INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

Bell & Howell Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

Genetic Structure And Molecular Ecology Of The North Atlantic Fin Whale, *Balaenoptera physalus* (Linnaeus, 1758).

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements of the degree of

Doctor of Philosophy

by

Martine Bérubé

Department of Natural Resource Sciences
Macdonald Campus of McGill University
Ste-Anne de Bellevue, Québec

Thesis committee:
Dave Sergeant, Ph. D.
Fred Whoriskey, Ph. D.
Urs Kuhnlein, Ph. D.
Finn Larsen, Senior Scientist

May 1998

National Library of Canada
Acquisitions and Bibliographic Services
395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque nationale du Canada
Acquisitions et services bibliographiques
395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

Our file Notre référence

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-50113-2

Canada

To

my parents, Irène and Jean Berchmans

ABSTRACT

Although biological data concerning the North Atlantic fin whale (*Balaenoptera physalus*) have been collected since the beginning of this century, much is still unknown regarding their biology and evolution. The aim of this study was to increase our knowledge of the evolution, population structure and molecular ecology of this species using modern molecular techniques. Fin whale samples were collected in six feeding areas from the North Atlantic Ocean, the Mediterranean Sea and the Sea of Cortez (North Pacific Ocean). To facilitate the molecular analyses, two new sexing techniques were developed, and several hyper-variable microsatellite loci were isolated from humpback whale (*Megaptera novaeangliae*) genomic DNA. The sex as well as the genotype at six microsatellite loci was determined and the first 288 nucleotides of the mitochondrial (mt) control region sequenced in 407 samples. Population genetic and phylogeographic analyses of mtDNA and nuclear loci supported the hypothesis that North Atlantic and North Pacific fin whales constitute separate populations. However, the degree of divergence did not correlate with the rise of the Panama Isthmus and suggested the occurrence of occasional gene flow between the two oceans. The analysis of the mtDNA indicated the existence of several separate populations in the North Atlantic Ocean and Mediterranean Sea. The mtDNA analysis identified eastern and western fin whale populations, both distinct from the Mediterranean Sea fin whales. The result of the microsatellite loci analysis revealed significant levels of heterogeneity only between the most distant areas. The observed difference in the relative level of divergence at mtDNA and nuclear loci was consistent with expanding populations not yet in "drift-mutation" equilibrium, which have diverged recently. This scenario was supported by the distribution of pairwise differences among the mtDNA nucleotide sequences. The sex ratio in the samples collected from Gulf of St. Lawrence fin whales was unexpectedly biased towards males. Subsequent comparisons with group composition indicated that large groups were mainly comprised of males. Finally, a fin/blue whale hybrid caught off Northwest Spain

was identified and additional molecular analyses were performed to determine the species identity of the parents.

RESUME

Bien que des données biologiques concernant les rorquals communs (*Balaenoptera physalus*) dans l'Atlantique Nord aient été amassées depuis le début de ce siècle, il existe beaucoup d'incertitudes concernant leur biologie et évolution. Le but de cette étude était d'augmenter notre connaissance sur l'évolution, la structure de population et l'écologie moléculaire de cette espèce en utilisant des techniques moléculaires modernes. Les échantillons de rorqual commun proviennent de six aires d'alimentation de l'Atlantique Nord, de la Méditerranée et de la Mer de Cortez (Océan Pacifique du Nord). Afin de faciliter les analyses moléculaires, deux nouvelles techniques de détermination du sexe ont été developpées et plusieurs microsatellites hyper-variables ont été isolés de l'ADN de rorqual à bosse (*Megaptera novaeangliae*). Le sexe ainsi que le génotype de six microsatellites a été determinés et les 288 premiers nucléotides de la région contrôle de l'ADN mitochondrial (mt) ont été séquencés dans 407 échantillons. Les analyses génétiques des populations et phylogéographiques de l'ADNmt et nucléaire supportent l'hypothèse que les rorquals communs de l'Atlantique et du Pacifique constituent des populations séparées. Cependant, le degré de divergence ne correspond pas avec l'élévation de l'Isthmus du Panama et suggère la présence de flux génique occasionnel entre les deux océans. L'analyse de l'ADNmt indique l'existence de plusieurs populations dans les régions Atlantique Nord et Méditerranée. L'analyse de ADNmt identifie une population orientale et occidentale de rorqual commun, les deux distinctes des rorquals communs de la Méditerranée. L'analyse des microsatellites détecte des niveaux significatifs d'hétérogénéité seulement parmi les localités les plus éloignées. La différence observée dans les niveaux relatifs de divergence des loci mt et nucléaire est conforme à ce qui est prévu dans les populations en expansion, pas encore dans un équilibre "dérive génique-mutation", lesquelles ont récemment divergées. Ce sénario est supporté par la distribution des différences entre les nucléotides des

séquences mt. Le sex-ratio dans les échantillons de rorquals communs du Golfe du St-Laurent était inopinément biaisé en faveur des males. La comparaison ultérieure de la compositon des groupes indique que les grands groupes comportent principalement des males. Finallement, un hybride entre un rorqual bleu et un rorqual commun capturé au large de la côte N. O. de l'Espagne est détecté et des analyses moléculaires supplémentaires sont entreprises pour déterminer l'identité d'espèce des parents.

CONTRIBUTION TO THE KNOWLEDGE OF CETACEAN BIOLOGY

The driving force behind most studies of fin whales has been the fact that they were subject to human exploitation, which necessitated the identification of sub-populations in order to define "stocks" for management and conservation purposes. However, it is still important to increase our insight into the biology and evolution of fin whales as they constitute a component of the North Atlantic ecosystem and are thus susceptible to the effects of the steadily increasing influence of human activities on the marine environment.

This study represents the first attempt to investigate the population structure of fin whales at an ocean-wide level using current molecular techniques. The analyses were performed on data collected from mitochondrial (mt) as well as nuclear DNA loci to assess population structure and gene flow within and between fin whales in the North Atlantic and North Pacific oceans. In the near absence of data, several hypotheses for the population structure have previously been proposed ranging from separate populations with or without overlapping ranges, a patchy continuum, or separate summer feeding "populations" which share a common breeding range. This study provides good evidence for a cetacean population structure model of several populations that are separate on the breeding as well as the feeding range contrary to what have been observed in other North Atlantic baleen whales, such as the humpback whale, *Megaptera novaeangliae*, and right whale, *Eubalaena glacialis*, where several separate feeding grounds share a common breeding range.

The thesis also includes technical breakthroughs, e.g., two new methods to determine the sex of cetaceans. These new methodologies are simple, fast and readily applicable to other mammalian orders with only minor additional work. Furthermore, some 350 microsatellite loci have been isolated from genomic humpback whale DNA. Some of these loci differ from previously isolated cetacean microsatellite loci by consisting of tri- and tetramer repeats which are rare in the eukaryote genome. The analysis of tri- and tetramer microsatellite loci usually generate clear and unambiguous data as opposed to most dimer

microsatellite loci, which are prone to "stutter bands". These highly variable loci represent a significant contribution to the existing cetacean microsatellite database and will facilitate future fine scale studies of kinship and life history.

The thesis has demonstrated that highly valuable biological information can be extracted from molecular population genetic data and has added significantly to our basic knowledge of the evolution and biology of the fin whale.

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

TABLES

FIGURES

ACKNOWLEDGMENTS

This thesis would not have been possible without the contributions from and the collaborations with a number of different people. First, I would like to thank my thesis committee; Drs. Dave Sergeant, Fred Whoriskey, Urs Kuhnlein, Finn Larsen and Rodger Titman. Thanks are also due to Dr. Mark Curtis who replaced Dr. Whoriskey during my comprehensive exam and for his time while visiting colleagues at the Zoological Museum at Copenhagen University. Thanks to the DNA Laboratory of the Department of Population Biology, University of Copenhagen, especially Professor Bent Christensen and Dr. Per J. Palsbøll. Dr. Per J. Palsbøll was the principal investigator of the "Cetacean Genetics Group", at the time located in the Department of Population Biology. Although not my official supervisor, in reality Dr. Palsbøll has served such a function concerning the molecular and population genetic aspects of the study. I also would like to thank my husband who gave me love, support, and trust, which helped me to go through the whole project. I do not want to forget Marie Kubecki, secretary of the Natural Resource Sciences Department, who has been the link between members of my committee, the McGill University administration and myself. Thank you to Dr. Greg Donovan who reviewed parts of the thesis and Dr. Phil J. Clapham who never said no to look at a manuscript and always believed that I could make it. Finally, Dr. Harry Smith, Editor of Molecular Ecology, and Dr. William F. Perrin, Editor of Marine Mammal Science, who gave the permission to include copyrighted material (Chapter 1, 2, 3 and 5) as part of my thesis.

During the thesis project, I participated in the collection of skin biopsy samples and behavioral data from fin whales in the Gulf of St. Lawrence and the Davis Strait while working for the Mingan Island Cetacean Study and the Greenland Nature Research Institute, respectively. However, a study such as "Evolution, Genetic Structure And Molecular Ecology Of The North Atlantic Fin Whale, *Balaenoptera physalus* (Linnaeus, 1758)" relies heavily upon collaboration. Obtaining skin biopsy samples from free-ranging fin whales requires an extensive knowledge of the animals and their behavior. Such expertise is only acquired by

spending numerous hours at sea in the company of the whales. Hence, the contribution from the people in the field is indispensable for a project such as this. Specifically these people are: Finn Larsen; Giuseppe Notarbartolo di Sciara; Richard Sears; Alex Aguilar; Jóhann Sigurjónsson; Jorge Urban-Ramirez, and Dan Dendanto, who are responsible for and/or have coordinated the sample collections. To ensure clear guidelines for the collaborations that constitute part of the work for this doctoral project a letter of agreement was formulated between my collaborators and myself (Bérubé 1994).

Letter of agreement:

This letter represents a formal request for samples for use in the following project:

EVOLUTION, GENETIC STRUCTURE AND MOLECULAR ECOLOGY OF THE NORTH ATLANTIC FIN WHALE, BALAENOPTERA PHYSALUS.

The aims of this study are detailed in the attached proposal.

For all the received samples DNA will be extracted. Excess DNA, and sample, will be returned upon the completion of the analyses.

Determination of the gender, analysis of the mitochondrial D-Loop and single copy nuclear DNA sequences (microsatellites) will be conducted for all samples.

The data obtained by the analyses will be made available for each of the contributing institutions after publication. The results of the gender determination will be forwarded as soon as they are available. The donating institution can use the gender for their purposes, as long as I am acknowledged by the donating institution in an appropriate and mutually agreed way.

Regarding publications; one person to each donating institution will be included as a co-author on a publication on the genetic structure of the North Atlantic fin whale population. In addition, a behavioral study focusing on some areas will be conducted based up on field data and

microsatellite analysis and, one person from each donating organization will be included on that paper. Drafts will be circulated for review to all co-authors prior to submission to a journal.

This study constitutes my Ph.D. project, from which the expected deadline is Spring-Summer 96. However, I will do my best to keep you updated regarding the progress of the work.

It is understood that during the analyses, the contributors will not provide fin whale samples to other molecular DNA laboratories for use in ways that potentially duplicate the intended research.

After completion of the publications, the donating institutions will be free to use the molecular data; in this case, I would expect either co-authorship or acknowledgment, as seems fair.

Finally, I will provide the institutions with tubes and conservation buffer. If there is a need for biopsy arrows, Finn Larsen, from Greenland Fisheries Research Institute, can be contacted at fax number: (45) 35 82 18 50

The laboratory analyses involved in this thesis have been conducted by myself (apart from the exceptions mentioned below) under the supervision of Dr. Per J. Palsbøll and the day-to-day advice from his two technicians; Christina Færch-Jensen and Hanne Jørgensen. In chapter one, the sexing of all 3570 biopsy samples was only possible, thanks to the combined effort of all members of the "Cetacean Genetics Group". The cloning of di-, tri-, and tetramer microsatellite loci from humpback whale genomic DNA (Chapter 2) was mainly conducted by Hanne Jørgensen and Per Palsbøll. My involvement was primarily at assisting in the isolation, sequencing of positive clones, and the optimization of the oligonucleotide primers' sets for several species of baleen whales. The cloning experience acquired during this experiment allowed me to become autonomous in subsequent work, which involved cloning of PCR fragments (data not presented).

Co-authorship on publications was granted based upon the participation in at least two of the following parts of a publication: ensuring funding for fieldwork,

collecting samples and behavioral information, assuring the laboratory funding, and participation in the laboratory analysis. Besides conceiving the project itself, the first author has also conducted the data analysis, most of the writing of the final paper and contributed to other aspects of the publication as the co-authors. Listed below are the acknowledgments for each chapter:

Chapter 1

Authors: Martine Bérubé and Per J. Palsbøll.

Reference: Bérubé & Palsbøll (1996a) Molecular Ecology **5**: 283-287 and Bérubé & Palsbøll (1996b) Molecular Ecology **5**: 602. (Erratum)

Contribution: Although M. Bérubé was the lead author, both authors of this paper have performed all aspects of the work quite evenly.

Acknowledgments: Furthermore, the authors would like to acknowledge the following people for supplying or assisting in getting the necessary samples: the participants of the international collaboration project YoNAH (Years of the North Atlantic Humpback whale), A. Aguilar (Department of Animal Biology, University of Barcelona), G. Notarbarolo di Sciara (Tethys Research Institute), R. Sears (Mingan Island Cetacean Study), D. Dendanto (College of the Atlantic), C. Christian Kinze (Zoological Museum, University of Copenhagen). M. Milinkovitch and M. Roy are thanked for their comments on the manuscript. We would also like to thank the following people for technical assistance; T. H. Andersen, P. Clapham, K. Blanksø Pedersen, L. Enriquez, H. Jørgensen, T. P. Feddersen, A. H. Larsen, P. Raahauge and R. Sponer.

Chapter 2

Authors: Per J. Palsbøll, Martine Bérubé, Anja Holm Larsen and Hanne Jørgensen.

Reference: Palsbøll *et al.* (1997) Molecular Ecology **6**: 893-895.

Contribution: P. J. Palsbøll was the primary author on this paper. He also ensured funding for the laboratory analysis, supervised very closely the cloning as well as

designed the oligonucleotide primers from the flanking region of the tri and tetramer microsatellite loci. M. Bérubé assisted H. Jørgensen during the cloning and in addition sequenced positive clones, optimized oligonucleotide primer sets for several species as well as analyzed a subset of microsatellite loci in more than 700 samples from several different baleen whale species. H. Jørgensen was the person performing the isolation microsatellite loci, sequenced clones, optimized oligonucleotide primers and analyzed six loci in more than 2,000 samples. A. Holm Larsen participated in the optimization of oligonucleotide primer set and performed the analyses of 6 microsatellite loci in 400 humpback whale samples.
Acknowledgments: In addition, Drs. M. Fredholm and K. Winterø are thanked for their encouragement and advice, which originally got the work started, Dr. P. Clapham for his editorial skills in correcting and improving the English, and finally, T. Feddersen and R. Sponer for technical assistance. Samples for this project were collected by Center for Coastal Studies, Allied Whale, Greenland Institute of Natural Resources, the Iceland Marine Research Institute, Mingan Island Cetacean Study, Inc., the Norwegian Marine Research Institute, the Whale Group at Memorial University, and Universidad de Baja California-Sur.

Chapter 3

Authors: Martine Bérubé, Alex Aguilar, Dan Dendanto, Finn Larsen, Giuseppe Notarbartolo di Sciara, Richard Sears, Jóhann Sigurjónsson, Jorge Urban-R., and Per J. Palsbøll.

Reference: Bérubé *et al.* Molecular Ecology (1998) 7(5): 585-599.

Contribution: F. Larsen, G. Notarbartolo di Sciara, R. Sears, A. Aguilar, J. Sigurjónsson, J. Urban-R., and D. Dendanto ensured the collection of biopsy sampling and the administration associated with it. Also, they have contributed comments to the final manuscript. M. Bérubé participated in the part of the fieldwork, ensured funding for the laboratory analyses, undertook the laboratory and data analyses as well as wrote the final paper. P. J. Palsbøll participated in and supervised most aspects of the work.

Acknowledgments: M. Bérubé would like to thank Dr. D. E. Sergeant (Ph.D. supervisor) for his encouragement and help during this work. The authors are indebted to: C. Færch-Jensen, H. Jørgensen and C. Berchok for technical assistance; P. J. Clapham, R. R. Hudson, and H. R. Siegismund for comments on the manuscript; and A. Arnason for facilitating samples and B. Christensen for support. The US Geographic Survey, University of Michigan (http://www.indo.com/distance/), made the Geod program available.

Chapter 4

Authors: Martine Bérubé, Richard Sears and Catherine Berchok.

Reference: Manuscript.

Contribution: M. Bérubé collected samples, performed the molecular as well as data analyses, and wrote the final manuscript. R. Sears was in charge of the sample collection and provided comments on the final manuscript. C. Berchok gathered the behavioral information and commented on the final manuscript.

Acknowledgments: The authors would like to thank the research team of the Mingan Island Cetacean Study Inc. and the Groupe de Recherche et d'Education du Milieu Marin for their help with the sample collection. Thanks to Dr. D.E. Sergeant, Dr. P.J. Clapham, D.K. Mattila, and F. Larsen for their help and good advice. The manuscript was much improved by comments from Dr. P.J. Palsbøll.

Chapter 5

Authors: Martine Bérubé and Alex Aguilar.

Reference: Bérubé & Aguilar (1998) Marine Mammal Science **14** (1): 82-98.

Contribution: M. Bérubé performed the laboratory and data analysis, and wrote the manuscript except for the sections regarding the morphological and biological information in "Material and Methods" and "Results" that were written by A. Aguilar. Furthermore, Dr. Aguilar added valuable comments to the manuscript.

Acknowledgments: E. Grau and M. Olmos took the morphometric measurements and participated in the fieldwork. The following scientists contributed answers to

the questionnaire, often with useful comments on the subject: J. Bannister, P. Best, S. G. Brown, M. Cawthorn, I. Christensen, R. Gambell, E. Grau, T. Haug, C. Lockyer, T. Martin, J. G. Mead, S. Ohsumi, D. W. Rice, K. Sekiguchi, and S. Wada. We thank Mingan Island Cetacean Study Inc. for the blue whale sample and C. Lockyer (Sea Mammal Research Unit) for her assistance in the age determination. P. Palsbøll and T. Pastor reviewed the draft manuscript. We also would like to thank Drs. W. Perrin, P. Clapham and A. Dizon who reviewed the final manuscript.

Financial support to the thesis project was provided by:

This research was funded by the International Whaling Commission, England; the Commission For Scientific Research In Greenland; the Danish Research Council; the Danish Research Academy, the EU Biotechnology Program; WWF-DK; the Greenland Home Rule, and Aage V. Jensen Charity Foundation, Denmark.

PREFACE

My involvement in the field of marine mammal science began in 1985 when I joined the Mingan Island Cetacean Study Inc. (1985-1992). There, I obtained my initial experience in fieldwork and subsequently trained new marine mammal biologists. My communication skills were improved by giving oral presentations and undertaking administrative work as well as writing research reports. I subsequently expanded my fieldwork experience while working as a consultant for organizations such as the Center for Coastal Studies (USA) and the Greenland Nature Research Institute (Denmark).

In the autumn of 1992, I initiated a Master of Science project at the Department of Renewable Resources, Macdonald Campus of McGill University under the supervision of Dr. Dave Sergeant (main supervisor) and Dr. Fred Whoriskey (co-supervisor). My project focused on the population structure of Gulf of St. Lawrence and off West Greenland fin whales using individual photo-identification and molecular genetic techniques. At the time, one of the major groups focusing on molecular studies of marine mammals was established at the Department of Population Biology, University of Copenhagen, under the direction of Dr. Per J. Palsbøll. As McGill and Copenhagen Universities had an established student-exchange program, it was a relatively simple process to become part of the "Cetacean Genetics Group" in Copenhagen and pursue the molecular analyses of the North Atlantic fin whale samples there. As new collaborations were established and more samples became available from other areas, an ocean-wide study became possible. With the support of my supervisors and both McGill and Copenhagen University, the analyses was extended to a doctoral project. In addition to the more academic aspects of my doctoral, the work has also taught me how to structure a project and how to acquire the funds necessary to carry out the work. I have also learned to appreciate the importance of establishing and maintaining good collaborations.

During the last five years, aspects of my thesis work have been presented at several international meetings including: Tenth Biennial Conference on the

Biology of Marine Mammals, Galveston, Texas (November 1993); Genetic Symposium on Marine Mammals, La Jolla, California (September 1994); Eleventh Biennial Conference on the Biology of Marine Mammals, Orlando, Florida (December 1995); Ninth Annual Conference of the European Cetacean Society, Lugano, Switzerland (February 1995); First Population Biology Conference, Copenhagen, Denmark (January 1995); International Whaling Commission Scientific Committee, in Dublin (May 1995) and Cambridge (May, 1996), and finally, World Marine Mammal Science Conference, Monaco (January 1998).

During the thesis project, I have been a guest student at the University of California in Irvine (UCI) at the Department of Ecology and Evolutionary Biology (1996-1997) besides the University of Copenhagen. I am presently located at the Free University of Brussels (1998) at the Department of Molecular Biology where my spouse (Dr. Per Palsbøll) has received funding to carry out research. The laboratory of Professor Richard R. Hudson at UCI is concerned with theoretical aspects of population genetic analyses. Needless to add that the two years spent at Professor Hudson's laboratory has proven highly valuable for my education and served as a very good introduction to theoretical population genetics.

Organization of the thesis

I have chosen to present my doctoral dissertation as a collection of papers, therefore the format observes the following *Faculty regulations*:

> Candidates have the option of including, as part of the thesis, the text of one or more papers submitted or to be submitted for publication, or the clearly duplicated text of one or more published papers. These texts must be bound as an integral part of the thesis.
>
> If this option is chosen, connecting texts that provide logical bridges between the different papers are mandatory. The thesis must be written in

such a way that it is more than a mere collection of manuscripts; in other words, results of a series of papers must be integrated.

The thesis must still conform to all other requirements of the "Guidelines for Thesis Preparation". The thesis must include: A Table of Contents, an abstract in English and French, an introduction which clearly states the rationale and objectives of the study, a comprehensive review of the literature, a final conclusion and summary, and a thorough bibliography or reference list.

Additional material must be provided where appropriate (e.g., in appendices) and in sufficient detail to allow a clear and precise judgment to be made of the importance and originality of the research reported in the thesis.

In the case of manuscripts co-authored by the candidates and others, the candidate is required to make an explicit statement in the thesis as to who contributed to such work and to what extent. Supervisors must attest to the accuracy of such statements at the doctoral oral defense. Since the task of the examiners is made more difficult in these cases, it is in the candidate's interest to make perfectly clear the responsibilities of all the authors of the co-authored papers.

Finally, the thesis has been divided in three parts that reflect the aspects of molecular ecology in which this project has contributed; 1) the tools, 2) genetic variation within and between the fin whale populations and 3) hybridization.

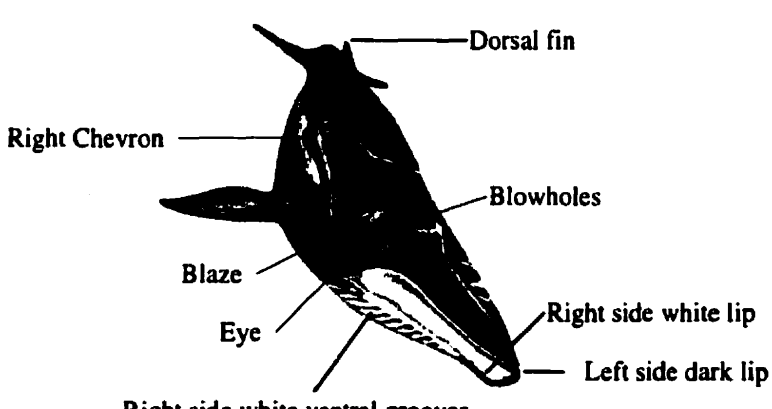
INTRODUCTION

The ability to recognize individual animals, using natural or "artificial" markings, has considerable advantages when studying the biology, estimating biological parameters and determining the population dynamics of animals (Stonehouse 1978). Appropriate methods for marking animals are not always easy, especially for species that are, for instance, aggressive, evasive, live in dense cover, or spend part of their time under ground or submerged. In particular, the study of cetaceans presents special difficulties, as they are often evasive, highly maneuverable, and spend most of their time under the water surface. In order to study the distribution and migration of cetaceans, different methods to identify animals have been developed. One of the first techniques employed to mark free-ranging cetaceans was the "Discovery marks" (Brown 1978), named after the British Antarctic "Discovery" Expeditions during the 1930s. Discovery marks are numbered stainless steel projectiles fired from shotguns into the whale, where they are embedded in and retained by the body musculature. The marking technique was efficient, but retrieval was only possible from dead animals, e.g., during whaling operations. The recovery rate was relatively low although novel and useful information was obtained in this manner (e.g., see Gaskin 1982).

In the early 1970's, photographic techniques to identify individual whales using their natural markings were successfully employed in several cetacean species, especially in species that spend most or part of the year in coastal waters and thus were easily accessible. Extensive catalogues of photo-identified individuals have been compiled for a number of species including; humpback whales, *Megaptera novaeangliae*, (Katona & Whitehead 1981), blue whales, *Balaenoptera musculus*, (Sears *et al.* 1987), killer whales, *Orcinus orca*, (Bigg *et al.* 1987), and fin whales, *B. physalus*, (Agler *et al.* 1990; Seipt *et al.* 1989). Photo-identification of individuals, along with long-term field observations, has generated valuable additional insight regarding migration patterns, biological parameters and behavior in cetaceans (Hammond *et al.* 1990).

Fin whales are identified photographically from individual variations in the distinct asymmetrical head coloration that is characteristic of this species. Specifically, the right side of the head, (as well as the lip and baleen plates) is white or pale gray, while the left side is dark. On the right side, the light coloration sweeps back to form a clear pattern called the blaze. In addition, behind the blowholes, arching down both sides of the back, there is a light region in the form of a "V" shape, known as the chevron (Figure i.1). The pigmentation pattern is used in conjunction with the shape of the dorsal fin and other marks along the flanks (such as scars) to identify individual fin whales. However, in some cases, the variation in the coloration pattern is limited and dependent on high-quality pictures (right angle, light, etc.). The reliability of repeated identifications of individual fin whales from photographs was tested among a group of cetacean researchers (Agler 1992). The results of the study showed that the quality of the picture as well as the experience of the person comparing the pictures clearly influenced the results (Agler 1992). In fact, given the inevitable variation in photographic quality, it has been shown that no methods that identify cetaceans using natural markings were completely unambiguous (e.g., humpback whales see Carlson *et al.* 1990). False negative or positive matches generated in this manner introduce various kinds of bias in the subsequent data analyses (e.g., see Hammond *et al.* 1990).

Figure i.1 Right side of the asymmetrical pigmentation pattern of the fin whale.



Note. Illustration by Hariett Corbett.

A recent study demonstrating the feasibility of genetic "tags" was based upon microsatellite analysis of 3,060 North Atlantic humpback whale skin samples (Palsbøll *et al.* 1997a). Genetic tagging is a generally applicable and unambiguous approach to identify individual animals (and thus whales). While the use of genetic markers to identify individual whales is a relatively novel approach (Amos & Hoelzel 1990), the application of genetic markers to investigate population or species differentiation began in the late 1960s with the introduction of enzyme electrophoresis (Hubby & Lewontin 1966). The proteins isolated from tissue or blood samples are electrophoresed through an acrylamide or starch gel matrix. The gel is subsequently stained for a specific enzyme with the substrate, which is coupled to a color reaction. Isoforms of a specific enzyme are distinguished by differences in electrophoretic mobility caused by an overall change in electric charge due to nucleotide substitutions that alter the amino acid sequence of the protein. This type of molecular analysis has been successfully applied to study several species of baleen whales e.g., fin whales, sei whales, minke whales and Bryde's whales (Danielsdottir *et al.* 1991; Wada & Numachi 1991). The data generated from these analyses made it possible to obtain information about genetic variation, population structure and levels of gene flow from a large number of natural populations (Lewontin & Hubby 1966). However, for large mammals the overall level of differentiation in allozyme studies among conspecifics was generally quite low.

In the 1980s, it became possible to detect changes in the nucleotide sequences at the level of the DNA itself. Such analyses normally detect a higher level of variation than enzyme electrophoresis since not only are non-synonymous nucleotide substitutions (nucleotide substitutions in coding DNA sequences that alter the amino acid sequence) detected but also synonymous nucleotide substitutions (e.g., nucleotide substitutions that do not alter the amino acid sequence). Contrary to enzyme electrophoresis, non-coding regions (DNA sequences that are not transcribed into mRNA that subsequently are translated into

the amino acid sequences that constitute proteins) can also be analyzed, which commonly evolve at a higher rate than coding regions. Analysis of DNA can be performed using either indirect methods which usually employ endonucleases (referred to as restriction fragment length polymorphisms, or RFLP) or directly by DNA sequencing. The analysis of mt DNA by RFLP (Avise *et al.* 1979) has been successfully employed to study fin whales in the central and eastern North Atlantic (Danielsdottir *et al.* 1992). Multi-locus DNA fingerprinting (which also is based upon RFLPs) (Jeffreys *et al.* 1985) allows simultaneous detection of allelic variation at a large number of hyper-variable nuclear loci, known as minisatellites. Minisatellites are DNA sequences, which contain tandemly arranged repeats of short DNA sequences (usually each from 10 to 60 nucleotides in length) and the variability arises from differences in the number of repeats at each locus. The repeats in different classes of minisatellites share a common core sequence, such that a radio-labeled probe containing this sequence can detect many minisatellite loci of the same class in a single experiment. DNA multi-locus fingerprint patterns differ between individuals (hence their name) and can thus be used for individual identification. However, the inability to identify which alleles belong to a single specific locus is undesirable, making it impossible to estimate parameters such as heterozygosity and detect deviations from Hardy-Weinberg (H-W) genotypic proportions. Another drawback of multi-locus fingerprinting is the fact that the banding patterns that constitute the raw data are influenced to some degree by the conditions during the electrophoresis. Hence, affirmative identifications of multiple samples from the same individual or identification of parent-offspring pairs requires that samples are run next to each other on the same gel. For large-scale analyses, such a requirement renders the technique highly impractical, despite attractive features in other regards. These difficulties were overcome partly with the introduction of single-locus minisatellite analysis, which enabled the scoring of alleles by their band sizes and consequently facilitated comparison across gels.

