtDNA) lineages have revealed cases of local divergence between intraspecific populations (Hoelzel & Dover 1991; Dowling & Brown 1993; Garcia-Martinez et al. 1995; Hoelzel et al. 1998a, b; Pichler et al. 1998). Genetic divergence has been shown to occur over small geographic scales with gene flow limited by either geographic barriers (Dowling and Brown 1993), resource heterogeneity (Hoelzel & Dover 1991; Hoelzel 1998; Hoelzel et

---

[*]Reprinted with permission from "Genetic structure, diversity and historical demography of New Zealand's dusky dolphin": by AD Harlin, CS Baker, RL Honeycutt TM Markowitz, and B Würsig, 2003. *Journal of Mammalogy* 84(2):702-707. 2003 by Allen Press, Inc.

al. 1998a, b), habitat specialization (Pichler et al. 1998), or feeding- or breeding-site fidelity (Shane et al. 1986; Baker et al. 1990, 1993; Duffield & Wells 1991).

The dusky dolphin, *Lagenorhynchus obscurus*, is a southern hemisphere species that generally prefers deep, temperate waters along the edge of continental shelves (Würsig et al. 1997). This preference for cooler mesopelagic waters has resulted in a discontinuous, regional distribution along the coasts of South Africa, South America, and New Zealand at approximately 42º southern latitude. In New Zealand, dusky dolphins occur around the North and South Islands from Stewart Island in the south, Hawke Bay in the north, and the Chatham Islands in the east (Würsig et al. 1997). Historical sightings suggest that groups are rare around the North Island but are common and abundant on the east coast of the South Island where deeper, cooler waters approach close to shore. This is especially true off the coasts of the Kaikoura and Otago Peninsulas, where large groups are seasonally common. Both Otago and Kaikoura have short continental shelves that bring deep, nutrient-rich waters close to shore. In Kaikoura, the rich waters of the subtropical convergence support large populations of fishes and squid (Boyd et al. 1999) that comprise the majority of the dusky dolphin diet (Cipriano 1992). Although abundant in these 2 regions, dusky dolphins are rarely seen along the shallow shelf waters of the Chatham Rise that separate Kaikoura from Otago and extend eastward 467 nautical miles to the Chatham Islands (Würsig et al. 1997). In contrast to the east coast, anecdotal evidence and casual observations during short coastal surveys (Bräger and Schneider 1998) suggest dusky dolphin groups on the west coast of the South Island are comparatively small and rare. The northern half of the west

coast, like the Chatham Rise, has a wide, shallow continental shelf. The pattern of increased abundance of dusky dolphins in the deep, coastal waters of New Zealand also is seen in Chile and Peru, where a narrow continental shelf and deep near shore waters appear to be correlated with large aggregations of dusky dolphins (Würsig et al. 1997).

Although the general distribution of New Zealand's dusky dolphins has been documented, little is known regarding population structure throughout the region. For instance, seasonal difference in group membership, size, behavior, and distribution support "summer" and "winter" groups in Kaikoura that are segregated temporally and behaviorally (Würsig et al. 1997). Preliminary photo-identification data show a seasonal trend to re-sightings of individuals in Kaikoura, with >50% occurring in the same season and 9 of 10 in the same or adjacent seasons (Markowitz 2004). Photo-identification data also support the conclusion that at least some individuals photographed in Kaikoura during the summer move 250 km north of Kaikoura to the Marlborough Sounds during winter and return again to Kaikoura the next summer (Markowitz 2004). These data support a general seasonal shift in distribution from the south in summer to north in winter in response to changes in water temperature (Gaskin 1968, 1972; Leatherwood 1991; Würsig et al. 1991; Cipriano 1992). Although there appears to be little overlap in group membership between summer and winter months in Kaikoura and Marlborough, no information is available on the genetic distinction of these seasonal cohorts. Furthermore, it is unknown if the presence of shallow waters (e.g., the Chatham Rise) and seasonal movement patterns have acted as physical and behavioral barriers to

dispersal. If gene flow were limited, division of dusky dolphins into genetically diagnosable populations along the New Zealand coast would be expected.

Intraspecific population structure often is examined by overlay of geography on haplotype trees generated from mtDNA sequence data (e.g., phylogeography, Avise 1995). However, this type of analysis alone can be misleading if sample size is not adequate to provide statistical significance or if sampling does not adequately cover the geographic range of the species (Templeton 1998). Allelic markers (e.g., microsatellites) can increase statistical power to detect a difference between populations, but often lack an evolutionary framework in which to examine historical processes (such as population fragmentation and range expansion) responsible for current patterns (Templeton et al. 1995; Templeton 1998). Furthermore, tests for partitioning of allelic variance (e.g., *Fst*, Wright 1931, 1943; AMOVA, Excoffier et al. 1992) can erroneously assess amount of gene flow, especially if populations have shared a recent ancestry (Templeton 1998). Nested clade analyses developed by Templeton et al. (1987, 1992, 1995) provide statistical tests for association of haplotype variation and geography in a phylogenetic framework. This approach uses a statistical parsimony network and the number of steps linking haplotypes to examine the spatial distribution of haplotypes (Templeton 1998). Statistically significant associations are interpreted using an inference key (Templeton 1998) based on patterns predicted by coalescence theory (Crandall and Templeton 1993; Castelloe and Templeton 1994). With this technique, it is possible to examine both population structure and history and to employ statistical tests of different predictions related to population range expansion and fragmentation.

Genetic differentiation can occur very rapidly when populations are in decline. In such cases, isolation due to reduction in range size in combination with accelerated extinction of DNA lineages (Avise et al. 1984) can result in greater differentiation between than within populations over short evolutionary time and on small geographic scales. Therefore, low levels of genetic diversity are often indicative of long-term past or contemporary population reduction. Tests for geographic range expansion with nested-clade analysis, although informative, do not allow interpretation of historical demographic patterns (e.g., population equilibrium or growth over evolutionary time) that accompanied the expansion. Additional tests of historical population demography, such as lineages-through-time (Nee et al. 1995) or analysis of pair-wise differences (i.e., "mismatch distribution", Slatkin & Hudson 1991), can add a demographic perspective to interpretation of population history and are potentially important to understanding intraspecific population patterns.

In this study, I investigate the geographical distribution of mtDNA control region haplotypes of the dusky dolphin in 4 main regions of their distribution along New Zealand's South Island (Kaikoura Peninsula, Otago Peninsula, Marlborough Sounds, and West Coast). Nested clade and lineages-through-time analyses are combined with standard phylogenetic approaches and information on natural history (e.g., from photo-identification and behavioral observations) to investigate changes in population size, patterns of genetic variation, and the overall influence of historical and demographic effects on current patterns of population structure in New Zealand.

Materials and methods

Sample collection

From 20 July 1997 to 26 July 2000, exfoliated skin was collected from bow riding dolphins with a scrub pad affixed to a wooden dowel ("skin swabbing") as described in Harlin et al. (1999) from 4 regions throughout New Zealand (Table 1, Fig. 1). Additional samples were from dolphins cast onto beaches that were collected for necropsy (Table 1). Vessel position with respect to the group was changed regularly to prevent repeated sampling of individuals. Observations of scars, pigmentation patterns, and dorsal fin notches suggested that the composition of bow-riding individuals was fluid, as dolphins frequently replaced one another from the larger group. Skin samples were stored in 20% dimethylsulfoxide (DMSO) solution saturated with salt (Amos and Hoelzel 1991). Sample size ranged from 14 to 106 per locality (Table 1, Fig. 1).

DNA extraction and amplification

Total genomic DNA was isolated from beach-cast specimens with a standard protease digestion and phenol-chloroform protocol (Sambrook et al. 1989). Skin swab samples were extracted with a similar protocol modified for accommodation of the sampling pad (Harlin et al. 1999) or with a silica-based spin-column procedure (Qiagen, Valencia, California). DNA was suspended in 100 ml of 1X TE (10 mM Tris, 1mM EDTA), pH 8.0. Negative controls were included in all extractions.

A fragment of approximately 500 base pairs of the 5' end of the mtDNA control region was amplified by the polymerase chain reaction (PCR, Saiki et al. 1988) for 169

samples from throughout New Zealand (Table 1). The PCR was performed in 50 μl reaction volume containing 10X Tris-HCl (pH 8.8), 2.5 mM $MgCl_2$, 200 mM deoxynucleotide triphosphates (dNTP's), 0.2 M of each oligonucleotide primers, and 1 unit of *Taq* DNA polymerase. PCR primers were those of Baker et al. (1996), tPro (5'-TCA CCC AAA GCT GRA RTT CTA-3') and Dlp5 (5'-CCA TCG WGA TTT CTT ATT TAA GRG GAA-3'). The PCR conditions included 94° C for 2 min followed by 35 cycles at 92° C for 30 s, annealing at 52° C for 30 s, and extension at 72° C for 30 s.

Table 1. Pairwise distances between 4 regions along South Island coast of New Zealand. KK=Kaikoura, OT=Otago, MS=Marlborough Sounds, WP=Westport, JB=Jackson Bay.

Locality	Pairwise distances (km)				Sample size	Geographic location
	KK	OT	MS	WC		
Kaikoura	-				106	042°28' S 173°33' E
Otago	550	-			26	045°50' S 171° 01' E
Marlborough Sounds	250	724	-		23	040°53' S 173°56' E
West Coast	650	1043	426	-	14	041°57' S 171°20' E

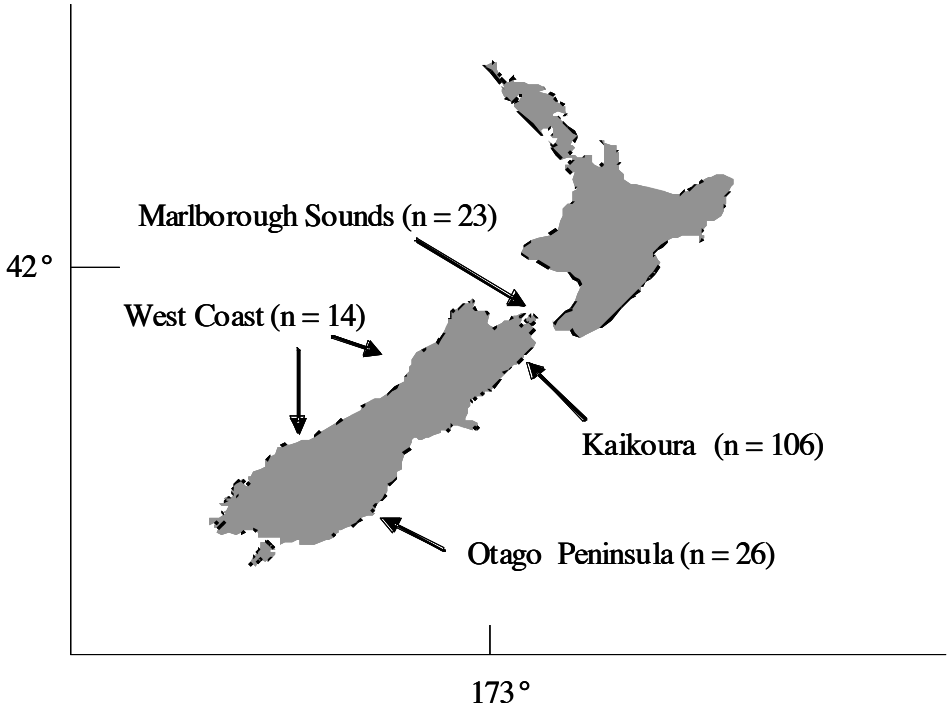


Figure 1. Sampling localities in New Zealand (n = sample size).

PCR products were visualized on a 1.6 % agarose and Tris-borate EDTA (TBE) gel matrix. Prior to sequencing, excess primers and dNTP's were removed using either High Pure (Roche Diagnostics Corporation, Indianapolis, Indiana) or Qiaquick (Qiagen, Valencia, California) spin columns. Amplified PCR products were sequenced with Big Dye termination chemistry (ABI Biosystems, Foster City, California) on an Applied Biosystems (ABI Biosystems, Foster City, California) automated sequencer (either model 373 or 377) following protocols of the manufacturer (Perkin-Elmer Biosystems, Wellesley, Massachusetts). Sequences were aligned using Clustal W software (version 1.6, Thompson et al. 1994). A minor adjustment was made to the sequence alignment with the MacClade software data editor (version 3.06, Maddison & Maddison 1992) to

incorporate a single insertion or deletion of 1 nucleotide. Forward and reverse strands were sequenced for all haplotypes.

Genetic evaluation of sampling design

Measures of population subdivision can be influenced by group structure if samples are not drawn at random from the population (Ross 2001). To assess the potential for sex-bias in the sampling regime, gender was determined for a total of seventy-eight individuals: thirty-six from Kaikoura, fourteen from Otago, eighteen from Marlborough Sounds, and ten from the West Coast. Gender determination was done by simultaneous amplification of the X-chromosome zinc-finger protein (ZFX) and the Y-chromosome sex-determining region (SRY) following the protocol of Banks et al. (1995) and is available from authors upon request.

To exclude the potential of sampling the same dolphin more than once, a combination of microsatellite allelic diversity and distinguishing marks on the flank, face, or dorsal fin were used to discriminate between individuals. Three DNA microsatellite loci (415/416, Schlötterer et al. 1991; EV 94, Valsecchi & Amos 1996; TAA-31, Palsbøll et al. 1997) were amplified for 25 individuals with the same mitochondrial DNA haplotype. Fluorescent dye-labeled primers were used for PCR amplification, and genotyping was conducted on an ABI Biosystems 373 automated sequencer. Allele sizes were determined by comparison to internal size standards with the software program GeneScan (version 1.1, ABI Biosystems, Foster City, California). In addition, 39 individuals were identified by sample collection in tandem with photographs of sampled individuals. None of the individuals that shared mtDNA

haplotypes had the same microsatellite allele profile. Comparison of photographs of sampled individuals revealed no inadvertent re-sampling of dolphins with unique markings. Therefore, it was assumed that no individual was sampled more than once.

Genetic analyses

Standard diversity indices, such as nucleotide diversity (π), haplotype diversity (*h*, Nei 1987), and number of polymorphic sites were estimated. Diversity indices were calculated for all 169 individuals. New Zealand-wide estimates of diversity were used in subsequent calculations of demographic statistics. All diversity indices were estimated via the Arlequin software package (version 2.0, Schneider et al. 2000).

Genetic differentiation between 4 regions around the South Island (Table 1, Fig. 1) was investigated with an analysis of molecular variance (AMOVA, Excoffier et al. 1992). One thousand non-parametric permutations were performed to test for significant differences in calculated ϕ_{sr} values. In addition to an AMOVA, an exact test of population subdivision based on haplotype frequencies (Raymond & Rousset 1995) was performed with the Arlequin software package (version 1.1, Schneider et al. 1997).

Both maximum parsimony and distance-based analyses of relationships among haplotypes were performed using PAUP* (version 4.063a, Swofford 1999). Parsimony analyses used the heuristic search option, tree bisection-reconnection branch swapping, and ten random additions. Given the large number of equally parsimonious trees (greater than 4,000), an unrooted neighbor-joining phylogram (Saitou & Nei 1987) was constructed using Tamura-Nei distances (Tamura & Nei 1993) corrected for variation in substitution rate among sites. The frequency of multiple substitutions per site was

assumed to follow a gamma distribution. A discrete method (Yang & Kumar 1996) with 8 rate classes was used to estimate the shape-parameter (alpha, Yang 1996) of the distribution by maximum likelihood evaluation of an uncorrected neighbor-joining phylogeny in the program PAUP*. Groups identified by neighbor-joining analysis were evaluated with five hundred bootstrap replicates (Felsenstein 1985), via heuristic search with tree bisection-reconnection and random addition of haplotypes.

The frequency, spatial distribution, and evolutionary history of the mtDNA control region haplotypes were used to test the null hypothesis of random geographic association of maternal lineages-through-time. This analysis of nested clades (Templeton et al. 1987, 1992, 1995) allows one to examine population structure and history, and to interpret the processes (range expansion, fragmentation, isolation by distance, etc.) that relate to current patterns of haplotype distribution. Prior to nested clade analysis, a haplotype network derived under the criteria of maximum parsimony is generated and subsequently nested into a series of clades separated by 1 substitution ("step") increments (Templeton et al. 1992). The average distance that clades (or haplotypes) within a nested level lie from the geographic center of distribution for all members of that clade that bear a particular haplotype are calculated for all members of a given x-step clade (clade distance, D_c) and the higher-level nested clade (nested clade distance, D_n) that contains the x-step clade. These distances (D_c and D_n), as well as the geographic distance between members of tip and interior clades (or haplotypes) that bear the same haplotype (interior to tip distances, I-T), are randomly permuted to create a null distribution for comparison to observed values. Historical processes were inferred in

this study with a standardized key (Templeton 1998) based on distances within clades that were significantly larger or smaller than expected at the $P \leq 0.10$ level.

A statistical parsimony network of the control region haplotypes was created with the program TCS (version1.13, Clement et al. 2000). Haplotypes in the 0-step network were nested in a hierarchical series of 1-step, 2-step, etc. clades until the entire network was nested into a single clade following the rules outlined in Templeton et al. (1992). The resulting nested network was used in nested clade analysis. Statistical tests of geographic association of nested clades were performed with the program GeoDis (version 2.0, Posada et al. 2000) with 10,000 replications. Because dolphins must use aquatic corridors to move between coastal regions, a pairwise distance matrix of the shortest distances between the centers of the 4 South Island regions was used to estimate D_c, D_n, and I-T distances (Table 1).

Population demography was examined with lineages-through-time analysis (Nee et al 1995). This method generates a plot of the log-transformed number of coalescence events per unit time in a phylogeny, and compares this to the frequency of cladogenesis expected from a population at equilibrium. The shape of the lineages-through-time curve indicates periods of either exponential growth, constant population growth, or decline, i.e., a convex curve indicates exponential growth; a concave curve suggests a constant rate of change over time. A log-likelihood ratio test (Felsenstein 1981, 1988) failed to find a significant difference between Tamura-Nei neighbor-joining topologies with and without a molecular clock constraint ($P = 0.14$); therefore, lineages-through-time analysis was performed under the assumption of an equal rate of evolution among

lineages. A plot of the number of lineages-through-time was generated from a phylogenetic tree reconstructed with the UPGMA ultrametric clustering algorithm in PAUP*, based on Tamura-Nei distances with correction for among site rate heterogeneity (alpha = 0.10). The probability (P_k) that a lineage at time t will have k bifurcations at the present was calculated for all lineages to test the null hypothesis of population stability. Lineages-through-time analysis was performed with the program End-Epi (version 1.0, Rambaut et al. 1994).

In addition to tests for constant population size, we calculated the effective population size of female dusky dolphins in New Zealand. Watterson (1975) showed that $\theta = 4N_e\mu$, where the parameter θ estimates nucleotide diversity (π, Nei 1987; Nei & Li 1979), N_e is the effective population size, and μ is the per-lineage nucleotide substitution rate per generation for a particular nucleotide sequence. Whereas nucleotide diversity is an estimate of θ, it holds true that at equilibrium $E(\pi) = \theta = 4N_e\mu$ for nuclear markers (Rooney 1998). It follows that $E(\pi) = \theta = 2N_f\mu$ for mitochondrial DNA where N_f is the number of breeding females.

In order to calculate female effective population size, dusky dolphin and harbor porpoise control region sequences were compared to estimate the mutation rate (μ) of the hypervariable I and flanking regions of the dusky dolphin lineage. The rate of nucleotide substitution per lineage per year (λ) was estimated with the equation $\lambda = d/2T$, where d was the number of nucleotide substitutions per site and T was the time since divergence (Li 1997). The per-lineage rate of nucleotide substitution per generation (μ) was obtained by multiplying λ by generation length. The average number

of nucleotide substitutions per site (d) was estimated with pairwise, gamma-corrected, Tamura-Nei distances (alpha = 0.10) between a harbor porpoise and 169 dusky dolphin control region sequences. The date of the Phocoenidae and Delphinidae divergence (T) was estimated from the fossil record to be between 1.0×10^7 and 1.1×10^7 years ago (Barnes 1985). The resulting range of λ was multiplied by a conservative estimate of dusky dolphin generation time of ten years to obtain μ. Nucleotide diversity (π) and μ were used to estimate a range of effective female population sizes for New Zealand dusky dolphins.

Results

Genetic diversity

Of the 169 dusky dolphin control regions examined, 62 polymorphic sites defined 76 haplotypes. Of these 76 haplotypes, 45 were found only once, and the most common (H) shared by 16 individuals. All substitutions were transitions except for 1 insertion or deletion and 3 transversions. Haplotype diversity was 0.972, and nucleotide diversity was 0.022 ± 0.011. Mean number of pair-wise differences was 10.3 ± 4.73. Nucleotide composition was typical for cetacean control region sequences (Hoelzel et al. 1991), with nucleotides A and T represented in greatest proportion (30.62% and 34.76%, respectively), followed by C (21.49%) and G (13.13%).

Seasonal sex ratios

Gender determination revealed a seasonal bias to group composition. The sex ratio was 1:1 for samples collected in Kaikoura (n = 36) and for all 78 samples combined

from all regions. Samples from Otago, Marlborough Sounds, and the West Coast deviated from the expected 1:1 ratio. In Otago and the West Coast, the ratio of females to males was 13:1 and 7:3, respectively and in Admiralty Bay, Marlborough Sounds, all 18 individuals examined were males. These skewed sex ratios coincide with the season in which the samples were collected; i.e., samples were collected in Otago and the West Coast during the breeding season (summer) when females and calves are known to be aggregated closer to shore (Würsig et al. 1997). Samples from the Marlborough Sounds were collected from all-male groups during the non-breeding season (winter).

Regional population structure

No significant partitioning of variance between the 4 geographic regions of the South Island was observed (ϕ_{sr} = -0.041, P = 0.13, 1,000 permutations), and an exact test (Raymond and Rousset 1995) revealed no evidence of population differentiation (P = 0.57). The neighbor-joining phylogram contained 2 subgroups with distinct branching patterns (Fig. 2) that corresponded to similar patterns in the nested clade network (Figs. 3, 4). Subgroup A (Fig. 2) and clade 5-1 (Fig. 4) were characterized by long branches and internodes and as many as twelve missing intermediates between haplotypes. In contrast, subgroup B (Fig. 2) had short internodes and branches that corresponded to clade 5-2 (Fig. 4) with few unsampled intermediates. A 50% bootstrap consensus of the neighbor-joining phylogeny failed to support relationships among haplotypes. Visual inspection of the neighbor-joining tree did not find an association of haplotypes with geographic locality.

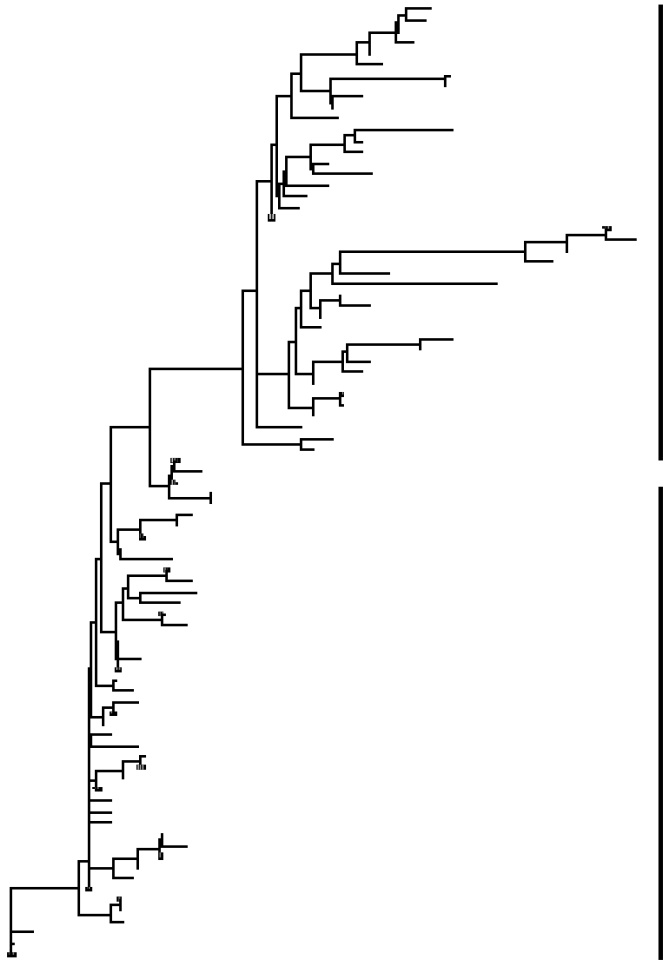


Figure 2. Unrooted neighbor-joining phylogram of 169 mtDNA control region haplotypes from dusky dolphins. Phylogram was generated using Tamura-Nei distances corrected for among-site rate heterogeneity ($\alpha = 0.10$). Branch-length scale is in units of substitutions per site. Subgroup A corresponds to haplotypes in clade 5-2 (Fig. 4) and is characterized by long branches and multiple unsampled intermediates. Subgroup B corresponds to clade 5-1 (Fig. 4). The short branches, internodes, and low frequency of missing intermediates in this group suggest recent population expansion.

Haplotype network and nested clade analysis

Figure 3 represents the 0-step statistical parsimony network of control region haplotypes. Three reticulations, or alternate equally parsimonious connections, were present within the 0-step haplotype network. Two of these reticulations (R1 and R2) were unambiguously resolved with the rules outlined by Templeton and Sing (1993) and Crandall and Templeton (1993) (Fig. 3). One reticulation (R3) could not be unambiguously resolved within the 0-step network, but subsequently was collapsed in the formation of 1-step clades (Fig. 4). The final network contained a series of 5 nested levels (Fig. 4). Nested chi-square contingency analyses found significant relationships among the spatial distribution of haplotypes for clades 3-2 ($P = 0.07$), 3-6 ($P = 0.009$), and 5-2 ($P = 0.09$). Furthermore, 10,000 permutations of clade (Dc) and nested clade (Dn) distances suggested that the distribution of haplotypes in New Zealand was significantly different from random within clades 2-23 (alpha ≤ 0.05; Fig. 4) and 1-38, 3-2, 3-8, 4-4, 5-1, 5-2 (alpha ≤ 0.10; Fig. 4). A summary of the nested-clade analysis results and interpretations are summarized in Table 2 and presented in Figure 5.

Historical demography

The presence of 2 convex curves in the log-transformed plot of lineages-through-time (Fig. 6A), separated by a smooth, horizontal curve, indicates at least 1 old, and perhaps another relatively recent, expansion with a period of stasis between periods of

Table 2. Summary of inferences regarding demographic events deduced from clades with significant nested clade values. Clade 2-23 is nested within 3-2, 3-2 within clade 5-2, and clade 3-8 within 5-1. TIP = Tip clade, INT = Interior clade. A greater-than or less-than symbol indicates a Dc or Dn value that is significantly larger or smaller than expected if haplotypes were distributed randomly. P-values indicate probability that the Dc or Dn estimated from these data were observed by chance. Inferences were drawn from the nested clade analysis interpretation key of Templeton (1998). The steps in the chain of inference can be examined by comparison to this key.

Clade	χ^2	Nested clades	D_c	D_n	Chain of inference	Demographic event
2-23	$P = 0.11$	1-34 (TIP) 1-33 (TIP) 1-32 (TIP) 1-36 (INT) I-T	579.2, >, $P = 0.01$ n.s.[A] n.s. n.s. -99.6, <, $P = 0.02$	349.3, >, $P = 0.07$ n.s. n.s. n.s.	1 YES, 2 NO, 11 YES, 12 NO	Contiguous range expansion
3-2	$P = 0.07$	2-23 (INT) 2-2 (TIP) 2-7 (TIP) I-T	245.4, <, $P = 0.06$ n.s. n.s. -132.5, <, $P = 0.09$	253.7, <, $P = 0.06$ n.s. 341.5, <, $P = 0.05$ -38.4, <, $P = 0.07$	1 YES, 2 NO, 11 YES, 12 NO	Contiguous range expansion
3-6	$P = 0.01$	n.s	n.s	n.s	No tip clades to infer process[B]	Non-random spatial distribution of haplotypes
3-8	n.s	2-12 (INT) 2-13 (TIP) 2-10 (TIP) 2-9 (TIP) I-T	n.s. n.s. n.s. n.s.	514.3, >, $P = 0.10$ n.s. n.s. n.s. 211.7, >, $P = 0.06$	1 YES, 2 NO, 11 NO, 17 YES, 4 YES, 9 NO	Past fragmentation

Table 2. Continued.

Clade	χ^2	Nested clades	D_c	D_n	Chain of inference	Demographic event
5-1	n.s	4-4 (INT)	268.9, <, $P = 0.05$	276.2, <, $P = 0.07$	1 YES, 2 NO, 11 YES, 12 NO	Contiguous range expansion
		4-2 (TIP)	408.7, >, $P = 0.10$	342.8, >, $P = 0.07$		
		I-T	-139.7, <, $P = 0.07$	-66.7, <, $P = 0.08$		
5-2	$P = 0.09$	4-1 (TIP)	n.s.	n.s.	1 YES, 2 NO, 11 NO, 17 YES, 4 YES, 9 NO	Past fragmentation
		4-5 (INT)	n.s.	329.2, >, $P = 0.08$		
		4-3 (TIP)	89.3, >, $P = 0.10$	n.s.		
		I-T		26.0, >, $P = 0.09$		

A. Indicates results are not significant at $P < 0.10$.

B. Results of chi-square were significant ($P < 0.05$), but lack of tip clades prevents nested clade inference

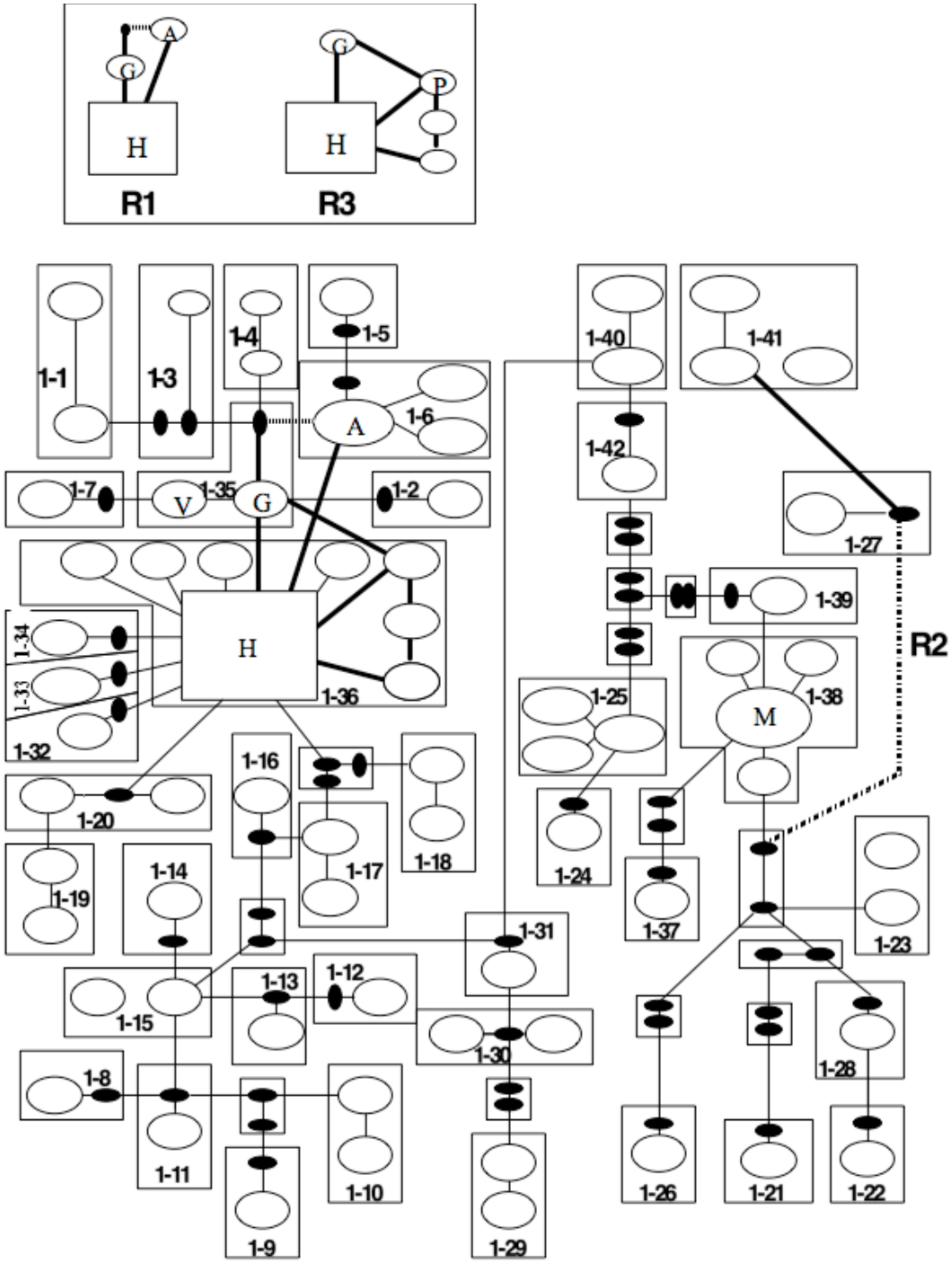


Figure 3. Nested clade network of haplotypes. Solid circles designate unsampled intermediates; nested haplotype groups are enclosed in rectangles. Level of nesting is given as 1-*x* for 1-step clades, 2-*x* for 2-step clades, etc., where *x* identifies individual clades. Size of shapes is relative to frequency. Reticulate connections are indicated by thick lines and are labeled R1, R2 and R3. Hashed branches indicate where R1 and R3 were broken. Insert at top left gives the details of R1 and R3.

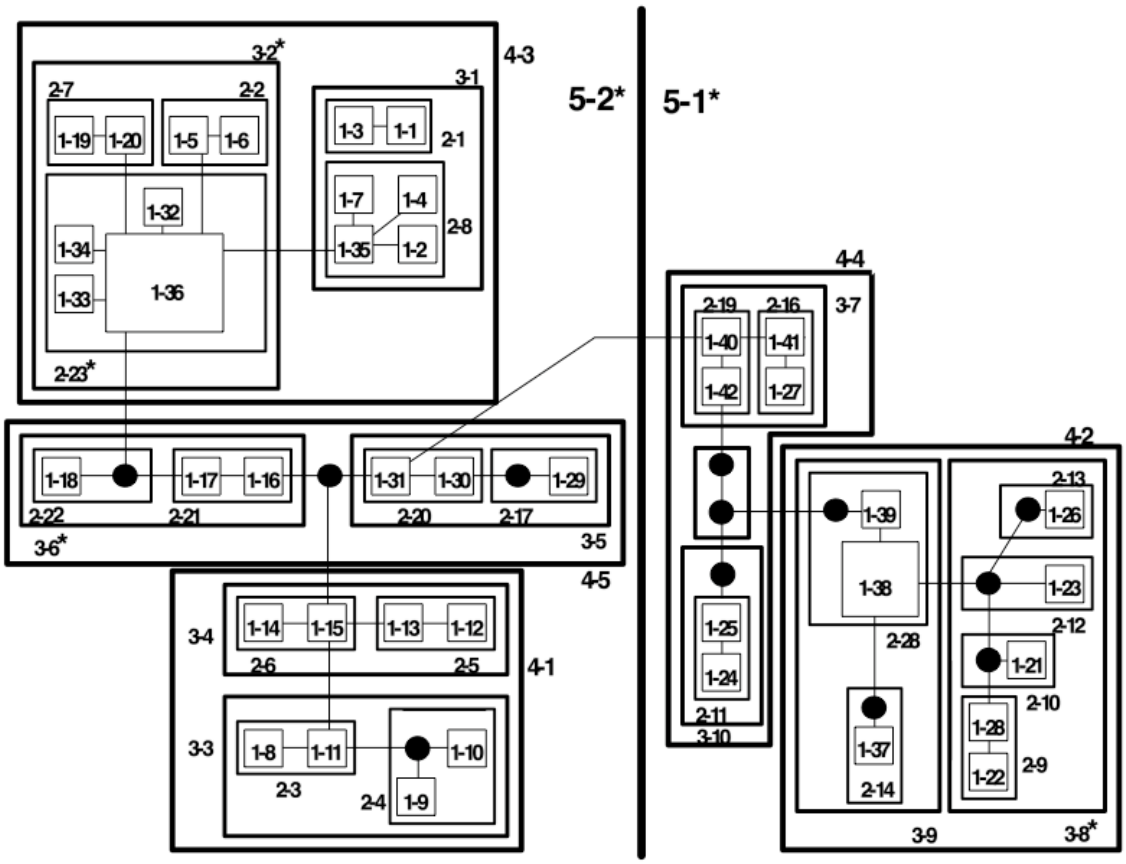


Figure 4. Continuation of nested clade network, Levels 1-5. Each solid circle indicates 4 mutational steps. Asterisks indicate clades with geographic distribution significantly different from random (see Table 2). Symbols as in Fig. 3.

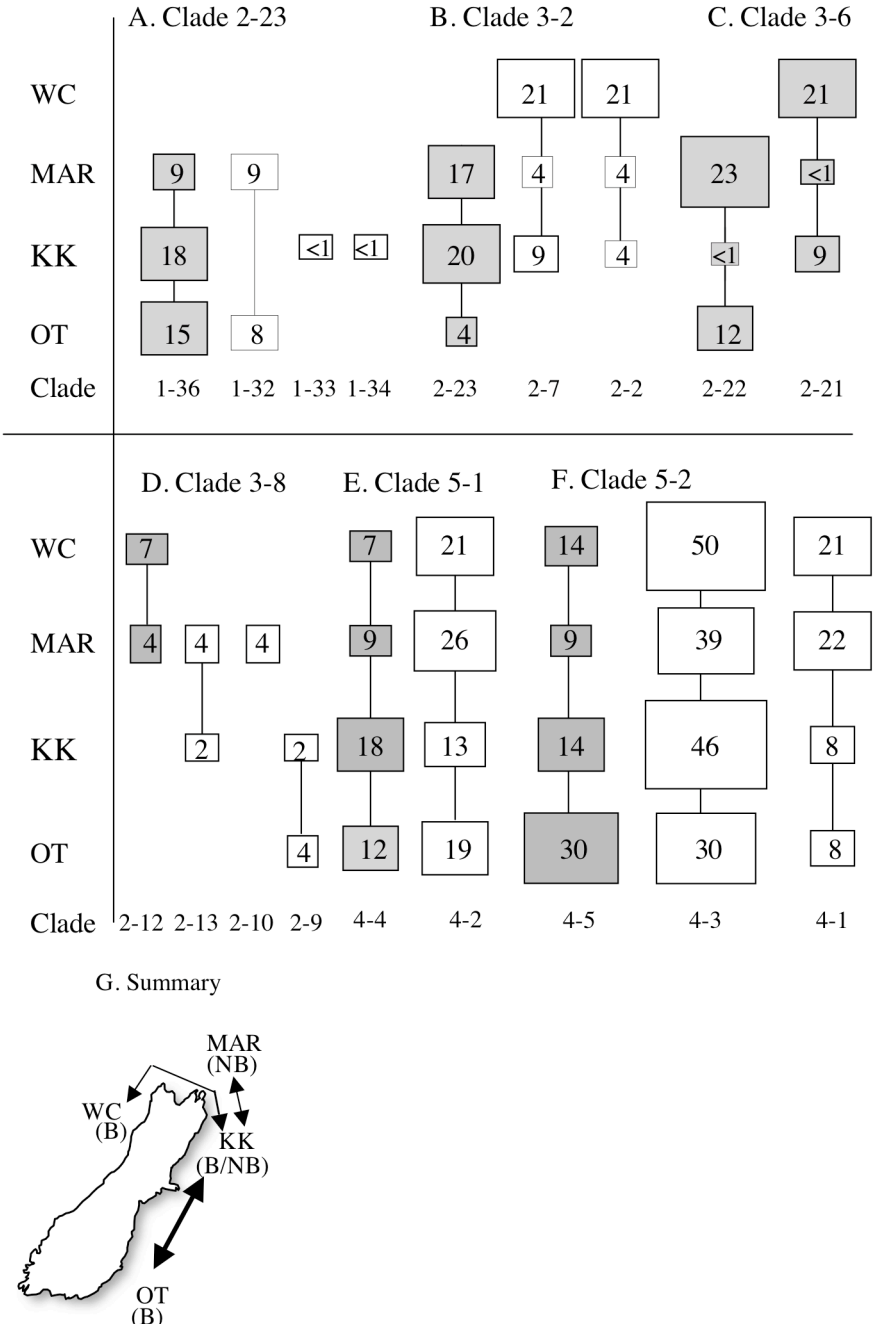


Figure 5. Summary of demographic events supported with nested clade analysis. A-F) Nested clades with significant evidence of a demographic events. Numbers within squares represent relative frequency of haplotypes, expressed as percentage of total number of samples collected in each region that are present in a given clade. Square size is proportional to relative frequency. Shaded squares are interior clades; open squares are tip clades. Distance between squares is not indicative of geographic distance. Clades 2-23, 3-2, 5-1 and 3-8, 5-2 represent 2 series of nested clades (Fig. 4). OT= Otago, KK = Kaikoura, MAR = Marlborough, WC = West Coast. C) Clade 3-6 showed significant patterns in the chi-square contingency test only. G) The summary of range expansion events represented in A-F. B = breeding season, NB = non-breeding season.

rapid growth. The rate of cladogenesis was greater than expected for all but 1 internal branch ($P_k < 0.001$, Fig. 6B). The within-lineage mutation rate (λ) for the 5' end of the dusky dolphin control region was estimated at 6.3 - 7.0 x 10^{-8} substitutions per site per year based on a 10-million-year divergence date. This estimate is slightly less than for humans (Vigilant et al. 1991 as modified by Nei 1992) and is several fold faster than the minimum estimate for delphinids (Hoelzel et al. 1991) and baleen whales (Baker et al. 1993, but see Rooney et al. 2001). Using mutation rates (μ) from 6.3 - 7.0 x 10^{-7}, the female effective population size was estimated to be between 15,700 and 17,500 individuals.

Discussion

Regional population structure and range expansion

Both AMOVA and the neighbor-joining haplotype phylogeny failed to demonstrate significant geographic structure relative to the distribution of mtDNA haplotypes throughout New Zealand. Although neighbor-joining analysis and tests of population structure failed to find genetic differentiation among localities in New Zealand, nested clade analysis (Templeton 1998) indicated significant association of haplotypes with geography within 2nd, 3rd, and 5th level clades (Table 2). This nonrandom distribution of haplotypes was interpreted with Templeton's inference key to be a result of both contiguous range expansion and population fragmentation (Table 2).

The pattern of haplotype distribution across nested levels supports a general north and south shift in relative frequency of haplotypes between regions (Fig. 5). All

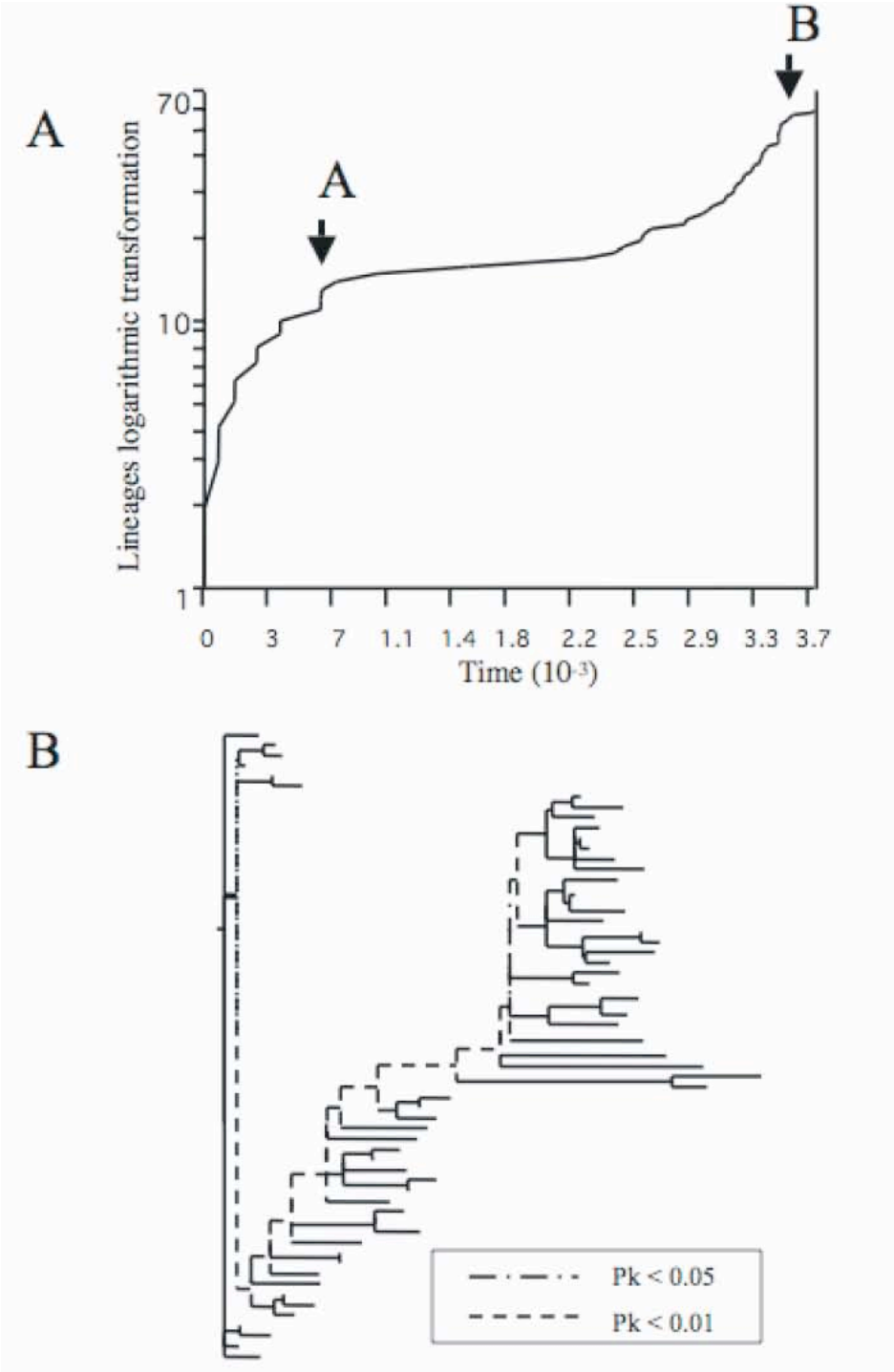


Figure 6. A) Graph depicting results of lineages-through-time analysis. The 2 convex points, A and B, suggest 2 population expansions (Nee et al. 1995). B) Dendrogram derived from the UPGMA clustering algorithm. Distances were estimated via the method of Tamura-Nei with gamma correction (alpha = 0.10) for among-site rate heterogeneity. Hashed branches indicate greater than expected rate of lineage diversification at $P < 0.001$

nested clades with significant evidence of range expansion and fragmentation contained haplotypes from Kaikoura (Fig. 5), which were placed within internal clades (Figs. 3, 4) more often than other geographic regions. Furthermore, the region furthest from Kaikoura (i.e., Otago) had haplotypes present more often in tip clades (Fig. 5). Given these results, coalescence theory predicts that Kaikoura was the center of a range expansion along the New Zealand coast (Crandall & Templeton 1993) either to Otago or to the Marlborough Sounds and the West Coast (Fig. 5). All of these results are related to the distribution and abundance of mtDNA haplotypes associated with Kaikoura and populations in other regions.

Taken at face value, the nested-clade analysis suggests geographic structure to the mitochondrial data, reflecting old events as seen by the level of nesting associated with many of these expansions and fragmentations. Nevertheless, there is an alternative explanation that relates more to the seasonal biology of the species. Dusky dolphins are well-known for being highly vagile and regularly make seasonal movements over long distances. There is evidence for this along the coasts of Argentina (Würsig & Bastida 1986) and New Zealand (Würsig et al. 1997). In New Zealand, photographic data indicate dusky dolphins make regular, 500 km round-trip migrations between Kaikoura and the Marlborough Sounds and longer, less frequent movements between Kaikoura and the West Coast (Markowitz 2004). In addition to seasonal migrations, dusky dolphins also alter group composition between breeding and non-breeding seasons. Gender determination suggests the Marlborough Sounds is a region for post-breeding adult males to over-winter. Furthermore, the majority of individuals in small groups in

Otago and the West Coast during the breeding season are female. Only in Kaikoura, where dolphins were sampled in all seasons, does the sex ratio equal the expected 1:1. If seasonal migrations north and south along the coastline result in temporal segregation of maternal lineages, these movements could produce a pattern of haplotype distribution that appears similar to range expansion, depending on the season and location where samples were collected. This would be especially true if seasonal migrations were temporary, long-distance movements of genetically defined demes (e.g., matrilineal), as observed in female humpback whales (*Megaptera novaeangliae*) during the breeding season (Baker et al. 1993). Further investigation into these seasonal patterns may provide important information regarding the behavioral biology of these dolphins along the New Zealand coast.

The inability of distance methods and tests for partition of genetic variance to detect the differentiation along the New Zealand coast supports Templeton's (1998) claim that such methods alone can be inadequate to evaluate intraspecific population structure. However, in cases where fine-scale temporal segregation of populations such as those due to seasonal changes in distribution results from nested clade analysis should be interpreted with caution. Conclusions derived from nested clade analysis will be misleading unless sample collection is designed to account for seasonal shifts in distribution. In cases where the seasonal movement patterns of a species are unknown, the interpretation of nested clade analysis results becomes unreliable.

Population expansion

An examination of the lineages-through-time plot suggests a period of demographic expansion within New Zealand (Fig. 6A). These results support a previous study by Harlin (1999), which interpreted a bimodal distribution of pair-wise differences between mtDNA control region haplotypes as support for potentially 2 expansion events in New Zealand (Rogers & Harpending 1992, Rogers 1995). The neighbor-joining tree, depicting relationships among haplotypes (Fig. 2), the nested clade networks (Figs. 3, 4), and the two convex curves in the lineages-through-time plot (Fig. 6B) allude to these events. For instance, in terms of haplotype groupings, there appears to be a set of haplotypes consisting of short branch lengths, short internodes, and few unsampled intermediates (subgroup B, Fig. 2; clades 5-2 Fig. 4). Another group of haplotypes (subgroup A, Fig. 2; clade 5-1, Fig. 4) is characterized by long branches, long internodes, and as many as 12 inferred intermediates between haplotypes. Coalescence theory predicts that clades are nested in a non-decreasing age series so that higher-level clades must be as old as or older than lower-level nested clades (Templeton 1998). The *Nature,* of these 2 subgroupings within the neighbor-joining phylogram and nested haplotype networks may reflect the relative timing of expansion events suggested by lineages-through-time analysis. For example, the "bush-like" branching pattern within subgroup B (Fig. 2) and 1-step haplotypes within nested clades 1-36 and 2-23 (Figs. 3, 4), and the lack of missing intermediates in clade 4-3 (Fig. 4) all indicate a relatively recent expansion. In contrast, the abundance of unsampled intermediates in

clade 5-1 (Fig. 4) is indicative of a much older expansion followed by differential lineage sorting.

Consideration of this hypothesis requires the investigation of broader geographic patterns of mtDNA haplotype diversity throughout the range of the dolphin species. Dusky dolphins are discontinuously distributed in a step-wise fashion around 42° S latitude, i.e., their 3 main regions of distribution (South America, South Africa, and New Zealand) are connected by a series of small islands. Sightings around these islands have been confirmed; however, the regularity with which dusky dolphins use these islands is unknown. Because of the potential for stepwise, long-distance movements between regions, the study of the historical demography of the species must include sampling in New Zealand, South America, and South Africa. Perhaps the pattern of haplotype relationships observed in New Zealand may reflect historical events related to exchanges between regions, as well as past and current restrictions to gene flow among these regions. This is an important consideration, especially when evaluating female effective population size from diversity estimates. For instance, the calculations of N_f from nucleotide diversity suggest an effective female population of 15,000 – 17,000 in New Zealand. The diversity levels calculated in this study might include remnants of a much larger, global population, resulting in an inflated estimate of N_f for New Zealand. Further insight into the demographic history of dusky dolphins in New Zealand could be achieved by examining the degree and timing of divergence among the 3 regions of dusky dolphin distribution (i.e., New Zealand, South Africa, and South America). Very little is known about the relationship between New Zealand dusky dolphins and

conspecifics in other regions of the world. The phylogeographic pattern of dusky dolphin origin and radiation in the Southern Hemisphere may allow us to pinpoint the timing of colonization of or dispersal from New Zealand, and thus place observed demographic patterns into a temporal perspective. If the separation from other regions occurred relatively recently, one would predict similar levels of diversity and degree of divergence among all 3 regions. A comparison of molecular diversity and phylogenetic relationships among South American, South African, and New Zealand dusky dolphins would provide further insight into the biogeographic and demographic history of dusky dolphins in New Zealand and throughout their range.

CHAPTER III

ASSESSMENT OF CHARACTER INTERACTION AND PHYLOGENETIC UTILITY OF MULTIPLE DATA PARTITIONS: RELATIONSHIPS AMONG DOLPHINS OF THE SUBFAMILY LISSODELPHININAE

Introduction

Nucleotide sequence data from both nuclear and mitochondrial genomes are a well-established source of characters for phylogeny reconstruction. Nevertheless, controversies remain regarding the treatment of molecular characters in phylogenetic analysis, including the differential weighting of characters (both *a priori* and *a posteriori*) (e.g., Kluge & Wolf 1993; Allard & Carpenter 1996; Kluge 1998; Allard et al. 1999; Grant & Kluge 2003), the adequacy of optimality criteria to employ models of complex evolutionary processes (e.g., Yang 1994, 1996; Bruno & Halpern 1999; Steel & Penny 2000), and the simultaneous analyses of multiple data partitions (i.e., "total evidence") (e.g., Kluge 1989; Miyamoto & Fitch 1995; Nixon & Carpenter 1996; DeSalle & Brower 1997; Baker & DeSalle 1997). For example, it has been argued that the combined effects of asymmetry in the substitution process and differential rates of evolution create process partitions that can compromise diagnosis of relationships when characters are either equally weighted under parsimony or used with maximum likelihood under a simple model of molecular evolution (Felsenstein 1978; de Quieroz et al. 1995; Yang 1996, 1997; Allard & Carpenter 1999; Sullivan & Swofford 2001). One

assumption associated with such an argument is that phylogenetic signal of a particular sequence erodes evenly as one proceeds from the tip of a phylogeny to deeper nodes within a phylogeny. Based on this assumption, one would predict portions of a gene or different genes that evolve at different rates to resolve relationships in a phylogeny commensurate with their degree of homoplasy and overall rate of evolution. In such cases, saturation of nucleotide substitutions or some variation in the evolutionary process among data partitions should create incongruence between different gene trees. The compromise position between the separate analysis followed by strict consensus method and the total evidence method is to test for combinability of datasets *a priori* (Bull et al., 1993; de Quieroz et al. 1995, Farris et al., 1995; Huelsenbeck et al., 1996, Barker & Lutzoni, 2002). Nevertheless, the robustness of such tests is unclear (Baker & DeSalle, 1997; Cunningham, 1997; Darlu & Lecointre, 2002; Dowton & Austin, 2002), and in cases where taxa being compared span a broad range of evolutionary time, separate analysis and consensus methods are doomed to provide poor resolution.

Previous studies demonstrate that data sets interact when combined (Nixon & Carpenter 1996; Baker & DeSalle 1997; Cognato & Vogler 2001; Damgaard & Cognato 2003) and that properties of data emerge that are not detectable when partitions are analyzed separately (Nixon & Carpenter 1996; Baker & DeSalle 1997; Gatesy et al. 1999; Cognato & Vogler 2001). These emergent properties can result in either data synergy (complementary interaction) or conflict among characters due to variations in evolutionary process of each data partition. Several methods have been proposed to quantify congruence and conflict in combined analyses (Gatesy et al. 1999; Wheeler

1999; Cognato & Vogler 2001; Sota & Vogler 2001; Damgaard & Cognato 2003; Lee & Hugall 2003). Although these measures are useful descriptors of inconsistency among analyses, they usually lack a statistical framework and are not comparable across trees derived from independent data partitions. In a recent review, Grant and Kluge (2003) argued that a total evidence approach, based on equally weighted characters and parsimony, is the most stringent means of evaluating the "explanatory power" of characters. Furthermore, an *a posteriori* analysis of various character partitions against a total evidence tree is part of a heuristic approach for evaluating the utility of different partitions to resolve relationships among lineages that potentially differ in terms of evolutionary rates and/or overall level of evolutionary divergence.