The development of the Polymerase Chain Reaction (PCR) technique in the late 1980s (Mullis & Faloona 1987; Saiki *et al.* 1988) introduced a new tool with a high potential for the entire field of molecular biology, including novel possibilities within the field of population genetics. The PCR technique enabled *in vitro* generation of millions of copies of a single DNA sequence in the course of a few hours. As opposed to RFLP-based techniques the PCR technique requires, in principle, only a few target molecules and thus minute amounts of DNA (e.g., extracted from degraded samples). The procedure involves repeated cycles of denaturation of the target DNA and the subsequent copies, followed by annealing of two oligonucleotide primers to the opposite strands of the target sequence and finally the replication of each complementary strand starting from the oligonucleotide primers, which are designed to anneal to a specific DNA sequence only (Saiki *et al.* 1988). The generation of the new complementary DNA strands starting from the oligonucleotide primers are synthesized by a DNA polymerase. The DNA polymerase typically used in PCR reactions is isolated from thermophilic bacteria and thus can withstand the multiple cycles of temperatures above 70 degrees Celsius contrary to e.g., native mammalian DNA polymerases. Each cycle of amplification exponentially increases the amount of the target DNA sequence.

Currently, three types of genetic markers are commonly employed in population genetic and molecular ecology studies at the intra-specific level: 1) analysis of changes in repeat number at microsatellite loci, 2) nucleotide sequence variation in the mtDNA control region, and 3) nucleotide sequence variation in nuclear single-copy genes. Microsatellites, also known as simple sequence length polymorphisms (Rassmann *et al.* 1991) or short tandem repeats (Schlötterer & Pemberton 1994), are similar to minisatellites (mentioned above) but with repeat motifs of only one to six nucleotides long. Similar to minisatellites, microsatellites are very common in the eukaryotic genome but are often prone to a high mutation rate. Microsatellite loci have an advantage over single-locus minisatellites in that the target DNA sequence typically is only 100-200

nucleotides long. Hence, a microsatellite locus can easily be amplified by PCR (see Chapter 2), avoiding the more time consuming Southern blotting technique (Southern *et al.* 1988) and subsequent hybridization which is necessary for single-locus minisatellites. The short length of the PCR products of amplified microsatellite loci enables electrophoresis through a high concentration polyacrylamide matrix and thus determination of the fragment length with the precision of a single nucleotide. Hence, the data can easily be digitally encoded (simply as the length of the fragments) and thus the need for subsequent tests on the same gel is unnecessary.

Changes in the nucleotide sequence of the mtDNA control region or nuclear single-copy genes can be detected by direct sequencing of PCR amplified fragments (whereas typically only the overall length is scored in microsatellite analyses). Hence, all kinds of possible changes, e.g., nucleotide substitutions, deletions and/or insertions, can be detected. Naturally, any DNA sequence can be analyzed, including coding regions (exons) and non-coding regions such as introns (non-transcribed sequences embedded in coding sequences). Although, some nuclear genes, e.g., α-lactalbumin, insulin and actin, have been studied in eukaryotes (Li *et al.* 1985; Vilotte *et al.* 1987), it is still relatively uncommon to include analyses of single-copy nuclear genes in studies of natural populations. The mtDNA control region, however, has become the standard marker in population genetic and evolutionary studies, mainly due to its unique maternal inheritance and its high evolutionary rate (Avise *et al.* 1979; Lansman *et al.* 1983).

Characteristics of the markers used in this study

Autosomal nuclear markers detected at the DNA level, e.g. by multi-locus or single-locus DNA fingerprinting, and microsatellite analysis, are inherited in a co-dominant Mendelian fashion, meaning that each allele has an equal probability of being transmitted to any given offspring. Thus, analysis of the data acquired can provide information on the degree of relatedness between individuals, paternity (Amos & Hoelzel 1990; Burke *et al.* 1991; Gill *et al.* 1985; Hoelzel & Amos 1988; Kuhnlein *et al.* 1989), population identity and level of genetic variability

(Amos & Hoelzel 1990; Bruford & Wayne 1993; Burke *et al.* 1991; Gill *et al.* 1985; Hoelzel & Amos 1988; Kuhnlein *et al.* 1989; Queller *et al.* 1993) in addition to genetic identification of individuals. Comprehensive summaries of the applications of microsatellite loci in population biology are given in Ashley & Dow (1994), Bruford & Wayne (1993) and, Schlötterer & Pemberton (1994).

Microsatellites are highly abundant, widely dispersed in most eucaryotic genomes (Tautz & Renz 1984) and generally have a high mutation rate (around 10^{-4} per locus per generation, Hughes & Queller 1993; Schlötterer & Tautz 1992; Tautz & Schlötterer 1994; Tautz 1989; Weber & May 1989). As for minisatellites, polymorphisms are due to differences in the repeat number among alleles at a locus, probably caused by slipped-strand mispairing (or single-strand slippage) during DNA replication facilitated by the repeated DNA sequence (Amos & Rubinsztein 1996; Ellegren *et al.* 1995; Schlötterer & Tautz 1992; Tautz 1989).

Relatively early, it was recognized that the infinite allele model (IAM) did not apply to microsatellite loci as the step-wise mode of evolution (simple addition or loss of repeats) could easily result in alleles of identical length that differed in descent (Valdes *et al.* 1993). One pre-requisite of the IAM is that new mutations generate new and unique alleles not previously existing in the population. In view of this, new statistics have been proposed (Feldman *et al.* 1997; Goldstein *et al.* 1995a; Goldstein *et al.* 1995b; Goodman 1997; Kimmel *et al.* 1996; Rousset 1996; Slatkin 1995) based on the step-wise mutation model (SMM) originally developed for electrophoretic data (Ohta & Kimura 1973). However, subsequent empirical studies have shown that the resulting genetic distances calculated from the new statistics are not consistent with the expectations from, e.g., geographic distance (Chapter 3; Valsecchi *et al.* 1997). One possible reason for this observation is that the mutation process in microsatellites is more complicated than suggested by the step-wise mutation model. Several studies have since demonstrated and suggested other mutational mechanisms such as multi-step mutations (Di Rienzo *et al.* 1994), directional mutation biases (Amos &

Rubinsztein 1996; Deka *et al.* 1994; Ellegren *et al.* 1995; Garza *et al.* 1995; Primmer *et al.* 1996), and insertion or deletion of non-repeat sequences or single nucleotide substitution (Angers & Bernatchez 1997; Estoup *et al.* 1995; FitzSimmons *et al.* 1995; Garza *et al.* 1995; Palsbøll *et al.* In press).

Another highly relevant issue to take into account during the analysis of microsatellite data is the presence of "null alleles". Null alleles arise if a mutation has occurred at the oligo-nucleotide priming site (the flanking region of the microsatellite locus) which subsequently prevents proper annealing and subsequent amplification of the specific allele. There are numerous reports of null alleles, e.g. such as in deer (*Cervus elaphus*) (Pemberton *et al.* 1995), bears (Ursidae) (Paetkau & Strobeck 1995) and minke whale (*B. acutorostrata*) (Van Pijlen *et al.* 1995; Chapter 2). The presence of null alleles can be assessed by inclusion of parent-offspring pairs and will show up as homozygous parent and offspring that do not share an allele (Larsen *et al.* 1996; Pemberton *et al.* 1995; Chapter 2). Null alleles will affect, for instance, the genotypic frequencies, and typically result in significant deviations from the expected H-W genotypic proportions under panmixis (random mating in a single population) due to an excess of homozygotes. In the present work on fin whale populations, deviations from H-W genotypic proportions at several loci were detected, but at different loci in the different populations, suggesting that the deviations from H-W proportions most likely do not stem from null alleles (Chapter 3).

In order to complement the information acquired from the analysis of the nuclear microsatellite loci, we sequenced 288bp of the mtDNA control region. The combined analysis of both bi-parental (microsatellite loci) and solely maternal (mtDNA) transmitted markers will, in principle, allow us to distinguish between maternally directed site fidelity to summer feeding grounds and the existence of separate sub-populations at summer feeding grounds (Chapter 3). Maternally directed site fidelity to summer feeding areas is common among cetaceans (Katona & Beard 1990; Knowlton *et al.* 1992) and thus analysis of maternally transmitted mtDNA loci alone does not provide sufficient information to discern

between the two previously mentioned hypotheses. Although mtDNA is considered as strictly maternally inherited in mammals (Hutchinson *et al.* 1974), evidence of some degree of bi-parental inheritance was detected in mouse hybrids (Gyllensten *et al.* 1991) and in invertebrates, e.g., marine mussels (Zouros *et al.* 1992). The mt control region is particularly of interest as it is the most rapidly evolving part of the mtDNA which in itself has a mutation rate that is 5 to 10 times higher than single-copy nuclear DNA (Aquadro & Greenberg 1982; Brown *et al.* 1979). Although mt control region (sometimes called the D-loop) is evolving at a high rate and there appear to be mutational "hot spots", it probably follows an IAM rather closely at the intra-specific level. In other words, each mutation is most likely to happen at a nucleotide position that is monomorphic in the population and thus results in a new and unique allele (or haplotype) not present in the population. The absence of recombination (due to the clonal maternal inheritance) implies that a genealogy of the sampled sequences can be estimated directly from the polymorphic sites and enables detection of past evolutionary events such as migration, bottlenecks and population divisions (Hoelzel 1994). In vertebrates, the mt control region spans between the genes that encode the tRNAs for Phenylalanine and Proline (Annex A) and is extremely variable in length (200 to 4000bp) (Brown 1985) and in nucleotide composition, compared to the remainder mt genome (Saccone *et al.* 1987; Saccone *et al.* 1991). Hoelzel *et al.* (1991) suggested the substitution rate in the mtDNA control region of cetaceans was one order of magnitude lower than the one found in the human mtDNA control region, estimated to 2.8 to 5 times the rates found in the rest of the human mt genome (Aquadro & Greenberg 1982). However, a recent study suggested that the rate of mtDNA substitution in cetaceans, in particular in sperm whales, is much faster than previously assumed (Lyrholm *et al.* 1996).

Collection of the samples

Although the analysis of hyper-variable DNA markers, such as microsatellite loci, has now dramatically changed the study of population biology for all species, these techniques were initially confined to those species for which tissue or blood

samples were easily obtainable. For cetaceans, tissue samples have been collected either during commercial and aboriginal whaling operations or more recently using biopsy-sampling techniques. During the last decade, biopsy equipment has become increasingly common, which usually comprises a crossbow (or modified gun) and an arrow with a modified stainless steel tip with a float molded to the arrow (Figure i.2A-C) (Aguilar & Nadal 1984; Hoelzel & Amos 1988; Lambertsen 1987; Mathews *et al.* 1988; Palsbøll *et al.* 1991; Patenaude & White 1995; Whitehead *et al.* 1990). This equipment has evolved to become a reliable tool for collecting skin biopsy samples from a variety of marine mammal species, particularly in conjunction with photo-identification studies. The biopsy tip itself is typically a cylinder (approximately 8mm in diameter and 25mm in length) with a sharp leading edge and barbs inside the cylinder to retain the sample.

Alternative successful approaches to obtain DNA from marine mammals include the collection of sloughed skin (Amos *et al.* 1992; Clapham *et al.* 1993b) and feces (Reed *et al.* 1997). Although the concentration of the total genomic DNA extracted from such samples is often lower than from the skin biopsy samples, it is still sufficient to perform PCR amplification of short DNA fragments (Amos *et al.* 1992; Clapham *et al.* 1993b; Reed *et al.* 1997). It should be taken into account, however, that biopsy samples not only allow samples for genetic analysis but also a small amount of blubber, which provides enough material to conduct toxicological (Gauthier *et al.* 1997; Marsili & Focardi 1996) and fatty acid analyses (Borobia *et al.* 1995). In addition, with groups of animals, it is often difficult to assess from which individual the sloughed skin or feces sample originates. Despite this uncertainty, these alternative methods have provided opportunities to collect samples either from species or individuals that can be hard to observe and/or approach or in areas where biopsy collection is not allowed.

Initially, concerns were expressed as to the possible effects of biopsy sampling on whales. Subsequent studies have shown that the effects appear to be minor and short lived, e.g., for right whales (Brown *et al.* 1991), belugas (Patenaude & White 1995), humpback whales (Palsbøll *et al.* 1991; Weinrich *et al.* 1991) and sperm

whales (Whitehead *et al.* 1990). The tissue samples used in this study were collected from past commercial whaling operations (off Iceland and Spain), aboriginal whaling (part of the West Greenland samples), and skin biopsies collected from free-ranging animals (Gulf of St. Lawrence, Gulf of Maine, Ligurian Sea, Sea of Cortez and West Greenland). For all biopsied individuals, photographs were taken in order to prevent and/or detect duplicate sampling. In addition, associated behavioral information was collected before, during, and after the biopsy attempt when possible. Most biopsies were collected from the area just below the dorsal fin on the right side (Figure i.2B). We recorded and conducted a preliminary analysis of the immediate response of fin whales in the Gulf of St. Lawrence to the biopsy attempts. Absence or low level reaction (sinking or rolling) was detected in 86% of the biopsy attempts irrespective if the whale was hit or not. No strong reactions were observed (e.g., full breach, tail breach, or quick acceleration), however, moderate responses (slight acceleration or tail flick) were observed in the remainder 14% of the attempts. In most instances, (some post-biopsy reactions were not recorded) the individual would return to the pre-biopsy behavior at their next surfacing. The individuals that reacted most strongly to the biopsy attempts were most often single individuals and the reactions appeared to be due more to the close encounter with the boat rather than the impact of the arrow. From this *very* preliminary work, it appears that biopsy-sampling has only minor and short-lived effects on the behavior of fin whales as seems to be the case for other baleen whales (Brown *et al.* 1991; Palsbøll *et al.* 1991; Patenaude & White 1995; Weinrich *et al.* 1991; Whitehead *et al.* 1990).

Figure i.2. Collection of samples.

A

B

C

Note. A) Crossbow with arrows. B) The pointer indicates the ideal target area for a biopsy, just below the dorsal fin. C) the stainless steel tip and a float molded to the arrow with the skin (not visible, in the tip) and blubber (part of it is visible at the edge of the tip) sample.

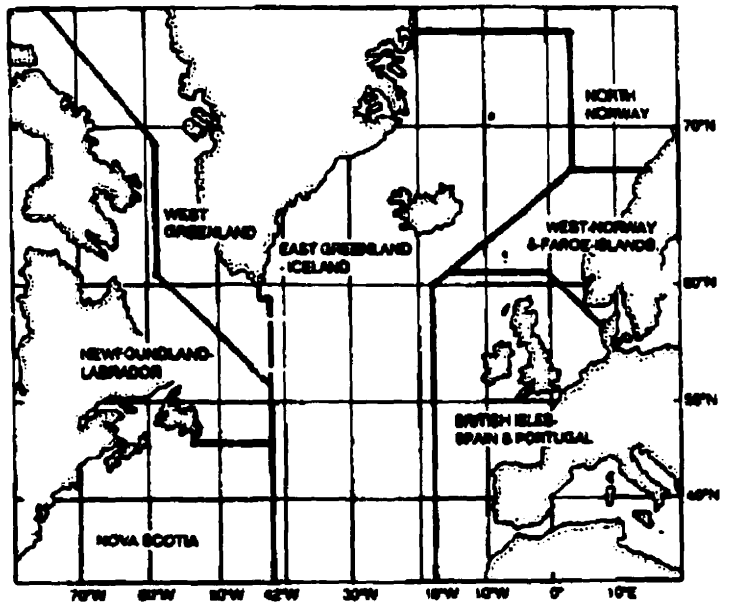
Objectives and hypotheses

Historically, the North Atlantic fin whale has been hunted over most of its North Atlantic range, with fisheries based in Newfoundland, West Greenland, Iceland, Britain, Norway and Spain. Fin whales have been a prime target of the whaling industry since the 1860's (Tønnessen & Johnsen 1982) when the Norwegians developed an explosive grenade harpoon fired from a steam-powered boat (Scarff 1977). Around 1930, the importance of fin whales in the whaling industry increased even more as the blue whale stocks became depleted. The fin whales from the eastern North Atlantic were heavily fished (Sergeant 1977) compared to the western North Atlantic where the fishery was rather episodic and finally ended when the two Canadian shore-based stations closed in 1972 (Mitchell 1972). Today, the only fin whales killed in the North Atlantic are aboriginal catches off West Greenland (the annual quota is set at 19 whales until year 2002, see Press release from the last International Whaling Commission (IWC) Scientific Committee meeting in October 1997, Monaco). This harvest constitutes a vital component of the local diet and economy in the West Greenland settlements.

In 1977, the North Atlantic fin whales were subdivided by the IWC into 7 management units or "stocks" to facilitate the management and allocation of catch quotas (Figure i.3 IWC 1992). The delineation of the seven management units

was based on limited and often inconclusive information; mainly the geographical distribution of catches and differences in the length frequency distribution in catches from different areas (IWC 1992). One of the main objectives of this study was to investigate the population structure of the North Atlantic fin whale, which would verify if the seven management units proposed by the IWC scientific committee have any biological meaning. More specifically, I wanted to discriminate among the three hypotheses regarding the population structure of the North Atlantic fin whales currently put forward: 1) a single panmictic population, 2) a continuum of sub-populations with limited gene flow, or 3) a series of discrete and genetically isolated stocks. The analysis of the genetic data obtained from the samples was performed in a hierarchical manner starting with comparisons across different oceans, and subsequently between distant locations within the North Atlantic Ocean. Analysis of both nuclear and mt loci was undertaken; more specifically the mtDNA control region and six microsatellite loci.

Figure i.3. "Stock" boundaries for the North Atlantic fin whales.



An often-significant variable in population studies is the sex of individuals, which in turn can affect, e.g., the reproductive success and social interactions

between individuals. Hence, the proportion of male to female fin whales in each sampling locality was estimated to assess if there were significant deviations from the 1:1 sex ratio found in most cetacean populations. If a significant deviation from parity was detected, we attempted to identify the source of the observed bias by including additional biological information collected during the biopsy sampling, such as, group size, position, and date. Finally, we combined the sex information with the mt and nuclear DNA data to test the possibility of sex-biased dispersal.

PART I. The Tools

Prerequisite to study the evolution and genetics of individuals, populations or species, are the means to quantity the variation between entities. The level of genetic variability, as revealed by molecular techniques (e.g., gel electrophoresis, RFLPs, microsatellite loci, mtDNA sequences, etc.), can be estimated from the gene frequencies and sequence composition of each allele in a population. Within a population, the combined action of mutation, recombination, gene flow, and genetic drift will determine the genotype (defined as the sum of genes of an organism) of the individuals.

Molecular techniques are constantly evolving. During this project, I applied a number of techniques, which needed to be refined or developed further during the course of my thesis work. The following two chapters describe (1) new molecular methods to determine the sex in cetaceans and (2) the isolation of new microsatellite loci from humpback whale genomic DNA.

Before applying the novel sex-determination technique, the Palsbøll *et al.* (1992) approach for sexing fin whales was used and optimized. The technique is based upon the PCR amplification of approximately 170 nucleotides of the SRY gene (located on the Y chromosome, thus observed only in males). The absence of PCR products can thus indicate either the absence of a Y-chromosome (e.g., a female sample) or failure of the PCR amplification. To differentiate between these two possibilities, an internal amplification control was included, namely the co-amplification of the autosomal α-lactalbumin gene (around 600 nucleotides) with the SRY gene. While this, in principle, is a reasonable approach, such co-amplification can result in differential amplification of one gene over the other, as discussed in Chapter 1. While using this technique, it became clear that the interpretation of the results could be problematic, which lead to the development of the new techniques described in Chapter 1. These new techniques circumvent the problem of differential amplification as well as being quick, efficient and reliable. The basic principle of multiplexing with three primers, with a single forward oligonucleotide primer that anneals to the Y- as well as the X-

chromosome, and two reverse oligonucleotide primers that anneal to either the Y- or the X-chromosome was novel at the time. The technique is applicable to all mammals, requiring only limited additional sequencing work for a new family/taxon.

The isolation and characterization of microsatellite loci was first described in 1989 (Tautz 1989; Weber & May 1989). When the present work was initiated, only a few cetacean microsatellite loci and sequences for the oligonucleotide primers had been published (Schlötterer *et al.* 1991). The "Cetacean Genetics Group" at the University of Copenhagen was (at the start of my thesis project) committed to performing microsatellite analysis of several thousand samples. Preliminary trials with yet unpublished dimer microsatellite loci isolated from humpback and sperm whale (Valsecchi & Amos 1996) yielded data of an unsatisfactory quality, mainly due to the occurrence of intensive "stutter bands" (amplification products that differ in length from the original allele, which are generated during the PCR amplification). Hence, it was decided to isolate microsatellite loci from genomic DNA isolated from humpback whale, focusing on microsatellite loci with tri- and tetramer repeat motifs, but at the same time screen for the more numerous dimer microsatellite loci. Tri- and tetramer microsatellite loci are known to generate far fewer or no stutter bands, but are much less common than the dimer microsatellite loci (Chapter 2).

Although many aspects of evolutionary process at microsatellite loci still are poorly understood, these are mainly of concern for estimations between relatively diverged populations or species, and pose less of a problem for analyses between individuals and close populations. Microsatellite analyses have been successfully applied in numerous population genetic studies (Ashley & Dow 1994; Bruford & Wayne 1993), phylogeny (Bowcock *et al.* 1994; Goldstein *et al.* 1995b; Takezaki & Masatoshi 1996), estimation of the degree of relatedness between individuals, (Blouin *et al.* 1996; Queller & Goodknight 1989; Richard *et al.* 1996a) and as genetic tags (Palsbøll *et al.* 1997a).

Chapters 1 and 2 have both been published in the scientific journal Molecular Ecology. Chapter 1, "Identification of sex in Cetaceans by multiplexing with three ZFY and ZFX specific sequences" is a short communication (Molecular Ecology, 1996, 5: 283-287 and Molecular Ecology, 1996, 5: 602) and chapter 2, "Primer for the amplification of tri- and tetramer microsatellite loci in cetaceans", is a Primer Note (Molecular Ecology,1997, 6: 893-895).

CHAPTER 1. IDENTIFICATION OF SEX IN CETACEANS BY MULTIPLEXING WITH THREE ZFY AND ZFX SPECIFIC SEQUENCES

Martine Bérubé [1,2] and Per Palsbøll [1].

1. Department of Population Biology, Zoological Institute, University of Copenhagen, Universitetsparken 15, Copenhagen Ø, DK-2100, Denmark.
2. Department of Natural Resource Sciences, Macdonald Campus, McGill University, Ste-Anne-de-Bellevue, Québec, H9X 3V9, Canada.

Keywords: Cetacea, Odontoceti, Mysticeti, ZFX, ZFY

Running title: Gender determination in cetaceans by multiplexing

Abstract.

We sequenced 540 nucleotides of the last exon in the ZFY/ZFX gene in two males and two females for eight cetacean species; four odontocetes (toothed whales) and four mysticetes (baleen whales). Based upon the obtained nucleotide sequences, we designed two sets of oligonucleotide primers for specific amplification of the ZFX and the ZFY sequence in odontocetes and mysticetes, respectively. Each primer set consisted of three oligonucleotides; one forward-oriented primer, which anneals to the ZFY as well as the ZFX sequence, and two reverse-oriented primers that anneal to either the ZFX or the ZFY sequence. The resulting two amplification products (specific for the ZFY and ZFX sequences) can be distinguished by gel-electrophoresis through 2 % NuSieve™. The accuracy of the technique was tested by determination of gender in 214 individuals of known sex. Finally, we applied the technique to determine the sex of 3570 cetacean specimens; 2284 humpback whales, 315 fin whales, 37 blue whales, 7 minke whales, as well as 592 belugas, 335 narwhals and 25 harbor porpoises.

Introduction.

The use of molecular genetics in conservation biology has increased steadily since the advent of the Polymerase Chain Reaction (PCR). PCR-based methods require only minute amounts of DNA, which in turn allow the investigator to collect the necessary tissue by remote sampling. Thus, many projects are based upon samples collected as skin biopsies (Aguilar & Nadal 1984), sloughed skin (Clapham *et al.* 1993b), hair (Taberlet & Bouvet 1992), blood (Arctander 1988) or feces (Höss *et al.* 1992) from which total-cell DNA for PCR-analysis is extracted. Consequently, the gender of the individual from which the sample originated is often unknown. Several PCR-based methods, which allow sex determination from such samples, have already been presented (Bradbury *et al.* 1990; Nakahori *et al.* 1991; Palsbøll *et al.* 1992; Richard *et al.* 1994; Sasi *et al.* 1991; Taberlet *et al.* 1993). When using the method presented by Palsbøll *et al.* (1992), we came across difficulties in amplifying the 1149 base pair ZFX/ZFY fragment when the DNA was degraded. To circumvent this problem, we amplified only the last 700 base pairs (of the same fragment).

While the reduction in the size of the target sequence increased the rate of successful amplifications, the subsequent digestion with the restriction endonuclease *Taq* I yielded no gender-specific restriction fragment pattern. Hence the assumption by Palsbøll *et al.* (1992) that the diagnostic *Taq* I restriction site was in the same position in cetaceans as in humans was incorrect. The findings reported here are the result of an attempt to further develop a simple and robust PCR-based method for identification of sex in cetaceans (whales, dolphins and porpoises). Our objectives were the identification of the correct sex in all species tested, the omission of restriction endonuclease digestion but providing a control for success of PCR amplification, and finally allowing for identification of sex for samples that yield degraded DNA extractions.

Materials and Methods.

Sequencing of ZFY/ZFX sequences.

DNA was extracted by standard procedures (Maniatis *et al.* 1982). ZFY/ZFX nucleotide sequences were obtained by direct sequencing of asymmetrically PCR-amplified DNA (Higuchi *et al.* 1988; Saiki *et al.* 1988). Initial symmetrical amplifications were performed in a total volume of 10 µl (conditions as described in Palsbøll *et al.* (1992), but 1 µM of each oligonucleotide primer). An initial 2 min denaturing step at 94°C was followed by 30 cycles of each 1 sec at 94°C, 1 sec at 55°C and 20 sec at 72°C on an Idaho Technology Air™ Thermo-Cycler. Asymmetrical amplifications were conducted under identical conditions except that the concentration of the limiting oligonucleotide primer was lowered to 0.01 µM and the reaction volume increased to 50 µl. Amplifications were carried out on a Techne™ Thermo-Cycler for 30 cycles of each: 1 min at 94°C, 1 min at 56°C and 1 min at 72°C. Blank controls were included with all amplification reactions to detect possible contamination. Sequencing was conducted following the manufacturer's instructions (Sequenase' Ver.2.0, US Biochemicals Inc.). The oligonucleotide primers ZFYX0606F (5'-ATA GGT CTG CAG ACT CTT CTA-3'), designed from the human ZFX (Schneider-Gadicke *et al.* 1989) and ZFYX1204R (Palsbøll *et al.* 1992) were used for amplification as well as sequencing. In addition, ZFYX0983R (5'-CTT ACA CCT AAA TGG AAG ATC C-3', also designed from the human ZFX (Schneider-Gadicke *et al.* 1989) was used as an internal sequencing oligonucleotide primer.

Sex determination.

Based upon the obtained sequences, we designed two sets of each, three oligonucleotide primers, for differential PCR amplification of the ZFY and ZFX sequence. Each set was specific for either odontocetes or mysticetes. Each set consisted of a forward-oriented oligonucleotide primer designed to anneal to the ZFY as well as the ZFX sequence. The 3' end, for each of the two reverse-oriented oligonucleotide primers, was placed at positions that were polymorphic

between the ZFX and the ZFY sequences. Hence each of the reverse-oriented primers would amplify (in combination with the forward-oriented primer) only the ZFY or the ZFX sequence. The size of each of the two amplification products differed sufficiently to be distinguished by gel electrophoresis through 2 % NuSieve™. Amplification was performed on a Techne™ Thermo-Cycler with the oligonucleotide primer combinations mentioned below (Results and Discussion; Design of sex-chromosome-specific oligonucleotide primers). Amplifications were carried out in 20 µl volume for 37 cycles (60 sec at 94°C, 60 sec at 52°C and 90 sec at 72°C) preceded by a 5 min denaturing step at 94°C. The amplification products were separated by electrophoresis at 200 volts through a 2% NuSieve™ gel (containing 0.05 µg/ml ethidium bromide) (Figure 1.1).

Results and Discussion.

ZFY/ZFX sequences.

In all, 540 nucleotides were sequenced of the last exon in the ZFY/ZFX gene (Schneider-Gädicke *et al.* 1989) for two females and two males, respectively, in eight cetacean species: four odontocetes (the harbor porpoise, *Phocoena phocoena*, the narwhal, *Monodon monoceros*, the beluga, *Delphinapterus leucas* and the sperm whale, *Physeter macrocephalus*), and four mysticetes (the minke whale, *Balaenoptera acutorostrata*, the fin whale, *B. physalus*, the blue whale, *B. musculus*, and the humpback whale, *Megaptera novaeangliae*). The ZFX sequences were determined directly from the homozygous (ZFX/ZFX) female samples. The ZFY sequences were obtained by deducting the ZFX sequence from the heterozygous (ZFX/ZFY) sequence in the male samples. In no instances did we observe intra-specific polymorphisms in the ZFY or ZFX sequence.

Design of sex-chromosome specific oligonucleotide primers.

Of the 540 nucleotides, we detected a total of 23 polymorphic sites among the 32 sequences. Five were gender-specific for all species; two only for odontocetes

and one only for mysticetes (Table 1.1). The sequences of the oligonucleotide primers used for sex determination are listed in Figure 1.2.

Sex determinations.

The sex identification method presented proved to be simple, fast as well as reliable. It was employed for sex-identification of 3570 cetacean specimens; 2284 humpback whales, 315 fin whales, 37 blue whales, 7 minke whales, 592 belugas, 335 narwhals and 25 harbor porpoises, respectively. The sex was correctly identified in 214 individuals of known sex; 152 humpback whales, 31 fin whales, 7 minke whales, 12 narwhals, 6 belugas and 6 harbor porpoises. The methodology relies on the assumption of no polymorphisms at the 3' end of the annealing sites for the applied oligonucleotide primers. That no substitutions were observed in any of the eight species analyzed, indicate that such substitutions are non-existent or at minimum very rare.

Of 152 humpback whale mothers (defined by the close association with a calf), one was identified as a male. In this case, errors in sample handling or contamination appear to present more plausible causes for such a 'male mother' than inaccuracy of the method.

As mentioned in the Introduction, we were unable to amplify the 1149 base pair ZFY/ZFX fragment (Palsbøll *et al.* 1992) in 211 samples where the DNA was highly degraded. With the presented method gender was determined for 93 % of these samples.

The described procedure circumvents digestion of the PCR amplification products with restriction endonucleases and provided an internal control for successful PCR amplification. Most PCR-based methods for gender-determination presented so far include an internal control (Bradbury *et al.* 1990; Nakahori *et al.* 1991; Palsbøll *et al.* 1992; Taberlet *et al.* 1993). Evidently such internal control is vital since absence of a Y-chromosome-specific amplification product could also be due to failure of amplification. The choice of target sequence acting as internal control is far from simple, since the simultaneous use of several oligonucleotide primers in a single PCR amplification can result in

preferential amplification of one target sequence. Most sex identification methods presented so far co-amplify a mitochondrial sequence as the internal control. However, for most types of tissue, cells contain many mitochondria, but only one nucleus and hence such approach adds to the overall risk of preferential amplification of the control sequence. Using a nuclear gene, for instance the ZFX (Nakahori *et al.* 1991; Palsbøll *et al.* 1992), as an internal control ensures an equal number of copies per cell. However, if the internal control target sequence is an autosomal gene, the ratio will still be 2:1 relative to the Y-chromosome target sequence. That such seemingly small differences change the outcome of the PCR amplification became evident when we used the 'odontocete' primer set for sex-identification in baleen whales. The intensity of the ZFY-specific product was much lower (compared to the ZFX-specific product) relative to when odontocete DNA was amplified. The reverse oligonucleotide primer specific for the odontocete ZFX exon (ZFX0947R) matches to the mysticete ZFY and ZFX exon. Hence, for each cycle, twice as many copies are synthesized of 245 bp (one from the ZFY exon and one from ZFX exon) than of the 212 bp ZFY-specific fragment.

In the presented procedure, we have placed the 3'end of each of the two reverse oligonucleotide primers at positions that have gender-specific substitutions. This ensures a 1:1 amplification (all other factors being equal) of the diagnostic Y-chromosome-specific sequence and the internal control sequence.

Tables.

Table 1.1. The polymorphic positions (relative to the cetacean consensus sequence in Figure 1.2).

	6 6 4	7 1 5	7 7 6	8 0 9	8 5 4	8 6 9	9 0 8	9 1 1	9 1 4	9 4 4	9 4 7	9 5 3	9 9 2	1 0 2	1 0 5	1 1 8	1 1 0	1 1 4	1 1 1
Odontocetes																			
ZFX																			
P. phocoena	T	T	C	A	C	G	C	G	T	A	C	A	G	C	A		A	C	G
D. leucas	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.		.	.	.
M. monoceros	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.		.	.	.
P. macrocephalus	.	.	.	.	.	C	.	.	C	G	.	.	A	T	.		G	.	.
ZFY																			
P. phocoena	G	.	.	.	T	.	.	.	.	.	T	G	.	T	.		.	.	.
D. leucas	G	.	T	.	.	.	.	.	.	.	T	G	.	T	.		G	.	.
M. monoceros	G	.	T	.	.	.	T	.	.	.	T	G	.	T	.		G	A	.
P. macrocephalus	.	.	T	G	.	C	.	A	.	G	T	G	A	T	.		G	.	.
Mysticetes																			
ZFX																			
M. novaeangliae	T	T	C	A	C	C	C	G	T	G	C	A	G	T	A		G	C	A
B. acutorostrata	.	.	.	.	.	C	.	.	.	G	.	.	.	T	G		G	.	.
B. musculus	.	.	.	.	.	C	.	.	.	G	.	.	.	T	.		G	.	.
B. physalus	.	.	.	.	.	C	.	.	.	G	.	.	.	T	.		G	.	.
ZFY																			
M. novaeangliae	.	G	T	G	.	C	.	.	.	G	.	.	.	T	.		G	.	.
B. acutorostrata	.	.	T	G	.	C	.	.	.	G	.	.	.	T	.		G	.	.
B. musculus	.	.	T	G	.	C	.	.	.	G	.	.	.	T	.		G	.	.
B. physalus	.	G	T	G	.	C	.	.	.	G	.	.	.	T	.		G	.	.

Note. Darker shading denotes substitutions that are gender-specific for all sequenced cetaceans, lighter shading for only odontocetes or mysticetes. The substitution in position 734 creates a recognition site for the endonuclease *Hph I* in males.

Figures

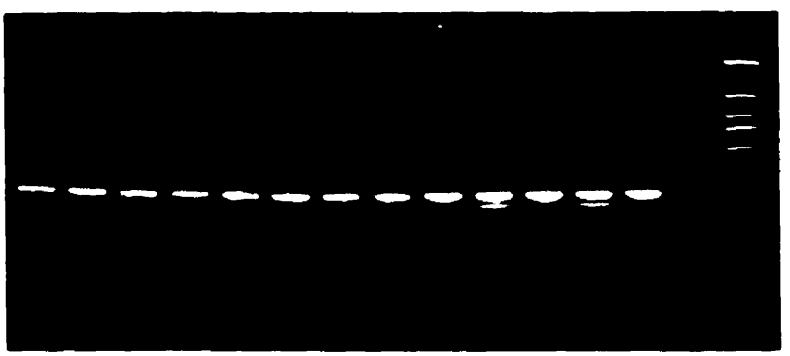
Figure 1.1. The ZFY and ZFX amplification products after gel electrophoresis.

A

M M M F M M M M M F F F F C Mk1

B

F M M F F F F M F M F M F C Mk2



Note: (A) Odontocete oligonucleotide primer set (ZFYX0606F, ZFY0791R, and ZFX0947R), (B) mysticete oligonucleotide primer set (ZFYX0606F, ZFY0776R, and ZFX0809R). Odontocete samples in A: are female (F) and male (M) beluga and for the mysticetes in B: female (F) and male (M) fin whales. Lanes' C denote the negative control amplification and lanes Mk1 and Mk2 are the markers, restriction endonuclease *HaeIII* digested lambda Phi174 DNA and restriction endonucleases *HinfI* and *HinfI/sspI* digested pUC 18 DNA, respectively. The long band (corresponding to the X-chromosome) is 383 bp in the odontocetes and 245 base pairs long in the mysticetes PCR amplifications. The short band (corresponding to the Y-chromosome) is 227bp for the odontocetes and 212 bp in the mysticetes PCR amplifications. The longer band observed in the male specimens is probably a PCR artifact (ex. recombination during the amplification).

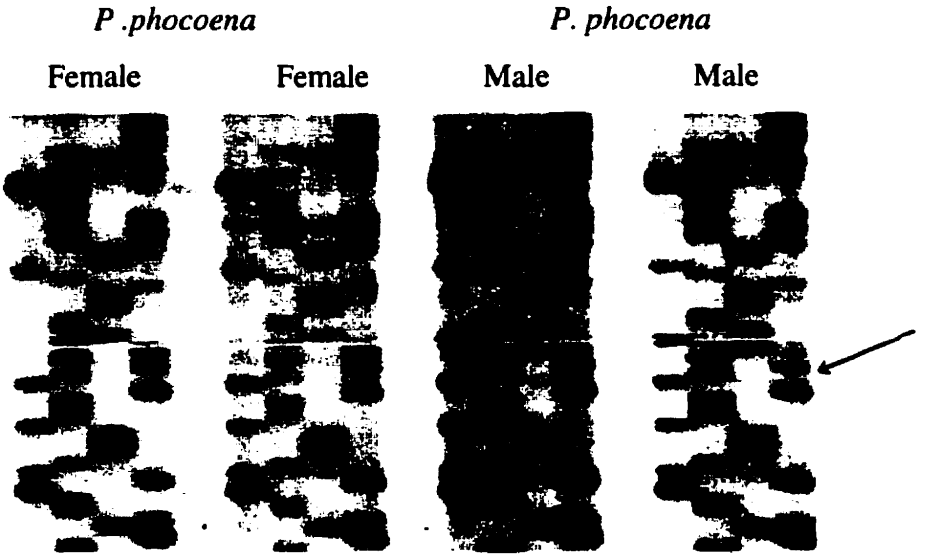
Figure 1.2. The human and the cetacean consensus ZFX sequences with positions and sequences of the used oligo-nucleotide primer combinations.

Figure 1.2 (continuous)

```
Odon.primers    5' AGAATATGGCGACTTAGAACG3'ZFX0947R   5' CTTACACCTAAATGGAAGATCC3'ZFYX0983R
Myst.primers
Odon.    zfx    tcgttctaagtcgccatattctctcagttcacacaaaagatcttccgtttaggtgtaagagatgtagaaagggattcaggcaacagaatgagcttaaaaagcatatgaagacacacagtg
Odon.    zfy    .t.....g.....................................................................t.....................................
Myst.    zfx    tcgttctaagtcgccatattctctcagttcacacaaaagatcttccgtttaggtgtaagagatgtagaaagggatttaggcaacagaatgagcttaaaaagcatatgaagacacacagtg
Myst.    zfy    ...............................................................................................................................
Human    zfx    ttgttctaagtcgccatattctctcagttcacacaaaggatcttccatttaggtgcaagagatgtagaaagggatttaggcaacagagtgagcttaaaaagcatatgaagacacacagtg 1065
Odon.primers
Myst.primers
Odon.    zfx    gtagaaaagtatatcagtgtgagtactgtgagtatagcactacagatgcctcaggctttaaacggcacgttatctccattcatacgaaagactatcc-3' 1162
Odon.    zfy    .c.....................................................................................t.
Myst.    zfx    gtaggaaagtatatcagtgtgagtactgtgagtatagcactacagatxc-tcaggctttaaacggcacgttatctccattcatacgaaagactatcc
Myst.    zfy    .c.....................................................................................t.
Human    zfx    gcaggaaagtgtatcagtgtgagtactgtgagtatagcactacagatgcctcaggctttaaacggcacgttatttccattcacacgaaagactatcctcatcggtgtgagtactgcaaga 1185

Primer                          5'CATTATGTGCTGGTTCTTTTCTG3'ZFYX1204R
Human    zfx    aaggcttccgaagaccttcagaaaagaaccagcacataatg-3'1226
```

Note. The underlined oligonucleotide primers were used for amplification and sequencing.

Figure 1.2 (continuous)



Note: Please note that all reverse-oriented oligo nucleotide primers (denoted by R) are complementary to the consensus sequence. Two versions of the ZFY specific odontocete oligo-nucleotide primer ZFY0791R were ordered, one (by mistake) with an incorrect 3' end sequence (5'-CAC-3'). When the mistake was discovered a second order was placed with the correct sequence. However only the 'incorrect' version (listed in Figure 1.1) would amplify the ZFY sequence when used in combination with the ZFX0947R oligo nucleotide primer. Two additional batches of the correct version (each from different manufacturers) were tested, but with similar results. We cannot offer any rational explanation for this occurrence. The shaded section (where the "incorrect and correct" ZFY0791R are located) of the consensus sequence is shown on the harbor porpoise *(P. phocoena)* autoradiogram. The arrow indicates the first nucleotide at the 5'end of the ZFY0791R oligo nucleotide primer on the male harbor porpoise sequence.

CHAPTER 2. PRIMERS FOR THE AMPLIFICATION OF TRI- AND TETRAMER MICROSATELLITE LOCI IN BALEEN WHALE.

Per J. Palsbøll[§], Martine Bérubé[§#] , Anja H. Larsen[§] and Hanne Jørgensen[§]

§ Department of Population Biology, University of Copenhagen, Universitetsparken 15, DK-2100 Copenhagen Ø, Denmark

Department of Natural Resources Science, Macdonald Campus at McGill University, Ste-Anne de Bellevue, Québec H9X 3V9, Canada

Keywords: VNTR, population genetics, whales, polymorphic loci

Corresponding author: Per J. Palsbøll, Department of Ecology and Evolutionary Biology, University of California, Irvine CA92717, USA. Phone: +1 714 824 3670, Fax: +1 714 824 2181, e-mail: ppalsbol@newt.bio.uci.edu

Cetaceans were among the first animal taxa from which microsatellite loci were isolated (Tautz *et al.* 1986). Later, Schlötterer *et al.* (1991) showed that oligonucleotide primers matching the flanking sequences of loci isolated from one species often would amplify the homologous sequence in other cetacean species as well. Since then, numerous research groups have isolated and analyzed variable microsatellite loci in cetaceans to address issues such as; social organization, population structure and mating systems (e.g., Amos *et al.* 1993; Larsen *et al.* 1996).

All but one (Richard *et al.* 1996b) cetacean microsatellite loci published so far, have a dimer repeat motif (e.g., TG). The products generated from amplification of dimer microsatellite loci often contain numerous and intense stutter bands; this frequently makes unambiguous determination of the genotype difficult, especially

distinguishing between hetero- or homozygotes when allele lengths differ by a single repeat (Edwards *et al.* 1991; Rassmann *et al.* 1991). Normally, such ambiguous samples are re-amplified until data of an adequate quality are obtained. While this presents only minor inconveniences for small sample sizes, the number of re-amplifications can reach an inconveniently high level when dealing with large samples.

The problem of ambiguous genotypes due to intense stutter bands can be overcome by one of two strategies: one can either isolate a large number of dimer loci and focus only on loci that generate less intense stutter bands; or isolate microsatellite loci with tri- or tetramer repeat motifs, which consistently produce less intense stutter bands (Tautz & Schlötterer 1994). Faced with the task of conducting microsatellite analyses of three to six loci in approximately 3,500+ samples collected from baleen whales, we chose to isolate a large number of dimer loci, but at the same time screen the genomic libraries for tri- and tetramer inserts as well. In this communication, we present the sequence for the oligo-nucleotide primers and their level of variation for the polymorph tri- and tetramer loci we isolated during this process.

The majority of the analyzed samples were collected from free-ranging whales (Palsbøll *et al.* 1991), while others were collected during aboriginal hunting or pre-moratorium commercial whaling operations. Total-cell DNA was extracted from all samples by standard procedures (Sambrook *et al.* 1989). Microsatellite loci were isolated from size-selected total-cell DNA extracted from humpback whale, *Megaptera novaeangliae*, using the approach described by Rassmann and colleagues (1991). Recombinant colonies (dot blotted to a ZetaProbe™ nylon membrane) were hybridized (under the conditions recommended by the manufacturer) with a variety of oligo-nucleotides (of different di-, tri-, and tetramer repeat motifs), each end-labeled with γ-P^{32}ATP using T4 kinase (Sambrook *et al.* 1989). Positive clones were isolated, grown over night, then the plasmid DNA was isolated (Sambrook *et al.* 1989), dot blotted and re-hybridized to the oligonucleotides. Clones that resulted in a strong signal after the second

round of hybridization were subsequently sequenced by the dideoxy chain termination method (Sanger *et al.* 1977).

Oligonucleotide primers matching the flanking regions were synthesized for a total of 18 inserts with tri- and tetramer repeat motifs, and each tested on a few (between four and 54) specimens from four species of baleen whales (minke whale, *Balaenoptera acutorostrata*, blue whale, *B. musculus*, fin whale, *B. physalus*, and humpback whale). Amplifications were carried out in 10µl volumes under the following conditions; 0.067M Tris-HCl (pH 8.8), 2mM $MgCl_2$, 16.6mM $(NH_4)_2SO_4$, 10mM β-mercaptoethanol, 0.5mM of each dNTP, 0.4 units AmpliTaq™ DNA polymerase, 1µM of each oligonucleotide primer, and 0.1 µg extracted DNA. Amplifications were performed on a Gene Machine™ Junior Thermal Cycler (USA/Scientific Plastics, Inc.), electrophoresed through a 2% NuSieve™ gel and the products visualized by staining with ethidium bromide.

Eleven of the above loci, which appeared polymorphic, were subsequently amplified with one of the oligonucleotide primers end-labeled with γ-P^{32}ATP (as described above). Reactions were performed as above, the only difference being that the end-labeled oligonucleotide primer was added in a 0.04 µM concentration. The amplification products were electrophoresed through a 5% standard denaturing polyacrylamide gel and autoradiography conducted at room temperature for 4 - 16 h. λ-M13 sequences, internal controls of known allelic composition as well as blank PCR control were included on each gel to enable detection of handling errors and contamination, respectively. Amplification profiles and oligonucleotide sequences are listed in Table 2.1 for the loci tested with end-labeled primers.

All 11 loci were polymorphic (two or more alleles) in at least one of the four baleen whale species. Between two to six loci were selected for the subsequent population analyses (Table 2.1). In the humpback whale samples, we tested the possible presence of null-alleles by analysis of 201 mother and calf pairs. Null-alleles were found at one locus GATA098, for which new oligonucleotide primers were designed and re-tested. The loci employed in the population analyses all

displayed high levels of variation (an average probability of identity of 0.045) enabling unique identification of individuals with five or six loci.

As reported for other species (Rothuizen *et al.* 1994), we found that the dimer microsatellite loci were 10 to 15 times more numerous in the cetacean genome than tri- and tetramer loci. Screening the same libraries for dimer loci (GT and AT repeat motifs) yielded ≈ 300 dimer loci. In agreement with earlier findings (Edwards *et al.* 1991), we found that the tri- and tetramer microsatellite loci generated stutter bands that were less intense than observed for dimer loci (Figure 2.1), which facilitated the timely (≈ 9 months) completion of the analysis of large numbers of samples (Table 2.1) on traditional "manual" equipment.

Tables

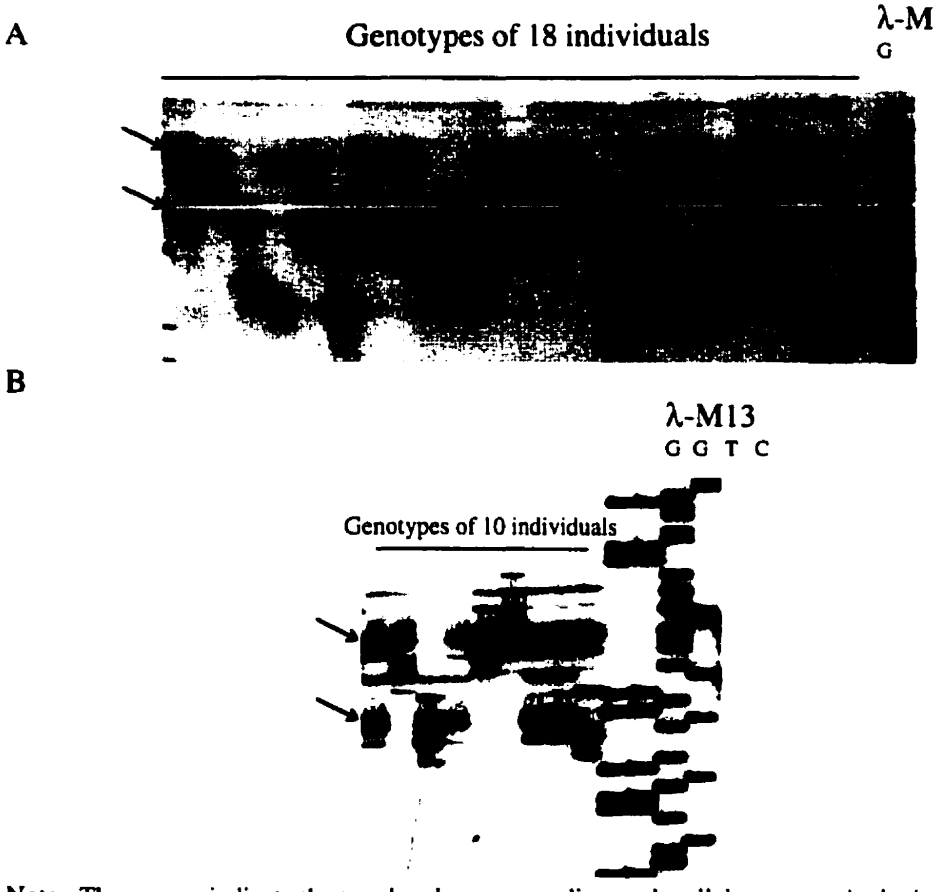
Table 2.1. Amplification and population data for tri- and tetramer microsatellite loci where more than one allele was detected.

Locus[1]	Primers sequences	°C[2]	*B. acutorostrata*		*B. musculus*		*B. physalus*		*M. novaeangliae*	
	5' 3'	Cycl.[3]	# alleles Size range[5]	*I*[6]	# alleles Size range	*I*	# alleles Size range	*I*	#alleles Size range	*I*
ACCC392	CTG ATG TTT GGT TGA TTA C	55	4		8	0.081	7		3	
	CTT CCC TCC ATC CAA GTA TTG	30	191-247		211-367		235-295		147-187	
GATA019	TGA TGA AAT CGG ACA CAC AGT	55							6	
	CTA TAA GGG AAA AGA ATC TGA	30							203-239	
GATA028	AAA GAC TGA GAT CTA TAG TTA	54	17	0.035	15	0.026	19	0.016	11	0.280
	CGC TGA TAG ATT AGT CTA GG	30	158-246		146-202		184-236		147-191	
GATA053	ATT GGC AGT GGC AGG AGA CCC	54			2		14	0.025	9	0.029
	GAC ACA GAG ATG TAG AAG GAG	30			192-196		180-220		178-210	
GATA098	TGT ACC CTG GAT GGA TAG ATT	54	8	0.074	8	0.072	8	0.043	8	0.103
	TCA CCT TAT TTT GTC TGT CTG	30	76-104		100-148		104-132		92-134	
GATA417	CTG AGA TAG CAG TTA CAT GGG	49	16	0.018	13	0.019	3		18	0.015
	TCT GCT CAG GAA ATT TTC AAG	30	213-253		181-253		251-271		193-293	
GGAT416	GAG ACC ACT GCA GGA ACA CAG	55					4			
	CAG AGG CTG ACT TTA TAC CAC	30					288-312			
GGAA520	TAG CAG AYC TGA GTT ATT TCC	54					18	0.017	27	0.034
	TAG CAT TTT AGT CTT GGG TGG	32					162-226		191-359	
GAAT400	GTC TGG AGC CAC TAC TCA GCC	55	2						3	
	AGA GCC CAG CAT CAC GGC TGG	30	151-163						167-183	
TAA023	CTC GCA CAG AAA TGA AGA CCC	55					7	0.074		
	AGA GCC TGA ACC AGA ACA AGG	30					82-100			
TAA031	AGA TCC TGC AAG CCG CAT CGG	54					3	0.610	15	0.034
	TCA CTT CCT ACT TTG ATG AGG	30					85-106		85-121	

[1] Letters denote the repeat motif and the three following digits identify the locus. [2] The annealing temperature in degrees Celsius. [3] Number of amplification cycles. [4] Number of samples analyzed. [5] In base pairs [6] Probability of identity (Paetkau & Strobeck 1994). The first of the two oligonucleotide primer sequences listed for each was end-labeled. The nucleotide sequences of each locus have been deposited to Genbank (Accession numbers; U93888-U93897).

Figures

Figure 2.1 Autoradiogram of genotypes of A) the tetramer locus GATA053, and B) the dimer GT011.



Note. The arrows indicate the two bands corresponding to the alleles present in the individual at that locus. The other bands on the autoradiograms are called stutter bands.

PART II. Genetic variation within and between fin whale populations.

Ideally populations are composed of groups of panmictic individuals, with each population in mutation-drift equilibrium, as well as genetically isolated from the other (Herskowitz 1977). The process of differentiation is well-explored under such "model-like" conditions and the degree of differentiation among populations is measured as the proportion of genetic variance within populations relative to the overall genetic variance. The degree of divergence can subsequently be converted into time (assuming a molecular clock) and correlated with e.g., geological events (Hartl & Clark 1989). Although simple in principle, the real-life situation is of course somewhat more complex as populations rarely behave in a genetically and ecologically tidy fashion. For instance, individuals migrate between populations and thus counteract genetic divergence and population sizes are usually not constant but fluctuate over time, which affects the rate of genetic drift. Finally, several cetacean species are only seasonally divided into sub-populations and do in fact constitute one panmictic population. All these aspects have to be taken into consideration when designing and interpreting the results of a population genetics study.

However, even if appropriate data and statistical analyses are employed, a frequent limitation is to obtain enough samples to ensure sufficient statistical power to discriminate between putative close sub-populations. Palsbøll (1994) explored the effect of sample size by simulations based upon humpback whale mtDNA sequences collected from two different populations. The study revealed that sample sizes of less than 40-50 individuals provided poor statistical power even between relatively divergent populations (Palsbøll 1994) a fact, which has to be taken into consideration when interpreting the results.

In this study, samples were obtained from five North Atlantic feeding areas; the Gulf of Maine (n=31), the Gulf of St. Lawrence (n=109), West Greenland (n=46), Iceland (n=33) and Spain (n=39), as well as one feeding area from the Mediterranean Sea (n=74). In addition, samples from the Sea of Cortez (n=75), in the Gulf of California (North Pacific Ocean), were used as a reference population.

Significant levels of differentiation were detected between the two oceans at all loci. While two populations were identified within the North Atlantic (the eastern and the western populations), more samples and/or loci are necessary for a reliable estimate of the degree of differentiation between adjacent sampling localities within the North Atlantic Ocean.

The sex ratio in fin whale populations had been estimated previously for many areas, such as, the Gulf of St. Lawrence, the Gulf of Maine, West Greenland, Iceland, Spain, the Mediterranean Sea, the Gulf of Alaska, and the Bering Sea (see Chapter 3 and Aguilar & Lockyer 1987; Kapel 1979; Mitchell 1974; Tarasevich 1967), and usually were found to be at parity. In our study, the sex ratio differed from parity only in the samples collected off Iceland and in the Gulf of St. Lawrence (Chapter 3). As samples from Iceland originated from past commercial whaling operations, we concluded that the sex bias towards females was probably a result of a whaling bias (as females are larger than males) (IWC 1992). However, in the Gulf of St. Lawrence, the samples were obtained as skin biopsies collected from free-ranging whales. A significant excess of males was detected and subsequent stratification of the samples by group size revealed that the source for the male excess was large groups, which mainly consisted of males. Unfortunately, more samples, behavioral information and better coverage of the Gulf of St. Lawrence are needed in order to reach more definitive conclusions. Chapter 3 is the pedestal of the thesis and is published as a full paper in the scientific journal Molecular Ecology (7(5):585-599), and chapter 4 is a manuscript intended to be submitted to the journal of Marine Mammal Science.

CHAPTER 3. POPULATION GENETIC STRUCTURE OF NORTH ATLANTIC, MEDITERRANEAN SEA AND SEA OF CORTEZ FIN WHALES, *BALAENOPTERA PHYSALUS* (LINNAEUS 1758): ANALYSIS OF MITOCHONDRIAL AND NUCLEAR LOCI.

M. Bérubé[1,2,*], A. Aguilar[3], D. Dendanto[4], F. Larsen[5], G. Notarbartolo di Sciara[6], R. Sears[7], J. Sigurjónsson[8], J. Urban-R.[9], and P. J. Palsbøll[1*]

[1]Institute of Zoology, Dept. of Population Biology, University of Copenhagen, DK-2100, Copenhagen, Denmark. [2]Dept. of Natural Resource Sciences, Macdonald Campus, McGill University, Ste-Anne de Bellevue, Qc, H9X 3V9, Canada. [3]Dept. of Animal Biology, Faculty of Biology, University of Barcelona, 08071 Barcelona, Spain. [4]Allied Whale, College of the Atlantic, Bar Harbor, Maine, 04609 USA. [5]Greenlandic Fisheries Research Institute, Nuuk, Greenland. [6]Tethys Research Institute, Acquario Civico, 20121 Milano, Italy. [7]Mingan Island Cetacean Study Inc., St-Lambert, Qc, J4P 1T3 Canada. [8]Program for Whale Research, Marine Research Institute, 121 Reykjavik, Iceland. [9]Dept. de Biologia Marina, Universidad Autonoma de Baja California Sur, La Paz, 23080 B.C.S., Mexico.

Key words: Microsatellites, mitochondrial control region, mtDNA, fin whale, gene flow, population identity.

Running head: Fin whale population genetic structure.

* Present address: Ecology and Evolutionary Biology, University of California, Irvine, CA 92697-2525, USA. Phone +1 714 824-8680, Fax +1 714 824-8121, E-mail martine@newt.bio.uci.edu

Abstract

Samples were collected from 407 fin whales, *Balaenoptera physalus*, at four North Atlantic and one Mediterranean Sea summer feeding areas as well as the

Sea of Cortez in the Pacific Ocean. For each sample, the sex, the sequence of the first 288 nucleotides of the mitochondrial (mt) control region and the genotype at six microsatellite loci were determined. A significant degree of divergence was detected at all nuclear and mt loci between North Atlantic/Mediterranean Sea and the Sea of Cortez populations. However, the divergence time estimated from the mt sequences was substantially lower than the time elapsed since the rise of the Panama Isthmus, suggesting occasional gene flow between the North Pacific and North Atlantic ocean after the separation of the two oceans. Within the North Atlantic and Mediterranean Sea, significant levels of heterogeneity were observed in the mtDNA between the Mediterranean Sea, the eastern (Spain) and the western (the Gulf of Maine and the Gulf of St. Lawrence) North Atlantic. Samples collected off West Greenland and Iceland could not be unequivocally assigned to either of the two areas. The homogeneity tests performed using the nuclear data revealed significant levels of divergence only between the Mediterranean Sea and the Gulf of St. Lawrence or West Greenland. In conclusion, our results suggest the existence of several recently diverged populations in the North Atlantic and Mediterranean Sea possibly with some limited gene-flow between adjacent populations, a population structure which is consistent with earlier population models proposed by Kellogg, Ingebrigtsen, and Sergeant (Ingebrigtsen 1929; Kellogg 1929; Sergeant 1977).

Introduction

The subcutaneous layer of adipose tissue, referred to as blubber, found in all extant mysticetes (baleen whales), has provided this group of mammals with the physiological means for extensive ranges of movement. The blubber acts not only as a highly efficient insulating barrier, but also for storage of excess energy (Brodie 1975), facilitating semi-annual migrations of more than 8,000 kilometers (Stone *et al.* 1990).

Given these characteristics, it is not surprising that spatial as well as temporal separation of feeding and breeding ranges has been observed in many mysticetes. Most mysticetes summer in high latitudes in cold water where the primary

production is high. The blubber acquired during the summer months provides sufficient energy reserves for the often extensive semi-annual migrations to lower latitudes where the primary production is low but the water temperature adequate for giving birth to insufficiently insulated calves. Mating also takes place in low latitudes, usually immediately following parturition.