Although the explanatory power of data is maximized in combined analyses, the increased resolution of a total-evidence phylogeny does not necessarily guarantee that these relationships reflect the true evolutionary history of taxa. A hypothesis is supported "if the critical evidence confers a greater degree of corroboration on it than any other hypothesis" (Grant & Kluge 2003, p. 383); therefore, in the case of phylogenetic analysis, support for a particular clade is "the degree to which critical evidence refutes competing hypotheses" (Grant & Kluge 2003, p. 383). Metrics such as bootstrap (Felsenstein 1985) and decay indices (Bremer 1994) provide a means to assess the relative strength of corroboration of a hypothesis provided by the data for particular nodes. For this reason, the examination of the dynamics of character interaction in relation to the degree of corroboration for a hypothesis of monophyly at a particular node can provide direction for further testing of hypotheses (Grant & Kluge 2003) and to

assess *a posteriori* the contribution and interaction of characters from various data partitions with respect to the total evidence phylogeny (Grant & Kluge 2003). For example, if the rates of evolution among data sets differ, one might predict that each heterogeneous process partition will contribute to the resolution of the combined analysis in accordance with its rate of evolution.

In this paper I use a total evidence tree, derived from four molecular markers (two nuclear and two mitochondrial) that vary in overall rates of evolution, to explore patterns of character divergence and the utility of these characters to diagnose relationships among species of dolphin in the genus *Lagenorhynchus*, subfamily Lissodelphininae (LeDuc et al. 1999). All members of this genus are distributed in temperate to cool waters in the North Pacific, South Pacific, and North Atlantic Oceans (Würsig et al. 1997, Cipriano 2002; Goodall 2002a,b; Kinze 2002), with one species in the North Pacific (*L. obliquidens*), two in the North Atlantic (*L. acutus, L. albirostris*), and three in the Southern Hemisphere (*L. obscurus, L. australis, L. cruciger*) (Fig. 7). Although color patterns, number of teeth, and the length of the rostrum in relation to the brain case (Gray 1828, 1846, Flower 1883, True 1889) have been used to support a monophyletic *Lagenorhynchus*, these characters are unreliable indicators of relationships within the family Delphinidae because they vary with sex and age among individuals (Flower 1893). Furthermore, the ubiquitous use of these morphological characters to define monophyletic groups of dolphins has resulted at one time or another in the placement of *Lagenorhynchus sp.* into at least 8 different genera (e.g., *Delphinus*, Gray

1828; *Electra*, Gray 1871; *Phocoena*, Peale 1848; *Tursio*, Gray 1866; *Leucopleurus*, Gray 1871; *Clymenia*, Gray 1868; *Sagmatius*, Cope 1866).

Recent molecular systematic studies of mitochondrial cytochrome *b* (cyt *b*) and control region sequences suggest that the genus *Lagenorhynchus* is a polyphyletic assemblage of antitropical and disjunctly distributed species (Cipriano 1997; LeDuc et al. 1999). The parsimony topology of LeDuc et al. (1999) derived from complete cyt *b* sequences suggests a division of the genus into *Lagenorhynchus* and *Leucopleurus* in the North Atlantic and *Sagmatius* in the Pacific and Southern Hemisphere (Fig. 7). However, their taxonomic revisions were based on weakly supported relationships. For example, their data provided strong phylogenetic support for a monophyletic subfamily Lissodelphininae, which contains *L. obscurus*, *L. obliquidens*, *L. cruciger*, and *L. australis* and excludes the North Atlantic species (LeDuc et al. 1999); however, the relationships among the *Lagenorhynchus* within the Lissodelphininae and the North Atlantic species were unresolved (Fig. 7).

This study is the first to use both nuclear and mitochondrial loci for the diagnosis of relationships within the family Delphinidae, and thus provides an unique opportunity to explore the utility of multiple process partitions in addressing the evolutionary history of this group. Furthermore, I provide statistical measures of character utility and interaction derived from detailed examination of character behavior within the total evidence topology and across all cladistic analyses.

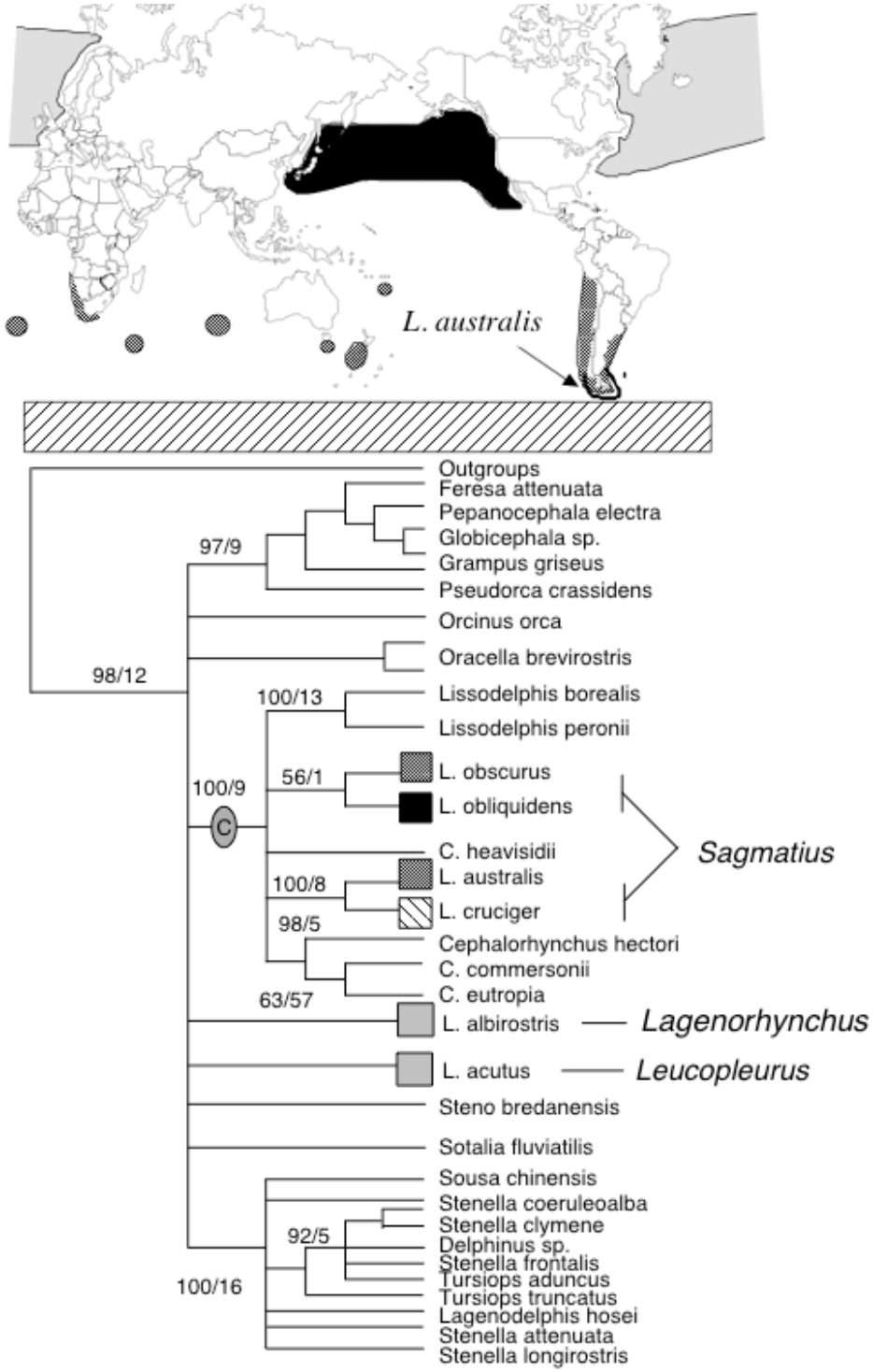


Figure 7. Distribution of the genus *Lagenorhynchus* and the phylogeny of Delphinidae derived from cytochrome *b* sequences (LeDuc et al. 1999). C = Lissodelphininae. Shaded blocks at terminal ends of branches correspond to distribution of species in the map above the phylogeny. Generic names suggested by the taxonomic revision of Le Duc et al. 1999 are indicated to the right of the phylogeny.

Materials and methods

Sampling

Tissues were obtained from a variety of sources including biopsy punches and skin swabs from living animals and post-mortem samples from beach-cast or net-caught individuals (Table 3). Skin swabs were collected following the noninvasive procedure of Harlin et al. (1999), preserved in either 90% ETOH or a solution of 20% dimethylsulfoxide (DMSO) saturated with salt, and stored at -20°C. When possible, DNA from the same individuals was used to amplify and sequence all genes. In some cases, preexisting sequences from GenBank were used to complete the data matrix (Table 3).

Fifty-two individuals from eight genera and 18 species from the family Delphinidae were examined (Table 3). Special attention was given to the genus *Lagenorhynchus* and other members of the subfamily Lissodelphininae. All 12 species of the putative subfamily and members of the genus *Lagenorhynchus* (i.e., *Lagenorhynchus* sp., *Cephalorhynchus* sp., and *Lissodelphis* sp.) are represented (Table 3). In addition, members of the *Tursiops*/*Delphinus*/*Stenella* clade identified by LeDuc et al. (1999), and the killer whale (*Orcinus orca*), a taxon thought to be the basal delphinid lineage (Barnes et al. 1985), were included. *Australophocoena dioptrica* and *Phocoena phocoena* from the family Phocoenidae were selected as outgroup taxa because of their sister-taxon relationship to the Delphinidae (Barnes et al. 1985; Cassens et al. 2000).

DNA isolation and amplification

Total genomic DNA was isolated using either a standard phenol-chloroform protocol (Maniatis et al., 1982) or a Qiagen DNeasy kit (Qiagen, Valencia, California). Isolated DNA was stored in tris-borate-EDTA (TBE) buffer at -20° C. The polymerase chain reaction (PCR) was used to amplify fragments of four genes including: (1) the complete mitochondrial cytochrome *b* (cyt *b*) gene (1040 nucleotides), (2) 474 base pairs (bp) of the mitochondrial control region, (3) 995 bp of nuclear DNA (nDNA) intron I of the muscle actin gene, and (4) 474 bp of the coding region of the nDNA recombination activator gene 2 (RAG2). External primer sets included: (1) cyt *b* - 766F (5'GAAAAACCAYCGTTGTWATTCAACT3') and 766R (5'GTTTAATTAGAATYTYAGCTTTGGG3'); (2) control region - tRNA-Pro and Dlp5 of Baker et al. (1996); (3) actin - *Lagenorhynchus* specific LagAct1 (5'GATTTGGTCCCTCTATGTCTCT3' and LagAct2 - 5'TACTTTTGAACTTGCCACCTAC3'). Actin primers were designed from published cetacean sequences (Palumbi & Baker 1994) and used to amplify the majority of ingroup taxa. Act1 (Palumbi & Baker 1994) and Act1385H (5'CTTGTGAACTGATTACAGTCC3') (Palumbi, unpublished) were used to amplify fragments for outgroup taxa and others that failed to amplify with the *Lagenorhynchus*-specific primers. RAG2 primers were the same as those reported by Murphy et al. (2001). PCR conditions were generally consistent across loci with adjustments made to

Table 3. List of taxa and sample size for each data partition.

Taxon	Cytb	Dloop	Actin	Rag2
	n	n	n	n
Lagenorhynchus obscurus	14[a]	14[a]	14[a]	14[a,r]
Lagenorhynchus obliquidens	6[b]	6[b]	6[q]	6[b]
Lagenorhynchus cruciger	1[h]	1[l]	1[g]	1[g]
Lagenorhynchus australis	3[a,i]	3[a]	3[a]	3[a]
Lagenorhynchus albirostris	5[b,c]	5[b,c]	5[b,c]	5[b,c]
Lagenorhynchus acutus	4[d]	4[d]	4[d]	4[d]
Cephalorhynchus hectori	2[j]	2[m]	2[g]	2[g]
Cephalorhynchus commersonii	1[b]	1[b]	1[b]	1[b]
Cephalorhynchus eutropia	2[k]	2[n]	0	0
Cephalorhynchus heavisidii	2[e]	2[e]	2[e]	2[e]
Lissodelphis peronii	1[b]	1[b]	1[b]	1[b]
Lissodelphis borealis	4[b]	4[b]	4[b]	4[b]
Delphinus delphis	2[c]	2[c]	2[c]	2[c]
Stenella attenuata	1[c]	1[c]	1[c]	1[c]
Tursiops truncatus	1[c]	1[c]	1[c]	1[c,s]
Australophocoena dioptrica	1[f]	1[f]	1[f]	1[f]
Phocoena phocena	1[f]	1[o]	1[f]	1[f]
Orcinus orca	1[f]	1[p]	1[f]	1[f]
TOTAL	52			

a. A. Harlin tissue collection
b. SWFSC National Marine Mammal Tissue Bank, La Jolla, California
c. Texas A&M Cooperative Wildlife Collection, Tissue Depository
d. R. L. Honeycutt tissue collection
e. M. Meyer, Sea Fisheries Research Institute, Capetown, South Africa
f. C. S. Baker, University of Auckland, Auckland, New Zealand
g. F. Pichler, University of Auckland, Auckland, New Zealand
h. GenBank Accession AF084068
i. GenBank Accession AF084069
j. GenBank Accession AF084071
k. GenBank Accessions AF084072, U13128
l. GenBank Accession AF084072
m. GenBank Accessions AF057997, AF057998
n. GenBank Accessions AF393555, AF393553
o. GenBank Accession U09694
p. GenBank Accession M60409
q. GenBank Accessions AF140826-AF140834
r. GenBank Accessions AF140832-AF140834
s. GenBank Accession AY011968

annealing temperatures. Approximately 1 -2 µl of DNA template were included in 50 µl PCR reactions containing the following: 5 µl each 10X Amplitaq PCR buffer (Perkin Elmer, Boston, Massachusetts), 25 mM MgCl, and 10 mM deoxynucleotide triphosphates (dNTP's, 2.5 mM each), 1 µl each of 10mg/ml bovine serum albumin (BSA), 10 uM of each primer, and 5 units (U) of Amplitaq (Perkin Elmer, Boston, Massachusetts) DNA polymerase. Thermocycler conditions were 94° C for 2 min followed by 35 cycles at 92° C for 30 s, annealing 30 s, and extension at 72° C for 30 s. Published annealing temperatures were used with the following exceptions: 765F/766R, 50° C; LagActin1/2, 58° C; Act1/1385H, 56° C. Amplicons were electrophoresed in 1.5% agarose-TBE (tris, boric acid, EDTA) and visualized under UV light, and prior to sequencing, excess oligonucleotides and dNTP's were removed with either Qiagen (Qiagen, Valencia, California) spin-columns or an Exonuclease I-Shrimp Alkaline Phosphotase (Exo-Sap) enzymatic procedure. Approximately 2 ng of cleaned PCR product per 100 bp of amplicon length was sequenced using ABI BigDye (Applied Biosystems, Foster City, California) cycle sequencing chemistry and an ABI 377 automated sequencer. All amplicons were sequenced in both directions. Internal primer pairs for cyt *b* (560, 5'-GCAACCCTAACACGATTCTTCG-3'; 610, 5'-CCAGTTTCGTGTAGGAATAATAGG- 3') and actin (Act5-L, 5'CCACTACTTTAGGCAG3'; M13Act5R-H, 5'TGTAAAACGACGGCCAGTCTGCCTAAACTAGTGG3' (S. Palumbi, unpublished) were used in sequencing reactions to obtain complete overlap in both directions.

Sequence alignment

Sequenced fragments were edited and compiled with the program Sequencher version 4.1. A consensus of sense and antisense strands for each individual and data partition were compiled and exported to MacClade vs. 4.05 (Maddison & Maddison, 1992). Sequences of the four data partitions were concatenated into a single string of nucleotide characters for the same individual when possible, or a combination of fragments from members of the same species. Cyt *b* and Rag2 contained no length variable regions, thus alignment of these fragments was trivial. Amino acid translations of the open reading frames of Rag2 and cyt *b* were examined for stop codons to verify sequence orthology. Actin fragments contained minor genus-specific insertions and deletions that were revealed by alignment in Clustal X (Thompson et al. 1994) with default parameters. The control region had one $(CT)_n$ length variable region of approximately 21 bp that was eliminated from analyses as primary homology could not be reliably assessed. All species were represented by fragments from the four data partitions with the exception of *C. eutropia,* which lacked actin and Rag2 fragments because I was not able to obtain tissue samples for this species (Table 3).

Treatment of gaps

The utility of gaps and their treatment in phylogenetic analyses is controversial (Vogler & Desalle 1994; van Dijk et al. 1999; Simmons & Ochoterena 2000, Simmons et al. 2001; Wheeler 1990,1995,1996). Several methods have been proposed to deal with indels in nucleotide sequence data, some of which advocate the inclusion of gaps in phylogenetic analyses as a unique (or 5th) state. There were 2 categories of gaps

(Simmons & Ochoterena 2000) that were revealed in the sequence alignment: (1) those that resulted from single base mutations, present in both non-coding regions; and (2) those that were > 1 nucleotide in length and had the same 5' and 3' termini. Some have advocated differential treatment of gaps of different "categories" in phylogenetic analysis. For example, Simmons and Ochotorena (2000) advocated that contiguous gaps with the same 5' and 3' termini should be considered one mutational event and coded as binary, presence/absence characters (Simmons & Ochoterena 2000). Previous studies suggest that gaps can be phylogenetically informative in both coding and non-coding regions (van Dijk 1999), thus all gaps were considered as 5th states in this study.

Heterozygosity of nuclear loci

Heterozygosity of actin intron I was observed for *Lagenorhynchus* (Hare et al. 2002). Therefore, all nucleotide ambiguities that resulted from two different, but equally strong, peaks on electropherograms were considered as evidence for potential heterozygous sites for actin and Rag2 fragments. All such positions were given IUPAC ambiguity codes with Sequencher version 4.01, and considered as ambiguous characters in subsequent phylogenetic analyses. PCR reactions consistently produced only one actin or Rag2 amplicon, which was subjected to a BLAST search to verify sequence orthology. In all BLAST searches, amplicons retrieved sequences from other delphinid or mammalian taxa as the closest match, providing further evidence for successful amplification of target loci.

Phylogenetic analysis

All phylogenies were generated with the program PAUP* version 4.0b10 (Swofford 1999) under the optimality criterion of parsimony, with all characters equally weighted and gaps treated as a 5th state. The number of taxa and characters precluded the use of exhaustive search options. Therefore, a heuristic search with 1000 random additions of taxa and 100 trees held at each replicate was initiated. Branch swapping was performed with tree-bisection-reconnection (TBR). These search settings were used for the combined data, and for each data partition independently. In addition, phylogenetic analyses were performed separately for mtDNA (cyt *b* and control region) and nDNA (actin and Rag2) data sets.

Clade support for the total evidence phylogeny was evaluated with bootstrap (Felsenstein 1985) and branch support (BS) (Bremer 1994) indices. The bootstrap procedure was replicated 1000 times, each as full heuristic search with 100 random addition sequences and TBR branch swapping with 10 trees held at each step. Bremer support (BS) indices for the simultaneous analysis (SA) tree were derived from a heuristic search of constraint trees created with the program TreeRot version 2b (Sorensen 1999). Heuristic search parameters were the same for the original SA analysis.

Character dynamics: phylogenetic utility

The behavior of combining characters from different process partitions was evaluated by examining the relative contribution, or utility, of data partitions to resolving relationships within the SA phylogeny, and the interactions of characters among the data

partitions when combined. Potential for saturation was investigated by comparing branch lengths under different models of nucleotide evolution. For each process partition, branch lengths of the MP SA tree were calculated under a simple model of equally probable substitution and equal base frequencies (i.e., Jukes & Cantor 1969), and a more complex maximum likelihood procedure with model parameters estimated from the data. Maximum likelihood model parameters were derived from a full likelihood evaluation of the SA tree under the assumption of empirical base frequencies. In addition to the 4 process partitions, branch length saturation plots also were constructed for the following character partitions: (1) nonsynonymous substitutions and 3rd position transversions (TV); (2) 3rd position transitions (TS), and (3) all TS from coding and non-coding regions (Cognato & Vogler 2001). Branch support indices were partitioned (i.e., "partitioned Bremer support", PBS) with the method of Baker and DeSalle (1997) to measure the relative contribution of each data set to node support. The larger the PBS is for a given partition at a particular node, the greater the relative contribution of that partition to the support of that node (Baker & DeSalle 1997; Baker et al. 1998). PBS values were divided by sequence length of each partition to standardize values for number of characters.

A number of methods have been proposed to quantify the contribution of a particular data partition to topological support in a combined analysis (Gatesy et al. 1999). These methods generally include a measure of the relative number of unambiguous character changes attributable to a given data partition (Gatesy et al. 1999; Lee 1999; Wheeler 1999; Lee & Hugall 2003). In this study, I propose methods that

summarize the contribution of a data partition to phylogenetic resolution by quantifying at each node a subset of unambiguous character changes that are consistent throughout the topology (i.e., CI = 1). Because these characters change once only and provide support for only one node, I refer to them as "perfectly consistent" (PC). I use the distribution of these characters in relation to the depth, or relative time since divergence, of nodes in the SA phylogeny as a measure of the utility of a data partition with respect to topological resolution at different levels of divergence. The null hypothesis is that each data partition contributes equally to the support of clades in the phylogeny with respect to relative age of coalescence. The number of PC changes at a node is likely related directly to its rate of evolution, and the frequency of PC changes might vary among data partitions within the same clade if the evolutionary history of partitions are independent (e.g., mtDNA vs. nDNA). Since there is only one "true" evolutionary history, deviations from the expected pattern, or conflicts among data partitions, can be attributed to variation in the evolutionary histories and rates of evolution of process partitions. Alternatively, congruent patterns of evolutionary change (i.e., frequency and rate of substitutions at particular node depths and among data partitions) would support the phylogeny as a hypothesis of taxic, not genic, evolution.

To test this hypothesis, the height of each node within the SA phylogeny was estimated by averaging the maximum likelihood branch lengths from each node to each OTU within a clade. All unambiguous character changes on the SA and independent phylogenies were traced and labeled with their consistency index in MacClade version 4.01 (Maddison & Maddison; 1992). A chi-square test for homogeneity of proportions

was used to determine if PC characters were evenly distributed within the SA phylogeny with respect to node depth. A Kruskal-Walis non-parametric statistic for multiple samples was used to test if the proportion of PC characters contributed by each data partition was consistent among all node depths. Additionally, the relative contribution of each partition to a node was examined by calculating the consistency index of unambiguous character changes at first, second, and third codon positions in the SA phylogeny. A chi-square was used test the null hypothesis of homogeneity of PC character changes among codon positions. All statistical tests were performed with SPSS version 11 statistical package. Null hypotheses were rejected if $P \leq 0.05$.

Character dynamics: data interaction

A Spearman's rank correlation of PBS values was used to determine the nature, and significance of interaction among data partitions in the SA topology. PBS values for data partitions at a particular node can be used as indicators of conflict among data partitions--positive values indicate support, negative values suggest conflict. A Spearman's rank correlation of PBS values from the SA phylogeny (Sota & Vogler 2001) was performed to measure the level and magnitude of character interaction on a node-by-node basis. A significantly positive correlation coefficient is evidence for synergistic character interaction; a negative correlation indicates significant conflict among data partitions. A non-significant correlation suggests that topological support is not associated with any combined pairs of data partitions (Damgaard & Cognato 2003). In this study, I propose a new measure of data conflict derived from the behavior of unambiguous synapomorphies in separate and combined analyses. The synapomorphy

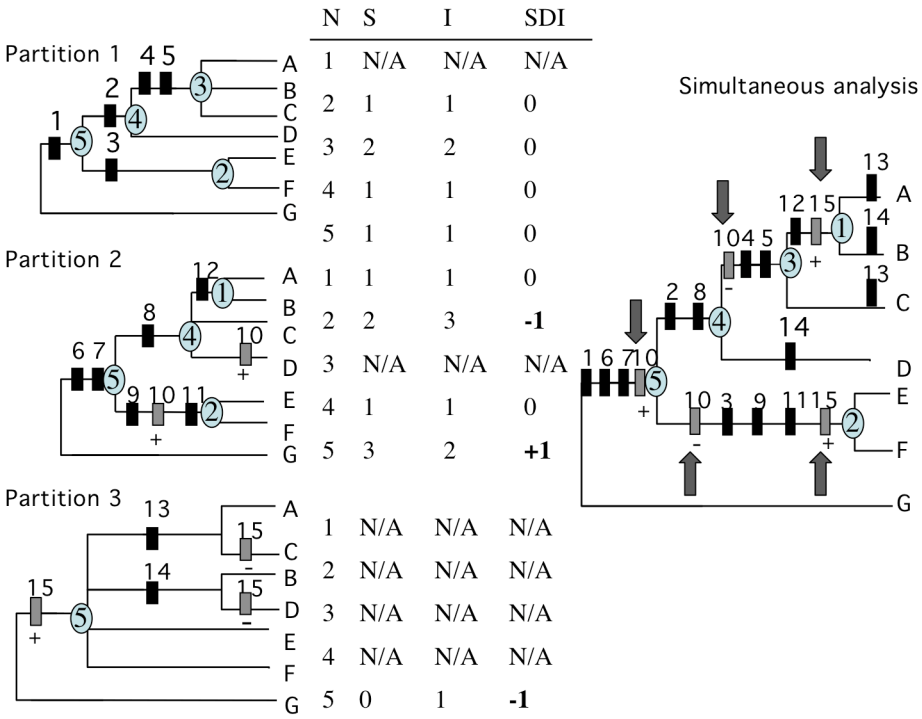
displacement index (SDI) is defined as the difference in the number of unambiguous character changes at a node in separate and combined analysis (Fig. 8). The number of unambiguous changes for each node was determined with MacClade version 4.01 (Maddison & Maddison 1992) by tracing and counting all unambiguous changes on one of the SA and IA parsimony trees. Only those nodes common to both phylogenies were considered. Negative values of SDI indicated the displacement, or loss, of unambiguous synapomorphies from a node as a result of combining data, whereas positive values suggested synergistic data interaction. The greater the magnitude of the SDI value, the greater the interaction among characters at that node. A sign test, a non-parametric analog of the t-test, was used to determine if the simultaneous analysis had a significant effect on the frequency of synapomorphy displacement.

Results

Simultaneous analysis and relationships within the Delphinidae

The SA of 2 nDNA and 2 mtDNA genes produced 4 equally parsimonious trees that differed only in their level of intraspecific resolution (Fig. 9). Of the 3053 characters in the combined data matrix, 485 were parsimony-informative (Table 4). The consistency of characters on the SA tree was relatively high (CI = 0.55), and improved when rescaled (RI = 0.80) (Table 4). Bootstrap and BS values were highest for clades at the base and tips, and lower for nodes in the central region of the SA topology (e.g., nodes 28, 24, 44, and 40, Fig. 9). Fourteen of 15 species-level clades had bootstrap

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
A	1	1	0	1	1	1	1	1	0	0	0	1	1	0	1
B	1	1	0	1	1	1	1	1	0	0	0	1	0	1	1
C	1	1	0	1	1	1	1	1	0	0	0	0	1	0	0
D	1	1	0	0	0	1	1	1	0	1	0	0	0	1	0
E	1	0	1	0	0	1	1	0	1	1	1	0	0	0	1
F	1	0	1	0	0	1	1	0	1	1	1	0	0	0	1
G	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0



N	S	I	SDI
1	N/A	N/A	N/A
2	1	1	0
3	2	2	0
4	1	1	0
5	1	1	0
1	1	1	0
2	2	3	**-1**
3	N/A	N/A	N/A
4	1	1	0
5	3	2	**+1**
1	N/A	N/A	N/A
2	N/A	N/A	N/A
3	N/A	N/A	N/A
4	N/A	N/A	N/A
5	0	1	**-1**

Figure 8. An example of SDI calculation with a simulated data set. The data matrix represents 3 concatenated character partitions for 7 taxa (A-G). Taxon G is the outgroup. Each partition was fabricated to depict one of 3 situations: 1) no change in the distribution of synapomorphies in the IA and SA topologies (SDI = 0 at all nodes); 2) a gain and a loss of synapomorphies at nodes in the SA topology; 3) a loss of a synapomorphy at a node in the SA. Characters were partitioned as follows: 1 = 1-5, 2 = 6-12, 3 = 13-15. Cladograms represent the strict consensus of all most parsimonious trees from independent analyses (IA). The 'simultaneous' cladogram was derived from an exhaustive search of the combined matrix. Solid bars indicate characters with CI = 1 in the IA; shaded bars in the IA or SA cladograms represent characters that are unambiguous at a given node but have a CI < 1, and whose position in the tree changes in SA. Each table to the right of cladograms summarizes the calculation of the synapomorphy displacement index (SDI), or the difference in the number of unambiguous character changes at individual nodes present in both the simultaneous (SA) and the IA. A negative SDI = character conflict or the loss of a synapomorphy at a node in the SA. A positive SDI = the gain of a synapomorphy or positive character synergy. A 'N/A' represents the case when the SDI cannot be estimated because a node (N) is not present in both IA and SA cladograms. A '+' = gain of a character; a '-' = loss of character. Thick arrows are used to indicate points in the SA where character placement changes when compared to IA.