'Typical' examples of species following the above migration pattern are the humpback whale, *Megaptera novaeangliae*, (Katona & Beard 1990) and the North Atlantic right whale, *Eubalaena glacialis*, (Knowlton *et al.* 1992) for which the winter and summer ranges are relatively well documented. For many mysticetes, however, the winter range has not yet been identified. Some species such as the Bryde's whale, *Balaenoptera edeni*, have a year-round tropical distribution (Best 1977) whereas other species, such as the minke whale, *B. acutorostrata*, and the fin whale, *B. physalus*, are observed in temperate waters during the summer as well as the winter months (Ingebrigtsen 1929; IWC 1992; Kellogg 1929; Sergeant 1977).

The fin whale is a cosmopolitan mysticete currently found in all major oceans. Being the second largest of all whales, the fin whale has been the target of commercial whaling operations, which have severely depleted their abundance in some oceans. The abundance of fin whales in the North Pacific Ocean has been reduced from about 42,000-45,000 to perhaps only 16,000 individuals (Barlow *et al.* 1995), whereas fin whales are still relatively numerous in the North Atlantic Ocean (approximately 56,000 individuals, Buckland *et al.* 1992a; Buckland *et al.* 1992b; IWC 1992).

While the distribution of fin whales in the North Atlantic has been described in general terms (Ingebrigtsen 1929; IWC 1992; Jonsgård 1966; Kellogg 1929; Sergeant 1977), little information is available regarding the population identity and population structure (IWC 1992). Ingebrigtsen (1929) discarded the hypothesis of a single common breeding ground based on the fact that fin whales were simultaneously observed in all regions of the North Atlantic and during the entire year. The movements of fin whales appeared to be mainly dictated by the availability of prey and the extent of ice (Ingebrigtsen 1929). Newly born calves

were observed during the winter months in high latitudes off northern Norway, indicating that warm and protected tropical waters were not a requirement for calving (Ingebrigtsen 1929) as presumed for, e.g., the humpback whale (Mitchell & Reeves 1983). Kellogg (1929) proposed that separate 'stocks' of fin whales had overlapping ranges, e.g., that the summer range for one stock of fin whales may constitute the winter grounds for another. Later, Sergeant (1977) postulated another pattern of distribution described as a "patchy continuum", mainly defined by areas of high primary production, and with a considerable degree of migration between adjacent areas. The distribution of mark-recaptures of tags, reported later (IWC 1992), was consistent with the above description of relatively limited ranges of movement. Of a total of 685 marked fin whales, 97 tags were later recovered. Only three of these 97 tags were recovered outside the area in which the whales were tagged: two animals marked in Nova Scotia were recovered in Newfoundland and one animal marked in Newfoundland turned up in Iceland nine years later (Sigurjónsson *et al.* 1991). The limited range of individual fin whales observed in the mark-recapture studies above was later confirmed by re-sightings of individual fin whales identified by their natural markings at summer feeding areas off eastern North America (Agler *et al.* 1990; Clapham & Seipt 1991).

Recently, population genetic studies of North Atlantic fin whales based upon analyses of isozymes (Arnason *et al.* 1991a; Danielsdottir *et al.* 1991; Danielsdottir *et al.* 1993) and mitochondrial (mt) DNA restriction fragment length polymorphisms (Danielsdottir *et al.* 1992) have been undertaken in an attempt to gain further insight into the population structure of North Atlantic fin whales. The above mentioned studies revealed significant differences in isozyme allele and mtDNA haplotype frequencies between samples from different geographic regions as well as between years off Iceland (Danielsdottir *et al.* 1991). Furthermore, significant intra-annual deviations from Hardy-Weinberg (H-W) proportion of genotype frequencies were observed in samples of fin whales taken off Iceland. Hence, the results of the genetic analyses undertaken so far indicate mixing of fin

whales from different populations on at least one summer feeding ground (off Iceland).

In this paper, we present the results of a population genetic study based upon the variation in the nucleotide sequence of the mt control region and allele frequencies at six nuclear microsatellite loci in 407 fin whale samples. The samples were collected at six different summer feeding areas in the North Atlantic and Mediterranean Sea. In addition, we included samples, as a reference population, from the Sea of Cortez in the North Pacific, where fin whales are observed throughout the year.

Considering the previous models regarding the population structure of the North Atlantic fin whales proposed by Kellogg, Ingebrigtsen, and Sergeant, we wanted to test the hypothesis that fin whales from different areas in the North Atlantic and Mediterranean Sea correspond to sub-populations between which there are limited amounts of gene flow. Accordingly, this should be evident as a positive correlation between the net inter-population genetic and geographic distances. Our study differs from previous analyses in its extensive geographic coverage as well as the use of hyper-variable nuclear and mt loci in each sample. Parallel analyses of Mendelian and maternally inherited loci are important especially in species such as marine mammals, which may display maternally directed philopatry (Larsen *et al.* 1996; Palumbi & Baker 1994).

Material and Methods

Sample collection

Samples were collected from five summer feeding areas in the North Atlantic Ocean: the Gulf of St. Lawrence, the Gulf of Maine, West Greenland, Iceland and the western coast of Spain. In addition, samples were obtained from fin whales in the Ligurian Sea (the Mediterranean Sea) and the Sea of Cortez (the North Pacific Ocean) (see Figure 3.1). The fin whales in the Sea of Cortez are assumed to represent a resident year-round population that is isolated from populations in the North Pacific (Gambell 1985; Leatherwood *et al.* 1988). All samples, except

those from Iceland and Spain, were obtained as skin biopsies taken from free-ranging whales. The Icelandic and Spanish samples were collected from fin whales taken during whaling operations prior to the international moratorium on commercial whaling. A few of the West Greenland samples were obtained from aboriginal catches.

Skin biopsies from free-ranging fin whales were collected using a crossbow and a bolt with a modified stainless steel tip and a float molded to the bolt (Palsbøll *et al.* 1991). The entire biopsy or, in some instances only the skin section, was conserved in a saturated NaCl solution with 20% dimethylsulphoxide (Amos & Hoelzel 1991). All samples were stored at either -20°C or -80°C pending analysis.

Laboratory analysis

DNA extraction and sex determination

Total-cell DNA was extracted using standard procedures (Sambrook *et al.* 1989) with cell lysis caused by addition of sodium dodecyllauryl sulfate, Proteinase *K* digestion, followed by phenol/chloroform/isoamyl alcohol extractions and finally precipitation with ethanol. The sex of each sample was determined as described in Bérubé & Palsbøll (1996a; 1996b).

Microsatellite loci

Five microsatellite loci (TAA023, GATA028, GATA053, GATA098, GGAA520) with either tri- or tetramer repeat motifs were amplified as described in Palsbøll *et al.* (1997). The letters in the microsatellite names identify the nucleotide sequence of the repeat motif followed by a serial number denoting the clone. The sixth locus, GT011, with a dimer repeat motif, was isolated from total-cell DNA extracted from a humpback whale as well. GT011 was amplified under similar conditions as the other loci using the oligonucleotide primers GT011F 5' CAT TTT GGG TTG GAT CAT TC 3' and GT011R 5' GTG GAG ACC AGG GAT ATT G 3'. The annealing temperature for locus GT011 was set at 62°C. The

amplification products were electrophoresed through a standard 5% denaturing polyacrylamide gel and visualized by autoradiography. The sequence of M13mp18, a blank reaction (all reagents except genomic DNA) as well as five control samples were included with each set of 43 samples to permit accurate sizing of the alleles, detection of contamination as well as loading errors.

Mitochondrial DNA control region sequences

The nucleotide sequence of the 5' end of the mt control region was determined for all samples by direct sequencing (Saiki *et al.* 1988). Symmetric double-stranded and subsequent asymmetric amplifications of the control region were performed as described in Palsbøll (1995). Blank reactions were added to detect possible contamination. Sequencing was performed following the manufacturer's instructions (Sequenase Ver. 2, US Biochemicals Inc.), using Bp15851F (Larsen *et al.* 1996) as sequencing primer.

Data analysis

Sex ratios

We used the χ^2 test, (Lindgren 1975 p296) to evaluate the significance of deviations from parity in the sex ratio within sampling areas.

Analysis of microsatellite loci

LEVEL OF POLYMORPHISM

The level of variation at nuclear loci was estimated as; the number of alleles per locus, the expected heterozygosity, and the probability of identity, *I* (Paetkau & Strobeck 1994).

TESTING FOR DEVIATIONS FROM H-W GENOTYPIC PROPORTIONS AND LINKAGE DISEQUILIBRIUM

Evaluations of possible deviations from the expected H-W genotypic frequencies and linkage disequilibrium were performed using Fisher's exact test and the Markov chain method implemented in GenePop version 1.2 (Raymond & Rousset 1995). Tests for H-W genotypic proportions were conducted for different partitionings stratified by sex, year, sampling area or oceanic origin. To correct for multiple simultaneous comparisons in the tests of H-W genotypic frequencies and linkage disequilibrium, sequential Bonferroni corrections were applied (Rice 1989) using a global significance level of 0.05.

HOMOGENEITY TESTS AND LEVELS OF DIFFERENTIATION

Several statistics were employed based upon either allele frequencies (GenePop Ver. 1.2, Raymond & Rousset 1995), or microsatellite-specific estimators of the variance of allele lengths, R_{ST} (Slatkin 1995) and differences in mean allele length, $(\delta\mu)^2$ (Goldstein *et al.* 1995a). The statistical significance of the observed values of R_{ST} and $(\delta\mu)^2$ was evaluated by comparison with 1,000 Monte Carlo simulations. As reported previously by Valsecchi *et al.* (1997) 'traditional' homogeneity tests based upon allele frequencies proved more powerful than R_{ST} or $(\delta\mu)^2$. Consequently, we employed the homogeneity test implemented in the GenePop version 1.2, which uses a Markov Chain method to estimate the values of Fisher's exact test (Raymond & Rousset 1995). Homogeneity tests between partitionings were performed as pairwise comparisons hierarchically starting within areas comparing; different months, males to females, different years, and finally between sampling areas, for each locus independently. Partitionings between which no heterogeneity was detected were pooled in subsequent tests. The overall level of significance for all six loci was evaluated as suggested by Sokal & Rohlf (1995). Sequential Bonferroni corrections were applied when performing multiple pair-wise comparisons using a table-wide significance level of 0.05 (Rice 1989).

The degree of genetic differentiation between partitionings of samples was estimated either as F_{ST} (Weir 1990) or Nei's D (Nei 1987). Estimates of F_{ST} and

confidence intervals across all six loci were obtained by bootstrapping over loci (Weir 1990).

Analysis of mtDNA

LEVELS OF POLYMORPHISM, HOMOGENEITY TESTS AND LEVELS OF DIFFERENTIATION

The degree of variation within samples was estimated as the nucleotide diversity (Nei 1987 p 256). Homogeneity tests between partitionings were conducted as described by Hudson *et al.* (1992), using the χ^2 statistic (which proved more powerful than either the K_{ST} or the H_{ST}). The level of statistical significance was estimated from 1,000 Monte Carlo simulations as the proportion of simulations in which similar or more extreme values of χ^2 were observed. Pairwise comparisons were performed in the same order and manner as described above for the microsatellite analysis. As in the microsatellite analysis, sequential Bonferroni corrections were applied when performing multiple pairwise comparisons (Rice 1989) using 0.05 as the table-wide level of significance.

The degree of differentiation between partitionings was estimated as either F_{ST} (with frequencies only using GenePop Raymond & Rousset 1995) or Nei's D (Nei 1987).

MTDNA GENEALOGY

Genealogies were estimated from the mt haplotypes employing the PHYLIP3.5c computer package by Felsenstein (1993). The phylogenies were rooted using the homologous sequence from a North Atlantic humpback whale (Palsbøll *et al.* 1995) as an outgroup. First, a phylogeny was estimated using DNAML with multiple jumbles (n= 10), to assess the phylogeny with the highest maximum likelihood value. Second, bootstrap values were obtained by generating 500 random samples (SEQBOOT) for each of which distance matrices were computed (Kimura's 2-parameter model, DNADIST) and a genealogy was estimated using the Neighbor-Joining method (Felsenstein 1993). Finally, a majority-rule consensus genealogy was calculated from the resultant 500 genealogies.

FREQUENCY DISTRIBUTION OF PAIRWISE GENETIC DISTANCES AMONG INDIVIDUALS

The distribution of the number of nucleotide substitutions between all pairs of individuals within sampling localities was computed (MEGA version 1.01, Kumar *et al.* 1993) and compared to the Poisson distribution (Slatkin & Hudson 1991) expected if the population had undergone exponential growth. The agreement of the observed distribution with the expected was evaluated using the χ^2 test (Lindgren 1975).

TIME SINCE DIVERGENCE OF NORTH ATLANTIC AND THE NORTH PACIFIC FIN WHALE POPULATIONS

The level of differentiation between the two oceans was estimated as Kimura's net inter-population genetic distances, d_A (Nei 1987 p 276). The time since divergence (T) was calculated as $T= d_A/2\lambda$ (Nei 1987 p277). We used an overall nucleotide divergence rate (λ) of 0.5-1.0% per Mya suggested by Hoelzel *et al.* (1991) and Baker *et al.* (1993) specifically for cetaceans. It should be noted, however, that this rate is probably an underestimate (see Lyrholm *et al.* 1996), which implies that T was likely to be an over-estimate.

Test of isolation-by-distance

Geographic distances between sampling localities were measured using the Geod computer program. The natural logarithm (ln) of the geographic distances (in kilometers) was plotted against Nei's D (Nei 1987 p220), estimated from the mt (denoted D_{MT}) as well as the microsatellite allele frequencies (denoted D_{MS}). To test the degree of correlation between D_{MT} or D_{MS} and the geographic distances, the Mantel test was used (Mantel 1967). The statistical significance of the observed test statistic was found by comparison with 1,000 random permutations of the geographical distance matrix, with a new test statistic estimated each time. The significance level was defined as the proportion of permutations in which the test statistic was equal to or more extreme than the observed value.

Phylogenetic relationship between areas

Nei's standard genetic distances for the mtDNA and microsatellite loci (denoted D_{MT} and D_{MS}, see above) between sampling areas were employed to estimate a phylogeny of the populations. The phylogeny was estimated using the Fitch-Margoliash program in the PHYLIP 3.5c package (Felsenstein 1993), with global rearrangements and multiple jumbles (n= 10). A majority-rule consensus phylogeny was estimated from 500 bootstrap replicates, either by re-sampling loci (microsatellite only) or haplotypes (e.g., microsatellite alleles or mtDNA haplotypes) within each population.

Results

Sex ratio

Two instances of statistically significant deviations from parity were observed; in the samples collected in the Gulf of St. Lawrence (2.4 males: 1 female; $\chi^2_{(1df)} = 16.33$, $P<0.0001$, n= 98) and off Iceland (0.4 male:1 female, $\chi^2_{(1df)} = 6.82$, $P<0.01$, n= 33) (Table 3.1). While the female-biased sex ratio detected in the Icelandic samples might be due to whalers' preference (females are larger than males), the overrepresentation of males in the Gulf of St. Lawrence was detected only in large groups (Chapter 4).

Microsatellite loci

Level of polymorphism

The total number of alleles per microsatellite locus ranged from eight to nineteen with an average of 10.7. In the North Atlantic/Mediterranean Sea the mean expected heterozygosity was estimated at 0.81 (range: 0.67-0.90) which was significantly higher than the estimate of 0.42 for the Sea of Cortez (range: 0.12-0.68). More than one copy of private alleles (an allele detected only in one area) were found only in the Sea of Cortez (Table 3.2).

The overall probability of identity (across all loci), in the North Atlantic Ocean was estimated at 1.14×10^{-8} yielding an expectation of 4×10^{-4} samples expected to have identical genotypes due to chance alone ({N x N-1}/2 pairwise comparisons among sampled whales (N) multiplied by 1.14×10^{-8}). The expected number of matches in genotype due to chance for the Mediterranean Sea was estimated at 2.7×10^{-5}. Among the North Atlantic and Mediterranean Sea samples a total of 21 pairs as well as a single incidence of three samples had identical sex, nuclear genotype and mt haplotype, all observed within the same sampling area. These were inferred to be duplicate samples collected from the same individual, because of the low expectation of a match by chance. Only in the Gulf of St. Lawrence and in the Gulf of Maine did we detect matches between years. All samples except one with identical genotypes were removed if present in the same partitioning, during the subsequent data analyses. As the expected number of pairs of samples with identical genotypes due to chance in the Sea of Cortez was estimated at 7.2, the observed 10 pairs and two trios with identical genotypes were not necessarily duplicate samples from the same individual. Indeed, in two cases, two samples with identical nuclear genotypes had a different sex or haplotype. As it was impossible to discern true from false matches, only one sample from each match with identical genotype, sex and mt haplotype was kept in the subsequent analysis. Thus, the 407 samples analyzed were inferred to represent 370 individual fin whales (Table 3.1).

Tests of Hardy-Weinberg genotypic proportions and linkage disequilibrium

The combined samples from the North Atlantic and the Mediterranean Sea as well as the combined samples from the North Atlantic only, deviated significantly from H-W genotypic proportions ($\chi^2_{(12df)}$ = infinity, $P < 0.0001$ in both cases). In addition, significant deviations from H-W genotypic proportions were detected in the 1989 Icelandic samples ($n = 24$, $\chi^2_{(12df)} = 30.0$, $P = 0.0028$) and in the Sea of Cortez, among the 1994 samples ($\chi^2_{(12df)} = 23.0$, $P = 0.0274$, $n = 42$). Although not statistically significant after applying the sequential Bonferroni test, deviations were also observed among the samples collected in the Gulf of St. Lawrence

($\chi^2_{(12df)}$= 24.8, P= 0.0156, n= 97) and off Iceland ($\chi^2_{(12df)}$= 25.8, P= 0.0116, n= 33). In the Sea of Cortez, no statistically significant deviations from H-W genotypic proportions were detected after exclusion of the samples with the two rare mt haplotypes (Bp50 and Bp51, Figure 3.2). The above deviations from the H-W genotypic proportions were all due to an excess of homozygotes and observed at two loci in the Gulf of St. Lawrence (GT011 and GGAA520) and at a single locus in the Sea of Cortez and Iceland (locus GATA028).

Of 45 tests for linkage disequilibrium, six loci tested in each of the three locations (North Atlantic Ocean, Mediterranean Sea, and Sea of Cortez), only six pairwise comparisons yielded statistically significant P values after application of sequential Bonferroni corrections. No single pair of loci in linkage disequilibrium occurred in more than one of the three locations suggesting the absence of physical linkage between loci.

Homogeneity tests

Homogeneity tests between sampling localities in the North Atlantic and Mediterranean Sea against the Sea of Cortez, yielded significant differences across all loci (range of F_{ST}: 0.234-0.280, Table 3.3). Only the two most northwestern sampling areas in the North Atlantic (Gulf of St. Lawrence and West Greenland) differed significantly in allele frequencies from the Mediterranean Sea (Table 3.3). Estimates of R_{ST} differed from zero only between the North Atlantic/Mediterranean Sea and the Sea of Cortez (range: 0.021-0.069, data not shown). The $(\delta\mu)^2$ estimates ranged from 0.345 to 13.669 with seemingly little correlation to the geographic distances (e.g., Gulf of Maine - Iceland: 4.477 and Gulf of Maine - Sea of Cortez: 3.983) (data not shown).

Mitochondrial DNA

Levels of polymorphism

The first 288 base pairs at the 5'end of the mt control region were sequenced in 402 samples. It was impossible to obtain unambiguous mtDNA sequences for two

Icelandic and three Mediterranean Sea samples. Thirty polymorphic sites were detected of which 29 were transitions and one a transversion. No insertion/deletion events were observed. The 30 polymorphic sites defined 51 unique haplotypes (Table 3.4). The majority of the samples collected in the North Atlantic and the Mediterranean Sea had mt sequences of haplotype number Bp02, Bp03, Bp04 or Bp08, accounting for 23 (6.3%), 104 (28.4%), 20 (5.4%) and 24 (6.6%) individuals, respectively. The mt haplotypes number Bp03 and Bp08 were found in all North Atlantic/Mediterranean Sea sampling localities. Ninety percent of the individuals from the Sea of Cortez were of haplotype number Bp49. In the North Atlantic Ocean private haplotypes (haplotype unique to a specific sampling locality) found in more than one individual were only observed in the Gulf of Maine (haplotype Bp23). In the Mediterranean Sea three private haplotypes were observed whereas all Sea of Cortez haplotypes were private (Figure 3.2).

In the North Atlantic, the overall nucleotide diversity was estimated at 0.0113 (SE= 0.0006). The estimates of the nucleotide diversity at all North Atlantic sampling localities were all within the same range. The nucleotide diversity of 0.0057 (SE= 0.0009) observed in the Mediterranean Sea samples was significantly lower than any of the observed values at North Atlantic sampling localities. The nucleotide diversity of 0.0007 (SE= 0.0002) estimated in the samples from the Sea of Cortez was exceptionally low, and significantly lower than in any other sampling localities.

Homogeneity tests

As noted above, the χ^2 statistic proposed by Hudson *et al.* (1992) detected more incidences of heterogeneity than either the sequence-based statistic K_{ST}, or the frequency-based H_{ST}. No statistically significant levels of heterogeneity were observed between sexes and years within any sampling area. However, significant levels of heterogeneity were detected between the Sea of Cortez and all the North Atlantic/Mediterranean Sea sampling localities. Likewise, statistically significant levels of heterogeneity were detected between the Mediterranean Sea and all North Atlantic samples. Within the North Atlantic, a significant degree of

differentiation was observed between the western (defined as the Gulf of St. Lawrence and the Gulf of Maine) and the eastern (defined as Spain) North Atlantic. Neither Iceland nor West Greenland could be assigned unambiguously to either the eastern or the western North Atlantic "population". The average level of differentiation (estimated as F_{ST}) between the Sea of Cortez and the remaining sampling areas was estimated at 0.51 (SE=0.058). Similarly, the mean degree of divergence between North Atlantic and Mediterranean Sea sampling areas was found to be 0.12 (SE=0.022). Within the North Atlantic, F_{ST} values ranged from zero to 0.036 (Table 3.3).

MtDNA haplotype genealogy

The two genealogies estimated by both the Neighbor-joining (Figure 3.2) and the maximum likelihood method (Figure 3.3), respectively, yielded similar topologies. No correspondence between the topology and geographic origin of the samples was observed except for the three haplotypes found in the Sea of Cortez, which was supported by a bootstrap value of 95%. Surprisingly, two North Atlantic haplotypes (Bp07 and Bp39 representing nine individuals) were clustered (bootstrap value of 62%) with the three Sea of Cortez haplotypes (Figure 3.2 and 3.3).

Frequency distribution of pairwise genetic distances

All observed distributions were significantly different from the expected Poisson distribution ($P<0.001$) (Figure 3.4). Nonetheless, Slatkin and Hudson (1991) found, by computer simulation experiments, that the distribution of pairwise differences almost never will show a statistically rigorous fit to a Poisson distribution even for populations (simulated) that have undergone exponential growth (Slatkin & Hudson 1991). However, uni-modality (as opposed to 2+ modes) of the frequency distribution indicates exponential population growth (Slatkin & Hudson 1991). The distributions plotted in Figure 3.3 reveal a cline from the western to the eastern part of the North Atlantic in which the frequencies

of pairwise differences gradually change from predominantly uni-modal (the Gulf of Maine and the Gulf of St. Lawrence) to multi-modal (Spain).

Divergence time

Kimura's net inter-population distance between the Sea of Cortez and the North Atlantic/Mediterranean Sea sampling localities ranged from 0.021 to 0.027, with a mean of 0.023. Assuming two rates of sequence divergence of 0.5% and 1.0% per Mya (Baker *et al.* 1993; Hoelzel *et al.* 1991), the approximate time of divergence was estimated at 1.15 Mya (range: 1.05-1.35 Mya) and 2.30 Mya (range: 2.10-2.70 Mya), respectively.

Isolation-by-distance

Figure 3.5 shows plots of Nei standard genetic distances (D_{MT} and D_{MS}, see above) against geographic distances (ln of kilometers) between the North Atlantic and Mediterranean Sea sampling localities. The Mantel tests revealed a statistically significant correlation between D_{MT} as well as D_{MS} and geographic distance when the Mediterranean Sea samples were included (PD_{MT}= 0.008, estimated coefficient of correlation (r)= 0.55 and PD_{MS}= 0.031, r= 0.41), and only between D_{MT} and the geographic distance when the Mediterranean Sea was excluded (PD_{MT}= 0.014, r= 0.70).

Phylogenetic relationship between sampling areas

Since the homogeneity tests as well as the genealogy of the mt haplotypes clearly distinguished the Sea of Cortez as separate from the remaining sampling localities, we used the Sea of Cortez as an outgroup for the estimation of the phylogenetic relationship between the sampling areas in the North Atlantic and Mediterranean Sea (Figure 3.6). None of the nodes were supported by bootstrap values above 50% (by re-sampling of nuclear alleles or mt haplotypes within each population). However, the mt as well as the nuclear phylogeny both had similar topologies and were consistent with the results obtained from the mt homogeneity tests.

Bootstrap values for all nodes above 50% were obtained for the nuclear-based phylogeny (when bootstrapping over loci) demonstrating that the topology was independent of the choice of locus.

Discussion

The hypothesis of absence of a definite, common breeding area for the North Atlantic fin whale (Ingebrigtsen 1929) as well as the population model of multiple sub-populations/stocks suggested by Kellogg (1929), are consistent with the results observed in our study. Evidence for the existence of separate sub-populations with limited gene-flow between adjacent populations should come from a positive correlation between the inter-population genetic and geographic distances. This was indeed the outcome of the Mantel tests (page 53). Although a few instances of significant levels of heterogeneity were observed at nuclear loci, we detected higher levels and more incidences of heterogeneity in mtDNA. Such a difference in the level of differentiation between the two genomes is expected either because of the lower divergence rate at nuclear loci relative to mt loci or because of male mediated gene flow among populations. The differences in divergence rates (all other factors being equal) are due to differences in the effective population size for each of the two genomes, which is four times larger for nuclear loci relative to mt loci. Many of the North Atlantic sampling localities were inaccessible to fin whales during the last glaciation some 18 000 years ago, (Eronen & Olander 1990), and were thus presumably colonized relatively recently (in evolutionary terms). From this it follows that the divergence of the North Atlantic populations presumably is in a relatively early stage and hence we would expect a higher degree of differentiation at mt loci relative to nuclear loci.

That some areas (e.g., Gulf of St. Lawrence and Gulf of Maine) only recently became accessible to fin whales, was further supported by the distribution of pairwise differences in the mt sequences among individual fin whales. As is evident from the genealogy (Figure 3.2 and 3.3), the vast majority of mtDNA haplotypes were separated by only a few substitutions. Hence, most lineages coalesce over a short time span, as expected in an exponentially growing

population (Slatkin & Hudson 1991). The distribution of pairwise differences among individual fin whales in the mt sequences in the western North Atlantic areas (Gulf of Maine and Gulf of St. Lawrence) were either uni-modal or nearly so (Figure 3.4), a finding which is consistent with exponential population growth (Slatkin & Hudson 1991). West Greenland and Iceland, which were not covered by the ice sheet but by continuous sea ice (Eronen & Olander 1990), had intermediate distributions. The more southern areas such as Spain and the Mediterranean Sea, which presumably would have been accessible to fin whales during the last glaciation, had pronounced multi-modal distributions of pairwise differences in the mt sequences.

A factor, which has to be considered when interpreting our results is that many cetaceans display maternally directed site-fidelity (e.g., Baker *et al.* 1986; Clapham & Seipt 1991). Previous humpback whale studies have shown that this behavioral trait can influence the distribution of genetic variation (Larsen *et al.* 1996; Palsbøll *et al.* 1995). Maternally directed site-fidelity to the 'summer' areas sampled during this study could generate differentiation at mt loci, even if all sampled whales constitute one panmictic population (Clapham *et al.* 1993a; Katona & Beard 1990). However, the significant deviation from H-W genotypic proportions observed in the combined North Atlantic samples, in the North Atlantic and Mediterranean Sea sample, as well as the significant level of heterogeneity at the nuclear loci between the Gulf of St. Lawrence/West Greenland and the Mediterranean Sea indicate the existence of non-random mating. Hence, a model assuming a panmictic population with maternally directed site fidelity to specific summer feeding areas is not consistent with our findings. Male mediated gene flow could also contribute to such a pattern of differentiation, however, to date there are no observations of sex-specific differences in dispersal for fin whales, as has been observed in, e.g., North Atlantic minke whales (Larsen & Øien 1988).

Even though neither the mt nor the nuclear phylogeny of populations (Figure 3.6) was well supported in terms of bootstrap values, it is noteworthy that the

topologies were congruent, especially when one considers the many possible branching patterns. The observed topology agreed with the results of the mt DNA homogeneity tests as well as the predictions from Kellogg's population model.

The level of divergence (estimated as Nei's D) at the nuclear loci observed in this study was similar to those reported by Danielsdottir and co-workers (Danielsdottir *et al.* 1991; Danielsdottir *et al.* 1993) who detected statistically significant differences in isozyme allele frequencies between the eastern Canadian coast, Iceland, Spain, and Norway as well as among years in Iceland. The average genetic distance among the North Atlantic sampling localities was estimated to be 0.048 (SE= 0.0114) in our study, whereas Danielsdottir and co-workers reported a value of 0.013 between Iceland and Spain and 0.060 between Norway and eastern Canada. However, the tests employed in the current study did not detect statistically significant differences in allele frequencies between most areas. This discrepancy in the significance of the results in the two studies may have a statistical rather than a biological origin, e.g., a result of the higher number of alleles detected in the microsatellite analysis relative to studies using isozymes, resulting in a reduced discriminatory power when comparing allele frequencies between areas.

The population structure of North Atlantic fin whales indicated by this study and previous work (Arnason *et al.* 1991a; Danielsdottir *et al.* 1992; Danielsdottir *et al.* 1991; Ingebrigtsen 1929; Kellogg 1929; Sergeant 1977) suggests that fin whale populations may be structured differently from those of humpback whales and northern right whales. Indeed, our results suggest sub-structure over even relatively short distances, and that different sub-populations may use the same feeding area, even within the same year as originally proposed by Kellogg (1929).

The levels of variation in the mtDNA estimated in our study (nucleotide diversity) correlated well with the current estimates of abundance. The latest estimate of abundance for the entire North Atlantic Ocean was 56,000 (Buckland *et al.* 1992a; Buckland *et al.* 1992b; IWC 1992), 3583 individuals (95% confidence interval: 2130-6027, in the Mediterranean Sea (Forcada *et al.* 1996),

and 297 in the Sea of Cortez (95% confidence interval: 217-376, Urban-R. 1996). The nucleotide diversity in the North Atlantic samples was estimated to be 0.0113, which was significantly higher than the estimates of 0.0057 and 0.0007 for the Mediterranean Sea and Sea of Cortez, respectively.

The nucleotide diversity estimated for the samples collected in the Sea of Cortez was among the lowest reported for cetaceans so far (Medrano-G. *et al.* 1995; Palsbøll *et al.* 1997b; Rosel *et al.* 1995). The observed nucleotide diversity was even lower than that of the highly endangered North Atlantic right whale (0.0026 in Schaeff *et al.* 1993).

The significant level of divergence observed at the mt and nuclear loci between the Sea of Cortez and the North Atlantic/Mediterranean Sea is not surprising. Such differences have been reported for the minke whale (*B. acutorostrata*) (Bakke *et al.* 1996; Van Pijlen *et al.* 1995) and the humpback whale (Baker *et al.* 1990; Baker *et al.* 1993; Baker *et al.* 1994; Valsecchi *et al.* 1997). However, the mt haplotypes unique to the Sea of Cortez clade were separated by only a few nucleotide substitutions from those found in the North Atlantic/Mediterranean Sea. In addition, two of the North Atlantic haplotypes were in the same well-supported clade as those from the Sea of Cortez. These findings were unexpected considering the large geographic distance and physical barriers separating these two oceans. The time of divergence was estimated at 1.2-2.3 Mya (and likely to be an overestimate of the divergence time, see Material and Methods), which is more recent than the rise of the Panama Isthmus some 3 Mya ago (Savage 1983). The most likely explanation for the unexpectedly low level of divergence is occasional gene-flow between oceans as observed in humpback whales (Baker *et al.* 1990; Baker *et al.* 1993; Baker *et al.* 1994; Palsbøll *et al.* 1995; Valsecchi *et al.* 1997). However, the clear separation of the haplotypes found in the Sea of Cortez from the remainder indicates that such gene-flow has not occurred recently (on an evolutionary time scale).

The fin whales in the Sea of Cortez probably constitute a year-round resident population (Leatherwood *et al.* 1988). However, the abundance of fin whales

increases during the winter and spring, suggesting that fin whales from the Pacific Ocean may visit the Sea of Cortez during part of the year (Tershy *et al.* 1990; Tershy *et al.* 1993). Consistent with this hypothesis, we detected a significant deviation from H-W genotypic proportions among the Sea of Cortez samples. If only individuals with the most common mt haplotype in the Sea of Cortez (Bp49, 90% of all the individuals) were considered, no such deviation from H-W genotypic proportions was detected. Hence, fin whales, which had one of the two rare mt haplotypes could represent migrants from the Pacific Ocean. Nonetheless, the two rare mt haplotypes (Bp50, Bp51) differ from the most common mt haplotype (Bp49) by only one substitution suggesting that the three mt haplotypes are very closely related. Since no information is available concerning the North Pacific Ocean fin whales in general, it is not possible to propose any definitive hypotheses.

As in the Sea of Cortez, the nucleotide diversity among the Mediterranean Sea samples was significantly lower than those estimated at North Atlantic sampling areas. It is unclear whether the fin whales observed in the Ligurian Sea during the summer are year-round residents of the Mediterranean that may winter in the southern part of this Sea (Duguy 1989) or winter in the eastern part of the North Atlantic. The notion of a year-round resident population was supported by the significant level of differentiation in the mtDNA between the Mediterranean Sea and Spain (Table 3.3). A similar degree of isolation between the North Atlantic and Mediterranean Sea has also been reported in the striped dolphin (*Stenella coeruleoalba*) (Archer 1996), a species with a migratory potential similar to fin whales. The lack of fin whale sightings in the Strait of Gibraltar (Duguy *et al.* 1988) and the observation of newborns during the summer in the Ligurian Sea (Cagnolaro *et al.* 1986; Notarbartolo di Sciara *et al.* 1996), also suggest the absence of major semi-annual migration of Mediterranean fin whales into the North Atlantic. In addition, blubber concentrations of persistent organochlorine pollutants are much higher in Mediterranean fin whales than in those from the Atlantic North-West coast of Spain (Marsili & Focardi 1996). However, no

significant deviations from H-W genotypic proportions were observed in the combined Mediterranean Sea and Spanish samples, nor did we detect any significant levels of heterogeneity at the nuclear loci. Therefore, with the current sample sizes, our study is consistent with the idea of a separate fin whale population in the Mediterranean; however, more data are needed to provide conclusive proof of separation.

In summary, the results from our North Atlantic fin whale study suggest "isolation by distance", which were in general agreement with the population model proposed in 1929 by Kellogg. He suggested that fin whales were subdivided into several 'stocks' with limited but sometimes overlapping ranges. Clear demonstration, from genetic data, of such fine-scaled structure will require a comprehensive collection and analysis of samples well represented in time and space. In addition, the level and distribution of the variation at the mt loci indicated recent population expansion in the western North Atlantic. Our study also indicated that hyper-variable microsatellite loci with many alleles may not be the optimal loci for population analyses, and that such studies will probably gain in statistical power if microsatellite loci of intermediate variability are analyzed.

Tables

Table 3.1. Number of samples (N) and individuals (n) per year and sex ratios from the Gulf of St. Lawrence (GSL), the Gulf of Maine (GM), West Greenland (WG), Iceland (IL), Spain (SP), the Ligurian Sea (IT), and the Sea of Cortez (SC) fin whales.

		Number of individuals (n) Years											
Areas	Sample sizes N	82	84	87	88	89	90	91	92	93	94	Sex ratio ♂/♀	Total n
GSL	109						10	29	40	14	5	**69/29**	98
WG	46				5		5	6	5	7	11	19/20	39
IL	33			9		24						**9/24**	33
SP	39	25	14									12/13 ¤	39
GM	31							20	8			17/11	28
IT	74							1	16	21	34	37/33 *	72
SC	75									19	42	38/23	61

Note. The bold values indicate significant difference between the proportion of males and females. '¤' means that only the 1982 samples from Spain have been considered for the sex ratio as the Spanish organization gave only females from the 1984 commercial whaling. *The ZFX-ZFY locus did not amplify in two individuals.

Table 3.2. Polymorphism at every locus for all sampling areas.

Locus-Allele size	GSL*	GM	WG	IL*	SP	IT	SC*
n	*97*	*28*	*39*	*33*	*39*	*65*	*61*
GATA028							
143	14	1	3	7	7	10	0
148	0	1	1	0	1	1	0
151	2	0	3	0	2	3	0
156	12	4	3	4	4	2	0
160	23	5	10	9	6	12	19
164	11	4	4	5	4	11	6
167	1	1	4	0	0	0	0
168	5	0	0	4	2	4	6
171	8	2	0	1	4	7	13
172	4	0	0	0	0	0	1
173	6	1	8	2	7	13	1
175	22	9	7	10	10	18	9
179	36	19	17	16	17	25	65
181	0	0	1	2	0	0	0
183	28	1	14	5	9	13	2
186	2	4	0	0	0	0	0
187	20	4	5	1	4	3	0
191	0	0	0	0	1	1	0
195	0	0	0	0	0	5	0
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
GT011							
67	19	6	10	4	8	17	0
69	18	4	5	9	6	12	0
71	0	0	0	2	0	0	3
73	20	6	7	7	11	11	2
75	29	9	14	12	11	20	0
77	29	10	18	9	18	29	76
79	55	16	18	16	20	28	38
81	23	5	8	7	4	11	0
83	1	0	0	0	0	0	2
91	0	0	0	0	0	0	1
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

Locus-Allele size	GSL*	GM	WG	IL*	SP	IT	SC*
n	*97*	*28*	*39*	*33*	*39*	*65*	*61*
GATA053							
139	0	1	0	0	0	0	0
143	18	0	7	3	4	2	0
144	47	18	23	26	28	57	0
148	6	1	2	4	6	3	1
156	12	1	1	1	2	1	0
158	18	7	1	4	2	8	0
159	0	0	0	0	1	0	115
160	33	10	26	14	14	20	2
164	31	11	13	10	9	17	1
167	11	2	3	1	6	8	0
168	17	5	4	3	5	11	0
175	1	0	0	0	0	1	0
176	0	0	0	0	1	0	3
A	10	9	9	9	11	10	5
PIC	0.74	0.75	0.72	0.67	0.70	0.75	0.52
I	0.04	0.07	0.08	0.08	0.06	0.09	0.79
TAA023							
41	59	14	20	22	22	31	1
44	1	1	1	0	1	0	0
47	9	1	35	3	3	6	0
50	73	22	9	29	33	49	67
53	22	6	11	2	2	9	51
56	21	11	1	9	15	24	3
59	9	1	3	1	2	9	0
A	7	7	7	6	7	6	4
PIC	0.85	0.80	0.77	0.76	0.81	0.74	0.42
I	0.11	0.11	0.12	0.16	0.11	0.10	0.33
GATA098							
63	6	0	0	0	0	0	0
67	46	17	18	13	21	27	0
71	34	12	16	14	14	27	79
75	55	13	27	21	19	24	0
79	34	8	16	11	17	23	43
83	13	4	2	5	4	20	0
87	5	2	1	2	0	6	0
91	1	0	0	0	3	1	0
A	6	5	5	5	6	7	2
PIC	0.77	0.73	0.75	0.77	0.77	0.76	0.40
I	0.07	0.09	0.10	0.09	0.08	0.09	0.40

Locus-Allele size	GSL*	GM	WG	IL*	SP	IT	SC*
n	*97*	*28*	*39*	*33*	*39*	*65*	*61*
GGAA520							
145	0	0	0	0	0	0	1
149	0	0	0	0	0	0	3
153	9	0	1	1	1	0	0
155	5	4	4	3	3	4	0
157	8	0	2	3	2	11	0
159	14	5	7	6	5	10	0
161	25	6	6	11	16	11	0
163	5	2	5	3	6	4	0
165	25	11	9	9	12	23	2
167	3	0	1	2	0	5	0
169	36	8	15	9	7	13	105
171	3	1	3	1	1	8	0
173	25	11	13	9	10	14	11
177	25	5	8	3	11	12	0
181	9	2	5	4	3	10	0
183	0	0	1	0	0	0	0
185	2	1	0	2	1	3	0
A	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
H	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
I	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
All loci							
H_e	0.83	0.81	0.81	0.80	0.81	0.82	0.42
I	7.32×10^{-9}	4.70×10^{-8}	3.15×10^{-8}	3.04×10^{-8}	2.17×10^{-8}	1.00×10^{-8}	2.60×10^{-3}

Note. The abbreviations for the areas are described in Table 3.1. "A" indicates the number of alleles, "H_e" the expected heterozygosity and "I" the probability of identity. Seven individuals from the IT and one from the GSL did not amplify successfully at all loci, therefore, they were not included in the microsatellite analysis. The star (*) denotes the population that deviates from the Hardy-Weinberg genotypic proportions.

Table 3.3. Degree of differentiation (F_{ST}) between areas and their significance for both microsatellites (above diagonal), and mtDNA (below diagonal) allele frequencies.

Areas	GSL	GM	WG	IL	SP	IT	SC
GSL		_	_	_	0.001	**0.007**	**0.234**
GM	_		_	_	_	_	**0.269**
WG	_	0.002		_	_	**0.006**	**0.258**
IL	_	_	_		_	_	**0.280**
SP	**0.033**	**0.036**	0.015	_		_	**0.270**
IT	**0.119**	**0.150**	**0.116**	**0.094**	**0.097**		**0.254**
SC	**0.422**	**0.516**	**0.486**	**0.509**	**0.519**	**0.601**	

Note. The abbreviations for the areas are described in Table 3.1. The F_{ST} values were estimated using GenePop (Raymond & Rousset 1995) for the microsatellite and the mtDNA alleles (haplotypes) frequencies. Bold values are significant at $P<0.001$ level.

Table 3.4. Different mtDNA haplotypes and their segregating sites for the 365 individuals.

MtDNA Hap#	13	18	25	27	56	67	78	86	87	96	103	104	132	140	162	178	180	195	200	213	222	237	253	255	258	263	276	280	282	283
Bp01	*G*	*A*	*C*	*G*	*T*	*A*	*T*	*C*	*T*	*C*	*T*	*T*	*G*	*C*	*T*	*T*	*T*	*G*	*T*	*T*	*T*	*A*	*A*	*A*	*G*	*T*	*C*	*A*	*C*	*T*
Bp02	.	.	.	A	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Bp03	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Bp04	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	T	.
Bp05	.	.	.	.	.	.	.	T	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	T	.
Bp06	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	T	G	T	.
Bp07	.	.	.	.	.	G	.	.	.	T	C	.	.	.	.	C	.	A	C	.	.	.	G	.	.	.	.	G	T	.
Bp08	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	T	G	T	.
Bp09	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	T	.
Bp10	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.
Bp11	.	.	.	.	.	.	.	T	C	.	.	.	.	.	.	.	C	.	.	.	.	.	.	A	.	T	G	T	.
Bp12	.	.	.	.	.	.	.	.	.	T	.	C	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.
Bp13	.	.	.	.	.	.	.	.	C	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	T	.
Bp14	.	.	.	.	.	.	.	T	C	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	T	G	T	.
Bp15	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.
Bp16	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	C	.	.	.	.	.	.	.	T	G	T	.
Bp17	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	T	C
Bp18	.	.	.	A	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.
Bp19	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	T	G	T	.
Bp20	.	.	.	.	.	.	.	.	.	T	.	C	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	G	.	.
Bp21	A	.	.	A	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Bp22	.	.	.	.	.	.	.	.	C	T	.	.	.	.	.	C	.	.	.	.	.	.	G	.	.	.	.	.	T	.
Bp23	.	.	.	.	.	.	.	T	.	T	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	T	G	T	.
Bp24	.	.	.	A	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.
Bp25	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.
Bp26	.	.	.	.	.	.	.	T	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Bp27	.	.	.	.	.	.	.	T	.	T	.	C	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.
Bp28	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	G	.	.
Bp29	.	.	.	.	.	.	.	T	.	T	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	T	.
Bp30	.	G	.	.	.	.	.	.	.	T	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	T	G	T	.
Bp31	.	.	.	.	C	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Bp32	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.
Bp33	.	.	.	.	.	.	.	.	.	T	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C
Bp34	.	.	.	A	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.
Bp35	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	C	.	.	.	.	.	.	G	.	.	.	.	.	T	.
Bp36	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C
Bp37	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	G	T	.
Bp38	.	.	.	A	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Bp39	.	.	.	.	.	G	.	T	.	T	C	.	.	.	.	C	.	A	C	.	.	G	.	.	.	.	.	G	T	.
Bp40	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	T	G	T	.
Bp41	.	.	.	.	.	.	.	T	C	T	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	G	T	.
Bp42	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	T	G	T	.
Bp43	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	G	T	C
Bp44	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	G	T	.
Bp45	.	.	.	.	.	.	.	T	C	T	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	T	G	T	.
Bp46	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.
Bp47	.	.	T	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Bp48	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.
Bp49	.	.	.	.	.	.	.	T	.	T	.	.	.	T	.	C	.	A	C	.	.	T	.	.	.	.	.	G	T	.
Bp50	.	.	.	.	.	.	.	T	.	T	.	.	.	T	C	C	.	A	C	.	.	T	.	.	.	.	.	G	T	.
Bp51	.	.	.	.	.	.	.	T	.	T	.	.	.	T	.	C	.	A	C	C	.	T	.	.	.	.	.	G	T	.
M. nov.	.	.	.	.	.	.	.	T	.	T	.	A	.	-	.	.	.	A	.	.	.	.	.	.	A	.	.	G	.	C

Note. The dots indicate the same bases and the positions of the segregating sites are indicated at the top of the table. Only the segregating sites present within the fin whales sequences are compared with the humpback whale (*M. nov.*).

Figures

Figure 3.1. Sampling areas.

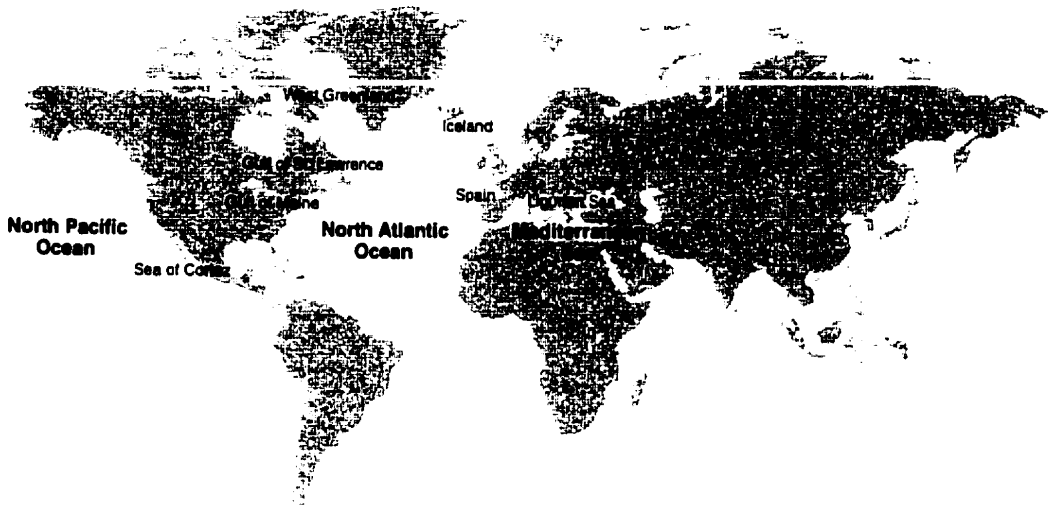
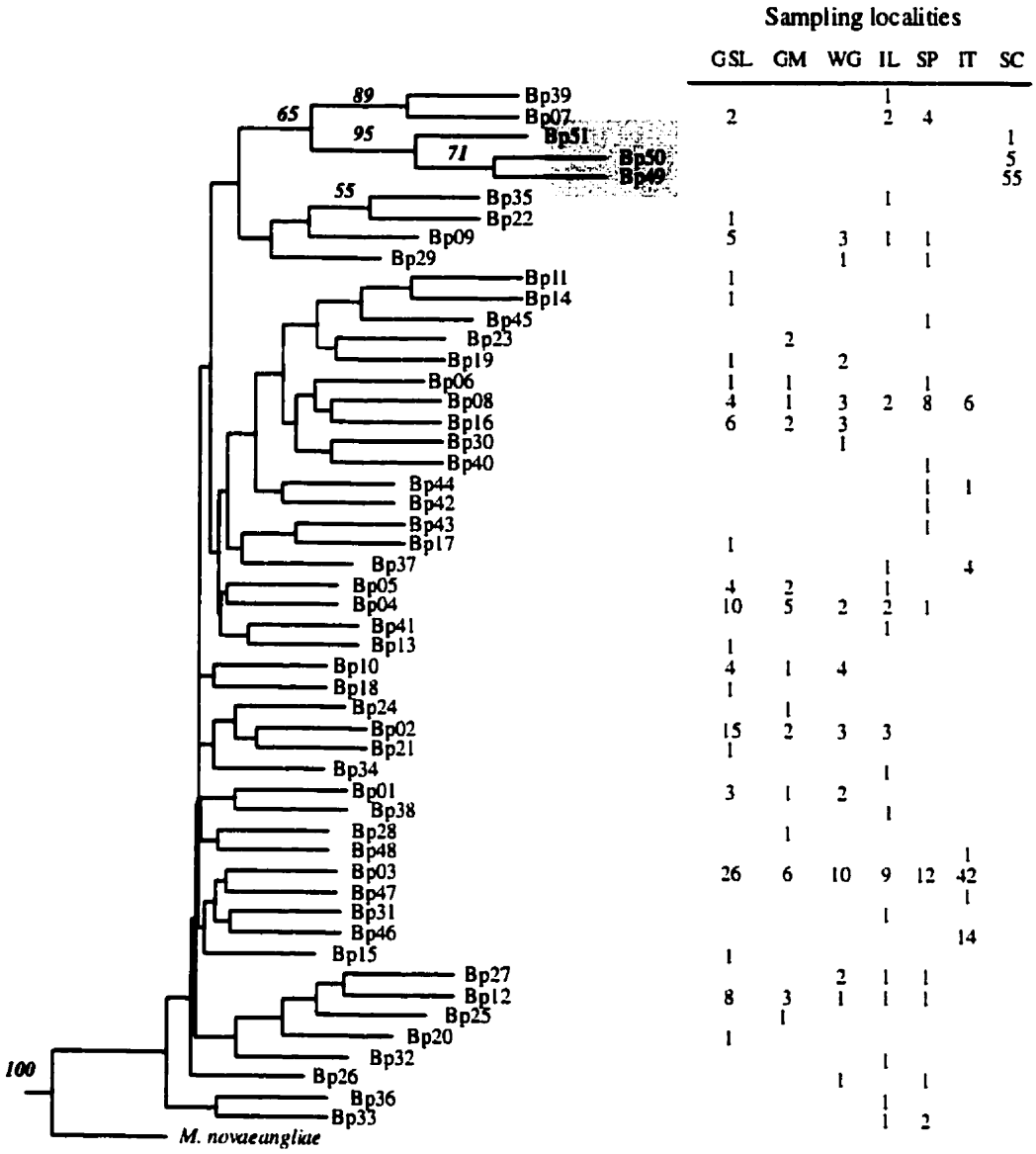


Figure 3.2. Majority rule consensus genealogy estimated from the mt haplotypes.

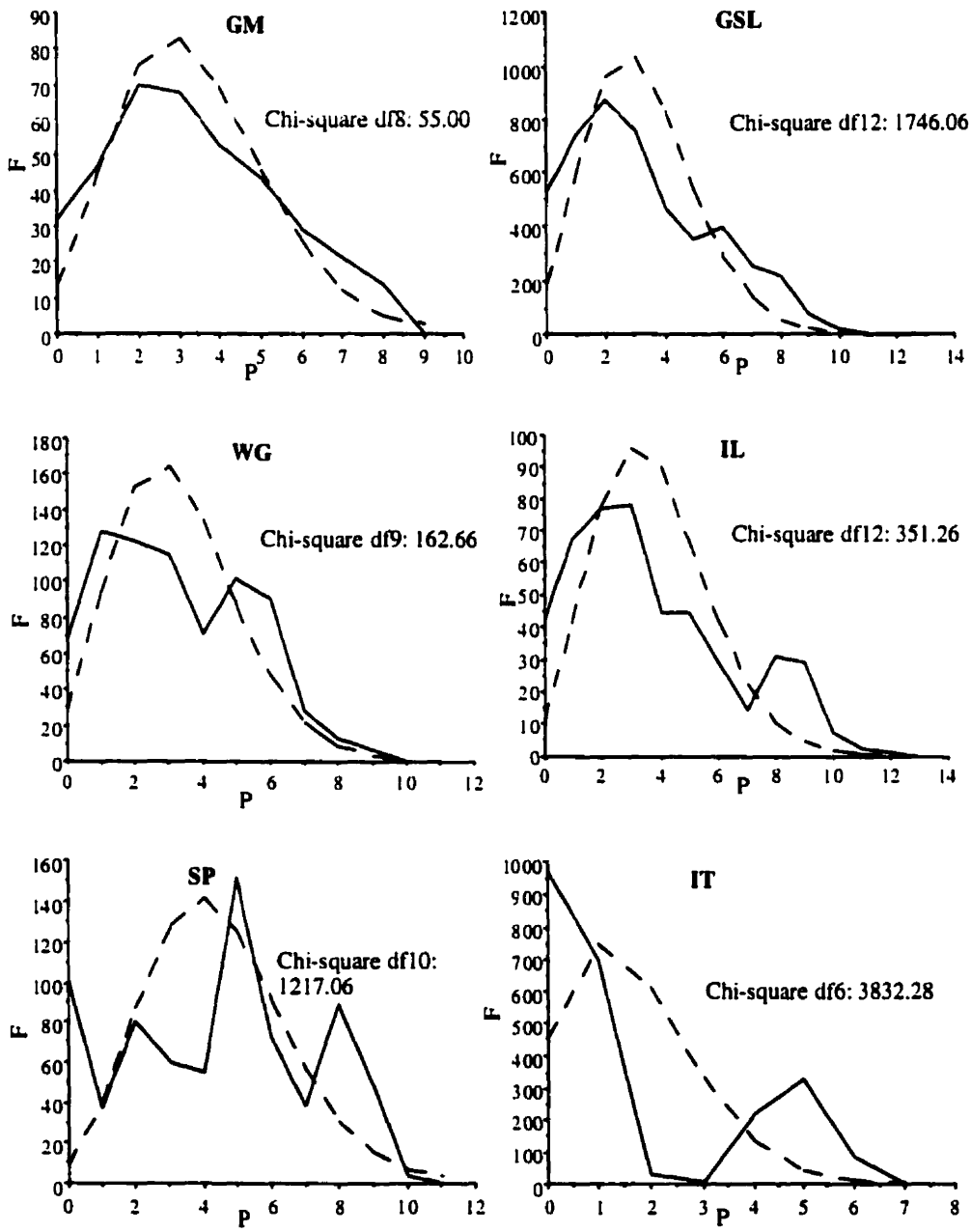


Note. The number at the branches denotes the mt haplotype number. Only bootstrap values above 50% are shown. *M. novaeangliae*: humpback whale.

Figure 3.3. Genealogy with the highest maximum likelihood value

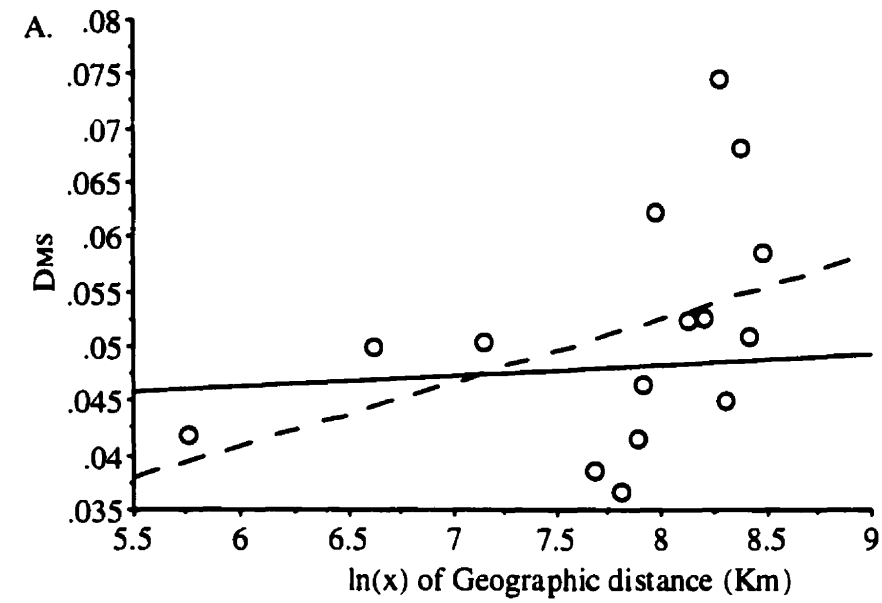
Note. The numbers at the end of the nodes are the haplotype number. *M. novaeangliae*: humpback whale.

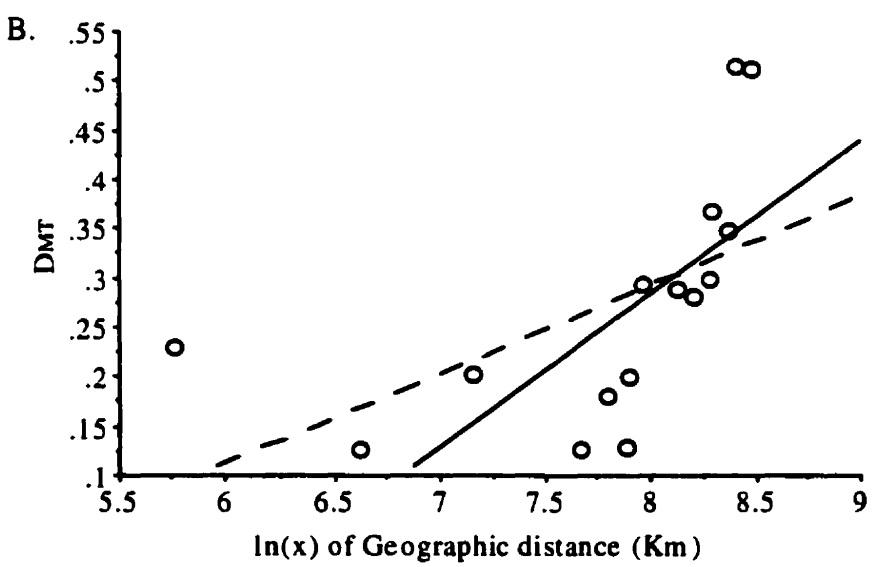
Figure 3.4. Observed distributions of pairwise difference (plain line) in base pair substitutions compared with the number that would be expected assuming a Poisson distribution (dashed line) for each of the North Atlantic areas and the Mediterranean Sea.



Note. The x-axis (P) is the number of substitutions, the y-axis (F), the frequencies of the pairwise comparisons. The abbreviations for the areas are described in Table 3.1.