Table 4. Summary of results from parsimony analysis of data partitions. SA = simultaneous analysis. MtDNA = cyt *b* + Dloop; nDNA = actin + RAG2. MP = Maximum parsimony; TL = tree length; CI = consistency index; RI = re-scaled consistency index; PI = parsimony informative.

Partition	Characters	MP Trees	TL	CI	RI	PI Characters
SA	3053	4	1328	0.55	0.799	485
DLOOP	444	9	424	0.451	0.757	126
CYTB	1140	2	700	0.523	0.823	280
ACTIN	995	36755	133	0.895	0.944	69
RAG2	474	310	15	1	1	7
mtDNA	1584	6	1150	0.514	0.793	406
nDNA	1469	22650	152	0.882	0.934	76

values >99% and high BS (5 to 44) (Fig. 9). The results provide unambiguous support for the monophyly of the subfamily Lissodelphininae (node 33, Fig. 9), and the polyphyly of the genus *Lagenorhynchus* (nodes 40, 27, and 23, Fig. 9). Several relationships were well-resolved in all 4 of the MP trees (100% consensus) with a BS value $\geq$1 (Fig. 9) despite low bootstrap support. These relationships included: (1) the monophyly of *L. albirostris* and *L. acutus* and the sister-group relationship of this clade to the *Stenella*/*Tursiops*/*Delphinus* clade (nodes 40 and 44, Fig. 9); (2) increased support for the sister taxa relationship of *L. obscurus* and *L. obliquidens* (node 23, Fig. 9); (3) the placement of *L. cruciger* and *L. australis* basal to the *Cephalorhynchus* (node 24, Fig. 9).

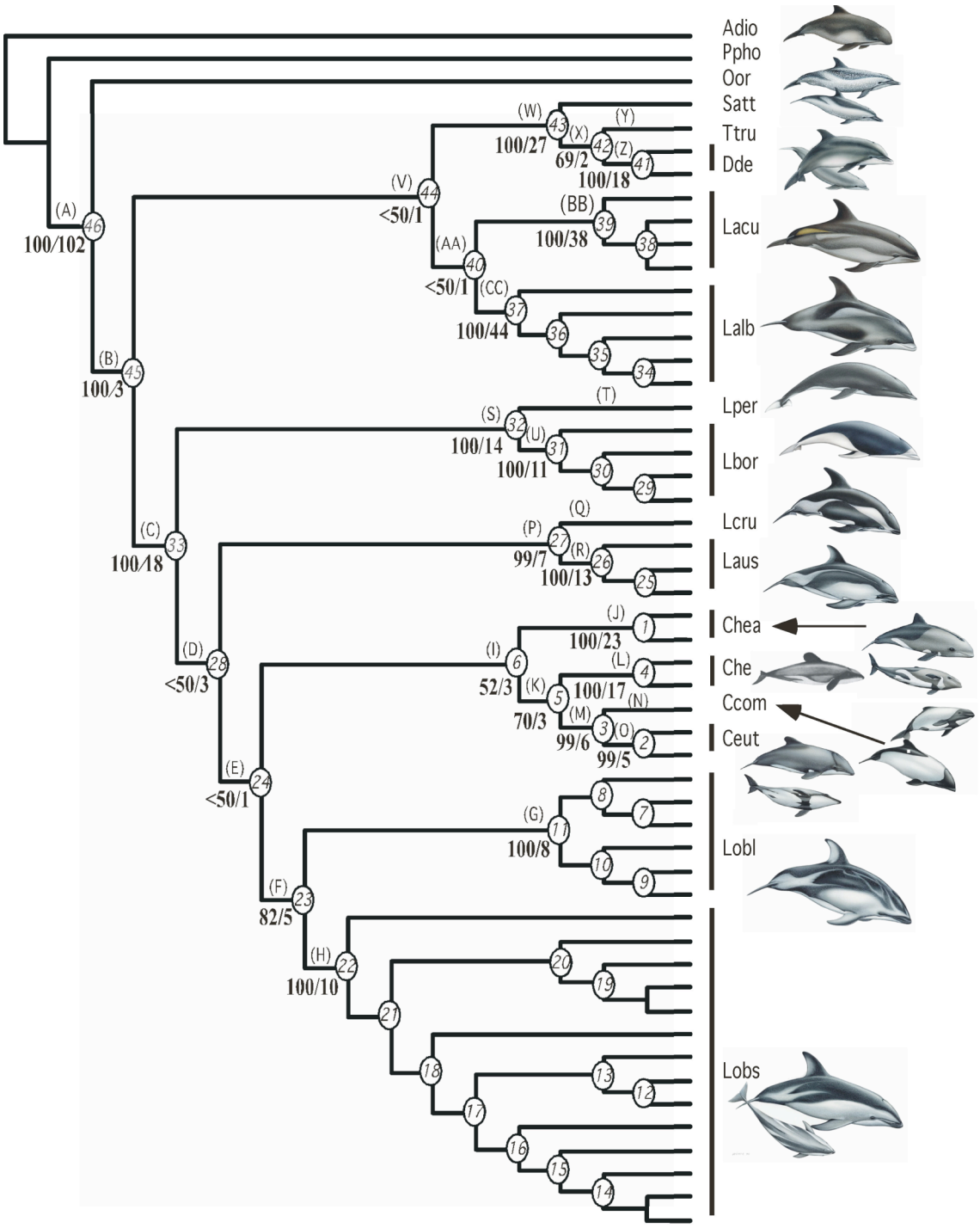


Figure 9. Cladogram derived from simultaneous analysis (SA) of 4 data partitions. Cladogram represents one of 4 equally parsimonious trees from the SA (Table 2). This topology differs from other equally parsimonious trees only in the placement of individuals within the *L. obscurus* clade. Numbers below branches indicate the bootstrap and branch support indices for nodes, respectively. Numbers in circles are used to identify individual nodes, and letters in parentheses designate branches referred to in the text.

Separate analyses

Gene trees for each of the 4 process partitions were consistently less resolved than the SA phylogeny. The mtDNA partitions had a greater number of parsimony informative characters and produced fewer MP trees than nDNA data (Table 4). Separate analysis of the cyt *b* partition produced topologies most similar to those of the SA analysis, with 100% consensus among all interspecific clades. There are several noteworthy structural similarities among the independent and SA topologies. For example, those clades with the highest BS values in the SA tree also were present in the nDNA and mtDNA trees (Fig. 10). Furthermore, monophyly of the family Delphinidae and the *Tursiops*/*Delphinus*/*Stenella* clades were recovered independently by mtDNA and actin partitions, and the nDNA data also recovered the monophyly of *L. acutus*, *L. albirostris*, and *C. hectori* (Fig. 10). In fact, the only source of incongruence among the IA and SA topologies was the IA of the control region -- many of the relationships towards the base of the Delphinidae in the control region phylogeny were not present either in the nDNA, cyt *b*, or SA phylogenies (Fig. 10).

Character dynamics: phylogenetic utility

Each of the four process partitions contributed in some degree to the resolution of the SA phylogeny. Saturation plots indicated that all substitution classes, including 3rd position transitions, contained phylogenetic information even at the deepest nodes (Fig. 11). When the BS was partitioned among nodes, the total PBS values for each data set were positive (Table 5). When standardized for sequence length, cyt *b* and the control

region contributed the largest proportion of PBS per nucleotide than either nDNA partition with Rag2 providing the least amount of total support (Table 5). On a node-by-node basis, the amount of localized topological support varied by data set and among regions of the SA phylogeny (Table 5, Fig. 12). For example, cyt *b* PBS values were consistently positive at all nodes, while the control region PBS values were positive at the tips of the tree but negative at the base (Table 5, Fig. 12). Similarly, the basal and terminal nodes had statistically more PC characters than those in the midsection of the phylogeny ($\chi^2 = 27.9$, df = 3, $P < 0.0001$), with cyt *b* consistently providing the largest proportion of PC characters at each node ($\chi^2 = 26.9$, df = 2, $P < 0.0001$) (Fig. 13). Despite the greater frequency and wider distribution of cyt *b* PC characters, the relative proportion of support contributed by each data partition did not statistically differ among levels of node heights in the phylogeny ($H = 69.9$, df = 3, $P < 0.0001$). Therefore, all data partitions were consistent in the relative proportion of support at different levels of the phylogeny, even if the magnitude of this support varied among partitions (Fig. 13). The results further suggest that nonsynonymous substitutions are proportionately less homoplastic (more consistent) than 3rd position transitions ($\chi^2 = 12.4$, df = 4, $P = 0.02$), and occurred more frequently and provided more topological support than substitutions at other codon positions.

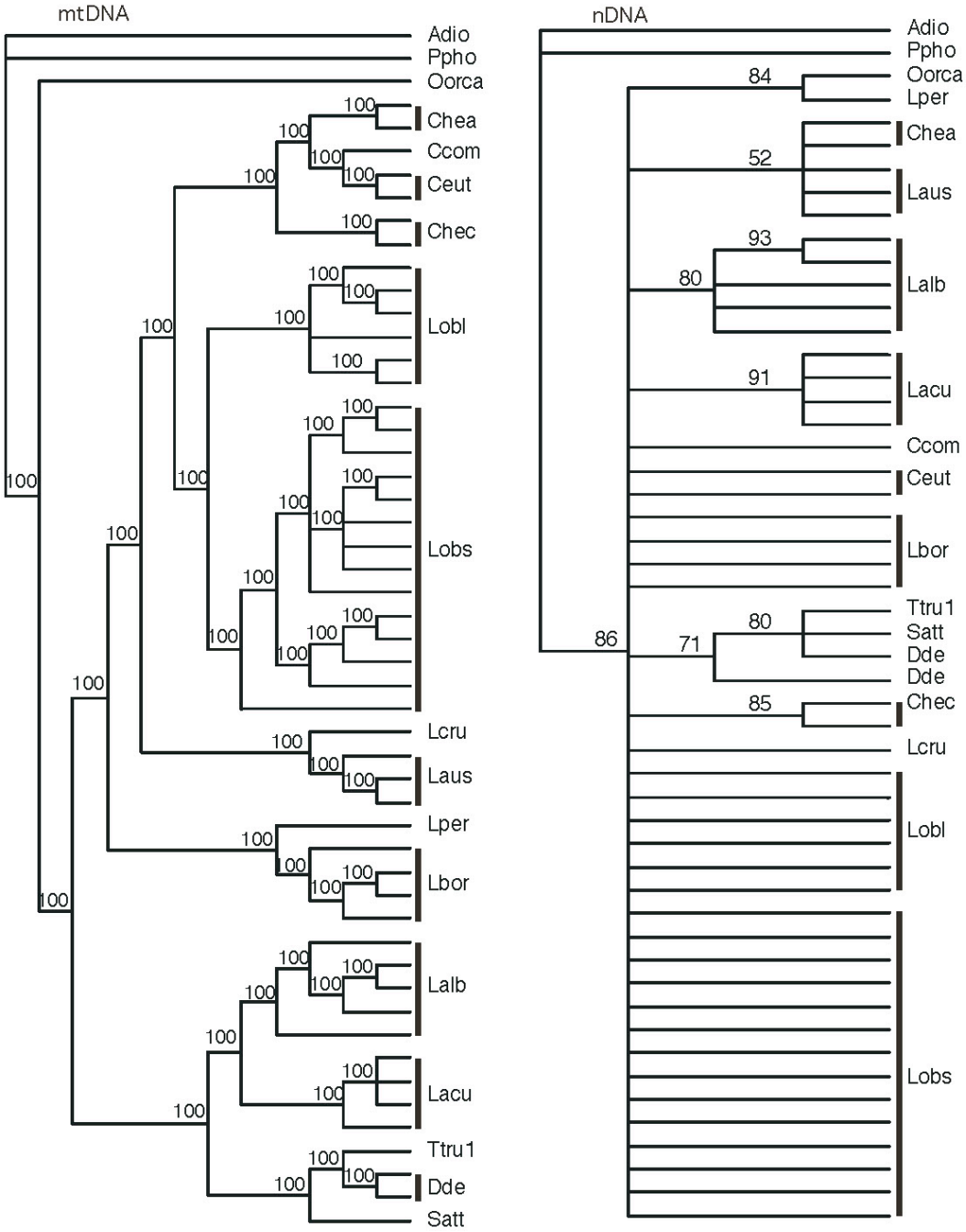


Figure 10. Cladograms derived from separate analyses of mitochondrial and nuclear genes, respectively. Numbers above branches represent majority-rule consensus values and are presented in lieu of branch support statistics in order to demonstrate the degree of consensus among most parsimonious trees.

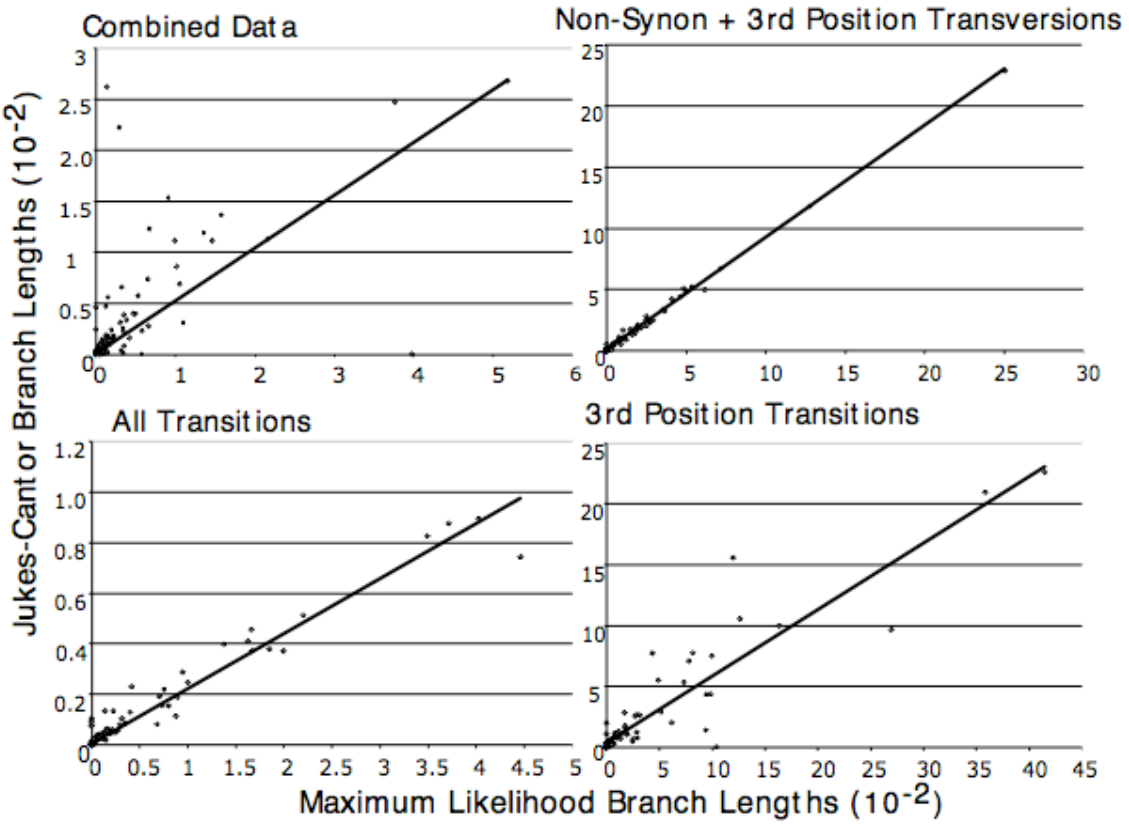


Figure 11. Saturation plots. Jukes-Cantor and maximum likelihood branch lengths were optimized for one of the four most-parsimonious trees from the SA analysis. Non-synonymous substitutions and 3rd position transitions were from both the Rag2 and cyt *b* data partitions, with majority of these from the cyt *b* gene. 'All Transitions' = all transitions from all data partitions, coding and non-coding, combined. 'Combined Data' = branch lengths derived from all substitutions from the combined data set.

Table 5. Partition branch support (PBS) for each node of the simultaneous analysis (SA) cladogram (Fig. 9). Node numbers correspond Fig. 9. Bremer support and % bootstrap for nodes in SA cladogram. <50 = bootstrap values less than 50%.

			Gene Partition			
		Bremer	Partition Bremer Support			
Node No.	%Bootstrap	Support	d-loop	Cyt b	Actin	RAG2
1	100	23	7	14	2	0
2	99	5	4	1	0	0
3	99	6	-1	7	0	0
4	100	17	11	6	0	0
5	70	3	6	3	-6	0
6	52	2.9	5.9	3.1	-6.1	0
7	91	4	4	0	0	0
8	91	3	3	0	0	0
9	77	2	1	1	0	0
10	<50	1	1	0	0	0
11	100	8	3	4	1	0
12	56	0.9	-1.8	5	-2.3	0
13	<50	1.1	-1.1	3.3	-1.1	0
14	51	1.4	1	0	0.4	0
15	<50	1.3	0.6	0	0.7	0
16	<50	1	0.6	0	0.4	0
17	<50	4	4	0	0	0
18	56	1	-2.5	8	-4.5	0
19	52	0.9	-1.8	5	-2.3	0
20	50	1.1	-0.8	2.7	-0.8	0
21	64	2	-0.7	2.9	-0.2	0
22	100	10	1	9	0	0
23	82	5	3.8	6.6	-5.4	0
24	<50	1	0	1	0	0
25	95	3	3	0	0	0
26	100	13	3	10	0	0
27	99	7	-1	8	0	0
28	<50	3	0.7	4.3	-2	0
29	<50	1	1	0	0	0
30	<50	1	1	0	0	0
31	100	11	3.5	7.5	0	0
32	100	14	-1	18	-3	0
33	<50	18	-2.5	13.5	7	0
34	56	1	1	0	0	0
35	73	2	1.5	1	-0.5	0
36	73	1	1	0	0	0
37	100	44	2.5	38	2.5	1
38	70	1	-1	2	0	0
39	100	38	7.5	28.5	1	1
40	<50	1	-5	6	0	0
41	100	18	4	14.5	-0.5	0
42	69	2	0	2	0	0
43	100	27	3	21	3	0
44	<50	1	-4	5	0	0
45	100	3	-0.8	4.9	-1.1	0
46	100	102	17.4	58.6	23	3
Totals		**417.6**	**82**	**325.4**	**5.2**	**5**

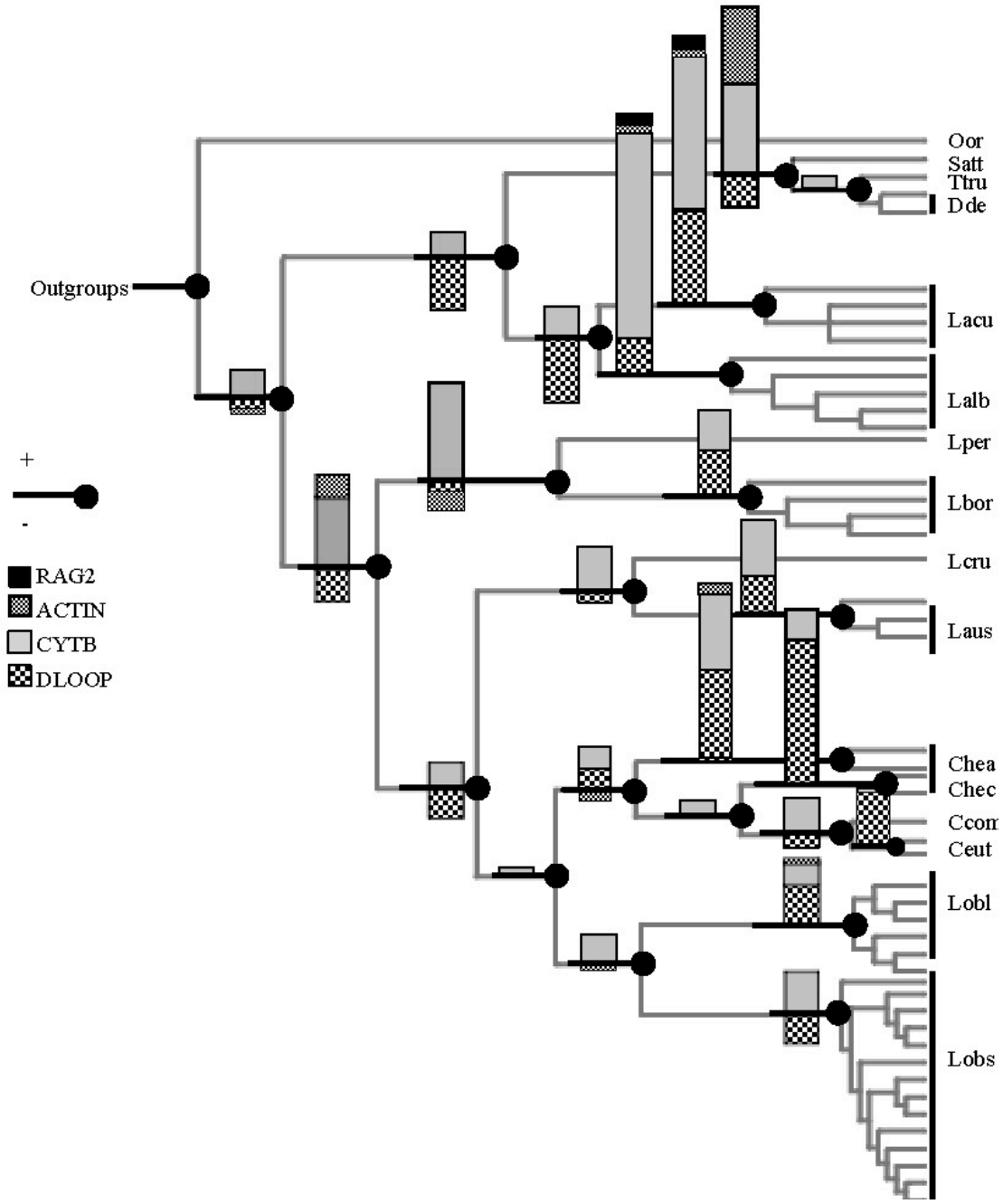


Figure 12. Distribution of branch support (BS) by character partition. At each node, bars represent the relative partitioned branch support (PBS) contributed by a given data partition to a node. Shaded bars above the branch indicate positive PBS values; below the branch are negative PBS values. PBS values were adjusted for sequence length and therefore represent a standardized contribution of each partition to node support.

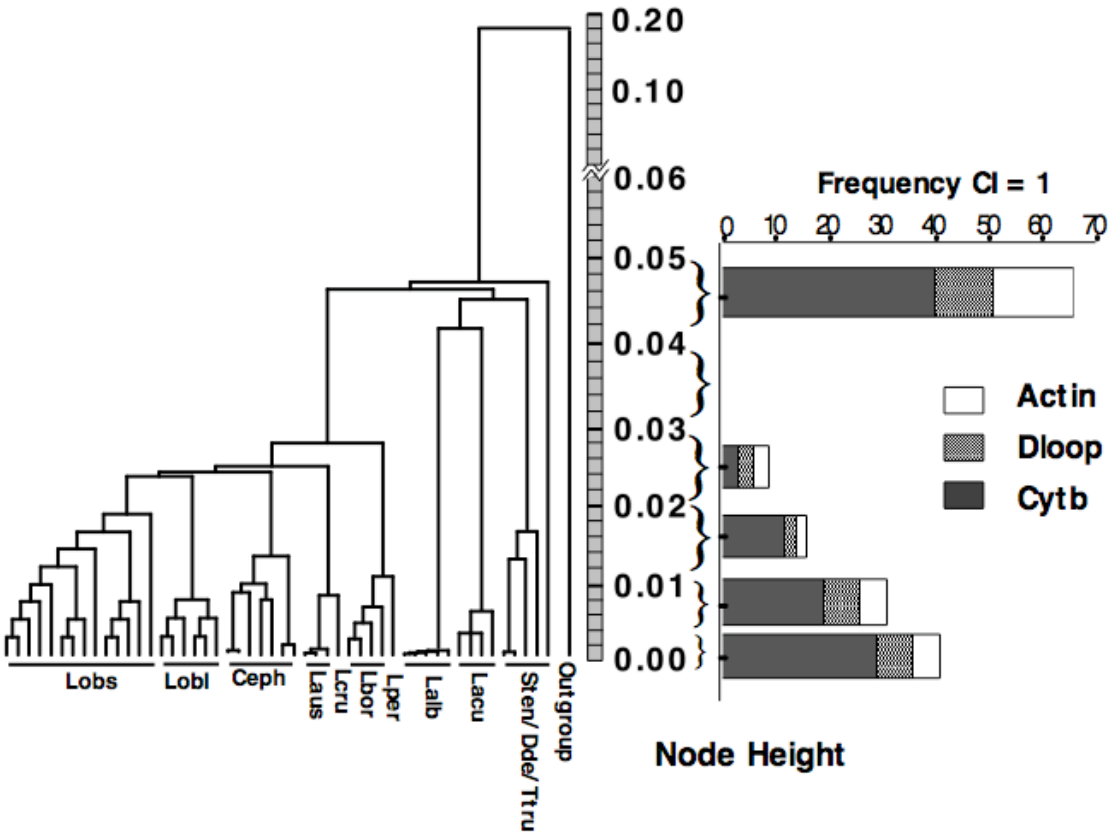


Figure 13. Distribution and relative frequency of perfectly consistent characters (PC) in the simultaneous analysis (SA) cladogram. The SA phylogeny is drawn to scale of node height derived from maximum likelihood optimized branch lengths. The number of PC characters in each of 5 node height categories, represented by bracketed values on the y-axis, was calculated for each data partition. The bar graph represents the sum of PC characters for each partition for each of the 5 node height categories. The number of PC characters contributed to support in the phylogeny was significantly different among node height categories. However, the relative number of PC characters was not significantly different among node height categories. RAG2 had only 5 parsimony-informative characters and was not included in statistical analyses. See text for statistical tests used and corresponding p-values.

Character dynamics: data interaction

Spearman's correlation analysis revealed a significantly positive relationship between Rag2 and the other data partitions ($\rho^2 = 0.432$, P = 0.003), but failed to find evidence for interaction between cyt *b*, actin, and control region partitions. However, the distribution of positive and negative PBS values within the SA topology suggested

localized character conflict among data partitions (Fig. 12). The greatest amount of dispersion in PBS indices was between the control region and other data partitions at the shortest nodes in the SA phylogeny, where the PBS values for the control region were consistently negative (e.g., nodes 28, 24, Figs. 1, 4). In addition, the SDI (displacement index) values for the control region were significantly negative (sign test, P = 0.007), with frequent loss of synapomorphies in simultaneous analysis (Table 6). The implications of these findings are two-fold: (1) The control region was responsible for the majority of data conflict. (2) The relatively weak support for some nodes was a result of this conflict. In contrast, cyt *b* did not demonstrate a significant pattern of gain or loss of synapomorphies (P = 0.13) due to SA. In fact, cyt *b* contributed one additional synapomorphy (SDI = 1) to 6 different nodes following SA (Table 6). The number of synapomorphies contributed by the nDNA data partitions did not change between independent and combined analyses (Table 6).

Table 6. Summary of support indices for branches in the simultaneous analysis (SA) and independent analyses (IA) of data partitions. Branches are designated by letters and correspond to those of Figure 2. 'PC' is the sum of the perfectly consistent (PC) characters on a particular branch in the SA. 'Sum PC partitions' = sum of PC characters for all data partitions for a node from the IA of each partition. 'RC' demonstrates the relative contribution of each partition to the support of a particular node, i.e., the total number of PC characters of a partition divided by the total number of PC characters in the SA. A '-' represents a node not present in both SA and IA, and were not included in calculations. 'UA' = unambiguous character changes (includes those with CI < 1.0). 'SDI' = Synapomorphy Displacement Index; 'PBS' = Partition Bremer Support and are calculated as described in the text.