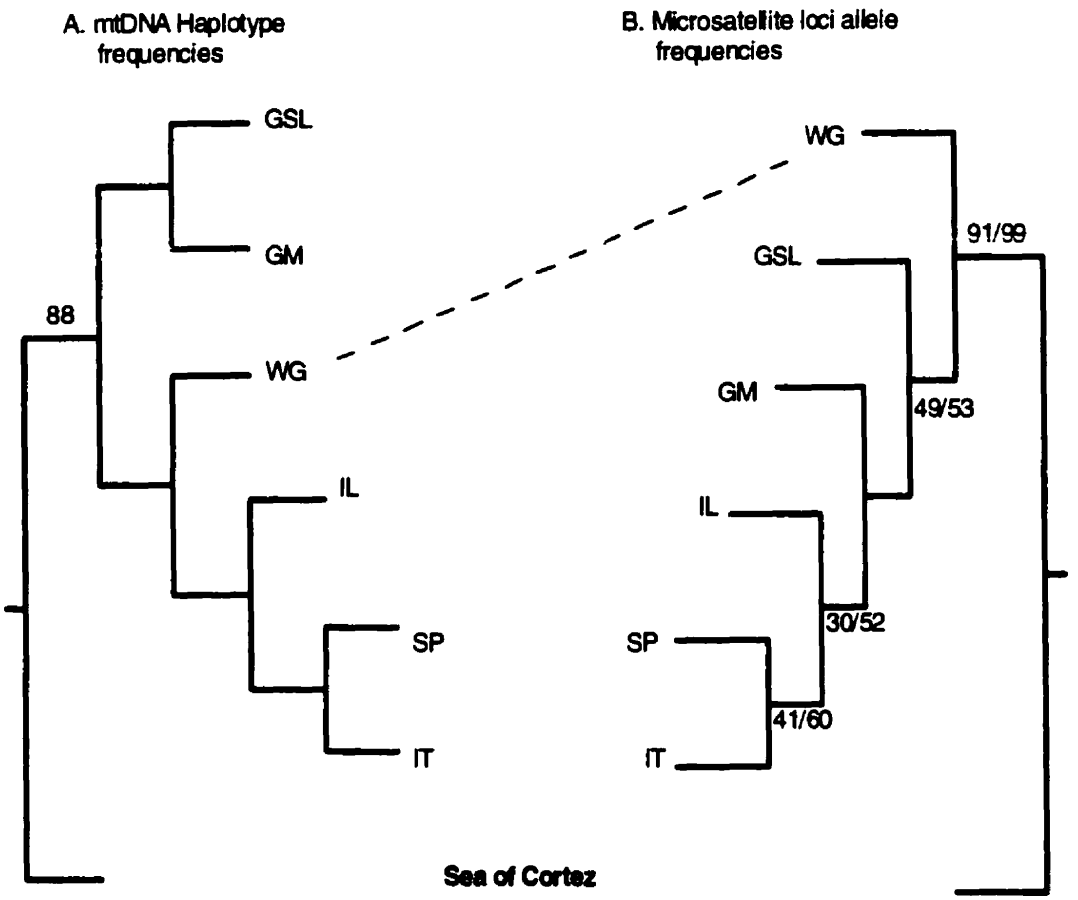
Figure 3.5. Plot of Nei's standard genetic distances calculated for (A) the microsatellite allele (D_{MS}) and (B) the mtDNA haplotype frequencies (D_{MT}) against the natural logarithm (ln) of the geographic distance (Km) between the North Atlantic sampling localities.





Note. The regression was performed between the geographic distance (see Materials and Methods) and all the North Atlantic and Mediterranean Sea localities (dashed regression line) as well as between the North Atlantic localities exclusively (plain regression line).

Figure 3.6. Majority rule consensus tree estimated from Nei genetic distances calculated between all localities.



Note. The sampling locality abbreviations are described in Table 3.1. Distances estimated from, A) the mtDNA haplotype frequencies (D_{MT}) and B) the six microsatellite loci's allele frequencies (D_{MS}). Bootstrap values (500 replicates) at the nodes were calculated by re-sampling alleles (first value in A and B) and then loci (B) within populations. The abbreviations for the areas are described in Table 3.1.

CHAPTER 4. MALE-BIASED SEX RATIO IN LARGE GROUPS OF THE GULF OF ST. LAWRENCE FIN WHALES *(BALAENOPTERA PHYSALUS)*

Martine Bérubé[¤*], Richard Sears[§], and Catherine Berchok[§]

¤ Population Biology Department, Zoological Institute, University of Copenhagen, Universitetsparken 15, DK-2100 Copenhagen, Denmark and Natural Resource Sciences Department, Macdonald Campus of McGill University, 21,111 Lakeshore Road, Ste-Anne de Bellevue, Quebec H9X 3V9, Canada.
§ Mingan Island Cetacean Study Inc., 285 Green St., Saint-Lambert, J4P 1T3 Québec, Canada.

*Corresponding address: Unit of Evolutionary Genetics, Dept of Molecular Biology, Free University of Brussels (ULB), CP 244, Bld du Triomphe, B-1050 Brussels, Belgium. Tel: +32 2 650 5427. Fax: +32 2 650 5421. Email: mberube@ulb.ac.be

Short title: Male biased sex ratio in fin whales.

Introduction

Fin whales (*Balaenoptera physalus*) are among the most common mysticetes (baleen whales) in the North Atlantic Ocean. Despite their high abundance and previous importance in commercial whaling operations, little is known about their migration and behaviour. The concentration of fin whales in the Gulf of St. Lawrence peaks during the summer, which constitutes the feeding season that extends from early June to late September (Sergeant 1977; Edds & Macfarlane 1987; Bérubé & Sears 1990;). During this period, fin whales are observed primarily along steep slope contours where biological productivity is high (Sergeant 1977).

Most information on the sex of fin whales originates from whaling data; this has permitted an assessment of the sex ratio in many areas, including the Gulf of

Alaska and the Bering Sea, West Greenland, and the northwestern and northeastern North Atlantic (Aguilar & Lockyer 1987; Kapel 1979; Mitchell 1974; Tarasevich 1967). For all of these regions, the sex ratio estimated from the catch data did not deviate significantly from parity.

Sex determination using photographs of sexually dimorphic characters is possible for a few cetacean species only, such as humpback whale (*Megaptera novaeangliae*) (Glockner 1983), killer whale (*Orcinus orca*) (Bigg *et al.* 1987), and right whale (*Eubalaena glacialis*) (Hamilton & Mayo 1990) but no such readily observable dimorphic characters have yet been identified in fin whales. The only means of identifying the sex of a fin whale in the field is if an adult is closely associated with a calf and presumed to be a female. Because of the difficulty in determining the sex of free-ranging whales, methods of molecular sexing using skin biopsies collected from free-ranging whales have become increasingly important for the study of cetaceans (Baker *et al.* 1991; Brown *et al.* 1991; Palsbøll *et al.* 1992; Medrano-G. *et al.* 1994; Richard *et al.* 1994; Brown *et al.* 1995; Bérubé & Palsbøll 1996b).

Although fin whales are characterized as a migratory species, they do not seem to congregate on a single breeding ground as is the case for e.g., North Atlantic humpback whales (Katona & Beard 1990; Clapham *et al.* 1993a). On the contrary, fin whales in the North Atlantic are believed to be sub-divided into several sub-populations (Bérubé *et al.* 1998). In all North Atlantic areas where fin whales have been studied so far (Tarasevich 1967; Mitchell 1974; Kapel 1979; Aguilar & Lockyer 1987; Bérubé *et al.* 1998), an even sex ratio has been observed except for two cases; in a sample of 33 Icelandic fin whales and among 109 skin biopsies collected from free-ranging fin whales in the Gulf of St. Lawrence samples (Bérubé *et al.* 1998). In this study, we report the findings from our investigation of the observed male-biased sex ratio in the Gulf of St. Lawrence. We analyzed the distribution of sex across three categories of groups: single individuals, pairs, as well as three individuals and more. A significant excess of males was detected only in large groups of fin whales and it is suggested that such

large groups could be equivalent to the competitive male groups known in humpback and northern right whales (Mattila *et al.* 1989; Stone *et al.* 1988).

Materials and Methods

Skin biopsies were collected off the northern shore of the Gulf of St. Lawrence between 1990 and 1994 in the period from June to November. The primary study area was the Mingan Island region (50°15'N, 64°10'W). However, a few samples were collected in the estuary between the Saguenay river (48°10'N, 69°45'W) and Sept-Iles (50°13'N, 66°22'W) as well as in the Strait of Belle Isle, (51°57'N, 55°25'W). The biopsy equipment consisted of a crossbow and a bolt with a modified stainless steel tip (Palsbøll *et al.* 1991). Attempts were made to biopsy all members of a group, whenever possible. The skin part of the biopsy was preserved in saturated NaCl with 20% DMSO (Amos & Hoelzel 1991) and stored at -20°C until DNA extraction. All sampled individuals and incompletely sampled individuals group were photographed and information including date, position (longitude and latitude) and group size were noted. A 'group' was defined as two or more whales swimming side by side within a body length, with coordinated surfacing, and diving pattern during the time of the observation. A mother/calf pair was defined as a group of two individuals of which one was estimated to be equal or less than half the body length of the accompanying adult.

Total-cell DNA was extracted from the skin biopsies using standard protocols (Sambrook *et al.* 1989). Sex was determined for all individuals as described by Bérubé and Palsbøll (1996a; 1996b). The genotype of six microsatellite loci (one dimer, one trimer and four tetramers) was determined for each sample as reported previously (Bérubé *et al.* 1998; Palsbøll *et al.* 1997).

We used a log-likelihood ratio test (Lindgren 1975 p296) to test the goodness of fit of the proportion of males to females against the 1:1 ratio observed in other areas. Various partitionings of the samples were tested, such as between years, seasons, and groups. Linkage between the microsatellite loci was tested using an implementation in the GenePop computer package version 1.2 (Raymond & Rousset 1995). The probability of identity at all analyzed loci for two unrelated

individuals by chance was estimated as the product of the allele frequencies, $I = \sum p_i^4 + \sum\sum\left(2p_i p_j\right)^2$, as suggested by Paetkau and Strobeck (1994).

Population was estimated using the Bailey's binomial model, $\hat{N} = n_1(n_2+1)/(m_2+1)$, with its standard deviation, $SD(\hat{N}) = \sqrt{v(\hat{N})}$ where

$v(\hat{N})$ is estimated by: $v(\hat{N}) = \dfrac{n_1^2(n_2+1)(n_2-m_2)}{(m_2+1)^2(m_2+2)}$

where n_1 is the number of samples in the first "tagging experiment" and n_2 is the number of samples in the second tagging, m_2 are the number of biopsy samples with the same genotypes in the first and second tagging. This estimate assumes that the proportion of biopsied to non-biopsied animals is constant during the experiment and that all animals within the sampling area have an equal probability of being biopsied (Seber 1982 p.61-64).

Results and Discussion

A total of 109 samples was collected and analyzed. Exact tests for linkage disequilibrium between microsatellite loci detected no significant incidences of linkage between loci. The probability of identity across all loci combined was estimated at 7.32×10^{-9}, thus the probability that two unrelated individuals have the same genotype across all loci is one in 140 million. Consequently, any two samples with identical genotypes across all loci were inferred to represent re-captures of the same individual. Among the 109 biopsies, eleven pairs of samples had identical genotypes. Only four of these re-captures were between different years, the remainder of re-captures were observed within the same year. The low recapture rate in itself suggests that we probably were sampling from a relatively large population. The estimates of abundance between years based upon these eleven re-captures yielded numbers ranging from 49 (SD= 23.9) to 1452 (SD= 825.5) individuals. The estimate obtained from the two years where most samples were collected (1991, n_1= 33 and 1992 n_2= 45, and m_2= 1), was 759 (SD= 183.67) fin whales.

Of the 98 individual fin whales sexed, 69 (70%) were male and 29 (30%) female. This ratio of 2.4 males for each female differs significantly from parity (χ^2_{1df}=16.33, P< 0.0001). The high number of males relative to females was observed in all years and months (range from 1.67 to 3.57 males for 1 female), except in 1993 (6 males and 6 females).

Brown *et al.* (1995) found a similar excess of males (male:female ratio of 2.4: 1) in humpback whale samples collected along the East Coast of Australia. However, these samples were collected in a 'migratory corridor', where whales were in transit between feeding and breeding grounds. The authors suggested that non-reproducing females might not migrate to the breeding grounds every year but remain on the feeding grounds during the winter. Apart from the study by Brown *et al.* (1995), three additional studies of humpback whales suggest an even overall sex ratio: from five feeding areas in the North Atlantic (Clapham *et al.* 1995; Palsbøll *et al.* 1995) and from the Hawaiian breeding range (a study based only on calves, Glockner-Ferrari & Ferrari 1990).

As mentioned above, the only estimates of sex ratio among fin whales reported so far originate from whaling data (Aguilar & Lockyer 1987; Kapel 1979; Mitchell 1974; Tarasevich 1967). It has been suggested that samples collected from commercial whaling operations could be biased toward males due to catch restrictions (e.g., on lactating females, Holt 1977) or females (as hunters might aim for larger individuals, which tend to be females). Despite these possible sources of bias, reported estimates of the sex ratio in fin whales calculated from whaling logbook data have yielded a 1:1 ratio of males to females. The collection of biopsies from free-ranging fin whales might be subject to bias as well, e.g., individuals of one sex are more approachable than the other. Hence, if females were shyer than males, this could explain the observed bias. However, with the data obtained from two other summer feeding grounds where samples have been collected as skin biopsies as well (the Gulf of Maine (n= 28) and the Ligurian Sea (n= 72), Bérubé *et al.* 1998), revealed no such male bias among samples (χ^2_{1df}= 1.28, P>0.05 and χ^2_{1df}= 0.13, P>0.05, respectively), or in the combined sample

(χ^2_{1df}= 1.02, P>0.05). These results suggest an absence of a sex bias in sampling technique used for this study.

Groups sizes in the Gulf of St. Lawrence ranged from one to fourteen individuals (n=76, mean=5.89, SE=1.39). More specifically, we sampled 23 single individuals (30.3%), 21 pairs (27.6%), and 32 (42.1%) groups of three or more individuals. All completely sampled trios (n=3) consisted of 2 males and 1 female, of which one trio was identified as a mother/calf and a male adult. Since the frequency of each group category in the Gulf of St. Lawrence is unknown, it is impossible to ascertain the existence of any sampling bias towards particular group sizes. However, as large groups are more visible (more blows at each surfacing) and far more approachable than e. g., pairs and single individuals, it is likely that there was a bias towards sampling from larger groups. Statistically significant deviations in the direction of males from the expected proportions of males and females (1:1) were found in both incompletely and completely sampled groups of three individuals and more (Table 4.1). In contrast, no significant deviation from the expected 1:1 sex ratio was observed in either categories of singles or pairs (Table 4.1) or the combined sample of singles and pairs (χ^2=0.81; P=0.3692).

Despite the fact that groups often gain and lose members, this bias towards male fin whales in large groups is of interest. In the North Atlantic, large groups dominated by males have been described for humpback whales, on the West Indies breeding ground (Competitive Groups, CG) (Clapham *et al.* 1993c; Mattila *et al.* 1989) as well as for the North Atlantic right whale on the Continental shelf of Nova Scotia (Surface Active Group, SAG) (Stone *et al.* 1988). These SAG or CG groups are most likely related to mating and courtship behavior, where males are in competition for access to a female. Two studies on the group composition of fin whales hunted in the Antarctic and North Pacific oceans in 1955-56 and 1963-64 included observations collected from the whaling operations during the spring migration to the feeding grounds and during the summer while at the feeding grounds (Nemoto 1964; Tarasevich 1967). In both studies (Nemoto 1964;

Tarasevich 1967), despite the possible whaling bias, groups consisting entirely of or dominated by males were larger and more common than groups consisting entirely of or dominated by females. In addition, contrary to humpback whales, males were found to be the nucleus of larger groups around which other whales grouped (Tarasevich 1967). In general, fin whales are mainly observed in groups of less than four individuals, mostly as singles or pairs (Chittleborough 1953; Giard 1996).

In order to verify if the high proportion of large groups observed in this study is an artifact of the sampling effort, an estimation of the relative proportions of the three group categories in the Gulf of St. Lawrence has to be undertaken. Nevertheless, our study demonstrates that 1) large groups of the Gulf of St. Lawrence fin whale are mainly dominated by males as found in the Antarctic and the Pacific Ocean and 2) these groups could constitute a male-male competition for access to females, as described for humpback and the northern right whales.

Tables

Table 4.1. Observed and expected (assuming a sex ratio of 1:1) numbers of male and female fin whales in the different categories of completely and incompletely sampled groups.

Group sizes			Single	Pair	Large group
Completely sampled groups	Sightings		23	4	4
	Observed	Male	14	4	12
		Female	9	4	3
	Expected	Male	11.5	4	7.5
		Female	11.5	4	7.5
	χ^2_{1df}		1.09	0	5.40
	P values		0.2971	1.0000	**0.0201**
Incompletely sampled groups	Sightings		-	17	30
	Observed	Male	-	12	34
		Female	-	5	11
	Expected	Male	-	8.5	23
		Female	-	8.5	23
	χ^2_{1df}		-	2.88	12.52
	P values		-	0.6959	**0.0004**

Note. Significant P values are in bold.

Appendix 4.1. Group composition of all biopsied fin whales in the Gulf of St. Lawrence between 1990 and 1994.

Years	1990		1991		1992		1993		1994	
Months/ Groups	Size	Members	Size	Members	Size	Members	Size	Members	Size	Members
June					1	**HF**	4	03F, 04M, X, X		
June					4	06M, 07M, **GM**, **IM**				
June					5	10M, **11M***, 13M, 14M, 15F				
June					10	**IM**, 18M, **HF**, X, X, X, X, X, X, X				
June					3	**GM**, X, X				
June					1	20M				
June					1	21M				
July			4	02M, X, X, X	1	30F	3	07M, 08M, X	4	101M, X, X, X
July			6	03M, X, X, X, X, X	1	38F	1	09F	3	102M, X, X
July			3	**DM**, 05M, X	6	103M, 102F, 104F, X, X, X	2	10F, X		
July			2	**EM**, X	1	106M				
July			1	07M	5	107M, X, X, X,X				
July			4	**EM**, 10M, 08F,X	2	108F, X				
July			1	12M	3	109M, 111M, 110F				
July					1	112F				
Aug.			3	15M, **DM**, 17F	3	66F, X, X	2	63F, X	4	30M, X, X, X
Aug.			1	**AM**	10	68M, 69M, 70M, X, X, X, X,X, X, X	14	64M, X, X, X, X, X, X, X, X, X, X, X, X, X		
Aug.			3	19M, X, X	2	71M, X	6	66M, X, X, X, X, X		
Aug.			1	**BM**	4	73M, X, X, X	2	**KM**, X		

Annex 4.1 (continuous)

Years		1990		1991		1992		1993		1994
Months/ Groups	Size	Members	Size	Members	Size	Members	Size	Members	Size	Members
Aug.			2	21M, X						
Aug.			1	23F						
Aug.			1	24M						
Aug.			1	25F						
Aug.			5	26F, X, X, X, X						
Aug.			1	31M						
Aug.			8	**CM**, 33M, **JF**, X, X, X, X, X						
Sept.	2	01M, X	4	35M, X, X, X	2	90M, X	2	70F, X	1	**KM**
Sept.	2	05F, 06M	4	53M, X, X, X	2	91M, X	2	71M, X	5	55F,X,X,X, X
Sept.			1	**JF**	5	94M, 95M, X, X, X	2	73F, 72M		
Sept.			2	60F, X	2	96M, 97M				
Sept.			4	106M, X, X, X	2	98M, X				
Sept.			2	110M, X	6	114M, X, X, X, X, X				
Sept.			2	112F, 113F	6	121M, X, X, X, X, X				
Oct.	1	08M							1	74F
Oct.	4	10F, 09M, X, X			1	**CM**				
Oct.	1	**AM**								
Oct.	4	13F, X, X, X								
Oct.	2	**BM**, X								
Oct.	1	26M								
Oct.	2	02M, X								

Note. M, denotes male, F, female, and X, unknown sex. All duplicate samples are in bold and their identification number has been changed to a letter followed by there sex (F or M). Only the individual note by * has been sampled twice during the same sighting.

Part III. Hybridization

The Spanish fin whale samples included in the analysis of population genetic structure in the North Atlantic were collected during commercial whaling operations prior to the whaling moratorium. During the analysis of the samples for the paper presented in Chapter 3 (Bérubé *et al.* 1998), it became apparent that one of the Spanish samples had a blue whale rather than a fin whale mtDNA sequence. Originally the whaling factory had reported the animal as a fin whale despite its several abnormal morphological traits. Further molecular analyses allowed us to conclude that the anomalous fin whale was a hybrid, born by a female blue whale, which had mated with a male fin whale.

The fin and blue whales belong to the family *Balaenopteridae* , which includes five sympatric species; the minke (*Balaenoptera acutorostrata*), the humpback (*Megaptera novaeangliae*), the sei (*B. borealis*) the fin (*B. physalus*) and the blue whale (*B. musculus*). The *Balaenopteridae* appear first in strata assigned to the late Miocene or early Pliocene, some 6-10 Millions years ago (Barnes *et al.* 1985). Within the family *Balaenopteridae*, five hybrids have been reported so far and all from crosses between blue and fin whales. The next chapter describes the fifth of these hybrids.

The frequency (estimated as the number of hybrids reported relative to the number of fin and blue whale samples analyzed) with which hybridization occurs between the fin and blue whales appears to be quite high, and yet both species remain intact. The manuscript is published as an article in the scientific journal, Marine Mammal Science volume 14 (1):82-98.

CHAPTER 5. A NEW HYBRID BETWEEN A BLUE WHALE, *BALAENOPTERA MUSCULUS*, AND A FIN WHALE, *B. PHYSALUS*: FREQUENCY AND IMPLICATIONS OF HYBRIDIZATION.

Martine Bérubé [1,2]* and Alex Aguilar [3].

1. Department of Population Biology, Zoological Institute, University of Copenhagen, Universitetsparken 15, Copenhagen Ø, DK-2100, Denmark.
2. Department of Natural Resource Sciences, Macdonald Campus, McGill University, Ste-Anne-de-Bellevue, Quebec, H9X 3V9, Canada.
3. Department of Animal Biology, Faculty of Biology, University of Barcelona, 08071 Barcelona, Spain.

*Present address: Department of Ecology and Evolutionary Biology, University of California, Irvine CA 92717. Tel: (714) 824-8680, fax: (714) 824-2181, e-mail address: martine@newt.bio.uci.edu.

Keywords: Hybrid, fin whale, blue whale, eastern North Atlantic
Short Title: Hybridization between fin and blue whales.

Abstract

A female hybrid between a fin (*Balaenoptera physalus*) and a blue whale (*B. musculus*) was caught in whaling operations in 1984 off N. W. Spain. Its coloration and body proportions were intermediate between those of a fin and a blue whale, although it was anomalously large (19.4m) when compared to fin whales of similar age (4 years). It was sexually immature, concomitant with its age but not its length if it was a fin whale. Molecular analyses revealed that the mother of the hybrid was a blue whale and the father a fin whale. Examination of data for the five fin-blue whale hybrids in the literature, plus other anecdotal reports, indicate that hybridization between these two species occurs in various geographic regions and is relatively frequent, notably in light of the absence of

reported hybrids between other mysticetes. Either species may act as father or mother, and there does not appear to be selection for a given sex among the hybrids. The reproductive capacity of these hybrids remains unknown, although incidence of reproductive impairment appears to be higher in hybrid males than in hybrid females.

Introduction

Genera in which inter-specific hybridization is frequent are usually composed of numerous species that have similar life histories and habitat requirements. In most cases the hybrids are sterile and thus do not challenge the status of the species involved. However, in some cases they may be able to reproduce, although the fitness of their offspring is usually low. The hybrid line eventually vanishes, and so these inter-specific hybrids do not represent a threat to the genetic constitution of the parental species (Mayr 1963).

In cetaceans, thirty-three cases of alleged hybridization have been described. Twenty-seven of them in the suborder *Odontoceti* have occurred among five species, with a bottlenose dolphin, *Tursiops truncatus*, being one of the parents in all cases. Another occurrence of hybrid involved *Delphinus capensis* and *Lagenorhynchus obscurus* (Reyes 1996). All but four occurred in captivity (Heide-Jørgensen & Reeves 1993; Shimura *et al.* 1986; Sylvestre & Tasaka 1985). Indeed, the alleged hybrid nature of the four wild specimens so far reported (Fraser 1940; Reyes 1996) could not be genetically established and their validity is questionable (Miyazaki *et al.* 1992; Spilliaert *et al.* 1991).

In the suborder *Mysticeti*, only four hybrids have been reported to date. All of them were found in the wild and were taken during commercial whaling operations, and in every case the parents involved were a blue (*Balaenoptera musculus*) and a fin whale (*B. physalus*). The first of these hybrids, taken in 1965 off Kodiak Island (Gulf of Alaska), was described by Doroshenko (1970), who then was able to substantiate the hybrid nature of the specimen only by its exceptional morphological traits. Later, three morphologically anomalous baleen

whales caught in the Icelandic whale fishery were later confirmed to be interspecific hybrids through molecular analysis of their nuclear and mitochondrial (mt) DNA. Spilliaert *et al.* (1991) described the first of them, a female caught in 1987. The molecular analysis was based on comparison of RFLP (restriction fragment length polymorphism) patterns of the complement component C4, restriction patterns of the mtDNA, and restriction hybridization patterns of the light balaenopterid DNA satellite. Later, Árnason *et al.* (1991b) analyzed two male hybrids, again produced by the crossing of a fin and a blue whale, that had been caught in 1983 and 1989. They sequenced the mtDNA control region and analyzed the liver esterases.

This study presents another case of hybridization between a fin whale and a blue whale. The anomalous whale was caught in 1984 by the whaling fleet operating off the western coast of Spain. We present here a detailed morphological and morphometric description of the hybrid, photographic documentation of its external aspect and coloration, and the results of molecular analysis of the mt (control region) and nuclear (α-lactalbumin) DNA.

Materials and Methods

On 14 September 1984 at 44°24'N-10°42'W, the anomalous whale (SP84075) was captured by the whalecatcher *IBSA III* in the presence of four animals identified as fin whales. The anomalous whale and another whale from the school, confirmed as a normal fin whale, were brought to the shore station of Caneliñas (N. W. Spain) where they were flensed on the following day. Data and samples of both animals were collected 14h *post mortem*.

Morphological and biological information

External morphometric data were collected following Norris (1961), modified when necessary to fit the shape and specific features of baleen whales. Description here of morphological features of the specimen is from direct observation at the time of examination and from photographs. Procedures for determination of age and reproductive status followed Lockyer (1984). Biological

data were compared with fin whale growth and reproductive parameters, as determined by Aguilar (1985), Aguilar and Lockyer (1987) and Aguilar *et al.* (1988). Morphological information, from North Atlantic blue whales, was extracted from Spilliaert *et al.* (1991). Fin whale data were compiled from the database of the University of Barcelona. Comparisons between the Caneliñas hybrid and the fin whale sample were carried out using a *t* test modified to deal with single cases *versus* a sample (Sokal & Rohlf 1995).

Molecular analysis

Blue whale DNA was obtained from a skin biopsy sample collected using a crossbow and an arrow with a modified stainless steel tip (Palsbøll *et al.* 1991) and molded float. The sample was preserved in saturated NaCl with 20% DMSO (Amos & Hoelzel 1991) and stored at -20°C. DNA from the fin whale and the anomalous whale (SP84075) was obtained from muscle samples collected at the whaling station and stored in a –20°C freezer. DNA was extracted using standard cell lysis and phenol-chloroform protocol (Sambrook *et al.* 1989).

Sex of the three individuals was molecularly determined using a method based on a set of three primers that are specific to either the ZFY or ZFX sequence found on the sex chromosomes (Bérubé & Palsbøll 1996a; 1996b).

A hybrid results from the combination of the two parental species; there is one chromosome set from the mother, species 1, and the other set from the father, species 2. In order to determine the maternal species identity of the anomalous whale, the first 299bp at the 5' end of the mt control region (D-loop) were determined by direct sequencing of asymmetrical amplified DNA (Higuchi *et al.* 1988; Saiki *et al.* 1988). MtDNA is predominantly maternally transmitted in mammals (Hutchinson *et al.* 1974), making it the ideal marker for addressing this question. The sequence of primers used for the amplification and the experimental conditions are summarized in Palsbøll *et al.* (1995). Evolutionary or genetic distances among the three individuals were estimated by using both the absolute number of nucleotide differences and the Kimura two-parameter distance (Kimura 1980).

In contrast to mtDNA, the α-lactalbumin gene, located in the cell nucleus, is inherited in a Mendelian fashion. Consequently, with the knowledge of the maternal species from the mtDNA identification, it can be used to identify the paternal species. The α-lactalbumin nucleotide sequences were obtained by direct sequencing of asymmetrically PCR-amplified DNA as described in Palsbøll *et al.* (1995), except that the initial symmetrical amplifications were performed in a volume of 35 µl with an initial 2 min. denaturing step at 94°C followed by 32 cycles, each of 15 sec at 94°C, 15 sec at 54°C and 30 sec at 72°C. Amplifications were performed on a Gene Machine® Junior Thermal Cycler (USA/Scientific Plastics). The oligonucleotide primers Lac1.R (5'-CTC ACT GTC ACA GGA GAT GT -3') and LacII.F (5'-CCA AAA TGA TGT CCT TTG TC -3') (Courtesy of Dr. Dave Irwin, University of Toronto) were used for amplification and sequencing. Because of the detailed morphological description of the hybrid, the molecular analysis at the nuclear level was reduced to only the comparison of the Caneliñas hybrid sequence to the blue and fin whale α-lactalbumin sequences.

From the 540 basepairs sequenced, a substitution (from G to T) in position 53 resulted in a loss or gain of a recognition site for the endonuclease *Fok* I (site GGATG) unique to the blue whale α-lactalbumin sequence (Figure 5.3). The difference in length was sufficient to be distinguishable on a 2% NuSieve™ gel.

Results

Morphological description

The Caneliñas hybrid was a 19.4m female. Figure 5.1 shows different views of the whale's external aspect. Table 5.1 describes its external characteristics and morphometrics compared to one of the Icelandic hybrids previously described (Spilliaert *et al.* 1991), as well as fin and blue whales of similar body length (range 18.5-20.5m).

The overall shape of the hybrid's body was quite similar to that of a fin whale. This included the dorsal fin, which was well defined, moderately high and falcate, as is common in this species. However, the wide rostrum, with the maxilla slightly

curved outward, resembled that of a blue whale. None of the 20 morphometric measurements taken showed significant differences from those of fin whales (Table 5.1). Unfortunately, the small size of the blue whale sample precluded a reliable statistical comparison with that species. The hybrid described by Spilliaert *et al.* (1991) appeared to have relatively large flippers, as is common in blue whales, but a relatively short rostrum, as is usual in fin whales.

Coloration of the Caneliñas hybrid was overall slate-gray on the back and flanks as is typical in fin whales. However, the hybrid had a central, roughly symmetrical white patch with some dark spotting in the ventral region. Although no distinct blaze and chevron could be observed as in a fin whale (Agler *et al.* 1990), the right side of the anterior part of the body was slightly paler than the left side. However, this asymmetry did not extend to the baleen plates, as invariably occurs in fin whales; the hybrid's plates were uniformly black with no striations, a pattern, which is very similar to that of a typical blue whale. Along the trunk the hybrid had some lamprey scars and remora-like marks, as is common in both fin and blue whales.

The other described fin-blue whale hybrids (Doroshenko 1970; Spilliaert *et al.* 1991) were also characterized by an overall slate-gray pigmentation of the back and flanks, the presence of dark spots in the ventral white patch, the absence of asymmetry in the pigmentation of the ventral cephalic region (although the back of the head was slightly lighter on the right side in some animals), and the shape and height of the dorsal fin, which resembled that of a fin whale. However, the baleen plates in the Caneliñas hybrid were all black, similar to those of a blue whale, while in the other two cases they were mixed, with some being black and others striated or even yellow.

Age, body length and reproductive status

The mid-section of the ear plugs was moderately pigmented as is typical of a fin whale. Based on the count of growth layer groups present in the core of the hybrid ear plug, it was estimated to be 4 years old. The mean body length of 21 four-year-old female fin whales caught during the period 1982-1985 was 17.4m (SD: 0.66,

range: 16.2-18m). Therefore, the Caneliñas hybrid was significantly larger ($p<0.01$) than fin whales of the same age inhabiting the waters off northwestern Spain and was about 1.4m longer than the largest four-year-old female fin whale recorded. Unfortunately, no comparable data are available for blue whales.

Reproductive information from the Caneliñas hybrid is detailed in Table 5.2. The ovaries were typical of a sexually immature individual, with reduced weight, small follicular size and absence of *corpora* of ovulation. Follicular abundance was high, indicating a potential for future ovulations. No fetus was found in the uterus. The uterine mucosa was pale in color and the uterine horns were small, confirming the absence of present or previous pregnancies. The mammary glands were thin, indicating that the whale had never lactated.

The mean age of sexual maturation in the northwestern Spain female fin whale population when the hybrid was caught has been estimated at 7.7 years (range: 4-12 years) (Aguilar *et al.* 1988), and the immature reproductive status of the Caneliñas hybrid is consistent with these age ranges. In contrast, the Caneliñas hybrid was significantly longer than 17.3m, the length, which would correspond to such an age according to the observed age-length growth curve for the fin whale population inhabiting waters off N. W. Spain (Aguilar & Lockyer 1987).

Molecular analysis

The molecular analysis confirmed that the Caneliñas hybrid was a female.

The difference between the sequence composition of the first 299bp of the mt control region of the blue whale and the Caneliñas hybrid was only three substitutions (Figure 5.2). The Kimura two-parameters genetic distance between the two individuals was 0.0105. However, in comparison with the fin whale sequence, the distance was 0.1392, equivalent to the calculated distance between the fin and blue whale samples (0.1348), or between a sei (sequence from Arnason and Gullberg (1993) and a blue or fin whale (0.2029 and 0.2058, respectively). Thus, the mother of the Caneliñas hybrid was in all likelihood a blue whale.

The paternal species identity, in this case a fin whale, was obtained by deducting the blue whale α-lactalbumin sequence from the SP84075 sequence. In view of

the morphological information, the resulting sequence was compared to a fin whale α-lactalbumin sequence, which was similar to that of the hybrid. The comparison of the α-lactalbumin sequences between the blue and the fin whale revealed five segregating sites. Since the α-lactalbumin sequence of the Caneliñas hybrid (SP84075) contained copies of each, the hybrid was the product of both species (Figure 5.3). An additional site was found in the Caneliñas hybrid at position 489. This is surprising, since no variation occurs at that site in other mysticete α-lactalbumin sequences (data not shown). More individuals of each of the two species should be sequenced to determine which one contains the substitution.

In order to ensure that the double bands observed in the nucleotide sequence were the result of the presence of the two sets of chromosomes and not of a gel or PCR artifact, we cut the double-stranded product with the restriction enzyme *FokI*. In the ≈ 600bp fragment of the α-lactalbumin amplified, one cutting site (position 381) was found in both species while a unique site (position 53) was present in the blue whale only. The pattern of digestion of the Caneliñas hybrid with the restriction enzyme *FokI* differs from those of the blue and fin whale (Figure 5.4b). One restriction fragment present in the Caneliñas hybrid is found in the blue whale but missing in the fin whale and another fragment in the Caneliñas hybrid is present in the fin whale but missing in the blue whale. The other restriction fragment (≈100bp) found in the blue whale as well as the Caneliñas hybrid is too small to be readily detectable on the picture. This method permitted us to confirm that the α-lactalbumin sequences of both the fin and the blue whale were present in the Caneliñas hybrid.

Discussion

As most mysticetes do, fin and blue whales undertake large-scale seasonal migrations that extend from high latitudes, where they feed, to temperate or tropical waters. In the latter areas, they feed little or fast while they engage in mating and calving activities. Migratory routes of baleen whales are usually

oceanic, thus they do not encounter physical barriers and have few constraints on their dispersal. Genetic divergence appears to derive primarily from geographical isolation produced by geological or oceanographic barriers, asynchronous migratory patterns (particularly between northern and southern hemisphere populations), and behavior (Hartl & Clark 1989; Palsbøll *et al.* 1995; Palumbi 1992).

The divergence of fin and blue whales has been suggested to be largely related to behavioral traits, particularly those associated with feeding and breeding (Spilliaert *et al.* 1991). Blue whales feed almost exclusively on euphausiids in all oceans. Fin whales are more euryphagous; although euphausiids are the main item in their diet, they also consume a variety of prey including schooling fish (Kawamura 1980; Sergeant 1977). Unfortunately, no data were recorded on stomach contents of the Caneliñas hybrid. Information on the reproductive behavior of both fin and blue whales is extremely limited and the precise locations of the breeding grounds of the two species remain unknown. However, at least during the feeding season, the two species often have sympatric distributions and, particularly when foraging on euphausiids, may form mixed schools. In the whaling grounds off northwestern Spain, blue whales are relatively rare and, when seen, they frequently associate with fin whales, the predominant mysticete species in the area. Indeed, in this region, fin whales only rarely form mixed schools with other cetaceans, and when they do, they associate either with odontocetes (e.g., striped dolphins and long-finned pilot whales) or with blue whales. Blue whales were sighted nine times between 1981 and 1987, and on five of these occasions fin whales were also present, indicating mixed schools of the two species (University of Barcelona database). No associations of either of the two species were recorded with sei, humpback or minke whales, the other mysticetes present in the area. Thus, overlapping distribution and feeding regimes, as well as the habit of forming mixed schools, undoubtedly bring opportunities for hybridization.

Both the blue and the fin whale belong to the genus *Balaenoptera* of the family Balaenopteridae, and they are often sympatric. The separation of the lineages is estimated to have occurred some 3.5 to 5 million years ago (Arnason & Gullberg 1993; Arnason *et al.* 1991b). Although some lineages of turtles, which have been separated for as long as 10 to 75 million years are known to produce viable hybrids (Karl *et al.* 1995), in mammals the oldest interspecific hybrids involve species that separated around six million years ago (Wilson *et al.* 1974). However, the estimated separation time for a number of other vertebrate species, which can hybridize is about 3 million years; an example is hybridization of the black-tailed and white-tailed deer (Carr & Hughes 1993). Therefore, fin and blue whales are among the most divergent mammal species that are still able to produce interspecific hybrids. A possible reason for the surprising hybridization ability of these two mysticetes, which are quite divergent relative to the mammalian evolutionary time scale, appears to be that cetaceans evolved very slowly at the molecular level. The karyotypes of most cetaceans are identical, 2n=44 (Arnason 1990), and molecular studies have shown a high level of similarity among them, particularly between the blue and fin whale (Arnason & Best 1991; Spilliaert *et al.* 1991).

Although the animal described here is only the fifth hybrid reported between these two species, the actual number of cases observed during the long history of whaling or during more recent whale research cruises may be higher. The whaling station at Blanford, Nova Scotia, in the summer of 1966 processed a 70-foot whale that was initially believed to be a fin-blue whale hybrid because of its black baleen plates and symmetrical pigmentation (Mead, *pers. comm.*). However, the whale was eventually considered to be a blue whale by the whaling inspector and is recorded as such in the statistics (Committee for Whaling Statistics 1967). In the case of the hybrid reported here, the whaling factory recorded the specimen as a fin whale despite its abnormal aspect. This decision was probably taken to avoid a sanction for the killing of a protected species. It is likely that the incidence of hybrids has been underreported in the whaling records because of

lack of appropriate scientific inspection and data recording during the earlier years and because of the desire to hide illegal catches after blue whales became protected.

Sightings of cetaceans made by skilled personnel during scientific cruises have on various occasions included apparently anomalous individuals, impossible to assign to any of the recognized species. For example, an anomalous animal was recorded in 1992 during a research cruise in the Gulf of Maine (Clapham, *pers. comm.*). Similarly, during an IDCR (International Decade of Cetacean Research) whale research cruise in the Antarctic during the winter of 1995-1996, an anomalous whale with coloration and body shape intermediate between those of a fin and a blue whale was recorded (Sekiguchi and Cawthorn, *pers. comm.*). Unfortunately, the lack of detailed examination or tissue samples from these specimens for molecular analyses prevented a positive identification.

We circulated a questionnaire requesting observations of potential cetacean hybrids in an attempt to evaluate their frequency in nature. It was distributed among 24 scientists that were personally involved in the inspection or collection of samples at whaling factories where fin or blue whales had been caught in the past. Most of those answering (15) were active during the period when the majority of the fin whale catches in this century were made (1945-1965). Only a few witnessed the blue whale fishery, which ended in the early 1950s (excluding illegal catches). Some answers (like the one described above) suggested that there might have been more hybrids caught. However, records were often imprecise, both in terms of the presence of hybrids and the total number of whales examined.

Although we have been unable to assess the hybridization rate, such events appear surprisingly common, perhaps in the range of one for every 500-1,000 "normal" fin whales. No similar calculations could be made in relation to blue whales. However, that the blue and fin whale constitute two distinct species indicate that these hybrids must have a very low level of fitness, as we do not observe the extensive hybridization found in other closely related species of mammals (Minkoff 1983; Wayne & Jenks 1991).

In the North Atlantic, both species were intensively harvested from the beginning of the 20th century. The massive blue whale, producing a large yield of marketable products, was the first target of the whalers and, due to excessive fishing pressure, its abundance decreased abruptly. In the 1930s blue whales became so scarce that the fishery shifted to the somewhat smaller, and therefore less productive, fin whale. Population numbers for the fin whale in the North Atlantic during the late eighties have been estimated at 56,000 individuals (Buckland *et al.* 1992a; Buckland *et al.* 1992b; IWC 1992) of which about 17,000 are thought to inhabit waters off northwestern Spain during the summer (Buckland *et al.* 1992a). No comparable information is available for blue whales, although it is generally accepted that, from an estimated initial population size in the North Atlantic of perhaps 15,000 in the pre-whaling era, numbers have dwindled to a current population of only a few hundreds (Klinowska 1991; Yochem & Leatherwood 1985). Using their natural markings and coloration pattern (Sears *et al.* 1990), identified a minimum of 203 individual blue whales in the North Atlantic. Except for an apparently healthy population in the eastern North Pacific, the situation concerning blue whales is similar in other oceans. It appears that the blue whale has been reduced to extremely low levels, while the fin whale remains relatively abundant (Klinowska 1991). Thus, the low number of mates for the blue whale might have increased the likelihood of interspecific breeding. If this were the case, the above observed hybridization rates would reflect a recent phenomenon produced by the rarity of a parental species rather than a natural situation. However, reports of anomalous whales that can be assumed to be hybrids preceded the period of intensive blue whale harvesting. In 1887, Cocks described a type of fin whale, which he called "bastard", because of its abnormal great length (Spilliaert *et al.* 1991).

The small number of hybrids that have been properly studied does not permit conclusions on the parental role played by each species. However, in three of the four individuals in which identity could be established, the blue whale acted as the mother and the fin whale as the father (Spilliaert *et al.* 1991 specimen, Árnason *et*

al. 1991b specimen, and the present case). Nonetheless, one cannot infer species-specific asymmetry in parental roles. Furthermore, since three of the five hybrids were females (Doroshenko 1970 specimen, Spilliaert *et al.* 1991 specimen, and the present case), and two were males (the two Árnason *et al.* 1991b specimens), there appears to be no selection for a given sex.

The reproductive capability of these hybrids remains unclear. From the three females described, one (the present case) was sexually immature, a condition consistent with its young age. Another (Doroshenko 1970) was sexually mature as evidenced by the presence of *corpora albicantia* in the ovaries. These corpora were assumed to originate from ovulations but not from pregnancy; based on this assumption, Doroshenko (1970) proposed that the whale, despite its ability to ovulate, was unable to become pregnant. This conclusion is disputable given the impossibility of distinguishing, in baleen whales, the *corpora* produced by ovulation from those that result from a successful pregnancy, even with detailed histological examination (Lockyer 1984). The third female described (Spilliaert *et al.* 1991) was carrying a backcrossed fetus, the viability of which was uncertain. In this case, the hybrid mother's ovaries had two *corpora*: one *corpus luteum* supporting the pregnancy, and a *corpus albicans* that was assumed to originate from a previous pregnancy, again a disputable conclusion.

Regarding the males, the two examined specimens were both sexually immature, an anomalous condition given the age (24 years) of at least one of them (Arnason *et al.* 1991b). Despite the small sample size, this might suggest a higher incidence of reproductive impairment in hybrid males than in hybrid females, substantiating Haldane's proposition that when only one sex is sterile or inviable in the offspring of species crosses, it is nearly always the heterogametic sex (Dobzhansky 1970). Thus, in most interspecific hybrids of mammals, deterioration of the hybrid lineage's is usually due to male sterility rather than to non-viability of descendants. Although the causes for the reduced fitness of the heterogametic sex are unclear, the effect may be explained by the fact that sex

chromosomes play a large role in post-zygotic mechanisms of reproductive isolation (Coyne 1992; Wu & Palopoli 1994).

The evolutionary implications of these hybrids, in some cases apparently fertile, is unclear. As mentioned above, the overlap of feeding habitats and the similarity at the molecular level favor hybridization between the two species. However, this would challenge the maintenance of the taxonomic integrity of the two species. Indeed, we should expect the occurrence of hybrids to be rare, as it potentially reduces the fitness of the parental species (Mayr 1963). Individuals of mixed ancestries, if not sterile, are usually less fit than their parents. Thus, the new gene combinations resulting from the hybridization are expected to be less well adapted to the environment or to interact poorly with those from the parental species, therefore reducing the incidence of hybrids in subsequent generations.

From these findings we conclude that i) hybridization between fin and blue whales occurs at relatively higher frequency than among other species of mysticetes, ii) there is no apparent asymmetry in mate choice, iii) hybrid males may be sterile, while females reach sexual maturity, although their actual capacity to produce viable offspring awaits confirmation, and iv) hybridization is geographically widespread and not restricted to a given area or population, as it has now been observed in both the North Atlantic and North Pacific oceans.

Tables

Table 5.1. External characteristics and morphometrics of the Caneliñas and the Spilliaert *et al.* (1991) hybrids, and of North Atlantic fin and blue whales of comparable body length (range 18.4-20.4m).

Measurement (expressed as the ratio to body length if not specified)	Present hybrid	Spilliaert *et al.* hybrid	Fin whales	Blue whales
Sex	♀	♀	♀	♂ & ♀
Body length (m)	19.4	21.0	19.5±0.78	-
Tip of snout to:				
Center of eye	19,5	19.1	19.9±0.69	21.1±0.90
End of gape	19.5	-	19.9±0.86	-
Center of ear hole	23.9	-	24.6±0.75	-
Center of blowhole	17.5	-	18.7±2.51	-
Anterior insertion of flipper	31.7	-	30.1±2.65	-
Tip of dorsal fin	74.2	-	74.9±2.41	-
Center of umbilicus	54.2	-	54.5±1.64	-
Center of genital slit	70.1	-	68.4±1.44	-
Center of anal slit	72.2	72.4	71.4±1.49	72.25
Anterior insertion to tip of flipper	11.3	-	11.4±0.53	-
Dorsal fin: height to length of base (ratio)	0.52	0.88	0.38±0.25	0.29
Height of dorsal fin (cm)	42	53	43.9±3.40	40
Height of dorsal fin	2.2	2.48	2.3±0.16	1.09±0.27
Width of tail flukes	21.4	-	20.5±1.15	-

Note. Fin whale data are from the database of the University of Barcelona. Blue whale data are from Spilliaert *et al.* (1991).

Table 5.2. Characteristics of the reproductive organs of the Caneliñas hybrid.

ANATOMICAL DATA	RIGHT SIDE	LEFT SIDE
Weight of ovary (g)	258	196
Abundance of follicles	High	High
Mean cross diameter of the seven largest follicles (mm)	4.5	5.4
Corpora of ovulation (no.)	0	0
Diameter of the uterine horns (mm)	60	60
Thickness of the mammary gland at its thickest (central) region (mm)	20	20

Figures

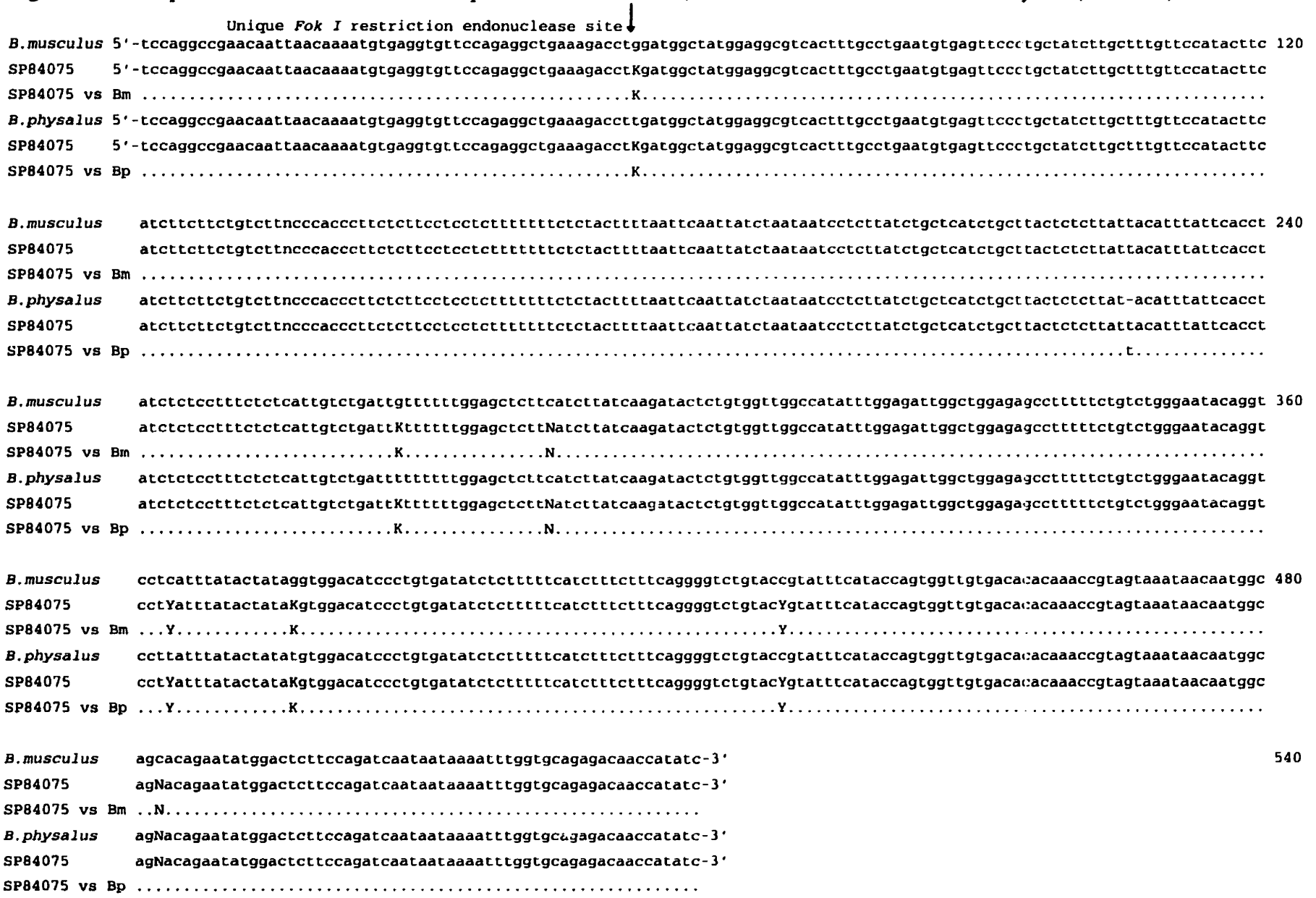
Figure 5.1. External morphology of the Caneliñas hybrid.

A) The baleen plates are uniformly black with no striation, as in blue whales. The creamish coloration of the first third of the baleen plates typical in fin whales is absent. The rostrum is wide with the maxilla slightly curved outwards, as is usual in blue whales. B) The ventral region has a roughly symmetrical white patch with some dark spotting. C) The dorsal fin is well defined and moderately high and falcate, similar to that of fin whales. The white coloration is due to the blubber being exposed by erosion of the skin during the hauling of the whale on the flensing platform, and it is not a pigmentation feature. D) The coloration of the lower jaws is symmetrical, as opposed to fin whales. The palate is black as in blue whales (in fin whales it is white-pink). E) The coloration of the anterior part of the belly is symmetrical, as opposed to fin whales, and has a well defined white patch, as opposed to blue whales. F) The posterior ventral region is white, with dark blazes running from the anal slit backwards.

Figure 5.2. The sequences of the first 299 nucleotides of the mt control region in the fin whale, the blue whale, and the Caneliñas hybrid (SP84075).

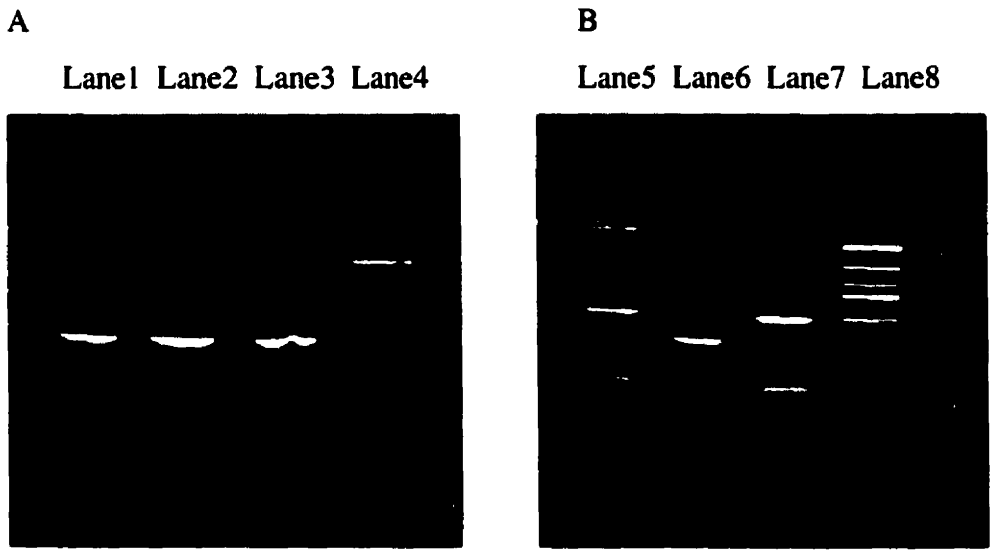
```
B.physalus  5'-aaaaa-gtatattgtacaataaccacaggaccacagtactatgtccgtattgaaaataac 59
SP84075     5'-aaacatgtatattgtacaataaccgcaaagccacagtactatgtccgtattaaaaaataa 60
SP84075 vs Bp  ...c.t.................g..aag.....................a....at.a
B.musculus     aaacatgtatattgtacaataaccgcaaagccacagtactatgtccgtattaaaaaataa 60
SP84075        aaacatgtatattgtacaataaccgcaaagccacagtactatgtccgtattaaaaaataa 60
SP84075 vs Bm  ............................................................

B.physalus     ttgccttattagatattattatgtaatccgtgcatgtatgtacttccacataa----tta 115
SP84075        ttatctcattacatattgttatgtacttcgtgcatgtatgtacttccccataaccagtta 120
SP84075 vs Bp  ..at..c....c.....g.......c.t..................c.....ccag...
B.musculus     ttatctcattacatattgttatgtacttcgtgcatgtatgtacttccccataaccagtta 120
SP84075        ttatctcattacatattgttatgtacttcgtgcatgtatgtacttccccataaccagtta 120
SP84075 vs Bm  ............................................................

B.physalus     atagcgtctttccatggg--tatgaacagatatacatgctatgtataattgtgcattcaa 173
SP84075        atcagtgttatccctgtgaatatgtatacatacacatgctatgtataattgtgcattcaa 180
SP84075 vs Bp  ..cagtgt.a...c..t.aa....t.t.c...c.........................
B.musculus     atcagtgttatccctgtgaatatgtatacatacacatgctatgtataattgtgcattcaa 180
SP84075        atcagtgttatccctgtgaatatgtatacatacacatgctatgtataattgtgcattcaa 180
SP84075b vs Bm  ...........................................................

B.physalus     ttatttccaccacgagcagttgaagctcgtattaaatttattaattttacatattacat 233
SP84075        ttatcttcaccacgagcagttaaagcccgtattaaatcttattaattttacatattacat 240
SP84075 vs Bp  ....c.t.............a....c.........c.....................
B.musculus     ttatcttcaccacgaacagttaaagcccgtattaaatttattaattttacatattacat 240
SP84075        ttatcttcaccacgagcagttaaagcccgtattaaatcttattaattttacatattacat 240
SP84075 vs Bm  ..............g.....................c.....................

B.physalus     aatatgt-attaatagtacaatagcgcatgttcttatgcatcctcaggtttatttaaat-3'291
SP84075        aatattttattaatagtacagtagtacatgttcttatgcatcctcaggtcaatttaaat-3'299
SP84075 vs Bp  .....t.t...........g...ta.....................ca........
B.musculus     aatattttattaatagtacagtagtgcatgttcttatgcatcctcaggtcaatttaaat-3'299
SP84075        aatattttattaatagtacagtagtacatgttcttatgcatcctcaggtcaatttaaat-3'299
SP84075 vs Bm  .......................a.............................
```

Figure 5.3. Comparison of the α-lactalbumin sequence of the fin whale, the blue whale and the Caneliñas hybrid (SP84075).



Note. "Y" indicates the presence of the nucleotides C and T; "K" for G and T; "N" for unknown; and dash (-) for a deletion.

Figure 5.4. Double-stranded amplification of the α- lactalbumin PCR products of the fin whale, the blue whale and the Caneliñas hybrid (SP84075) (A) and their digestion with *Fok I* restriction endonuclease (B).



Note. Lanes 1 and 5 are PCR products from a fin whale, 2 and 6 from a blue whale and, 3 and 7 are from the hybrid SP84075. Lanes 4 and 8 are the marker pUC 18digested with *HinfI* and *HinfI/sspI*. The numbers correspond to the marker's fragment length.

OVERALL CONCLUSION

The biota of the oceans has long been considered an endless resource for mankind to exploit. The exploitation of the large whales has become a classic example that these resources are finite, indeed. As the abundance of many species has decreased, the need to understand fundamental aspects of cetacean biology and distribution has become more and more urgent. Basic, but missing information, such as abundance, population identity and migratory routes, is the pre-requisite to formulate a biologically sound and worthwhile management strategy. Our understanding of cetacean biology and how marine mammals interact with their environment is still surprisingly poor but understandable given their submarine existence. Whales inhabit an open environment with few physical barriers and thus have wide ranges of movement. Still, whales are often found in relatively well-defined and predictable areas on the continental shelf characterized by high biological productivity. Most mysticetes are presumed to migrate between feeding and breeding areas, as their abundance on the feeding grounds typically decreases during the autumn. However, it is actually only for a few mysticetes (e.g., humpback and right whales) that the location of winter habitat is known. Currently, models proposed for mysticete population structure are basically variations of two general models; either a single panmictic population which segregates on the summer feeding grounds, e.g., North Atlantic and North Pacific humpback whales, (Baker *et al.* 1993; Clapham *et al.* 1993a; Palsbøll *et al.* 1995) and North Atlantic right whales (Knowlton *et al.* 1992); or separate panmictic populations, which share a common feeding ground, e.g., minke whales (Bakke *et al.* 1996; Palsbøll 1989) and southern hemisphere humpback whales (Gaskin 1982).

The results obtained from the population genetics analysis of North Atlantic fin whales conducted for this thesis provide compelling evidence for a third cetacean population model: several panmictic populations within the same ocean basin but with separate feeding and breeding grounds. More specifically, this study showed that indeed the Sea of Cortez (North Pacific Ocean) and the North Atlantic and the

Mediterranean Sea fin whales constitute different and genetically isolated populations. However, as with humpback whales (Baker *et al.* 1990; Baker *et al.* 1993; Palsbøll *et al.* 1995), occasional gene flow between oceanic populations appears to have occurred in fin whales, as well. Furthermore, our data showed that the North Atlantic Ocean fin whale was divided in at least two "populations"; an eastern and western North Atlantic population, each of which was separate from the Mediterranean fin whale population. The fact that the Mediterranean Sea fin whales constitute a population that is separate from the North Atlantic has immediate management implications for the relatively small Mediterranean population. The relevant institutions have ensured that this finding has been brought to the attention of the Italian authorities. The higher level of genetic structure detected at the mt loci, relative to the nuclear loci, could be due to the difference in the effective population size of the two genomes or natal homing to specific feeding grounds.

For recently diverged populations, as appears to be the case for most of the North Atlantic samples in this study, the main cause of differentiation is genetic drift and not only mutation. Mutation needs not only generate 'new' alleles but such new alleles will also have to increase in frequency (or add to the frequency of 'another' allele, in case of microsatellites). Hence a mutation has to occur first, and subsequently drift has to change the frequency of the 'new' allele. All other factors being equal an increased mutation rate will of course add to the rate of divergence. However, in the case of mt versus nuclear loci all factors are not equal and since the effective population size of the mtDNA is only one fourth of that of the nuclear DNA, the effects of genetic drift act more rapidly on mtDNA than nuclear DNA. This is presumably one of the reasons why so many studies based upon mtDNA have been successful in identifying population structure, and why many subsequent analyses of nuclear loci fail to detect the same extent of structure. Unfortunately, many studies fail to recognize the differences in the effective population size of the two genomes and the population genetic consequences of this fact. In order to account for the lack of differentiation at nuclear loci, many studies invoke male-mediated gene flow, which are predicted

to have the same effect. Unfortunately, the population genetic consequences of the relative difference in effective population size of the two genomes and/or male-mediated gene flow have not yet been explored or described in sufficient detail, e.g., by simulations to determine exactly what the predictions of the two models are.

If the difference of divergence rate for the two genomes is an explanation for the observed lack of population differentiation at the nuclear loci, we would expect that larger sample sizes and/or the analysis of more loci would reveal significant levels of heterogeneity at nuclear loci as well among close and adjacent North Atlantic sampling localities (e.g., between the eastern North Atlantic and the Mediterranean Sea). This prediction was supported by simulations with the current data set.