	COMBINED		PARTITIONS																			
	Sum PC		Cytb					Dloop					Actin					RAG2					
Branch	PC	Partitions	CDI	UA	PC	RC	SDI	PBS	UA	PC	RC	SDI	PBS	UA	PC	RC	SDI	PBS	UA	PC	RC	SDI	PBS
A	64	59	5	67	35	0.59	0	58.6	23	6	0.10	-1	17.4	15	15	0.25	0	23	3	3	0.05	0	3
B	6	6	0	12	2	0.33	0	4.9	7	4	0.67	-1	-0.8	-	-	-	-	-1.1	-	-	-	-	-
C	5	5	0	12	1	0.20	0	14.0	7	1	0.20	-1	-2.5	7	3	0.60	0	-2	-	-	-	-	-
D	3	3	0	4	2	0.67	1	1.0	2	1	0.33	0	0.7	-	-	-	-	0	-	-	-	-	-
E	0	0	0	1	0	0.00	0	1.0	-	-	-	0	0	-	-	-	-	0	-	-	-	-	-
F	0	1	-1	4	0	0.00	1	6.6	5	1	1.00	-2	3.8	-	-	-	-	-5.4	-	-	-	-	-
G	0	0	0	5	0	0.00	-1	4.0	4	0	0	-1	3.0	1	0	0	0	1	-	-	-	-	-
H	0	0	0	7	0	0.00	1	2.9	2	0	0	-1	1.0	-	-	-	-	-0.2	-	-	-	-	-
I	1	1	0	3	0	0.00	1	3.1	3	1	1.00	1	5.9	1	0	0	0	-6.1	-	-	-	-	-
J	5	5	0	16	4	0.80	0	14.0	6	1	0.20	0	7.0	1	0	0	0	2	-	-	-	-	-
K	0	0	0	3	0	0.00	0	3.0	4	0	0	0	6.0	1	0	0	0	-6	-	-	-	-	-
L	9	9	0	15	6	0.67	0	6.0	4	1	0.11	0	11.0	2	2	0.22	0	0	-	-	-	-	-
M	3	3	0	7	3	1.00	0	7.0	1	0	0	0	-1.0	-	-	-	-	0	-	-	-	-	-

Table 6. Continued

	COMBINED		PARTITIONS																			
	Sum PC		Cytb					Dloop					Actin					RAG2					
Branch	PC	Partitions	CDI	UA	PC	RC	SDI	PBS	UA	PC	RC	SDI	PBS	UA	PC	RC	SDI	PBS	UA	PC	RC	SDI	PBS
O	1	1	0	1	0	0.00	0	1.0	4	1	1.00	0	4.0	-	-	-	-	0	-	-	-	-	-
P	1	3	-2	10	1	0.33	0	8.0	2	0	0	-1	-1.0	2	2	0.67	0	0	-	-	-	-	-
R	4	4	0	8	3	0.75	0	10.0	3	1	0.25	0	3.0	-	-	-	-	0	-	-	-	-	-
S	5	5	0	12	4	0.80	1	18.0	8	1	0.20	-3	-1.0	-	-	-	-	-3	-	-	-	-	-
T	1	2	-1	10	1	0.50	0	-	7	1	0.50	0	-	-	-	-	-	-	-	-	-	-	-
U	3	3	0	7	3	1.00	1	7.5	6	0	0	0	3.5	-	-	-	-	0	-	-	-	-	-
V	1	1	0	5	1	1.00	0	5.0	-	-	-	0	-4.0	-	-	-	-	0	-	-	-	-	-
W	10	10	0	15	8	0.80	0	21.0	2	0	0	-1	3.0	7	2	0.20	0	3	-	-	-	-	-
X	4	5	-1	9	4	0.80	0	2.0	3	1	0.20	-1	0	-	-	-	-	0	-	-	-	-	-
Z	4	4	0	15	3	0.75	0	14.5	4	1	0.25	-1	4.0	-	-	-	-	-0.5	-	-	-	-	-
AA	3	3	0	13	2	0.67	0	6.0	3	1	0.33	-1	-5.0	-	-	-	-	0	-	-	-	-	-
BB	14	15	-1	40	8	0.53	0	28.5	21	6	0.40	-4	7.5	4	1	0.07	0	1	1	1	0.07	0	1
CC	18	19	-1	39	13	0.68	0	38.0	12	2	0.11	1	2.5	4	3	0.16	0	2.5	1	1	0.05	0	1
TOTAL	165	167	-2	340	104	0.62	5	278.1	143	31	0.19	-23	68	45	28	0.17	0	11.7	5	5	0.03	0	5

Discussion

Combined analysis and character dynamics

Several studies have presented methods to quantify the interaction of characters in combined analyses (e.g., Baker et al. 1998; Gatesy 1999; Damgaard & Cognato 2003). These measures are of generally two types: (1) topological indices that measure the change in tree length or structure ("topological congruence") (e.g., Miyamoto & Fitch 1995; Naylor & Brown 1998), and (2) measures of change in the amount of support at a particular node (e.g., Gatesy 1999, Cognato & Vogler 2001, Damgaard & Cognato 2003). The reliability of these measures and their philosophical foundation has

been questioned. In a recent review, Grant & Kluge (2003, pg. 409) advocated that *a posteriori* analyses of character partitions are heuristic only when "based on the results of the total-evidence analysis," and that there is "a great potential for the development of heuristic methods of *a posteriori* analysis of sets of characters." In this study, I have attempted to heuristically examine with statistical analyses the contributions and interactions of data partitions in the resolution of relationships within the family Delphinidae.

Variation in the expected amounts of phylogenetic support within a phylogeny can be investigated to determine the degree of interaction among characters. In this study, each data partition provided phylogenetic resolution as expected, i.e., nuclear genes had significantly fewer parsimony-informative characters than mitochondrial genes (Tables 2 and 4), and nuclear genes tended to provide the most support at the base of the tree and mitochondrial genes at more distal nodes (Table 5 and Fig. 13). For example, 60% of RAG2 and 54% of actin PC substitutions were at the base of the phylogeny (Table 5) when compared to 20% and 33% for the control region and cyt *b*, respectively. This pattern is expected for nuclear and mitochondrial genes that reveal as much as 10-fold differences in rates of evolution in mammals (Brown et al. 1979). One also might expect that the proportion of nodal support contributed by a particular data partition should not change with the addition of data unless there were conflicts among characters. However, because of variation in rate and patterns of substitution, a particular process partition analyzed separately is likely to interact to some degree with characters from other partitions, and the strength of this interaction is related to the

degree of homoplasy in each data set. In this study, I have no statistical evidence for data interaction, with the exception of positive interaction between Rag2 and other partitions. The statistical tests for data interaction were not significant, and there was no strong indication for saturation of either 3rd codon transitions or overall transitions in the combined data set (Fig. 11). Nevertheless, it is apparent from the distribution of positive and negative PBS values at each node that the contribution of the control region to phylogenetic stability was more consistently negative than other partitions (Fig. 12). In fact, 43% of nodes in Fig. 12 had a negative PBS value for the control region, while only 17% of nodes had negative PBS values for actin. In contrast, negative PBS negative values for both Rag2 and cyt *b* were rare (Fig. 12, Table 5), and nearly all of the negative PBS values for actin and the control region were coupled with positive PBS values from cyt *b* (Fig. 6, Table 5). Furthermore, unambiguous character changes from the control region partition were more likely to be shifted or displaced in the combined analysis (SDI, Table 5). This suggests that cyt *b* overall, and actin secondarily, provided the most positive support for the SA, and the most conflict among datasets was due to interaction with the control region. This pattern is further demonstrated by the phylogenies derived from independent analysis of each character partition. The phylogenies of actin and Rag2 are considerably less resolved but are never in conflict with the cyt *b* phylogeny; only the control region phylogeny contradicts that of cyt *b*. Despite the evidence that cyt *b* contributes the most to node support in the SA phylogeny, the lack of resolution of the cyt *b* phylogeny of LeDuc et al. (1999; Fig. 7) demonstrates that the analysis of nuclear and mitochondrial process partitions provided greater resolution of the delphinid

phylogeny than cyt *b* alone. Therefore, this study demonstrates unequivocally the explanatory power of combining data sets, despite evidence for localized character conflict.

Under the assumption of a molecular clock, the distribution of informative characters or substitutions within a phylogeny (i.e., branch lengths) is expected to be uniform with respect to a particular data partition. That is to say, one would expect no variation in the number of substitutions at different node heights in the phylogeny for a given data partition. These data do not indicate this pattern (Fig. 13). The number of perfectly consistent characters (or non-homoplasious substitutions) was statistically different among data partitions (Table 6). The lack of uniformity in the rate of evolution within each data partition is not surprising, given the multitude of evidence against a constant rate of substitution within and among genes (Li & Tanimura 1987; Allard & Honeycutt 1991; Martin & Palumbi 1993; Honeycutt et al. 1995; Sullivan et al. 1995; Yang 1995; Eyre-Walker & Gaut 1997; Gaut et al. 1997; Xia 1998; Adkins et al. 1996; 2001). What is more surprising, however, is that the pattern of variation in the number of substitutions within the phylogeny is consistent among partitions for all node height categories (Fig. 13). This suggests that, although each data partition varies significantly in the amount of information contributed to phylogenetic resolution, the relative contribution of each data partition at a given node height was the same (Fig. 13). The number of substitutions of each data partition was not consistent across the phylogeny, but they are consistently inconsistent with respect to each other. The implications of these results are two-fold. First, I suggest that the consistent lack of uniformity in the

number of informative characters among data partitions provides strong evidence for shifts in the rate of diversification of delphinid lineages since their divergence from the Phocoenidae around 10 mya. Second, the rapid reduction in the number of PC in the middle of the phylogeny and the corresponding short branch lengths and large number of lineages indicates that a shift in diversification rate occurred most recently at the base of the subfamily Lissodelphininae (Fig. 13). Given that 8 of the 10 species in this subfamily are found only in the Southern Hemisphere, this may indicate diversification of lineages south of the equator following equatorial transgression from the north.

Taxonomic implications

The total evidence phylogeny (Fig. 9) has several taxonomic implications. First, the results support the monophyly of the Lissodelphininae (Fig. 9), and therefore concur with Le Duc et al. (1999) that the genus *Lagenorhynchus* is polyphyletic. Le Duc et al. (1999) proposed that the generic name *Lagenorhynchus* remain with *L. albirostris*, the type specimen of the genus (Gray 1828), and that *L. acutus* be placed in it's own genus *Leucopleurus* (Gray 1866). In contrast to Le Duc et al. (1999), 100% of the MP trees (n = 4) from the SA analysis in this study recovered the monophyly of *L. acutus* and *L. albirostris* (Fig. 9, node 40), and supported their close relationship with *the Delphinus/Stenella/Tursiops* clade (Fig. 9, node 44). I therefore propose that both *L. acutus* and *L. albirostris* retain their current generic designation.

Second, the addition of mtDNA and nDNA process partitions produced a topology with 100% consensus for the paraphyly of *Lagenorhynchus* within the subfamily Lissodelphininae (Fig. 9). The sister-group relationship of *L. obliquidens* and

L. obscurus and the monophyly of the *Cepharlorhynchus* have been supported by several molecular phylogenetic studies (Cipriano; 1997; Le Duc et al. 1999; Pichler et al. 2001; Hare 2002). In addition, the parsimony analysis of Le Duc et al. (1999) recovered the monophyly *L. cruciger* and *L. australis*, but their analysis did not resolve the evolutionary relationship of this clade with the *L. obscurus/L. obliquidens* clade (Fig. 7). The evidence concurs with the results of Le Duc et al. (1999) that *L. cruciger*/*L. australis* are monophyletic and should be placed into a separate genus, *Sagmatius*, first described from *Sagmatius amblodon* (Cope 1866) and later synonymized with *L. australis* (Kellogg 1941). However, the results suggest that *L. obscurus* and *L. obliquidens* form a monophyletic group independent of *L. cruciger* and *L. australis*.

Molecular data have helped to resolve some of the problems associated with dolphin taxonomy, but it is apparent that issues related to variation in the rate of evolution within the family continue to make the resolution of some relationships problematic. These data support the taxonomic revision of the genus *Lagenorhynchus*, but the branches at the base of the subfamily Lissodelphininae are relatively short and contain significantly fewer PC nucleotide substitutions than other regions of the SA phylogeny (Fig. 13). Given the weak corroboration of the hypothesis of the polyphyly of *Lagenorhynchus*, I do not propose a unique generic name for the *L. obscurus/L. obliquidens* clade at this time. However, it is very evident that these species are not part of *Lagenorhynchus* sensu Gray (1828), and the relationship of these taxa to *Sagmatius* and *Cephalorhynchus* warrants further investigation. The next step in this study would be to combine morphology and DNA sequence characters in a total evidence analysis.

For example, Fraser and Purves (1960) found distinct differences among genera of dolphins based on a suite of characters in the sinuses related to the structure and function hearing. A re-evaluation of these characters and additional molecular loci for all members of the Delphinidae might provide resolution of evolutionary relationships for problematic taxa.

CHAPTER IV

CATCHING THE DRIFT: PHYLOGEOGRAPHY OF THE DUSKY DOLPHIN (*LAGENORHYNCHUS OBSCURUS*) AND THE WEST-WIND DRIFT HYPOTHESIS

Introduction

A rapid rate of coalescence, uniparental inheritance, and lack of recombination (Brown 1979; Giles et al. 1980) enhance the usefulness of mtDNA as the molecule of choice for studies of intraspecific phylogeography (Avise 1987). Despite these genetic attributes, mtDNA provides a single gene perspective depicting the genetic history of the female lineage only (Moritz 1994). Clearly, a decipherment of the phylogeographic history of a species, especially as it relates to population processes associated with demography, dispersal patterns, population structure, etc., requires the use of multiple markers inherited both uniparentally and biparentally (Sugg et al. 1996; Hoelzer et al. 1998; Ross et al. 2001). Recently, multiple genetic loci (both mitochondrial and nuclear) have been used to address questions related to historical patterns of evolutionary divergence within and between species (Bernardo et al. 1993; Palumbi & Baker 1994; Templeton 2002; Antunes et al. 2002). In terms of phylogeography (Avise et al. 1987) the addition of neutral nuclear markers provides an opportunity to distinguish between alternative explanations for inferred population history (Hare 2001; Zhang & Hewitt 2003). In theory, neutral nuclear loci coalesce 3 times more slowly on

average than mtDNA genes (Palumbi et al. 2001), thus providing more temporal depth for studying the phylogeography of recently diverged intra- and inter-specific lineages (Hare 2001; Zhang & Hewitt 2003).

Here I utilize a combination of nuclear and mitochondrial DNA sequences to infer the phylogeographic history of the dusky dolphin (*Lagenorhynchus obscurus*) (Gray 1846), a meso-pelagic, small-bodied species occurring in cool, temperate waters in parts of the Southern Hemisphere (Fig. 14). Although its range is large, the dusky dolphin's distribution is discontinuous and represented by large populations along the coasts of New Zealand, Peru, Argentina, and South Africa where deep, temperate waters approach close to shore. In New Zealand and along the west coast of South America a narrow continental shelf and deep near-shore waters appear to support large aggregations that feed on meso-pelagic fishes and squid (Cipriano 1992; Würsig et al. 1997).
The disjunct, yet broad distribution coupled with patterns of morphological and genetic (exclusively mtDNA) divergence have resulted in the proposal of several hypotheses related to the phylogeogeographic history of dusky dolphin. First, the west-wind drift hypothesis relates patterns of population divergence in both the dusky dolphin and the dolphin genus *Cephalorhynchus* to the eastward flowing, virtually uninterrupted current system in the Southern Ocean that carries cool water around the Antarctic. According to Pichler et al.'s (2001) mtDNA study, dolphins appear to have speciated along continental coastlines following dispersal with the west-wind drift. Second, Hare et al. (2002) used mitochondrial and nuclear genes to examine the demographic history of *L. obscurus* and *L. obliquidens*, and concluded that the dusky dolphin likely arose in the

Southern Hemisphere from a propagule of >10 *L. obliquidens* approximately 0.74 million years ago. However, their analyses lacked data from Argentina, and assumed that *L. obscurus* from New Zealand and Peru adequately represented overall genetic variation in the species. They suggested the lack of lineage sorting (e.g., Avise et al. 1984) within nuclear gene trees to be the result of large effective population size in both species. Nevertheless, an alternative explanation for this pattern is that an incomplete sample of lineage diversity has lead to an under-representation of the number of nodes at the distal portion of the phylogenies, thus mimicing a slowdown in lineage coalescence (Nee 1994).

Third, the morphological and molecular (mtDNA) discontinuity between Peru and the other populations of dusky dolphin have prompted several hypotheses. Although studies by both Van Warebeek (1995) and Cipriano (1997) imply an early divergence for the Peruvian stock, its overall placement relative to other populations and the recency of its divergence is still controversial. For instance, based on skull size and similarities with the Northern Hemisphere species, *L. obliquidens*, Van Warebeek (1993) suggested that the Peruvian population is ancestral to the small South African and New Zealand forms, thus implying a Pacific origin for the dusky dolphin. In contrast to Van Warebeek's hypothesis, analysis of complete cytochrome *b* sequences suggested a more basal position of Atlantic lineages in the gene tree of the dusky dolphin (Cassens et al. 2003); however, their results were ambiguous regarding the placement of the root within Atlantic lineages. Finally, aside from Peru, the historical and recent pattern of gene flow among the remaining three regions is unclear. Cassens et al. (2003) revealed significant

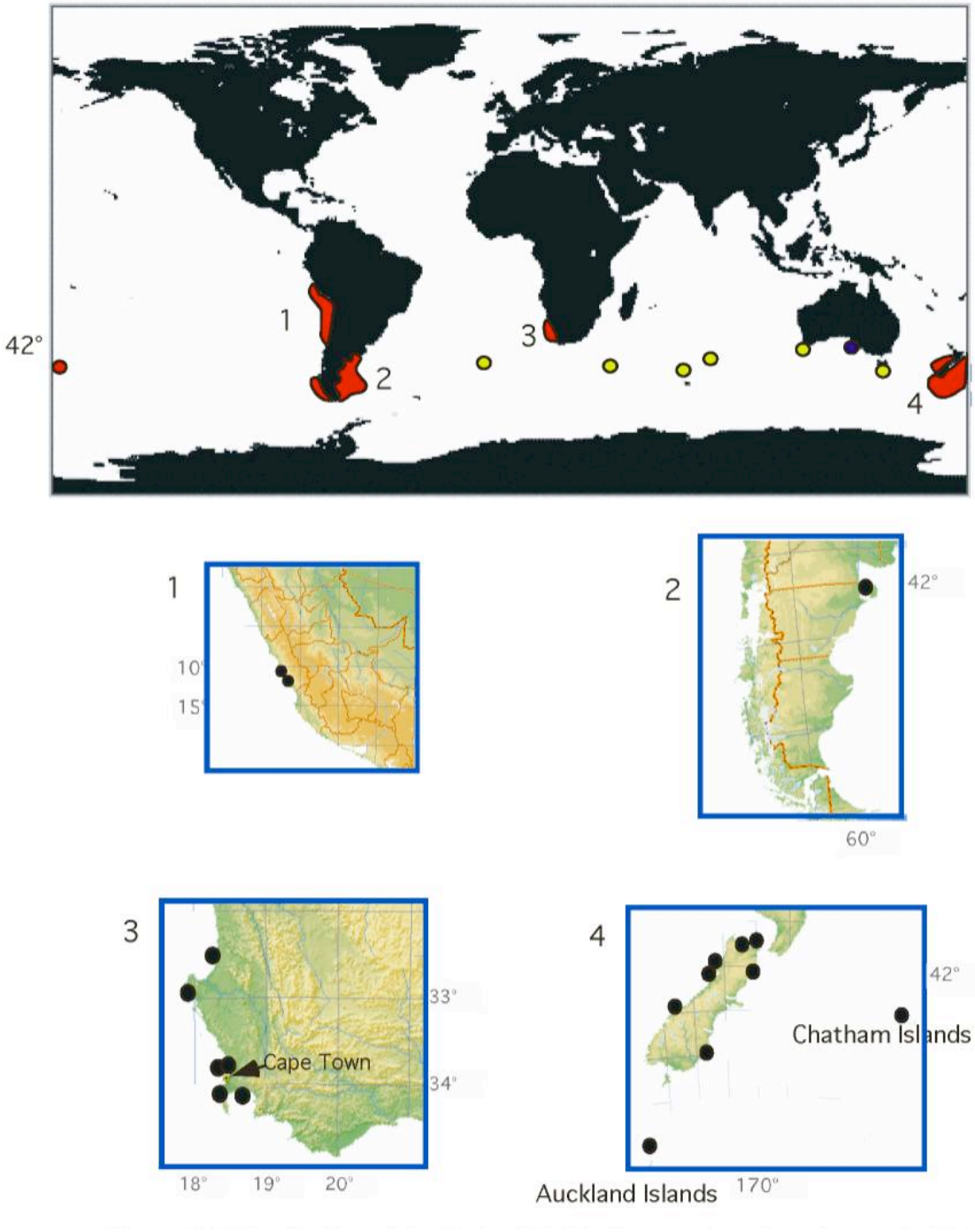


Figure 14. Distribution of the dusky dolphin (*Lagenorhynchus obscurus*). Red areas indicate regions of known concentrations along continental coastlines, and yellow dots demarcate oceanic islands with at least one confirmed sighting of the species (from Van Warebeek et al. 1995). Boxed inserts depict localities from which samples were collected for this study.

discontinuity between Southwest Africa and Argentina, yet found no obvious geographic partitioning of mtDNA variation. Although focused on New Zealand, Harlin et al. (2003) reported two distinct clades, one of which appeared to represent a more recent lineage expansion with the other possibly representing more ancestral mitochondrial lineages. If true, then this implies ongoing coalescence in the New Zealand stock as a result of restricted gene flow. This hypothesis can be tested with more global sampling.

Clearly, the above hypotheses cannot be addressed adequately until all four biogeographic areas (especially New Zealand) containing the dusky dolphin have been adequately incorporated into a detailed analysis. Furthermore, an assessment of recent and historical patterns of gene flow requires the use of both nuclear and mitochondrial markers, and these markers should be analyzed using similar methods. This study provides a global phylogeographic perspective by incorporating a large number of New Zealand samples along with samples from the other three biogeographic areas. It utilizes both nuclear and mitochondrial markers and tests hypotheses related to: 1) the origin of dusky dolphin populations; 2) the phylogeographic history of those populations; and 3) the degree to which recent and historical processes have shaped current patterns of genetic divergence within and between populations.

Materials and methods

Sampling, DNA extraction, amplification, and sequencing

Tissues of *L. obscurus* were obtained from a variety of sources including biopsy punches and skin swabs from living animals, and post-mortem samples from either

beach-cast or net-caught individuals. Skin swabs were collected following the noninvasive procedure of Harlin et al. (1999), preserved in either 90% ETOH or a solution of 20% dimethylsulfoxide (DMSO) saturated with salt, and stored at -20°C. When possible, DNA from the same individuals was used to amplify and sequence all gene regions. The number of individuals examined from each of the four geographic regions (New Zealand, Peru, Argentina, and South Africa) varied for each of the three gene partitions (Table 7). Tissues from Peru were not available for this study, but sequences from Cassens et al. (2003) were acquired from GenBank to complete the geographic sampling (Accession numbers AY257126-AY257161).

Total genomic DNA was isolated using either a standard phenol-chloroform protocol (Maniatis et al., 1982) or a Qiagen DNeasy kit (Qiagen, Valencia, California). Isolated DNA was stored in tris-borate-EDTA (TBE) buffer at -20° C. The polymerase chain reaction (PCR) was used to amplify fragments of two mitochondrial genes and one nuclear gene: (1) the complete mitochondrial cytochrome *b* (CYTB) gene (1040 nucleotides), (2) 474 base pairs (bp) of the mitochondrial control region (CR), (3) 995 bp of nuclear DNA (nDNA) intron I of the muscle actin gene (ACT1). External primer sets included: (1) cyt *b* - 766F (5'GAAAAACCAYCGTTGTWATTCAACT3') and 766R (5'GTTTAATTAGAATYTYAGCTTTGGG3'); (2) control region - tRNA-Pro and Dlp5 of Baker et al. (1996); (3) actin - *Lagenorhynchus* specific LagAct1 (5'GATTTGGTCCCTCTATGTCTCT3' and LagAct2 - 5'TACTTTTGAACTTGCCACCTAC3'). ACT1 primers were designed from published cetacean sequences (Palumbi and Baker 1994; Hare et al. 2002). PCR conditions were

Table 7. Summary of diversity indices for cytochrome b (Cytb), the control region (Dloop) from the mitochondrial genome, and actin intron 1 (actin) from the nucleus among geographic localities. AR = Argentina, SA = South Africa, NZ = New Zealand, PE = Peru. n = number of mitochondria or chromosomes examined (for mtDNA and nDNA, respectively). *h* = haplotypic diversity, π = nucleotide diversity. S = number of segregating sites for mtDNA regions, or the number of actin haplotypes present in each geographic region. Standard errors of estimates (where applicable) are in parentheses.

Partition	Region	n	*h*	π	S
Cytb	AR	34	0.93 (± 0.030)	0.0046 (± 0.003)	32
	SA	40	0.91 (± 0.027)	0.0047 (± 0.003)	34
	NZ	23	0.91 (± 0.052)	0.0063 (± 0.003)	37
	PE	78	0.71 (± 0.048)	0.0024 (± 0.001)	18
	Total	175			
Dloop	AR	16	0.91 (± 0.048)	0.0066 (± 0.004)	10
	SA	67	0.99 (± 0.005)	0.0170 (± 0.009)	47
	NZ	172	0.98 (± 0.005)	0.0187 (± 0.010)	82
	Total	255			
actin	AR	14	0.70 (± 0.095)	-	4
	SA	44	0.76 (± 0.049)	-	7
	NZ	48	0.82 (± 0.028)	-	7
	Total	116		-	

generally consistent across loci with adjustments made to annealing temperatures. Approximately 1 -2 µl of DNA template were included in 50 µl PCR reactions containing the following: 5 µl each 10X Amplitaq PCR buffer (Perkin Elmer, Boston, Massachusetts), 25 mM MgCl, and 10 mM deoxynucleotide triphosphates (dNTP's, 2.5 mM each), 1 µl each of 10mg/ml bovine serum albumin (BSA), 10 uM of each primer, and 5 units (U) of Amplitaq (Perkin Elmer, Boston, Massachusetts) DNA polymerase. Thermocycler conditions were 94° C for 2 min followed by 35 cycles at 92° C for 30 s, annealing 30 s, and extension at 72° C for 30 s. Published annealing temperatures were

used with the following exceptions: 765F/766R, 50° C; LagActin1/2, 58° C. Amplicons were electrophoresed in 1.5% agarose-TBE (tris, boric acid, EDTA) and visualized under UV light, and prior to sequencing, excess oligonucleotides and dNTP's were removed with either Qiagen (Qiagen, Valencia, California) spin-columns or an Exonuclease I-Shrimp Alkaline Phosphotase (Exo-Sap) enzymatic procedure. Approximately 2 ng of cleaned PCR product per 100 bp of amplicon length was sequenced using ABI BigDye (Applied Biosystems, Foster City, California) cycle sequencing chemistry and an ABI 377 automated sequencer. Internal primer pairs for cyt *b* (560, 5'-GCAACCCTAACACGATTCTTCG-3'; 610, 5'-CCAGTTTCGTGTAGGAATAATAGG- 3') and actin (Act5-L, 5'CCACTACTTTAGGCAG3'; M13Act5R-H, 5'TGTAAAACGACGGCCAGTCTGCCTAAACTAGTGG3' (S. Palumbi, unpublished) were used in sequencing reactions to verify sequence ambiguities.

Data analysis

Sequence alignment and haplotype determination. Sequenced fragments were edited and compiled with the program Sequencher v. 4.1 (Gene Codes Corporation, Ann Arbor, Michigan). A consensus of sense and antisense strands for each individual and data partition were compiled and exported to MacClade vs. 4.05 (Maddison and Maddison, 1992). Amino acid translations of the open reading frames of CYTB were examined for stop codons to verify sequence orthology. Neither ACT1 nor CYTB contained any insertions or deletions. The CR had a single indel that was accommodated in the alignment with a minor alignment adjustment. ACT1 sequences were subjected to a

BLAST search to verify sequence orthology. In all BLAST searches, amplicons retrieved sequences from other delphinid or mammalian taxa as the closest match, providing evidence for successful amplification of the target locus.

Heterozygous nucleotide positions of the ACT1 intron were determined objectively with the program Mutation Surveyor (SoftGenetics, State College, Pennsylvania). This software identifies with 99% accuracy heterozygous nucleotide positions based on an algorithm that incorporates measures of peak height, peak intensity, and background noise from ABI electrophereogram files. Putative heterozygous positions identified with Mutation Surveyor were cross-verified by examination of raw electrophereogram data in Sequencher. Only cross-verified sites were used in subsequent determination of allelic variation of the locus. Haplotype and genotype frequencies of ACT1 were determined statistically with the program PHASE version 2.1 (Stephens et al. 2001; Stephens and Donnelly 2003). Three independent chains of 1000 interactions each were performed with other parameters as default. Only those haplotypes and genotypes resolved with > 90% posterior probabilities were used in subsequent analyses.

Phylogenetic analyses. The statistical parsimony method of Templeton et al. (1993) as implemented in the program TCS version 1.18 (Clement et al. 2000) was used to construct separate haplotype networks for each of the three data partitions. The statistical parsimony method was chosen because of its ability to infer missing intermediate haplotypes within the network, which is more appropriate for examining

intra-specific relationships than standard phylogenetic methods that assume bifurcating lineages (Crandall et al. 1994; Posada & Crandall 2001). Furthermore, reticulations, or alternative (homoplasious) parsimonious connections among haplotypes, are permitted with this method, which allows one to test hypotheses regarding the complex historical processes that shape populations. The final network gives a visual representation of the number of mutations separating any two haplotypes. Mitochondrial DNA haplotypes were joined to the network within 95% parsimony confidence limits (Clement et al. 2000). Because of its slower rate of coalescence, parsimony confidence limits were expanded to 90%. Similarly, *L. obliquidens* actin haplotypes from GenBank sequences (accessions AF14026-AF14031) were included in the statistical parsimony network to examine the pattern of allelic diversity at the base of the *L. obliquidens/L. obscurus* divergence. Reticulations in the actin parsimony network were observed and attributed either to lack of lineage coalescence, homoplasious substitutions, or recombination among sites. Of these, recombination is the most problematic for inferring evolutionary history from nuclear loci. Therefore, the likelihood of recombination among haplotypes was evaluated with the program Lamarc version 1.2.2 (Kuhner et al. 2004).

The root of the *L. obscurus* haplotype network was determined using the outgroup method and a phylogenetic analysis of CYTB sequences in a Bayesian framework (Rannala & Yang 1996). Bayesian analyses were performed with Mr. Bayes version 3.0 (Huelsenbeck & Ronquist 2001). *Lissodelphis borealis*, a basal member lissodelphine subfamily (Le Duc et al. 1999), was used as the outgroup taxon, and *L. obliquidens*, the sister taxon to *L. obscurus* (Le Duc et al. 1999), was included as an

intermediate divergence point within the lissodelphine lineage. The Bayesian Metropolis-coupled Markov Chain Monte Carlo (MCMCMC) analysis consisted of two independent searches of 25 x 10^6 steps (1 cold, 3 heated chains to insure adequate mixing) with the best-fit model of substitution for CYTB determined *a priori* with Modeltest version 3.06 (Posada & Crandall 1998) as priors. The program Tracer version 1.1.1 (Rambaut & Drummond 2004) was used to trace convergence and determine the burn-in point of chains. Tracer also depicts the posterior probability density distribution from the searches, with a summary of the mean and standard deviations of log likelihood values. Trees prior to convergence were eliminated and a majority-rule consensus tree of the remaining topologies was constructed in PAUP* version 4.0b10 (Swofford 1999). The clade at the *L. obliquidens/L. obscurus* split in the Bayesian topology was considered the most probable root of *L. obscurus* divergence.

Diversity indices. Global and population haplotype (i.e., gene diversity) (*h*) and nucleotide diversity (π) (Nei 1987) and the average number of substitutions between sequences were calculated with Arlequin version 2.00 (Schneider et al. 2000). Standard errors of these estimates were calculated with 1000 random permutations of the data. Distances among haplotypes were inferred with models of substitution derived with Modeltest. DnaSP version 4.0 (Rozas et al. 2003) was used to determine the number of polymorphic sites (S) was determined from aligned sequence data.

Population divergence and geographic structure. An analysis of molecular variance (AMOVA) (Excoffier et al. 1992) and Wright's F-statistics were calculated with the program Arlequin version 2.00 (Schneider et al. 2000) to determine amount of genetic variation within and among geographic regions. The significance of these statistics was determined by comparison of observed values to the distribution of values derived from 1000 random permutations of the data. The "case-control" option in the program PHASE was used to group ACT1 haplotypes by geographic region and to subsequently test the null hypothesis of random association of haplotypes among localities. Test for Hardy-Weinberg equilibrium of ACT1 genotypes was performed in Arlequin version 2.00 (Schneider et al. 2000).

Results

Genetic diversity and parsimony networks

Levels of diversity varied according to patterns expected of mitochondrial and nuclear loci, with the CR having the highest levels of variation, followed by CYTB and ACT1, respectively (Table 7). Although the largest sample size was examined for Peru, levels of haplotype diversity ($h = 0.71$) were lower than those seen for both the CR and CYTB ($h > 0.90$) in all other regions (Table 7). Nucleotide diversity of the CR was consistently higher than CYTB within all geographic regions, and New Zealand demonstrated the highest level of diversity for all data partitions (Table 7). Interestingly, the number of segregating sites of the CR in each geographic locality increased consistently with sample size; however, the number of CYTB segregating sites (and

haplotypic diversity) was lowest in Peru despite the larger number of individuals sampled (Table 7).

The haplotype networks of CR and CYTB consistently recovered both a unique Peruvian clade separated by several missing intermediate haplotypes, and at least one cluster of haplotypes at the distal tip of the network restricted mainly to New Zealand (Figs. 15, 16). Within CR and CYTB networks New Zealand lineages were notably separated from the Peruvian clade by a mixture of haplotypes from South Africa, Argentina, New Zealand, and several missing intermediates (Figs. 15, 16). For both mitochondrial data partitions, the networks lacked a central dominant haplotype (Fig. 15;"CMI", Cassens et al. 2003), and the majority of the most frequent haplotypes appeared to be clustered at the ends of the networks (Figs. 15, 16). The Bayesian analysis placed the root of the CYTB phylogeny along a lineage of haplotypes found in South America and New Zealand (Fig. 15). Furthermore, clusters of haplotypes defined by the statistical parsimony analysis were supported with ≥ 90% posterior probability, supporting the pattern observed within the networks (Fig. 15). The likelihood of recombination among ACT1 haplotypes was insignificant, likely due to the relatively few polymorphic sites examined and their close placement (within 500 bp) in the intron. Reticulations were therefore attributed to a lack of lineage coalescence. The ACT1 network (Fig. 17) was dominated by three common haplotypes (a, b, l) that were at the center of the network and present in varying frequencies in all four biogeographic regions and in *L. obliquidens*. The tips of the network contained haplotypes present in lower frequency and with more restricted geographic distributions. In 3 of the 5 tip

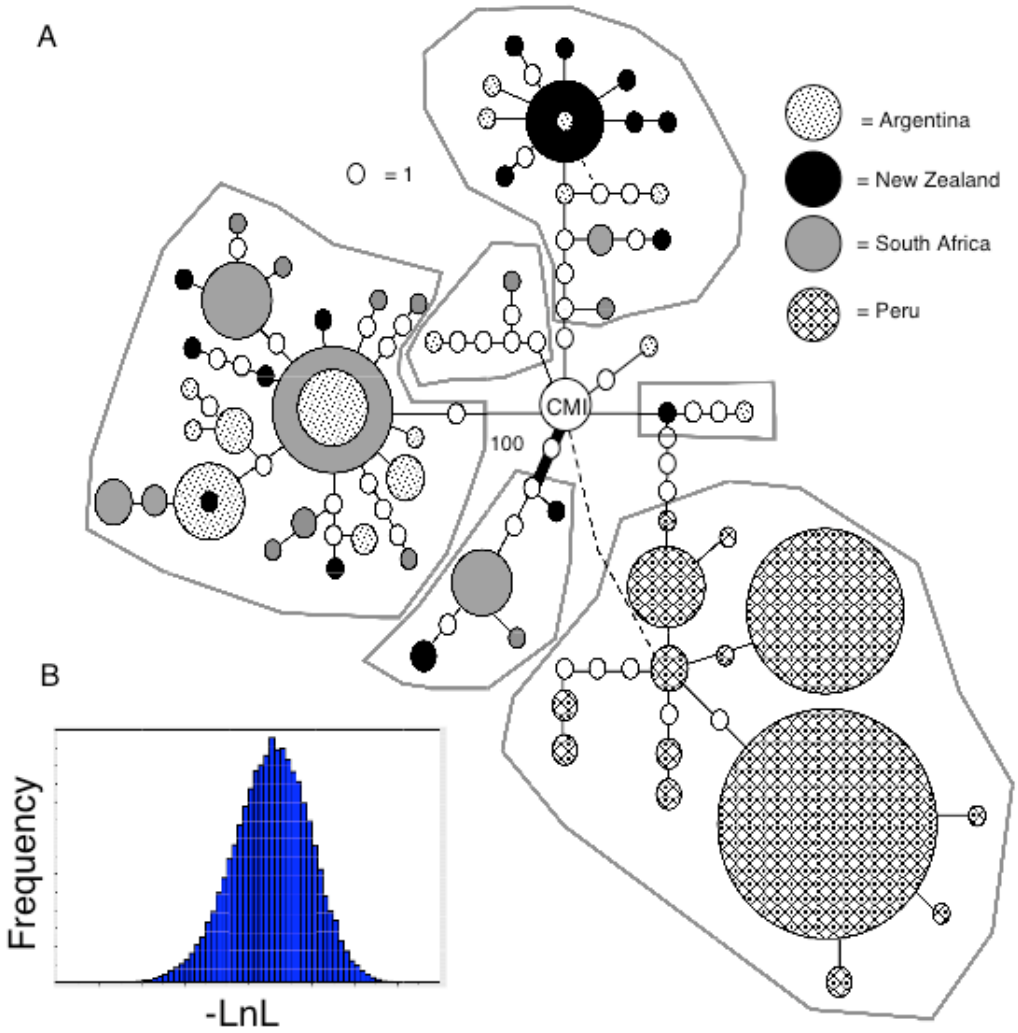


Figure 15. A) Statistical parsimony network of *Lagenorhynchus obscurus* cytochrome *b* haplotypes. Size of the circles is proportional to haplotype frequency. Unshaded circles represent inferred missing intermediates. Gray polygons group haplotypes that form monophyletic groups with a Bayesian posterior probability of 90% or greater. The 100% probable root of the phylogeny is indicated by the thick branch. 'CMI' = central missing intermediate (Cassens et al. 2003). B) The posterior probability density distribution of the Bayesian phylogenetic analysis that determined the root of the network.

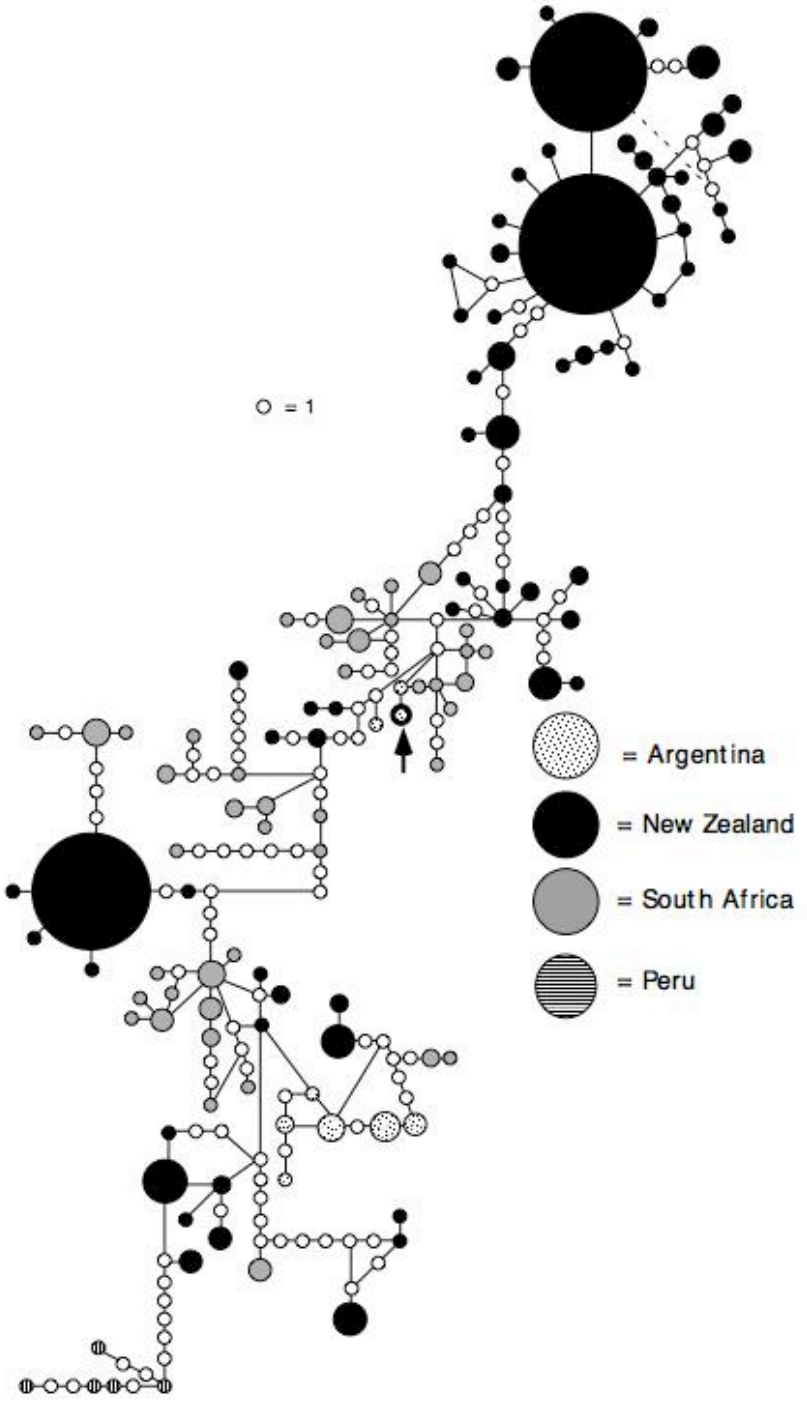


Figure 16. Statistical parsimony network of control region haplotypes. Size of shaded circles is proportional to haplotype frequency, and unshaded circles represent inferred missing intermediates.

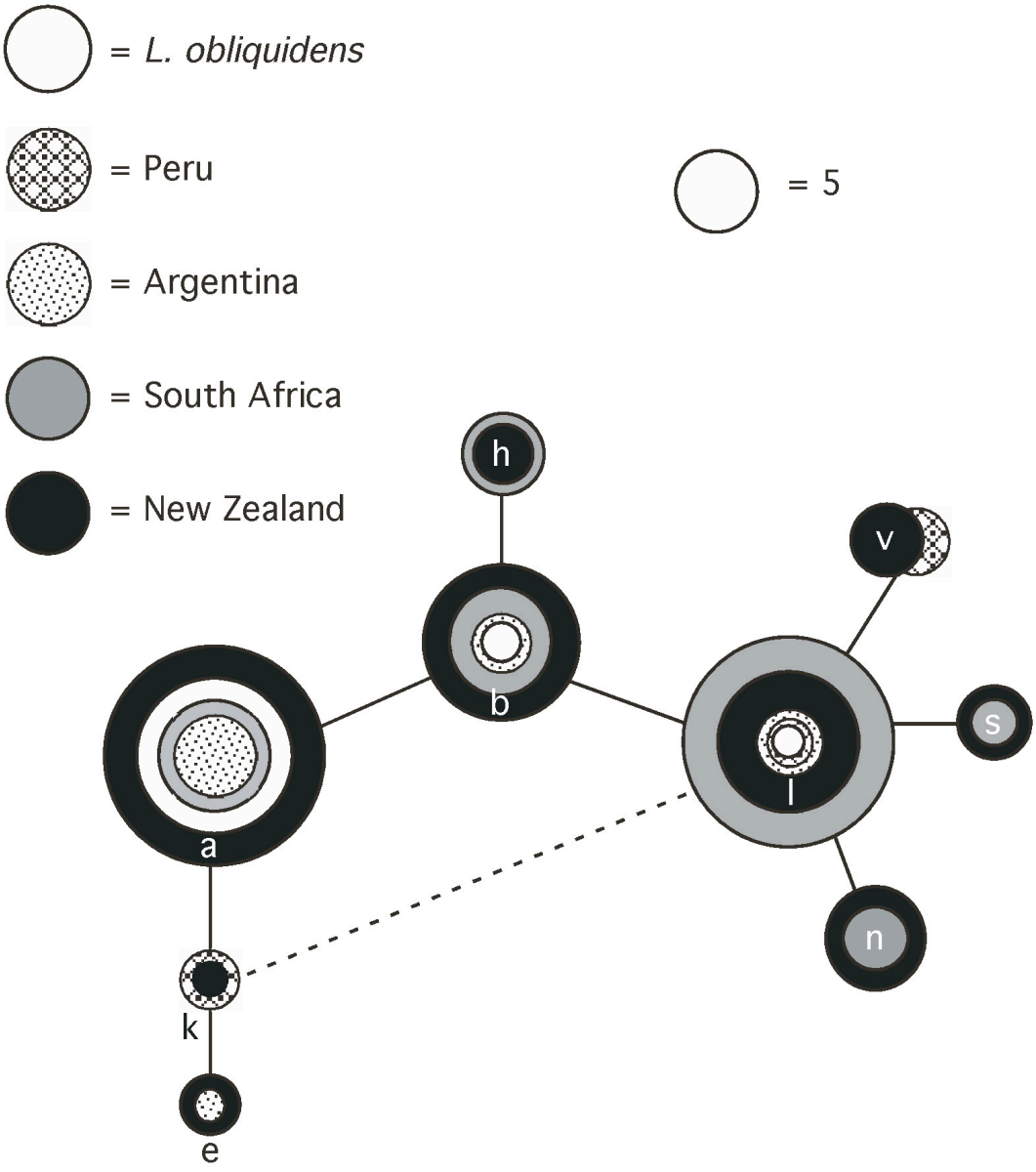


Figure 17. Parsimony network of actin haplotypes. Haplotype frequencies are proportional to sample size measured by comparison to the scale at the top right. Frequencies of haplotypes found in multiple regions are nested and shaded accordingly or offset for visualization if frequencies were equal. The dotted line indicates one alternative equally parsimonious connection among haplotypes (i.e., reticulation).

clades, haplotypes were common to South Africa and New Zealand. It is noteworthy that in only one instance was a haplotype from Argentina found at the periphery of the network (haplotype 'e'), and this haplotype was also sampled in New Zealand.

Haplotypes were not randomly distributed among geographic localities. For each of the three genetic markers, AMOVA and F_{st} values revealed significant partitioning of variation among Peru, Argentina, New Zealand, and South Africa (Table 8). The null hypothesis of random association of ACT1 haplotypes with geographic locality was rejected ($P = 0.04$), thus providing additional support from a nuclear locus for genetic structure among regions. Detailed examination of the frequency and distribution of haplotypes among geographic regions clearly depicted the source of the high levels of divergence among localities (Tables 3, 4). For example, only three haplotypes of 175 CYTB haplotypes and one haplotype of 255 CR sequences were found in multiple regions, but never common to more than two of the four regions (Table 9). Peruvian haplotypes were not found in any other locality for either the CYTB or CR (Table 9). The distribution of nuclear ACT1 intron diversity was similar in pattern to that of the mtDNA, but with weaker demarcation among regions. For example, haplotypes 'a' and 'l' were found in South Africa, Argentina, and New Zealand but at different frequencies (Table 10). Haplotypes 'n', 's', and 'h' were shared by South Africa and New Zealand, but haplotype 'e' was present in Argentina and New Zealand, and absent in South Africa (Table 10). The frequency of genotypes with >90% posterior probability also were not equal among geographic localities, but H-W equilibrium of genotype frequencies was

rejected in all regions but New Zealand ($P > 0.05$). Therefore, although the pattern of intra-regional genotypic variation was not in conflict with results derived from haplotypic data, dispersion in genotype frequency was not considered as reliable evidence of population divergence.

Table 8. Genetic divergence among *L. obscurus* populations. A. Analysis of molecular variance tables for the mitochondrial cytochrome b (Cytb) and control region (Dloop) and nuclear actin intron 1. Populations used in each analysis for each of the genetic partitions are indicated in parentheses with the following designations: NZ = New Zealand, AR = Argentina, SA = South Africa, PE = Peru. Phist values and their significance levels are indicated in bold. B. Estimates of pairwise divergence among geographic regions for mitochondrial and nuclear data partitions. For Cytb and Dloop, numbers above the diagonal represent the average number of substitutions between groups corrected for sample size. Values on the diagonal are the average pairwise differences within each group. For Cytb, Dloop, and actin, numbers below the diagonal are pairwise Fst values based on haplotype frequencies. Asterisks indicate statistically significant values. * = $P < 0.01$; ** = $P < 0.001$

A. AMOVA

	Source of variation	d.f.	Sum of Squares	Variance	% variation	**Phist**
Cytb	Among populations	3	260.92	2.09 Va	48.95	
						0.49**
NZ, AR, SA, PE	Within populations	171	373.42	2.18 Vb	51.05	
	Total	174	634.34	4.28		
Dloop	Among populations	2	126.88	0.99 Va	19.23	
						0.19**
NZ, AR, SA	Within populations	252	1042.77	4.14 Vb	80.77	
	Total	254	1169.65	5.12		
actin	Among populations	2	8.06	0.11	22.59	
						0.23**
NZ, AR, SA	Within populations	103	40.77	0.39	77.41	
	Total	105	48.83	0.50		

B. Pairwise divergence

	NZ	AR	SA	PE
NZ	**7.18**	1.59*	2.42**	5.92*
AR	0.21**	**5.23**	0.28**	5.07*
SA	0.29**	0.05*	**9.20**	5.22*
PE	0.61**	0.59**	0.59**	**2.70**

	NZ	AR	SA
NZ	**8.82**	3.42**	1.7**
AR	0.28**	**3.11**	1.72**
SA	0.16**	0.18**	**8.01**

	NZ	AR	SA
NZ	-		
AR	0.23**	-	
SA	0.21**	0.26**	-

Table 9. Absolute frequency of haplotypes among regions for *L. obscurus* mitochondrial cytochrome b (Cytb), and control region (Dloop). AR = Argentina, SA = South Africa, NZ = New Zealand, PE = Peru. 'n' represents the sample size.

Cytb

HAP	AR	SA	NZ	PE	Total
HAP1		8	8		16
HAP2		2		7	9
HAP3		3		1	4
HAP4		3			3
LA14		2			2
LOA6		1			1
LA70		1			1
LA81		3			3
LA82		1			1
LoA9		1			1
LA18		1			1
LA19		1			1
AR01		1			1
AR03		1			1
AR05		1			1
AR07		1			1
AR15		1			1
AR09		1			1
AR14		1			1
HAP6			8		8
HAP7			3		3
LA10			1		1
LA12			1		1
LA14			1		1
LA15			2		2
LA17			1		1
LA21			1		1
LoA3			2		2
LoA4			4		4
LoA5			1		1
S985			1		1
SA03			1		1
SA04			1		1
SA08			1		1
SA12			1		1
SA15			1		1
SA21			1		1
HAP5				2	2
K090				1	1
K092				1	1
K223				1	1
K089				1	1
K000				1	1
LO03				1	1
LO05				1	1

Dloop

HAP	AR	SA	NZ	PE	Total
AR01	2				2
AR02	2				2
AR05	3				3
AR07	4				4
AR12	1				1
AR13	1				1
AR14	1				1
AR16	1				1
SAF01		3			3
SAF02		2			2
SAF03		2			2
SAF04		4			4
SAF05		1			1
SAF06		1			1
SAF07		1			1
SAF08		1			1
SAF09		1			1
SAF11		2			2
SAF12		1			1
SAF13		2			2
SAF14		1			1
SAF15		1			1
SAF16		1			1
SAF17		1			1
SAF18		1			1
SAF19		1			1
SAF20		2			2
SAF21		1			1
SAF23		1			1
SAF26		2			2
SAF28		1			1
SAF29		1			1
SAF30		1			1
SAF32		1			1
SAF33		2			2
SAF34		1			1
SAF35		1			1
SAF38		1			1
SAF39		1			1
SAF40		3			3
SAF41		1			1
SAF43		1			1
SAF44		1			1
SAF45		1			1
SAF47		1			1

Table 9. Continued.

Cytb

HAP	AR	SA	NZ	PE	Total
NZ11			1		1
NZ12			1		1
NZ13			1		1
NZ07			1		1
NZ09			1		1
NZ08			1		1
LoP1				1	1
LP21				14	14
LP22				1	1
LP31				7	7
LP32				1	1
LP41				2	2
LP42				2	2
LP51				39	39
LP52				1	1
LP53				1	1
LP54				2	2
LP61				3	3
LP62				2	2
LoP7				2	2

Dloop

HAP	AR	SA	NZ	PE	Total
SAF48			1		1
SAF49			1		1
SAF52			1		1
SAF54			1		1
SAF55			1		1
SAF58			1		1
SAF61			2		2
SAF63			1		1
SAF64			1		1
SAF65			1		1
SAF66			1		1
SAF67			1		1
SA982			1		1
SA993			1		1
SF997			1		1
nzarg		**1**		**1**	2
HAPH				17	17
HAPA				10	10
HAPM				10	10
HAPD				6	6
HAPI				5	5
HAPL				5	5
HAPN				5	5
HAPQ				5	5
HAPE				4	4
HAPF				4	4
HAAA				3	3
HAPC				3	3
HAPO				3	3
HAPW				3	3
HAPZ				3	3
HABB				2	2
HACC				2	2
HADD				2	2
HAPB				2	2
HAPG				2	2
HAPJ				2	2
HAPK				2	2
HAPP				2	2
HAPR				2	2
HAPS				2	2
HAPT				2	2
HAPV				2	2
HAPX				2	2
HAPY				2	2
0T014				1	1
AB03				1	1
AB05				1	1
AB09				1	1
AB094				1	1
AB11				1	1
AB12				1	1
AB25				1	1
AB63				1	1

Table 9. Continued.

Cytb

HAP	AR	SA	NZ	PE	Total

Dloop

HAP	AR	SA	NZ	PE	Total
AB80			1		1
AB83			1		1
AB92			1		1
JB04			1		1
K011			1		1
K015			1		1
K053			1		1
K064			1		1
K068			1		1
K070			1		1
K076			1		1
K079			1		1
K089			1		1
K100			1		1
K124			1		1
K129			1		1
K160			1		1
K161			1		1
K173			1		1
K188			1		1
K199			1		1
K208			1		1
K210			1		1
K213			1		1
K219			1		1
K223			1		1
K325			1		1
K329			1		1
K400			1		1
K401			1		1
K404			1		1
K405			1		1
K406			1		1
K456			1		1
K457			1		1
K463			1		1
K465			1		1
LOB3			1		1
NZ08			1		1
NZ12			1		1
OT01			1		1
OT016			1		1
OT019			1		1
OT08			1		1
OT09			1		1
OT12			1		1
P019			1		1
PU15			1		1
A26				1	1
178				1	1
179				1	1

Discussion

Paleoceanography and the west-wind drift hypothesis

Dusky dolphin distribution is well-correlated with a band of sea surface temperatures (SST) between 8° and 16° C (Fig. 18A), the range of temperatures tolerated by dusky dolphins in New Zealand between winter and summer months (Würsig et al. 1997). This band of SST is associated with regions of high net primary productivity and the continental upwelling of nutrients (Fig. 18B). The cyclic nature of the movement of layers of the oceans in this region produces a predictable distribution of resources in space and time, providing continual support for large aggregations of dusky dolphins feeding on the deep scattering layer abundant with myctophid fishes and squid (Cipriano 1992; Würsig et al. 1997). It is highly probable that the contemporary and historical distribution of the dusky dolphin in the Southern Hemisphere is directly

Table 10. Frequency and geographic distribution of actin intron haplotypes. AR = Argentina, SA = South Africa, NZ = New Zealand, PE = Peru.

Allele	AR	SA	NZ	PE	Total
l	3	19	10	2	34
a	7	9	13		29
b	3	4	8		15
e	1		2		3
k		1		1	2
h		4	2		6
n		4	7		11
s		3	4		7
v			2	3	5
Total	14	44	48	6	73

related to the changes in the availability and abundance of resources with changes in oceanographic conditions.

The west-wind drift hypothesis (WWD) relates patterns of divergence in the dusky dolphin and the dolphin genus *Cephalorhynchus* to the eastward flowing, virtually uninterrupted current system in the Southern Ocean that carries cool water around the Antarctic. Pichler et al.'s (2001) phylogenetic study of mtDNA lineages suggested that *Cephalorhynchus* speciated along continental coastlines following dispersal with the WWD, beginning in the eastern Atlantic (*C. heavisidii*) and moving eastward to South America (*C. eutropia* and *C. commersoni*) and New Zealand (*C. hectori*). Dolphin affinity for near-shore river effluences, small group size, and high levels of site fidelity (Bräger & Schneider 1998) are thought to have contributed to the diversification of endemic species following the colonization of each region. In contrast, the dusky dolphin is a wide-ranging, meso-pelagic species that forms large social aggregations capable of long-distance movements. Nevertheless, congruent patterns of distribution of the dusky dolphin with *Cephalorhynchus* species suggest that the same dispersal mechanisms, namely the WWD, were responsible for dispersal and colonization events that resulted in the contemporary distribution of dusky dolphin populations (Pichler et al. 2001; Cassens et al. 2003).

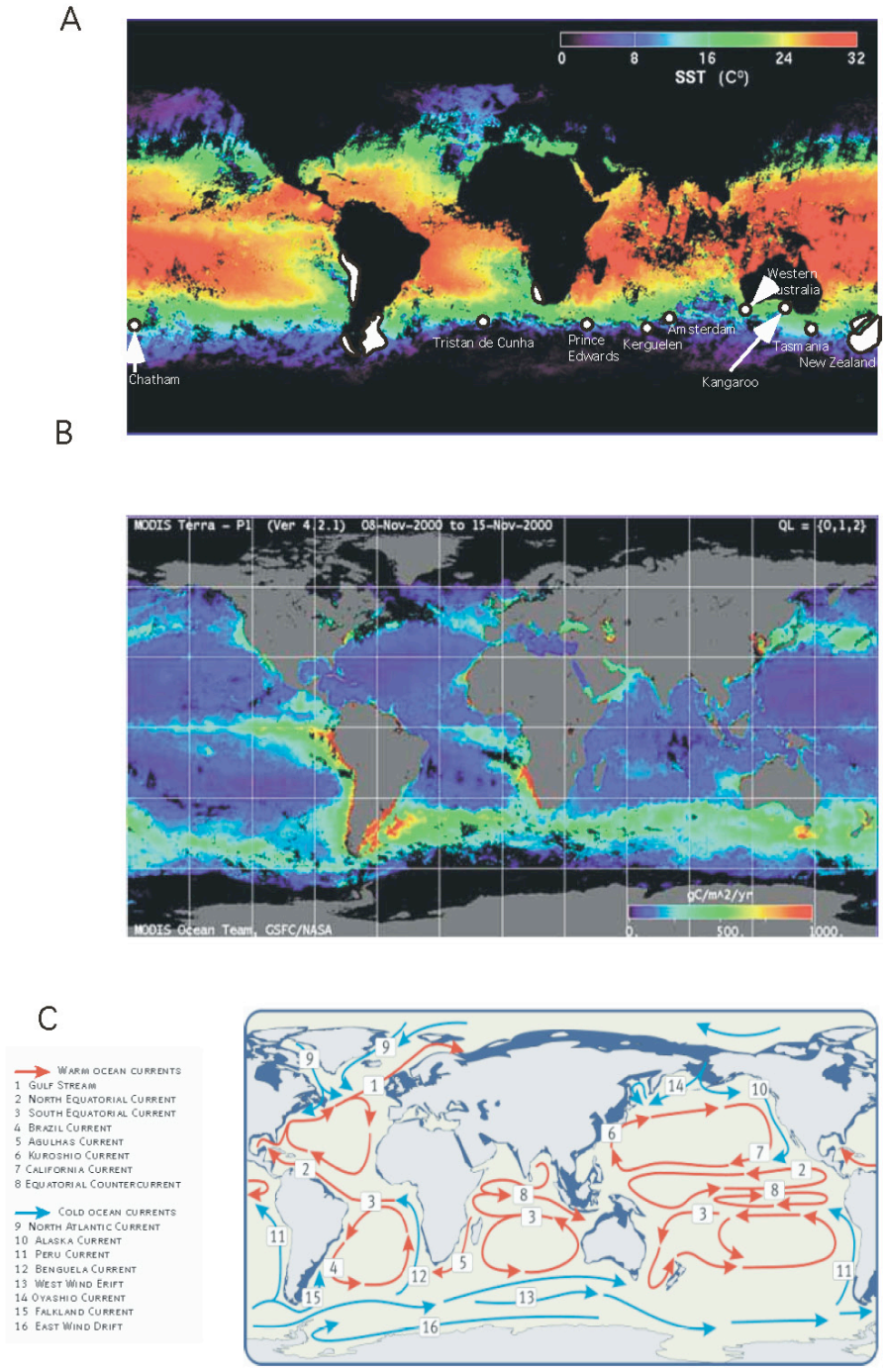


Figure 18. World-wide distribution of dusky dolphins in relation to oceanographic conditions. A. Average sea surface temperatures. B. Net primary productivity measured from the combined effects of available solar radiation, sea surface temperature, and levels of chlorophyll a. C. Direction and temperature of ocean currents.

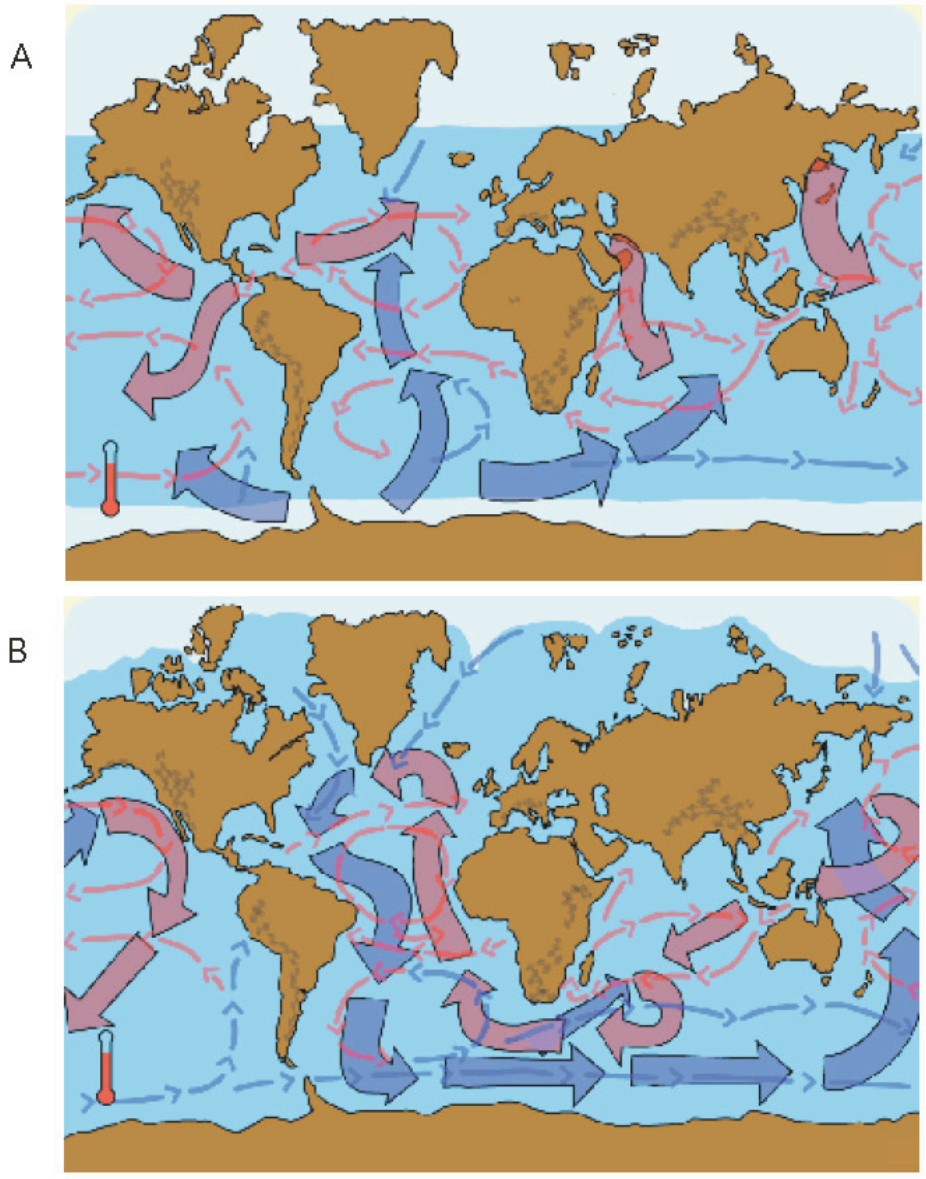


Figure 19. Changes in ocean current patterns and temperatures in the last 2 million years. A) The eastward-flowing west wind drift was interrupted during the Pleistocene as a result of Antarctic ice sheet expansion, virtually closing the connection between Pacific and Atlantic ocean basins. The interruption of current flow in the Southern Ocean spurred the formation of the contemporary Gulf Stream. B) Restoration of the circum-Antarctic current occurred in the Holocene with the re-opening of Drake's passage as glacial ice sheets receded, restoring uninterrupted cold water surface currents in the Southern Ocean.

The WWD is a contemporary oceanographic phenomenon. Oceans of the Southern Hemisphere were quite different during the last glacial maximum (LGM) (CLIMAP 1976; Gersonde et al. 2003; Niebler et al. 2003;Paul and Schäfer-Neth 2003). For example, during the Pleistocene, the formation and extension of ice sheets northward had profound global effects on the distribution of surface sea temperatures (SST) in the range preferred by *L. obscurus*. In general, SST patterns were shifted as much as 5 degrees of north latitude (Gersonde et al. 2003), and the northward shift in Antarctic ice constricted Drake's passage at the LGM, as recently as 25,000 years ago (CLIMAP 1976; Gersonde 2003; Paul and Schäfer-Neth 2003). This altered flow of the Gulf Stream and interrupted the west-wind drift in the Southern Ocean (Fig. 19A, B) (CLIMAP 1976; Gersonde 2003; Paul and Schäfer-Neth 2003). Of significant importance to the phylogeographic history of the dusky dolphin is the isolation of the Pacific and Atlantic Oceans during the constriction of Drake's passage and northward shift of temperate waters.

The isolation of Atlantic and Pacific Ocean basins provides a plausible explanation for the relatively recent, but extensive isolation of Peruvian dusky dolphins from other regions. The simultaneous divergence of Peru from all remaining regions is corroborated with measures of population structure (Table 8). Peru CYTB haplotypes are equally divergent from all lineages (Fst ≈ 0.59; Table 8). The more rapidly-evolving CR region suggests that the most recent link existed between Peru and Argentina (Fst = 0.16; Table 8). Dusky dolphins are currently found around Tierra del Fuego and southern Chile, but there is an unexplainable gap in the distribution of the species north

and south of the Strait of Magellan, which may be a remnant of historical isolation of a more northerly distributed Peruvian population that is maintained by ties to feeding localities.

If the closure of Drake's passage is the source of divergence of Peruvian dusky dolphins, one would predict levels of divergence congruent with the timing of the vicariant event. On average, Peruvian dusky dolphins are separated from other regions by 5 nucleotide substitutions in the CYTB gene (Table 8). It is difficult to calibrate the divergence of intra-specific lineages, due to the stochastic nature of the coalescent process at shallow levels of divergence. However, if I assume that the rate of divergence of CYTB of dusky dolphins is similar to that of other mammals, I would expect around 2.0% divergence per million years. This would date the divergence of the Peruvian dusky dolphins from all other regions around 200,000 years ago, well within the boundaries of the Pleistocene and its series of glaciations (CLIMAP 1976; Gersonde 2003; Paul and Schäfer-Neth 2003). Along a similar vein, New Zealand became isolated from South Africa around 100,000 ybp, from Argentina ~ 50,000 ybp, and the most recent divergence between South Africa and Argentina occurred approximately 12,000 ybp.

Phylogeographic patterns and population history

The genetic data provide evidence for paleoceanographic conditions playing a role in the phylogeography of the dusky dolphin, but the patterns observed do not conform to expectations of the WWD hypothesis. Given the vagile nature of other pelagic dolphin species with large societies, one would not necessarily expect to see

genetic structure of the dusky dolphin associated with passive dispersal mechanisms such as the WWD. The results indicate at least some active dispersal within and among geographic regions with a genetic signature of reciprocal exchange among ocean populations. New Zealand ACT1, CYTB, and CR region haplotypes are shared independently with South Africa and Argentina, indicating that relationships among regions are not geographically linear along an east/west corridor joining ocean basins.

The pattern of divergence suggests some isolation-by-distance, as seen in the lower Fst values between South Africa and Argentina and higher levels of divergence between New Zealand and other regions, but the relationships among Atlantic and Indo-Pacific populations are better described as triangular in pattern, with connections between New Zealand/Argentina, New Zealand/South Africa, and South Africa/Argentina. This triangular pattern may indicate either the presence of a geographically centralized ancestral source population, or populations intermediate to the three regions that have yet to be described. For example, eastern South America, South Africa, and New Zealand are loosely connected by a series of small islands, and the southern coast of Australia (Fig. 14) within the band of temperate ocean waters (Fig. 18A). These islands are used at least periodically by dusky dolphins (Van Warebeek et al. 1995; Gill et al. 2000), and even if frequented by a few individuals, the islands are potential stepping-stones for genetic exchange across the South Pacific and Indian Oceans. Nevertheless, high levels of haplotype diversity within Australia, New Zealand, and South Africa, suggest a very large effective population size of dusky dolphins (Hare et al. 2003), which is supported by the large aggregations of more than 2000 individuals

near Kaikoura, New Zealand (Würsig et al. 1997). Given the abundance of dusky dolphins in other regions, it seems unlikely that large pods off the coast of Australia would remain undocumented.

Parsimony networks and locality of origin

The Bayesian analysis of both mitochondrial sequences placed the root of the dusky dolphin lineage at the base of a branch containing haplotypes from South Africa and New Zealand (Fig. 15). Under the assumption of coalescence theory, the formation of nearly distinct New Zealand lineages in the CYTB and CR networks supports an older age of divergence of this population from other regions (Figs. 15, 16). Furthermore, the results indicate that the geographic isolation of Peruvian maternal lineages is complete (Figs. 15, 16). Between more terminal clades containing Peruvian and New Zealand lineages resides an agglomeration of haplotypes from Argentina and South Africa and multiple missing intermediate haplotypes. This phylogeographic pattern, along with the higher frequency of CYTB haplotypes shared among these two regions, suggests a more recent association between South Africa and South America across the Atlantic Ocean. The presence of multiple missing intermediates, including the central missing intermediate in the CYTB network ("CMI", Fig. 15) suggests that the dusky dolphin's current geographic distribution is the remnant of a much larger population distributed among Indian, Atlantic, and Pacific Ocean basins.

The pattern of parsimonious relationships among ACT1 haplotypes (Fig. 17) also supports a larger, Indian/Atlantic Ocean population as the source of dusky dolphin lineages. For example, the three most common ACT1 haplotypes (a, b, l) are the most

common in New Zealand and South Africa, and also are present in the sister taxon *L. obliquidens* (Fig. 17). Coalescence theory predicts that haplotypes in the greatest frequency and at the center of a haplotype network are older than those in low frequency at the tips (Templeton, Castelloe). Therefore, in addition to the placement of the root, the structure of the network can be used to infer relative age of clades and relationships among geographic regions. The presence of haplotypes from New Zealand at both the center and the periphery of the ACT1 network suggests an older age for the New Zealand population. This also is supported by greater levels of divergence between New Zealand and other regions, with the lowest levels of divergence between South Africa and Argentina (Table 8).

Contemporary gene flow among geographic areas

Levels of divergence measured by Fst and AMOVA do not indicate high levels of contemporary exchange among regions. These values support the findings of Cassens et al. (2003) in that significant population structure among Peru, Argentina and South Africa, with the greatest level of divergence among Peru and other regions (Table 8). Furthermore, the higher degree of isolation between New Zealand and Atlantic populations supports either a lack of migration across the Indian Ocean, or recent dispersal not detectable by the genetic signature of either nDNA or mtDNA loci. A previous study of New Zealand CR variation by Harlin et al. (2003) presented two large maternal lineages, one with short branch lengths and clusters of common haplotypes indicative of a range expansion. The other lineage was dominated by haplotypes separated by several missing intermediates, indicative an older population that has loss

lineages over time. Given the evidence for isolation and the relative age of the New Zealand population derived from mitochondrial and nuclear loci, the presence of the clade with evidence for population expansion represents a period of growth in New Zealand following isolation, and the older clade is the remnant of a larger, more globally distributed population. Nevertheless, despite their potential to disperse, the significant levels of differentiation among regions at different geographic scales indicate that the dusky dolphin is currently isolated in four ocean basins, with the most similar populations being South Africa and Argentina, followed by New Zealand and Peru. Furthermore, as suggested by Cassens et al. (2003), each of these ocean basin populations represents independent management units. In New Zealand, where dusky dolphins are the focus of a thriving swim-with-dolphin tourism industry, and in Peru where incidental mortality is thought to have resulted in a reduction of genetic variation, the findings of this study have direct conservation implications.

CHAPTER V

SUMMARY

Objective 1: Patterns of variation within New Zealand

Although generally limited to coastal waters of South America, South Africa, and New Zealand, dusky dolphins (*Lagenorhynchus obscurus*) show high potential for dispersal over large distances. In New Zealand, photographic identification data indicate a seasonal shift in residency of dolphins between Kaikoura and the Marlborough Sounds as well as changes in group size and behavior. The effect of this seasonal variation on the genetic structure of New Zealand's dusky dolphins was examined by sequencing a 473 base-pair fragment of the mitochondrial DNA control region for 169 individuals from 4 regions along the New Zealand coast. A neighbor-joining phylogeny and an analysis of molecular variance did not support genetic subdivision between regions (ϕ_{sr} = -0.041, P = 0.13). However, nested-clade analysis demonstrated significant evidence for contiguous range expansion and fragmentation along the New Zealand coast. Seasonal movement patterns from Kaikoura to either Otago or the Marlborough Sounds and West Coast are presented as an alternative explanation. New Zealand-wide diversity indices and rate of substitution among sites were used to estimate effective female population size. Lineages-through-time analysis was employed to test hypotheses of population growth. Structure of the neighbor-joining phylogeny, the nested haplotype network, and results of the lineages-through-time analysis suggest that the New Zealand dusky dolphin population underwent at least 1, if not 2, historical population expansions.

Objective 2: Molecular systematics of the genus *Lagenorhynchus*

Assessment of character utility to resolve phylogenies has received extensive debate in terms of *a priori* and *a posteriori* analyses. It is evident that *a posteriori* inspection provides the best heuristic approach when results of a total-evidence analysis are considered. A total-evidence analysis of two mitochondrial and two nuclear DNA regions that vary in overall rates of evolution, was performed to explore patterns of character divergence and the utility of these characters to diagnose relationships among species of dolphins in the genus *Lagenorhynchus*, subfamily Lissodelphininae. Character utility was quantified with several indices (e.g., partition branch support) including new indices introduced in this study. These indices quantify for each data partition at each node a subset of unambiguous character changes that are perfectly consistent throughout the topology (i.e., CI = 1). Together these measures provide evidence of nodal support, which is defined as the degree of corroboration for a hypothesis of monophyly. The simultaneous analysis produced four equally-parsimonious trees that differed only in the resolution of intraspecific relationships, and provided strong evidence for the monophyly of the subfamily Lissodelphininae and the polyphyly of the genus *Lagenorhynchus*. Through *a posteriori* examination of several process partitions, I provide statistical measures of character utility derived from details of character behavior within the total evidence phylogeny. Nodal support was not evenly partitioned among gene regions. Mitochondrial gene regions provided

significantly more support to nodes than nuclear genes throughout the phylogeny irrespective of node height. However the relative proportion of perfectly consistent characters contributed by each data partition was not significantly different among node-height categories. This proportionality in support among nuclear and mitochondrial data partitions regardless of node height was interpreted as evidence for rapid diversification at the base of the subfamily Lissodelphininae where branches were relatively short and weakly corroborated. Overall, mitochondrial genes had a lower proportion of unambiguous character changes that were perfectly consistent, indicating a higher degree of homoplasy than nuclear data partitions. Specifically, the control region exhibited significant conflict and contributed fewer synapomorphies among nodes near the base of the phylogeny. This conflict was attributed to the relatively few total characters that supported these nodes. Overall the total evidence phylogeny was well supported at terminal and basal nodes but weakly supported at the middle nodes. This pattern may be attributed to differential rates of species diversification. These results have taxonomic implications. Revision of *Lagenorhynchus* is needed and would only include *L. acutus*, *L. albirostris*. *Sagmatius* would be resurrected for *L. cruciger, L. australis*. A new generic name is warranted for *L. obscurus, L. obliquidens* only with the addition of character support for this group.

Objective 3: Phylogeography of *L. obscurus*

This study examined the phylogeographic history of the dusky dolphin with nuclear and mitochondrial sequence data. Complete cytochrome *b*, portions of the

control region, and intron 1 of the muscle actin gene were sequenced from dolphins in all four major biogeographic regions of their distribution: western South Africa, New Zealand, Argentina, and Peru. Measures of population structure indicate significant partitioning of haplotypic variation among all regions for all loci. Highest levels of divergence were found between Peruvian dolphins and those from other regions. Bayesian analysis of cytochrome *b* haplotypes, patterns of variation among regions, and statistical parsimony networks suggested that the dusky dolphin originated in the Atlantic or Indian Oceans. The distribution of haplotypic variation indicated that New Zealand contains the remnants of historical diversity in the species, as well as evidence for recent population expansion. Dusky dolphin distribution is correlated with temperate waters in the Southern Hemisphere. The paleoceanography of the Southern Ocean had profound effects on the distribution of temperate ocean currents. This study presents a plausible explanation for the patterns of distribution of dusky dolphin variation in relation to historical oceanography, including the constriction of Drake's passage as a vicariant event that isolated Peru from other regions. Furthermore, gene flow among Atlantic and Indian Oceans was affected by the northward shift of up to 5 degrees of latitude temperate waters, closing corridors for dispersal between these ocean basins multiple times in the Plio-Pleistocene. Current levels of divergence indicate low to non-existent gene flow among contemporary populations.

REFERENCES

Adkins RM, Gelke EL, Rowe D, Honeycutt RL (2001) Molecular phylogeny and divergence time estimates for major rodent groups: evidence from multiple genes. *Molecular Biology and Evolution,* **18**, 777-791.

Adkins RM, Honeycutt RL, Disotell TR (1996) Evolution of eutherian cytochrome c oxidase subunit II: heterogeneous rates of protein evolution and altered interaction with cytochrome *c*. *Molecular Biology and Evolution*, **13**, 1393-1404.

Allard MW, Carpenter JM (1999) On weighting and congruence. *Cladistics*, **12**, 183-198.

Allard MW, Farris JS, Carpenter JM (1999) Congruence among mammalian mitochondrial genes. *Cladistics,* **15**, 75-84.

Allard MW, Honeycutt RL (1991) Ribosomal DNA variation within and between species of rodents, with emphasis on the genus *Onychomys*. *Molecular Biology and Evolution,* **8**, 71-84.

Amos B, Hoelzel AR (1991) Long-term preservation of whale skin for DNA analysis. *Report of the International Whaling Commission (Special Issue),* **13**, 99-103.

Antunes, A, Templeton AR, Guyomard, R, Alexandrino, P (2002) The role of nuclear genes in intraspecific evolutionary inference: genealogy of the transferring gene in the brown trout. *Molecular Biology and Evolution*, **19**, 1272-1287.

Avise, JC (1995) Mitochondrial DNA polymorphism and a connection between genetics, and demography of relevance to conservation. *Conservation Biology*, **9**, 686-690.

Avise JC, Arnold J, Ball RM, Bermingham E, Lamb T, Neigel JE et al. (1987) Intraspecific phylogeography: the mitochondrial bridge between population genetics and systematics. *Annual Review of Ecology and Systematics,* **18**, 489-522.

Avise, JC, Neigel JE, Arnold J (1984) Demographic influences on mitochondrial DNA lineage survivorship in animal populations. *Journal of Molecular Evolution*, **20**, 99-105.

Baker CS, Cipriano F, Palumbi SR (1996) Molecular genetic identification of whale and dolphin products from commercial markets in Korea and Japan. *Molecular Ecology*, **5**, 671-685.

Baker CS, Palumbi SR, Lambertsen RH, Weinrich MT, Calambokidis J et al. (1990) Influence of seasonal migration on geographic distribution of mitochondrial DNA haplotypes in humpback whales. *Nature,* **344**, 238-240.

Baker CS, Perry A, Bannister JL, Weinrich MT, Abernethy RB, 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 RH, DeSalle R (1997) Multiple sources of character information and the phylogeny of Hawaiian drosophilids. *Systematic Biology,* **46**, 654-673.

Baker RH, Yu X, DeSalle R (1998) Assessing the relative contribution of molecular and morphological characters in simultaneous analysis trees. *Molecular Phylogenetics and Evolution,* **9**, 427-436.

Bakke I, Johansen S, Bakke Ø, El-Gewely MR (1996) Lack of population subdivision among the minke whales (*Balaenoptera acutorostrata*) from Icelandic and Norwegian waters based on mitochondrial DNA sequences. *Marine Biology*, **125**, 1-9.

Banks JD, Levine KF, Syvanen M, Theis J, Gilson A (1995) DNA tissue gender typing. *Standard operating procedure of the Wildlife Forensic Laboratory of the California Department of Fish and Game, Sacramento, California*.

Barker FK, Lutzoni FM (2002) The utility of the incongruence length difference test. *Systematic Biology,* **51**, 625-637.

Barnes LG (1985) Evolution, taxonomy and antitropical distributions of the porpoises (Phocoenidae, Mammalia). *Marine Mammal Science,* **1**, 149-165.

Barnes LG, Domming DP, Ray CF (1985) Status of studies on fossil marine mammals. *Marine Mammal Science,* **1**, 15-53.

Bernardi G, Sordino P, Powers DA (1993) Concordant mitochondrial and nuclear DNA phylogenies for populations of the teleost fish *Fundulus heteroclitus*. *Proceedings of the National Academy of Sciences, USA*, **90**, 9271-9274.

Boyd PW, LaRoche J, Gall M, Frew R McKay R. (1999) Phytoplankton dynamics in oceanic waters of south New Zealand. *Journal of Geophysical Research*, **1043**, 13391-13404.

Bräger S, Schneider K (1998) Near-shore distribution and abundance of dolphins along the West Coast of the South Island, New Zealand. *New Zealand Journal of Marine and Freshwater Research*, **32**, 105-112.

Bremer K (1994) Branch support and tree stability. *Cladistics,* **10**, 295-304.

Brown WM, George M, Jr., Wilson AC (1979) Rapid evolution of animal mitochondrial DNA. *Proceedings of the National Academy of Science, USA*, **76**, 1967-1971.

Bruno WJ, Halpern AL (1999) Topological bias and inconsistency of maximum likelihood using wrong models. *Molecular Biology and Evolution,* **16**, 564-566.

Bull JJ, Huelsenbeck JP, Cunningham CW, Swofford DL, Waddell PJ (1993) Partitioning and combining data in phylogenetic analysis. *Systematic Biology,* **42**, 384-397.

Burridge CP (2002) Antitropicality of Pacific fishes: molecular insights. *Environmental Biology of Fishes*, **65**, 151-164.

Carbone I, Kohn LM (2001) A microbial population-species interface: nested cladistic and coalescent inferences with multilocus data. *Molecular Ecology*, **10**, 947-963.

Cassens I, Van Warebeek K, Best PB, Crespo EA, Reyes J (2003) The phylogeography of dusky dolphins (*Lagenorhynchus obscurus*): a critical examination of network methods and rooting procedures. *Molecular Ecology*, **12**,1781-1792.

Castelloe J, Templeton, AR (1994) Root probabilities for intraspecific gene trees under neutral coalescent theory. *Molecular Phylogenetics and Evolution*, **3**, 102-113.

Cipriano F (1992) *Behavior and occurrence patterns, feeding ecology, and life history of dusky dolphins* (Lagenorhyncus obscurus) *off Kaikoura, New Zealand*. Ph.D. Dissertation, University of Arizona, Tuscon, Arizona.

Cipriano F (1997) Antitropical distributions and speciation in dolphins of the genus *Lagenorhynchus*: a preliminary analysis. Pp. 305-316 in *Molecular Genetics of Marine Mammals* (AE Dizon, SJ Chivers, WF Perrin, eds.) International Whaling Commission, Cambridge, Massachusetts.

Cipriano F (2002) Atlantic white-sided dolphin. Pp. 49-51 in *The Encyclopedia of Marine Mammals* (WF Perrin, B Würsig, JGM Thewissen, eds.). Academic Press, New York.

Clement M, Posada D, Crandall KA. (2000) TCS: a computer program to estimate gene genealogies. *Molecular Ecology,* **9**, 1657-1660.

CLIMAP (1976) The surface of ice-age Earth. *Science*, **191**, 1131-1136.

Cognato AI, Vogler AP (2001) Exploring data interaction and nucleotide alignment in a multiple gene analysis of *Ips* (Coleoptera: Scolytinae). *Systematic Biology,* **50**, 758-780.

Cope ED (1866) Third contribution to the history of the Balaenidae and Delphinidae. *Proceedings of the National Academy of Sciences, Philadelphia*, **18**, 293-300.

Crandall KA, Templeton AR (1993) Empirical tests of some predictions from coalescent theory with applications to intraspecific phylogeny reconstruction. *Genetics,* **134**, 959-969.

Crespo EA, Pedraza SN, Coscarella M, Carcia NA, Dans SL, et al. (1997) Distribution and school size of dusky dolphins, *Lagenorhynchus obscurus* (Gray, 1828), in the Southwestern South Atlantic Ocean. *Reports of the International Whaling Commission*, **47**, 693-697.

Cunningham CW (1997) Can three incongruence tests predict when data should be combined? *Molecular Biology and Evolution,* **14**, 733-740.

Damgaard J, Cognato AI (2003) Sources of character conflict in a clade of water striders (Heteroptera: Gerridae). *Cladistics,* **19**, 512-526.

Darlu P, Lecointre G (2002) When does the incongruence length difference test fail? *Molecular Biology and Evolution,* **19**, 432-437.

Davies JL (1963) The antitropical factor in cetacean speciation. *Evolution* **17**, 107-116.

de Quieroz A, Donoghue MJ, Kim J (1995) Separate versus combined analysis of phylogenetic evidence. *Annual Review of Ecology and Systematics*, **26**, 657-681.

DeSalle R, Brower AVZ (1997) Process partitions, congruence and the independence of characters: inferring relationships among closely related Hawaiian *Drosophila* from multiple gene regions. *Systematic Biology,* **46**, 751-764.

Dowling TE, Brown WM (1993) Population structure of the bottlenose dolphin (*Tursiops truncatus*) as determined by restriction endonuclease analysis of mitochondrial DNA. *Marine Mammal Science,* **9**, 138-155.

Dowton M, Austin AD (2002) Increased congruence does not necessarily indicate increased phylogenetic accuracy: the behavior of the incongruence length difference test in mixed-model analyses. *Systematic Biology,* **51**, 19-31.

Duffield DA, Wells RS (1991) The combined application of chromosome, protein and molecular data for the investigation of social unit structure and dynamics in *Tursiops truncatus*. *Report of the International Whaling Commission (Special Issue)*, **13**, 155-169.

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

Eyre-Walker A, Gaut BS (1997) Correlated rates of synonymous site evolution across plant genomes. *Molecular Biology and Evolution,* **14**, 455-460.

Farris JS, Kallersjo M, Kluge AG, Bult C (1995) Testing significance of incongruence. *Cladistics,* **10**, 315-319.

Felsenstein J (1978) Cases in which parsimony or compatibility methods will be positively misleading. *Systematic Zoology*, **27**, 410-410.

Felsenstein J (1981) Evolutionary trees from DNA sequences: a maximum likelihood approach. *Journal of Molecular Evolution*, **17**, 368-376.

Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* **39**, 783-791.

Felsenstein J (1988) Phylogenies from molecular sequences: inference and reliability. *Annual Review of Genetics,* **22**, 521-565.

Flower WH (1883) On the characters and divisions of the family Delphinidae. *Proceedings of the Scientific Meeting of the Zoological Society of London,* 467-513.

Fraser FC (1966) *Whales, Dolphins, and Porpoises* (KS Norris, ed.), University of California Press, Berkeley.

Fraser FC, Purves PE (1960) Hearing in cetaceans. *Bulletin of the British Museum of Natural History, Zoological Series*, *Volume 7.*

Garcia-Martinez J, Barrio E, Raga JA, and Latorre A (1995) Mitochondrial DNA variability of striped dolphins (*Stenella coeruleoalba*) in the Spanish Mediterranean Sea. *Marine Mammal Science,* **11**, 185-199.

Gaskin DE (1968) Distribution of delphinidae (Cetacea) in relation to sea surface temperatures off eastern and southern New Zealand. *New Zealand Journal of Marine and Freshwater Research*, **2**, 527-534.

Gaskin DE (1972) *Whales, dolphins and seals*. Heinemann Educational Books, London, United Kingdom.

Gatesy J, O'Grady P, Baker RH (1999) Corroboration among data sets in simultaneous analysis: hidden support for phylogenetic relationships among higher-level artiodactyls taxa. *Cladistics*, **15**, 271-313.

Gaut BS, Clark LG, Wendel JF, Muse SV (1997) Comparisons of the molecular evolutionary process at *rbcL* and *ndhF* in the grass family Poaceae. *Molecular Biology and Evolution,* **14**, 769-767.

Gersonde R, Abelmann A, Brathauer U, Becquey S, Bianchi C, et al. (2003) Last glacial sea surface temperatures and sea-ice extent in the Southern Ocean (Atlantic-Indian Sector): A multiproxy approach. *Paleoceanography*, **18**, 6.1-6.18.

Giles RE, Blanc H, Cann HM, Wallace DC (1980) Maternal inheritance of human mitochondrial DNA. *Proceedings of the National Academy of Science, USA*, **77**, 6715-6719.

Gill PC, Ross GJB, Dawbin WH, Wapstra H (2000) Confirmed sightings of dusky dolphins (*Lagenorhynchus obscurus*) in southern Australian waters. *Marine Mammal Science*, **16**, 452-459.

Goodall N (2002a) Hourglass dolphin. Pp. 583-585 in *Molecular Genetics of Marine Mammals* (AE Dizon, SJ Chivers, WF Perrin, eds.) International Whaling Commission, Cambridge, Massachusetts.

Goodall N (2002b) Peale's dolphin. Pp. 890-891 in *Molecular Genetics of Marine Mammals* (AE Dizon, SJ Chivers, WF Perrin, eds.) International Whaling Commission, Cambridge, Massachusetts.

Grant T, Kluge AG (2003) Data exploration in phylogenetic inference: scientific, heuristic, or neither. *Cladistics,* **19**, 379-418.

Graves JE (1998) Molecular insights into the population structures of cosmopolitan marine fishes. *Journal of Heredity* **89**, 427-437.

Gray JE (1828) *Specilegia Zoologica*. Treuttel, Wurtz, and Co., London.

Gray JE (1866) *Catalogue of Seals and Whales in the British Museum*. Taylor and Francis, London, 402 pp.

Gray JE (1871) *Supplement to the Catalogue of Seals and Whales in the British Museum*. British Museum of Natural History, London, 103 pp.

Hare MP (2001) Prospects for nuclear gene phylogeography. *Trends in Ecology and Evolution*, **16**, 700-706.

Hare MP, Cipriano F, Palumbi SR (2002) Genetic evidence on the demography of speciation in allopatric dolphin species. *Evolution*, **56**, 804-816.

Harlin, A D (1999) Dusky dolphins of Kaikoura, New Zealand: behavioral effects of genetic sampling and analysis of population structure. M.S. thesis, Texas A&M University, College Station, Texas.

Harlin AD, Markowitz T, Baker CS, Würsig B, Honeycutt RL (2003) Genetic structure, diversity, and historical demography of New Zealand's dusky dolphin (*Lagenorhynchus obscurus*). *Journal of Mammalogy*, **19**, 217-223.

Harlin AD, Würsig B, Baker CS, Markowitz. T (1999) Skin swabbing for genetic analysis: application on dusky dolphins (*Lagenorhynchus obscurus*). *Marine Mammal Science,* **15**, 409-425.

Hodell DA, Kanfoush SL, Shemesh A, Crosta X, Charles CD, et al. (2001) Abrupt cooling of Antarctic surface waters and sea ice expansion in the South Atlantic sector of the Southern Ocean at 5000 cal yr B.P. *Quaternary Research*, **56**, 191-198.

Hoelzel AR (1998) Genetic structure of cetacean populations in sympatry, parapatry, and mixed assemblages, implications for conservation policy. *Journal of Heredity*, **89**, 451-458.

Hoelzel AR., Dalheim M, and Stern SJ (1998a) Low genetic variation among killer whales (*Orcinus orca*) in the eastern North Pacific and genetic differentiation between foraging specialists. *Journal of Heredity*, **89**, 121-128.

Hoelzel AR, Dover GA (1991) Genetic differentiation between sympatric killer whale populations. *Heredity*, **66**, 191-195.

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

Hoelzel AR, Potter CW, Best PB (1998b) Genetic differentiation between parapatric 'nearshore' and 'offshore' populations of the bottlenose dolphin. *Proceedings of the Royal Society of London, Series B*, **265**, 1177-1183.

Hoelzer GA, Wallman J, Melnick DJ (1998) The effects of social structure, geographical structure, and population size on the evolution of mitochondrial DNA: II. Molecular clocks and the lineage sorting period. *Journal of Molecular Evolution*, **47**, 21-31.

Honeycutt RL, Nedball MA, Adkins RM, Janecek LL (1995) Mammalian mitochondrial DNA evolution- a comparison of the cytochrome *b* and cytochrome *c* oxidase II genes. *Journal of Molecular Evolution*, **40**, 260-272.

Huelsenbeck JP, Bull JJ, Cunningham CW (1996) Combining data in phylogenetic analysis. *Trends in Ecology and Evolution*, **11**, 152-157.

Huelsenbeck JP, Ronquist F (2001) MrBayes: Bayesian inference of phylogeny. *Bioinformatics*, **17**, 754-755.

Hughes AL, Nei. M (1989) Nucleotide substitution at major histocompatibility complex class II loci: evidence for overdominant selection. *Proceedings of the National Academy of Sciences, USA,* **86**, 958-962.

Jukes TH, Cantor CR (1969) Evolution of protein molecules. Pp. 21-132 in *Mammalian Protein Metabolism* (HN Munro, ed.). Academic Press, New York.

Kellogg R (1941) On the identity of the porpoise *Sagmatius amblodon*. *Zoological Series of the Field Museum of Natural History*, **27**, 293-311.

Kimura M (1968) Evolutionary rate at the molecular level. *Nature*, **217**, 624-626.

King JL, Jukes TH (1969) Non-Darwinian evolution. *Science*, **164**, 788-798.

Kinze CC (2002) White-beaked dolphin. Pp. 1332-1334 in *The Encyclopedia of Marine Mammals* (WF Perrin, B Würsig, JGM Thewissen, eds.) Academic Press, New York.

Kluge AG (1989) A concern for evidence and a phylogenetic hypothesis of relationships among *Epicrates* (Boidae, Serpentes). *Systematic Zoology*, **38**, 7-25.

Kluge AG, Wolfe AJ (1993) Cladistics: what's in a word? *Cladistics,* **9**, 183-199.

Kuhner MK, Yamato J, Beerli P, Walkup E, Sloan J, et al. (2004) LAMARC version 1.2.2, http://evolution.gs.washington.edu/lamarc.html. University of Washington.

Lea DW, Pak DK, Peterson LC, Hughen KA (2003) Synchroneity of tropical and high-latitude Atlantic temperatures over the last glacial termination. *Science*, **301**, 1361-1364.

Leatherwood S (1991) Dolphins of the genus *Lagenorhynchus* in the tropical South Pacific. *Marine Mammal Science,* **7**, 194-196.

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.

Lee MSY (1999) Measuring support for phylogenies: the "proportional support index". *Cladistics*, **15**, 173-176.

Lee MSY, Hugall AF (2003). Partitioned likelihood support and the evaluation of data set conflict. *Systematic Biology,* **52**, 15-22.

Li, W-H. 1997. *Molecular evolution*. Sinauer Associates, Inc., Sunderland, Massachusetts.

Li W-H, Tanimura W (1987) The molecular clock runs more slowly in man than in apes and monkeys. *Nature*, **326**, 93-96.

Lindberg DR (1991) Marine biotic interchange between the northern and southern hemispheres. *Paleobiology*, **17**, 308-324.

Maddison WP, Maddison DR. (1992) *MacClade: analysis of phylogeny and character evolution*. Sinauer and Associates, New York.

Maniatis T, Fritsch EF, Sambrook J (1982) *Molecular cloning: A laboratory manual*. Cold Spring Harbor, New York.

Markowitz, TM (2004) *Social structure of the dusky dolphin*. Ph.D. Dissertation, Texas A&M University, College Station, Texas.

Martin AP, Palumbi SR (1993) Body size, metabolic rate, generation time, and the molecular clock. *Proceedings of the National Academy of Sciences, USA,* **90**, 4087-4091.

Messier W, Stewart C-B (1997) Episodic adaptive evolution of primate lysozymes. *Nature*, **385**, 151-154.

Mitchell E (1970) Pigmentation pattern evolution in delphinid cetaceans: an essay in adaptive coloration. *Canadian Journal of Zoology*, **48**, 717-740.

Miyamoto MM, Fitch WM (1982) Testing species phylogenies and phylogenetic methods with congruence. *Systematic Biology,* **44**, 64-76.

Miyazaki N, Shikano C (1997) Preliminary study on comparative skull morphology and vertebral formula among the six species of the genus *Lagenorhynchus* (Cetacea: Delphinidae). *Mammalia*, **61**, 573-587.

Moritz C (1994) Applications of mitochondrial DNA analysis in conservation: a critical review. *Molecular Ecology,* **3**, 401-411.

Murphy WJ, Eizirkik E, Johnson WE, Zhang YP, Ryder OA, et al. (2001) Molecular phylogenetics, and the origin of placental mammals. *Nature*, **409**, 614-618.

Naylor GJP, Brown WM. (1998) *Ampioxus* mitochondrial DNA, chordate phylogeny, and the limits of inference based on comparisons of sequences. *Systematic Biology,* **47**, 61-76.

Nee S (1994) Extinction rates can be estimated from molecular phylogenies. *Philosophical Transactions of the Royal Society of London, Series B*, **266**, 695-700.

Nee S, Holmes EC, Rambaut A, Harvey PH (1995) Inferring population history from molecular phylogenies. *Philosophical Transactions of the Royal Society of London, Series B*, **349**, 25-31.

Nei M (1987) *Molecular Evolutionary Genetics*. Columbia University Press, New York.

Nei M (1992) Age of the common ancestor of human mitochondrial DNA. *Molecular Biology and Evolution*, **9**, 1176-1178.

Nei, M, Li W-H (1979) Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proceedings of the National Academy of Science, USA*, **76**, 5269-5273.

Nieber H-S, Arz W, Donner B, Mulitza S, Patzold J, et al. (2003) Sea surface temperatures in the equatorial and South Atlantic Ocean during the Last Glacial Maximum (23-19 ka). *Paleoceanography*, **18**, 14.1-14.12.

Nixon KC, Carpenter JM (1996) On simultaneous analysis. *Cladistics*, **12**, 221-241.

Ohta T (1973) Slightly deleterious mutant substitution in evolution. *Nature*, **246**, 96-98.

Os'kina NS (2001) Changes in the annual temperature values of the World Ocean in the Late Pleistocene. *Oceanology*, **41**, 744-747.

Palsbøll PJ, Bérubé M, Larsen AH, Jorgenson J (1997) Primers for the amplification of tri- and tetramer microsatellite loci in cetaceans. *Molecular Ecology*, **6**, 893-896.

Palumbi SR, Baker CS (1994) Contrasting population structure from nuclear intron sequences and mtDNA of humpback whales. *Molecular Biology and Evolution*, **11**, 426-435.

Palumbi SR, Cipriano F, Hare M (2001) Predicting nuclear gene coalescences from mitochondrial data: the three-times rule. *Evolution*, **55**, 859-868.

Paul A and Schäfer-Neth C (2003) Modeling the water masses of the Atlantic Ocean at the Last Glacial Maximum. *Paleoceanography*, **18**, 3.1-3.26.

Peale TR (1848) United States exploring expedition, during the Years 1838-1842 under the command of Charles Wilkes, U.S.N., Volume 8, *Mammalia and Ornithology*, 338 pp.

Pichler FB, Dawson SM, Slooten E, Baker CS (1998) Geographic isolation of Hector's dolphin populations described by mitochondrial DNA sequences. *Conservation Biology*, **12**, 666-682.

Pichler FB, Robineau D, Goodall RNP, Meyer MA, Olavarria C, et al. (2001) Origin and radiation of southern hemisphere coastal dolphins (genus *Cephalorhynchus*). *Molecular Ecology*, **10**, 2215-2223.

Posada D, Crandall KA (1998) Modeltest: testing the model of DNA substitution. *Bioinformatics*, **14**, 817-818.

Posada D, Crandall KA (2001) Selecting models of nucleotide substitution: an application to Human Immunodeficiency Virus 1 (HIV-1*).* *Molecular Biology and Evolution*, **18**, 897-906.

Posada D, Crandall KA, Templeton AR (2000) GeoDis: a program for the cladistic nested analysis of the geographical distribution of genetic haplotypes. *Molecular Ecology*, **9**, 487-488.

Rambaut A, Drummond A (2004) Tracer version 1.1.1, http://evolve.zoo.ox.ac.uk/software, University of Oxford, United Kingdom.

Rambaut A, Harvey PH, Nee S, Holmes E (1994*)* Endemic-Epidemic, Phylogenetic process analysis, version 10. Department of Zoology, University of Oxford, United Kingdom.

Rannala B, Yang Z (1996) Probability distribution of molecular evolutionary trees: a new method of phylogenetic inference. *Journal of Molecular Evolution*, **43**, 304-311.

Raymond M, Rousset F (1995) An exact test for population differentiation. *Evolution*, **49**, 1280-1283.

Rogers AR (1995) Genetic evidence for a Pleistocene population explosion. *Evolution*, **49**, 608-615.

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

Rooney AP (1998) *Assessment of recent population bottlenecks, a case study of the Bering-Chukchi-Beaufort Seas stock of bowhead whales* (Balaena mysticetus). Ph.D. Dissertation, Texas A&M University, College Station, Texas.

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

Rooney AP, Zhang J (1999) Rapid evolution of a primate sperm protein: relaxation of functional constraint or positive Darwinian selection? *Molecular Biology and Evolution,* **16**, 706-710.

Rosel PE, Dizon AE, Haygood MG (1995) Variability of the mitochondrial control region in populations of the harbour porpoise, *Phocoena phocoena*, on inter-oceanic and regional scales. *Canadian Journal of Fisheries and Aquatic Science,* **52**, 1210-1219.

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

Ross KG (2001) Molecular ecology, of social behaviour: analyses of breeding systems and genetic structure. *Molecular Ecology*, **10**, 265-284.

Rozas J, Sanchez-DelBarrio JC, Messeguer X, Rozas R (2003) DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics*, **19**, 2496-2497.

Saiki, K, Gelfand DH, Stoffel S, Scharf S, Higuchi R, et al. (1988) Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science,* **239**, 487-491.

Saitou, N and M Nei. (1987) The neighbour-joining method, a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution*, **4**, 406-425.

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

Schierup MH, Hein J (2000) Consequences of recombination on traditional phylogenetic analysis. *Genetics*, **156**, 879-881.

Schlötterer C, Amos B, Tautz D (1991) Conservation of polymorphic simple sequence loci in cetacean species. *Nature,* **354**, 63-65.

Schneider S, Roessli D, Excoffier L (2000) Arlequin: a software for population genetic data analysis. Version 2.00. Genetics, and Biometry Lab, Department of Anthropology, University of Geneva.

Shane SH, Wells R, Würsig B (1986) Ecology, behavior, and social organization of the bottlenose dolphin, a review. *Marine Mammal Science*, **2**, 34-63.

Shyue S-K, Hewett-Emmett D, Sperling HG, Hunt DM, et al. (1995) Adaptive evolution of color vision genes in higher primates. *Science*, **269**, 1265-1267.

Simmons MP, Ochoterena H (2001) Gaps as characters in sequence-based phylogenetic analyses. *Systematic Biology,* **49**, 369-381.

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

Sorenson M D (1999) TreeRot, version 2. Boston University, Massachusetts.

Sota T, Vogler VP (2001) Incongruence of mitochondrial and nuclear genes in the carabid beetles *Ohomopterus*. *Systematic Biology,* **50**, 39-59.

Steel M, Penny D (2000) Parsimony, likelihood, and the role of models in phylogenetics. *Molecular Biology and Evolution,* **17**, 839-850.

Stephens M, Donnelly P (2003) A comparison of Bayesian methods for haplotype reconstruction. *American Journal of Human Genetics*, **73**, 1162-1169.

Stephens M, Smith NJ, Donnelly P (2001) A new statistical method for haplotype reconstruction from population data. *American Journal of Human Genetics*, **68**, 978-989.

Stewart C-B, Schilling JW, Wilson AC (1987) Adaptive evolution in the stomach lysozymes of foregut fermenters. *Nature*, **230**, 401-404.

Stillman JH, Reeb CA (2001) Molecular phylogeny of eastern Pacific porcelain crabs, genera *Petrolisthes* and *Pachycheles*, based on the mtDNA 16S rDNA sequence:

phylogeographic and systematic implications. *Molecular Phylogenetics and Evolution*, **19**, 236-245.

Sugg DW, Chesser RK, Dobson FS, Hoogland JL (1996) Population genetics, meets behavioral ecology. *Trends in Ecology and Evolution*, **11**, 338-342.

Sullivan J, Holsinger KE, Simon C (1995) Among-site rate variation and phylogenetic analysis of 12S rRNA in sigmodontine rodents. *Molecular Biology and Evolution*, **12**, 998-1001.

Sullivan J, Swofford DL (2001) Should we use model-based methods for phylogenetic inference when we know that assumptions about among-site rate variation and nucleotide substitution pattern are violated? *Systematic Biology,* **50**, 723-729.

Swofford DL (1999*)* PAUP* Phylogenetic analysis using parsimony (*and other methods). Sinauer and Associates, New York.

Tamura K, Nei M (1993) Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Molecular Biology and Evolution*, **10**, 512-516.

Templeton, AR (1998) Nested clade analyses of phylogeographic data: testing hypotheses about gene flow and population history. *Molecular Ecology*, **7**, 381-397.

Templeton, AR (2002) Out of Africa again and again. *Nature*, **416**, 45-51.

Templeton AR, Boerwinkle E, Sing CF (1987) A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and

DNA sequence data. I. Basic theory and an analysis of alcohol dehydrogenase activity in *Drosophila*. *Genetics,* **117**, 343-351.

Templeton AR, Crandall KA, Sing CF (1992) A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. *Genetics,* **132**, 619-633.

Templeton AR, Routman E, Phillips CA (1995) Separating population structure from population history: a cladistic analysis of the geographical distribution of mitochondrial DNA haplotypes in the Tiger salamander, *Ambystoma tigrinum*. *Genetics,* **140**, 767-782.

Templeton AR, Sing CF (1993) Cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping IV. Nested analyses with cladogram uncertainty and recombination. *Genetics*, **134**, 659-669.

Templeton AR, Weiss KM, Nicherson D.A, Boerwinkle E, Sing CF (2000) Cladistic structure within the human lipoprotein lipase gene and its implications for phenotypic association studies. *Genetics*, **156**, 1259-1275.

Thompson JD, Higgins DG, Gibson TJ (1994) Clustal W. Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties, and weight matrix choice. *Nucleic Acids Research*, **22**, 4673-4680.

True FW (1889) A review of the family Delphinidae. *Bulletin of the United States National Museum (Smithsonian Institution)*, No. 36. Smithsonian Institution, Washington D.C., 191 pp.

Valsecchi E, Amos W (1996) Microsatellite markers for the study of cetacean populations. *Molecular Ecology*, **5**, 151-156.

van Dijk MAM, Paradis E, Catzeflis FM, de Jong WW (1999) The virtues of gaps: xenarthran (Edentate) monophyly supported by a unique deletion in alpha-crystallin. *Systematic Biology*, **48**, 94-106.

Van Warebeek K (1992) Records of dusky dolphins, *Lagenorhynchus obscurus* (Gray, 1828) in the Eastern South Pacific. *Beaufortia*, **43**, 45-61

Van Warebeek, K (1993) Geographic variation and sexual dimorphism in the skull of the dusky dolphin, *Lagenorhynchus obscurus* (Gray, 1828). *Fishery Bulletin*, **91**, 754-774.

Van Warebeek K, Van Bree PJH, Best PB (1995) On the identity of *Prodelphinus petersii* (Lutken, 1889) and records of dusky dolphin *Lagenorhynchus obscurus* (Gray, 1828) from the southern Indian and Atlantic Oceans. *South African Journal of Marine Science*, **16**, 25-35.

Van Warebeek K, Würsig B (2002) Pacific white-sided dolphin and dusky dolphin. Pp. 859-861 in *The Encyclopedia of Marine Mammals* (WF Perrin, B Würsig, JGM Thewissen, eds.) Academic Press, New York.

Vigilant L, Stoneking M, Harpending H, Hawkes K, Wilson AC (1991) African populations and the evolution of mitochondrial DNA. *Science*, **253**, 1503-1507.

Vogler AP, DeSalle R (1994) Diagnosing units of conservation management. *Conservation Biology*, **8**, 354-63.

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

Wheeler W (1999) Measuring topological congruence by extending character techniques. *Cladistics*, **15**, 131-135.

White BN (1986) The isthmian link, antitropicality and American biogeography: distributional history of the atherinopsinae (Pices: Atherinidae). *Systematic Zoology,* **35**,176-194.

Wright S (1931) Evolution in Mendelian populations. *Genetics,* **16**, 97-159.

Wright S (1943) Isolation by distance. *Genetics,* **28**, 114-138.

Würsig B, Bastida R (1986) Long-range movement and individual associations of two dusky dolphins (*Lagenorhynchus obscururs*) off Argentina. *Journal of Mammalogy*, **67**, 773-774.

Würsig B, Cipriano F, Slooten E, Constantine R, Barr K, et al. (1997) Dusky dolphins (*Lagenorhyncus obscurus*) off New Zealand: status of present knowledge. Pp. 715-722 in *Report of the International Whaling Commission*, International Whaling Commission, Cambridge, Massachusetts.

Würsig, B, Cipriano F, Würsig M (1991) Dolphin movement patterns information from radio and theodolite tracking studies. Pp. 79-111 in *Dolphin Societies* (K Pryor and KS Norris, eds.). University of California Press, Berkeley, California.

Würsig B, Jefferson TA (1990) Methods of photo-identification for small cetaceans. *Report of the International Whaling Commission, Special Issue*, **12**, 43-52.

Würsig B, Würsig M, Cipriano F (1989) Dolphins in different worlds. *Oceanus*, **32**, 71-75.

Xia X (1998) The rate heterogeneity of nonsynonymous substitutions in mammalian mitochondrial genes. *Molecular Biology and Evolution*, **15**, 336-344.

Yang Z (1996) Among-site rate variation and its impact on phylogenetic analyses. *Trends in Ecology and Evolution*, **11**, 367-372.

Yang Z (1997) How often to wrong models produce better phylogenies? *Molecular Biology and Evolution,* **14**, 105-108.

Yang Z, Kumar S (1996) Approximate methods for estimating the pattern of nucleotide substitution and the variation of substitution rates among sites. *Molecular Biology and Evolution*, **13**, 650-659.

Zhang, D-X, Hewitt GM (2003) Nuclear DNA analyses of populations: practice, problems and prospects. *Molecular Ecology*, **12**, 563-584.

VITA
April Dawn Harlin

I. Personal Information

Address: 2419 Pintail Loop, College Station, Texas, USA, 77845
Birthdate: October 19, 1970, St. Louis, Missouri.

II. Education

B.S., University of California at Davis, June 1996
M.S., Texas A&M University, May 1999

III. Fellowships

- National Science Foundation, Dissertation Enhancement, 2003-2004
- Tom Slick Graduate Student Fellow, Texas A&M University, 2001-2002
- L.T. Jordan Fellow, Texas A&M University, 1999
- Willie May Harris Graduate Fellow, Texas A&M University, 1996-1997
- President's Undergraduate Research Fellow, University of California, Davis, 1995

IV. Publications

Markowitz, T. M., A. D. Harlin, B. Würsig, and C. McFaddin. **2004**. Dusky dolphin foraging habitat: overlap with aquaculture in New Zealand. *Aquatic Conservation* 14:133-149.

Cognato, A. I., A. D. Harlin, and M. Fisher. **2003**. Genetic structure among pinyon pine bark beetles (*Ips confusus*). *Journal of Environmental Entomology* 32(5): 1262-1270.

Harlin, A. D., C. S. Baker, R. L. Honeycutt, T. M. Markowitz, and B. Würsig. **2003**. Genetic structure, diversity, and historic demography of New Zealand's dusky dolphin. Journal of Mammalogy 84(2): 702-707.

Markowitz, T.M., A. D. Harlin, and B. Würsig. **2003**. Digital photography improves efficiency of individual dolphin identification. Marine Mammal Science. 19(1): 217-223.

Harlin, A. D., B. Würsig, C. S. Baker, and T. M. Markowitz. **1999**. Skin swabbing for genetic analysis: Application on dusky dolphins (Lagenorhynchus obscurus). Marine Mammal Science 15:409-425.

Harlin, A. D. **1999**. *Dusky Dolphins of Kaikoura, New Zealand: Behavioral Responses to Genetic Sampling and Analysis of Population Structure*. M.S. thesis, Texas A&M University, College Station, Texas A&M University.