In most mysticetes studied so far, an even sex ratio has been found on feeding grounds (Aguilar & Lockyer 1987; Clapham *et al.* 1995; Kapel 1979; Mitchell 1974; Palsbøll *et al.* 1995; Tarasevich 1967). Unexpectedly, the fin whales from the Gulf of St. Lawrence showed a significantly male biased sex ratio. Although no definite conclusions could be drawn, the male bias was confined to groups of three individuals and more. Evidence in support of this observation comes from earlier studies of group composition in baleen whales from the Antarctic and North Pacific Ocean (Nemoto 1964; Tarasevich 1967). However, additional samples are necessary to make more affirmative conclusions about fin whale social organization.

It is obvious that molecular approaches possess a great potential for the study of natural populations, especially microsatellite analyses. Microsatellite loci are ideal for analysis of kinship, because of their elevated mutation rate. However, the accuracy of such kinship analyses is highly contingent upon a proper and extensive sample collection. To conduct a detailed analysis of a natural population, a high proportion of the individuals have to be sampled in order to increase the chance that the sample contains close relatives. With the addition of the mt control region sequence and the sex of individuals, it is in principle

possible to determine the nature of the parent-offspring relations. Once parent-offspring relations have been determined, such data can be analyzed including the sighting history, group composition and behavioral class of the individual, and provide a novel and detailed insight into the mating system, social organization, philopatry and the degree of dispersal in fin whales.

Finally, although microsatellites appear to be ideal Mendelian markers for the study of natural populations, it should be kept in mind that the mutation process is still relatively poorly understood. There are ongoing analyses, which address these issues, which in turn might expand or limit the utility of these loci e.g., as reliable estimators of the degree of divergence between populations.

The rapid developments in molecular techniques continuously bring us new tools to better estimate and describe the level of diversity within and between the species and their populations. Nevertheless, this gain in knowledge is not followed by a similar increase in conservation measures ensuring that the biological diversity is secured in the future as well. For example, the oceans have until recently been considered an inexhaustible resource, for which reason many species, including marine mammals, are now classified as endangered. One natural evolutionary process that influences genetic diversity is hybridization. As we saw, in mysticetes, several hybrids between fin whales and blue whales have now been reported (see Chapter 5 and Arnason *et al.* 1991b; Doroshenko 1970; Spilliaert *et al.* 1991). Mysticetes are very similar genetically, e.g., the level of genetic difference between the fin and blue whale is similar to what is observed between *Homo*, chimpanzee (*Pan troglodytes, P. paniscus*) and gorilla (*Gorilla gorilla*) (Arnason & Gullberg 1993). Although we were incapable of assessing the hybridization rate properly in this study, such events appear surprisingly common between fin and blue whales. However, that fact that the blue and fin whale constitute two distinct species indicate that the fitness of these hybrids must be low, or we would have observed more extensive introgression as observed in other closely related species of mammals (Minkoff 1983; Wayne & Jenks 1991).

One cause of the serious decline in abundance observed for many cetacean species is the extensive whaling during the first half of this century. Some species, such as the humpback whale, seem to have recovered well and are relatively abundant today. However, whales are at the end of a food chain and thus susceptible to changes in the abundance and accumulation of pollutants in their prey. The continued growth of human populations will put an even higher demand on the oceans in the future. Hence, even though whaling today is relatively tightly regulated, this does not imply that conservation of cetaceans is secure.

LITERATURE CITED

Agler, B. A. 1992. Testing the reliability of photographic identification of individual fin whales (*Balaenoptera physalus*). *Report of the International Whaling Commission* **42**, 731-737.

Agler, B. A., Beard, J. A., Bowman, R. S., Corbett, H. D., Frohock, S. E., Hawvermale, M. P., Katona, S. K., Sadove, S. S. & Seipt, I. E. 1990. Fin whale (*Balaenoptera physalus*) photographic identification: methodology and preliminary results from the western North Atlantic. *Report of the International Whaling Commission Special Issue* **12**, 349-356.

Aguilar, A. 1985. Biología y dinámica poblacional del rorcual común (*Balaenoptera physalus*) en las aguas atlánticas ibéricas. Ph. D. dissertation. University of Barcelona, 487pp.

Aguilar, A. & Lockyer, C. H. 1987. Growth, physical maturity and mortality of fin whales (*Balaenoptera physalus*) inhabiting the temperate waters of the north-east Atlantic. *Canadian Journal of Zoology* **65**, 253-264.

Aguilar, A. & Nadal, J. 1984. Obtención de biopsias hipodérmas de cetáceos en libertad. *Investigación Pesquera* **48**, 23-29.

Aguilar, A., Olmos, M. & Lockyer, C. H. 1988. Sexual maturity in fin whales (*Balaenoptera physalus*) caught off Spain. *Report of the International Whaling Commission* **38**, 317-322.

Amos, B. & Hoelzel, A. R. 1990. DNA fingerprinting cetacean biopsy samples for individual identification. *Report of the International Whaling Commission Special Issue* **12**, 79-85.

Amos, B., Schlötterer, C. & Tautz, D. 1993. Social structure of pilot whales revealed by analytical DNA profiling. *Science* **260**, 670-672.

Amos, W. & Hoelzel, A. R. 1991. Long-term preservation of whale skin for DNA analysis. *Report of the International Whaling Commission Special Issue* **13**, 99-104.

Amos, W. & Rubinsztein, D. C. 1996. Microsatellites are subject to directional evolution. *Nature Genetics* **12**, 13-14.

Amos, W., Whitehead, H., Ferrari, M. J., Glockner-Ferrari, D. A., Payne, R. & Gordon, J. 1992. Restrictable DNA from sloughed cetacean skin; its potential for use in population analysis. *Marine Mammal Science* **8**, 275-283.

Angers, B. & Bernatchez, L. 1997. Complex evolution of a salmonid microsatellite locus and its consequences in inferring allelic divergence from size information. *Molecular Biology and Evolution* **14**, 230-238.

Aquadro, C. F. & Greenberg, B. D. 1982. Human mitochondrial DNA variation and evolution: analysis of nucleotide sequences from seven individuals. *Genetics* **103**, 287-312.

Archer, F. I., II. 1996. Morphological and Genetic Variation of Striped Dolphins (*Stenella coeruleoalba*, Meyen 1833). Doctoral dissertation. Scripps Institution of Oceanography, UCSD, 187 pp.

Arctander, P. 1988. Comparative studies of avian DNA by restriction fragment length polymorphism analysis: Convenient procedures based on blood samples from live birds. *Journal für Ornithologie* **129**, 205-216.

Arnason, A., Danielsdottir, A. K., Spilliaert, R., Sigurdsson, J. H., Palsdottir, A., Duke, E. J., Joyce, P., Groves, V., Trowsdale, J. & Jonsdottir, S. 1991a. A brief review of protein and DNA mark studies in relation to the stock identity of fin whales (*Balaenoptera physalus*) from Iceland and Spain. *Report of the International Whaling Commission* **42**, 701-705.

Arnason, A. & Gullberg, A. 1993. Comparison between the complete mtDNA sequences of the blue and the fin whale, two species that can hybridize in nature. *Journal of Molecular Evolution* **37**, 312-322.

Arnason, U. 1990. Phylogeny of marine mammals - evidence from chromosomes and DNA. In *Chromosomes today*, vol. 10 (ed. K. Fredga, B. A. Kihlman & M. D. Bennet), pp. 267-278.

Arnason, U. & Best, P. B. 1991. Phylogenetic relationships within the Mysticeti (whalebone whales) based upon studies of highly repetitive DNA in all extant species. *Hereditas* **114**, 263-269.

Arnason, U., Spillaert, R., Palsdottir, A. & Arnason, A. 1991b. Molecular identification of hybrids between the two largest whale species, the blue whale (*Balaenoptera musculus*) and the fin whale (*B. physalus*). *Hereditas* **115**, 183-189.

Ashley, M. V. & Dow, B. D. 1994. The use of microsatellite analysis in population biology: Background, methods and potential applications. In *Molecular ecology and evolution: Approaches and applications* (ed. B. Schierwater, B. Streit, G. P. Wagner & R. DeSalle), pp. 185-201. Basel, Switzerland: Birkhäuser Verlag.

Avise, J. C., Giblin-Davidson, C., Laerm, J., Patton, J. C. & Lansman, R. A. 1979. Mitochondrial DNA clones and matriarchal phylogeny within and among geographic populations of the pocket gopher, *Geomys pinetis*. *Proceedings of the National Academy of Science, USA* **76**, 6694-6698.

Baker, C. S., Herman, L. M., Perry, A., Lawton, W. S., Straley, J. M., Wolman, A. A., Kaufmann, G. D., Winn, H. E., Hall, J. D., Reinke, J. M. & Ostman, J. 1986. Migratory movement and population structure of humpback whales (*Megaptera novaeangliae*) in the central and eastern North Pacific. *Marine Ecology Progress Series* **31**, 105-119.

Baker, C. S., Lambertsen, R. H., Weinrich, M. T., Calambokidis, J., Early, G. & O'Brien, S. J. 1991. Molecular genetic identification of the sex of humpback whales (*Megaptera novaeangliae*). *Report of the International Whaling Commission Special Issue* **13**, 105-111.

Baker, C. S., Palumbi, S. R., Lambertsen, R. H., Weinrich, M. T., Calambokidis, J. & O'Brien, S. J. 1990. Influence of seasonal migration on geographic distribution of mitochondrial DNA haplotypes in humpback whales. *Nature* **344**, 238-240.

Baker, C. S., Perry, A., Bannister, J. L., Weinrich, M. T., Abernethy, R. B., Calambodikis, J., Lien, J., Lambertsen, R. H., Urban Ramirez, J., Vasquez, O., Clapham, P. J., Alling, A., O'Brien, S. J. & Palumbi, S. R. 1993. Abundant mitochondrial DNA variation and world-wide population structure in

humpback whales. *Proceedings of the National Academy of Science, USA* **90**, 8239-8243.

Baker, S. C., Slade, R. W., Bannister, J. L., Abernethy, R. B., Weinrich, M. T., Lien, J., Urban, J., Corkeron, P., Calambodikis, J., Vasquez, O. & Palumbi, S. R. 1994. Hierarchical structure of mitochondrial DNA gene flow among humpback whales *Megaptera novaeangliae* world-wide. *Molecular Ecology* **3**, 313-327.

Bakke, I., Johansen, S., Bakke, Ø. & El-Gewely, M. R. 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.

Barlow, J., Brownell, R. L., DeMaster, D. P., Forney, K. A., Lowry, M. S., Osmek, S., Ragen, T. J., Reeves, R. R. & Small, R. J. 1995. U.S. Pacific Marine Mammal Stock Assessments. *NOAA Technical Memorandum NMFS* **SWFSC-219**.

Barnes, L. G., Domning, D. P. & Ray, C. E. 1985. Status of studies on fossil marine mammals. *Marine Mammal Science* **1**, 15-53.

Bérubé, M., A. Aguilar, D. Dendanto, F. Larsen, G. Notarbartolo-di-Sciara, R. Sears, J. Sigurjónsson, J. Urban-Ramirez and P. J. Palsbøll. In press. Population genetic structure of North Atlantic, Mediterranean Sea and Sea of Cortez fin whales, *Balaenoptera physalus* (Linnaeus, 1758): analysis of mitochondrial and nuclear loci. Molecular Ecology 7(5): 585-99.

Bérubé, M. 1994. Sample request agreement. La Jolla: Southwest Fisheries Science Center. Working document. GSID 31

Bérubé, M. & Palsbøll, P. J. 1996a. Identification of sex in Cetaceans by multiplexing with three *ZFX* and *ZFY* specific primers. *Molecular Ecology* **5**, 283-287.

Bérubé, M. & Palsbøll, P. 1996b. Erratum of Identification of sex in cetaceans by multiplexing with three ZFX and ZFY specific primers. *Molecular Ecology* **5**, 602.

Bérubé, M. & Sears, R. 1990. Analyses des données d'observation de rorquals dans le Golfe du St-Laurent et comparaison interannuelle de leur abondance entre 1983 et 1989. Longue-Pointe de Mingan, Québec: Peches et Océans Canada. Contrat # FP715-9-2320

Best, P. B. 1977. Two allopatric forms of Bryde's whale off South Africa. *Report of the International Whaling Commission Special Issue* **1**, 10-35.

Bigg, M. A., Ellis, G. M., Ford, J. K. & Balcomb, K. C. 1987. Killer whales: A study of their identification, genealogy and natural history in British Colombia and Washington State., pp. 79. Nanaimo, British Colombia: Phantom Press.

Blouin, M. S., Parsons, M., Lacaille, V. & Lotz, S. 1996. Use of microsatellite loci to classify individuals by relatedness. *Molecular Ecology* **5**, 393-401.

Borobia, M., Gearing, P. J., Simard, Y., Gearing, J. N. & Béland, P. 1995. Blubber fatty acids of finback and humpback whales from the Gulf of St. Lawrence. *Marine Biology* **122**, 341-353.

Bowcock, A. M., Ruiz-Linares, A., Tomfohrde, J., Minch, E., Kidd, J. R. & Cavalli-Sforza, L. L. 1994. High resolution of human evolutionary trees with polymorphic microsatellites. *Nature* **368**, 455-457.

Bradbury, M. W., Isola, L. M. & Gordon, J. W. 1990. Enzymatic amplification of a Y chromosome repeat in a single blastmore allows identification of the sex of preimplantation mouse embryos. *Proceedings of the National Academy of Science, USA* **87**, 4053-4057.

Brodie, P. F. 1975. Cetacean energetics, an overview of intraspecific size variation. *Ecology* **56**, 152-161.

Brown, M. R., Corkeron, P. J., Hale, P. T., Schultz, K. W. & Bryden, M. M. 1995. Evidence for a sex-segregated migration in the humpback whale (*Megaptera novaeangliae*). *Proceedings of the Royal Society, London, Series B.* **259**, 229-234.

Brown, M. W., Helbig, R., Boag, P. T., Gaskin, D. E. & White, B. N. 1991. Sexing beluga whales (*Delphinapterus leucas*) by means of DNA markers. *Canadian Journal of Zoology* **69**, 1971-1976.

Brown, S. G. 1978. Whale marking techniques. In *Animal marking: Recognition marking of animals in research.* (ed. B. Stonehouse), pp. 71-80. Macmillan Press Ltd.

Brown, W. M. 1985. The mitochondrial genome of animals. In *Molecular Evolutionary Genetics* (ed. R. J. MacIntyre), pp. 95-130. New York: Plenum Press.

Brown, W. M., George, M., Jr. & Wilson, A. C. 1979. Rapid evolution of animal mitochondrial DNA. *Proceedings of the National Academy of Science, USA* **76**, 1967-1971.

Bruford, M. W. & Wayne, R. K. 1993. Microsatellites and their application to population studies. *Current Opinion in Genetics and Development* **3**, 939-943.

Buckland, S. T., Cattanach, K. L. & Gunnlaugsson, T. 1992a. Fin whale abundance in the North Atlantic, estimated from Icelandic and Faroese NASS-87 and NASS-89 data. *Report of the International Whaling Commission* **42**, 645-651.

Buckland, S. T., Cattanach, K. L. & Lens, S. 1992b. Fin whale abundance in the eastern North Atlantic, estimated from Spanish NASS-89 data. *Report of the International Whaling Commission* **42**, 457-460.

Burke, T., Hanotte, O., Bruford, M. W. & Cairns, E. 1991. Multilocus and singlelocus minisatellite analysis in population biological studies. In *DNA fingerprinting: Approaches and Applications* (ed. G. D. T. Burke, A.J. Jeffreys & R. Wolff), pp. 154-167. Switzerland: Birkhäuser Verlag Basel.

Cagnolaro, L., Cozzi, B., Magnaghi, L., Podesta, M., Poggi, R. & Tangerini, P. 1986. Su 18 cetacei spiaggiati sulle coste italiane dal 1981 al 1985: rilevamento biometrico ed osservazioni necroscopiche (Mammalia: Cetacea). *Atti Soc. Ita. Sci. Nat. Mus. Civ. St. Nat. Milano* **127**, 79-106.

Carlson, C. A., Mayo, C. A. & Whitehead, H. 1990. Changes in the ventral fluke pattern of the humpback whale (*Megaptera novaeangliae*), and its effects on matching; evaluation of its significance to photo-identification research. *Report of the International Whaling Commission Special Issue* **12**, 105-111.

Carr, S. M. & Hughes, G. A. 1993. Direction of introgressive hybridization between species of North American deer (*Odocoileis*) as inferred from mitochondrial-cytochrome-b sequences. *Journal of Mammalogy* **74**, 331-342.

Chittleborough, R. G. 1953. Arial observations on the humpback whale, *Megaptera nodosa* (Bonaterre), with notes on other species. *Australian Journal of Marine and Freshwater Research* , 219-226.

Clapham, P. J., Baraff, L. S., Carlson, C. A., Christian, M. A., Mattila, D. K., Mayo, C. A., Murphy, M. A. & Pittman, S. 1993a. Seasonal occurrence and annual return of humpback whales in the southern Gulf of Maine. *Canadian Journal of Zoology* **71**, 440-443.

Clapham, P. J., Bérubé, M. & Mattila, D. K. 1995. Sex ratio of the Gulf of Maine humpback whale population. *Marine Mammal Science* **11**, 227-231.

Clapham, P. J., Palsbøll, P. J. & Mattila, D. K. 1993b. High-energy behaviors in humpback whales as a source of sloughed skin for molecular analysis. *Marine Mammal Science* **9**, 213-220.

Clapham, P. J., Palsbøll, P. J., Mattila, D. K. & Vasquez, O. 1993c. Composition and dynamics of humpback whale competitive groups determined by molecular analysis. *Behavior* **122**, 182-194.

Clapham, P. J. & Seipt, I. E. 1991. Resightings of independent fin whales, *Balaenoptera physalus*, on maternal summer ranges. *Journal of Mammalogy* **72**, 788-790.

Committee for Whaling Statistics. 1967. *International Whaling Statistics* **59**, 55.

Coyne, J. A. 1992. Genetics and speciation. *Nature* **355**, 511-515.

Danielsdottir, A. K., Duke, E. J. & Arnason, A. 1992. Mitochondrial DNA analysis of North Atlantic fin whales (*Balaenoptera physalus*) and comparison with four species of whales: sei (*B. borealis*), minke (*B. acutorostrata*), pilot

(*Globicephala melaena*) and sperm whales (*Physeter catodon*). Reykjavik: International Whaling Commission. SC/F91/F17

Danielsdottir, A. K., Duke, E. J., Joyce, P. & Arnason, A. 1991. Preliminary studies on the genetic variation at enzyme loci in fin whales (*Balaenoptera physalus*) and sei whales (*Balaenoptera borealis*) from the North Atlantic. *Report of the International Whaling Commission Special Issue* **13**, 115-124.

Danielsdottir, A. K., Sigurjónsson, J., Mitchell, E. & Arnason, A. 1993. Report on a pilot study of genetic variation in North Atlantic fin whales (*Balaenoptera physalus*): International Whaling Commission. SC/44/NAB16

Deka, R., Shriver, M. D., Yu, L. M., Aston, C. E., Chakraborty, R. & Ferrell, R. E. 1994. Conservation of human chromosome 13 polymorphic microsatellite $(CA)_n$ repeats in chimpanzees. *Genomics* **22**, 226-230.

Di Rienzo, A., Peterson, A. C., Garza, J. C., Valdes, A. M., Slatkin, M. & Freimer, N. B. 1994. Mutational processes of simple-sequence repeat loci in human populations. *Proceedings of the National Academy of Science, USA* **91**, 3166-3170.

Dobzhansky, T. 1970. *The Genetics of the Evolutionary Process*. New York, N. Y.: Columbia University Press.

Donovan, G. P. 1991. A review of IWC stock boundaries. *Report of the International Whaling Commission Special Issue* **13**, 39-68.

Doroshenko, N. V. 1970. A whale with features of fin whale and blue whale. *Tinro* **70**, 255-257.

Duguy, R. 1989. Les mammiferes marins de la Mediterranee occidentale. *Bulletin de la Société Zoologique de France* **114**, 89-92.

Duguy, R., Aguilar, A., Casinos, A., Grau, E. & Raga, J. A. 1988. Etude comparative des échouages de cétacés sur les Côtes Méditérranéennes de France et d'Espagne. *Miscelanea Zoologica* **12**, 339-345.

Edds, P. L. & Macfarlane, J. A. F. 1987. Occurrence and general behavior of balaenopterid cetaceans summering in the St. Lawrence Estuary, Canada. *Canadian Journal of Zoology* **65**, 1363-1376.

Edwards, A., Civitello, A., Hammond, H. A. & Caskey, C. T. 1991. DNA typing and genetic mapping with trimeric and tetrameric tandem repeats. *American Journal of Human Genetics* **49**, 746-756.

Ellegren, H., Primmer, C. R. & Sheldon, B. C. 1995. Microsatellite 'evolution': directional bias? *Nature Genetics* **11**, 360-362.

Engels, B. 1992. *Amplify for analyzing PCR experiments*. 1.2B. University of Wisconsin, Madison, WI.

Eronen, M. & Olander, M. 1990. On the world's ice ages and changing environments. Helsinki: Nuclear waste commission of Finnish power companies. YJT-90-13

Estoup, A., Tailiez, C., Cornuet, J.-M. & Solignac, M. 1995. Size homoplasy and mutational processes of interrupted microsatellites in two bee species, *Apis mellifera* and *Bombus terrestris* (Apidae). *Molecular Biology and Evolution* **12**, 1074-1084.

Feldman, M. W., Bergman, A., Pollock, D. D. & Goldstein, D. B. 1997. Microsatellite genetic distances with range constraints: analytic description and problems of estimation. *Genetics* **145**, 207-216.

Felsenstein, J. 1993. *PHYLIP*. 3.5c. Seattle: Manual and program available from Joe Felsenstein, University of Washington, WA.

FitzSimmons, N. N., Moritz, C. & Moore, S. S. 1995. Conservation and dynamics of microsatellite loci over 300 million years of marine turtle evolution. *Molecular Biology and Evolution* **12**, 432-440.

Forcada, J., Aguilar, A., Hammond, P., Pastor, X. & Aguilar, R. 1996. Distribution and abundance of fin whales (*Balaenoptera physalus*) in the western Mediterranean sea during the summer. *Journal of Zoology, London* **238**, 23-34.

Fraser, F. C. 1940. Three anomalous dolphins from Blacksod Bay, Ireland. *Proceedings of the Royal Irish Academy* **45, Section B**, 413-455 and plates XXXII-XXXVIII.

Gambell, R. 1985. Fin whale *Balaenoptera physalus* (Linnaeus,1758). In *Handbook of Marine Mammals*, vol. 3, pp. 171-192. London: Academic Press Inc.

Garza, J. C., Slatkin, M. & Freimer, N. B. 1995. Microsatellite allele frequencies in humans and chimpanzees, with implications for constraints on allele size. *Molecular Biology and Evolution* **12**, 594-603.

Gaskin, D. E. 1982. *The Ecology of Whales and Dolphins*. London and Portsmouth, New Hampshire: Heinemann.

Gauthier, J. M., Metcalfe, C. D. & Sears, R. 1997. Chlorinated organic contaminants in blubber biopsies from Northwestern Atlantic balaenopterid whales summering in the Gulf of St. Lawrence. *Marine Environmental Research* **44**, 201-223.

Giard, J. 1996. Evaluation de l'impact des activités d'observation sur le comportement de plongée des rorquals communs, *Balaenoptera physalus*, de l'Estuaire du Saint-Laurent à l'aide de la télémétrie VHF. Mémoire de Maîtrise. Biologie Université Laval, 53 pp.

Gill, P., Jeffreys, A. J. & Werrett, D. J. 1985. Forensic application of DNA "fingerprints". *Nature* **318**, 577-579.

Glockner, D. A. 1983. Determining the sex of humpback whales (*Megaptera novaeangliae*) in their natural environment. In *Communication and Behavior of whales. AAS Selected Symposia Series* (ed. R. Payne), pp. 295-329. Boulder, Colorado: Westview Press.

Glockner-Ferrari, D. A. & Ferrari, M. J. 1990. Reproduction in the humpback whale (*Megaptera novaeangliae*) in Hawaii waters, 1975-1988: the life history, reproductive rates and behavior of known individuals through surface and underwater photography. *Report of the International Whaling Commission Special issue* **12**, 161-168.

Goldstein, D. B., Linares, A. R., Cavelli-Sforza, L. L. & Feldman, M. W. 1995a. An evaluation of genetic distances for use with microsatellite loci. *Genetics* **139**, 463-471.

Goldstein, D. B., Ruiz Linares, A., Cavalli-Sforza, L. L. & Feldman, M. W. 1995b. Genetic absolute dating based upon microsatellites and the origin of modern humans. *Proceedings of the National Academy of Science, USA* **92**, 6723-6727.

Goodman, S. J. 1997. RST Calc: a collection of computer programs for calculating estimates of genetic differentiation from microsatellite data and determining their significance. *Molecular Ecology* **6**, 881-885.

Gyllensten, U., Wharton, D., Josefsson, A. & Wilson, A. C. 1991. Paternal inheritance of mitochondrial DNA in mice. *Nature* **352**, 255-257.

Hamilton, P. K. & Mayo, C. A. 1990. Population Characteristics of Right Whales *Eubalaena Glacialis* Observed in Cape Cod and Massachusetts Bays USA 1978-1986. *Report of the International Whaling Commission Special Issue* **12**, 203-208.

Hammond, P. S., Mizroch, S. A. & Donovan, G. P. 1990. *Individual recognition of cetaceans: use of photo-identification and other techniques to estimate population parameters*. Report of the International Whaling Commission Special Issue. Cambridge: International Whaling Commission.

Hartl, D. L. & Clark, A. G. 1989. *Principles of population genetics*. Massachusetts: Sinauer Associates, Inc.

Heide-Jørgensen, M. P. & Reeves, R. R. 1993. Description of an anomalous monodontid skull from West Greenland: A possible hybrid? *Marine Mammal Science* **9**, 258-268.

Herskowitz, I. H. 1977. *Principles of genetics*. London: Collier Macmillian.

Higuchi, R., Beroldingen, C. H. v., Sensabaugh, G. F. & Erlich, H. A. 1988. DNA typing from single hairs. *Nature* **332**, 543-546.

Hoelzel, A. R. & Amos, W. 1988. DNA fingerprinting and 'scientific' whaling. *Nature* **333**, 305.

Hoelzel, A. R., Hancock, J. M. & Dover, G. A. 1991. Evolution of the cetacean mitochondrial control region. *Molecular Biology and Evolution* **8**, 475-493.

Hoelzel, R. A. 1994. Genetics and ecology of whales and dolphins. *Annual Review of Ecology and Systematics* **25**, 377-399.

Holt, S. J. 1977. Question about the sex-ratio in catches of rorquals. *Report of the International Whaling Commission* **27**, 138-140.

Höss, M., Kohn, M., Pääbo, S., Knauer, F. & Schröder, W. 1992. Excrement analysis by PCR. *Nature* **359**, 199.

Hubby, J. L. & Lewontin, R. C. 1966. A molecular approach to the study of genic heterozygosity in natural populations. I. The number of alleles at different loci in *Drosophila pseudoobscura*. *Genetics* **54**, 577-594.

Hudson, R. R., Boos, D. D. & Kaplan, N. L. 1992. A statistical test for detecting geographic subdivision. *Molecular Biology and Evolution* **9**, 138-151.

Hughes, C. R. & Queller, D. C. 1993. Detection of highly polymorphic microsatellite loci in a species with little allozyme polymorphism. *Molecular Ecology* **2**, 131-137.

Hutchinson, C. A. I., Newbold, J. E., Potter, S. S. & Edgell, M. H. 1974. Maternal inheritance of mammalian mitochondrial DNA. *Nature* **251**, 536-538.

Ingebrigtsen, A. 1929. Whales caught in the North Atlantic and other seas. *Conseil International pour l'Exploration Scientific du Mer* **56**, 1-26 (finback only, 12-17).

IWC. 1992. Report of the comprehensive assessment special meeting on North Atlantic fin whales. *Report of the International Whaling Commission* **42**, 595-644.

Jeffreys, A. J., Wilson, V. & Thein, S. L. 1985. Individual-specific 'fingerprints' of human DNA. *Nature* **316**, 76-79.

Jonsgård, Å. 1966. The distribution of Balaenopteridae in the North Atlantic Ocean. In *Whales, Dolphins, and Porpoises.* (ed. K. S. Norris), pp. 114-124. Berkeley and Los Angeles: University of California Press.

Kapel, F. O. 1979. Exploitation of large whales in West Greenland in the Twentieth Century. *Report of the International Whaling Commission* **29**, 197-214.

Karl, S. A., Bowen, B. W. & Avise, J. C. 1995. Hybridization among the ancient mariners: Characterization of marine turtles hybrids with molecular genetic assays. *Journal of Heredity* **86**, 262-268.

Katona, S. K. & Beard, J. A. 1990. Population size, migrations, and feeding aggregations of the humpback whale, *Megaptera novaeangliae*, in the western North Atlantic Ocean. *Report of the International Whaling Commission Special Issue* **12**, 295-305.

Katona, S. K. & Whitehead, H. P. 1981. Identifying humpback whales using their natural markings. *Polar Record* **20**, 439-444.

Kawamura, A. 1980. A review of food of balaenopterid whales. *Scientific Reports of the Whales Research Institute, Tokyo* **32**, 155-197.

Kellogg, R. 1929. What is known of the migration of some of the whalebone whales. *Annual Report of the Smithsonian Institution* **1928**, 467-494.

Kimmel, M., Chakraborty, R., Stivers, D. N. & Deka, R. 1996. Dynamics of repeat polymorphisms under a forward-backward mutation model: within- and between-population variability at microsatellite loci. *Genetics* **143**, 549-555.

Kimura, M. 1980. A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution* **16**, 111-120.

Klinowska, M. 1991. *Dolphins, porpoises and whales of the world.* Gland: IUCN.

Knowlton, A. R., Sigúrjonsson, J., Ciano, J. N. & Kraus, S. D. 1992. Long-distance movements of North Atlantic right whales, *Eubalaena glacialis*. *Marine Mammal Science* **8**, 397-404.

Kuhnlein, U., Dawe, Y., Zadworny, D. & Gavora, J. S. 1989. DNA fingerprinting: a tool for determining genetic distances between strains of poultry. *Theoretical and Applied Genetics* **77**, 669-672.

Kumar, S., Tamura, K. & Nei, M. 1993. *MEGA: Molecular Evolutionary Genetics Analysis*. 1.01. University Park: The Pennsylvania State University.

Lambertsen, R. H. 1987. A biopsy system for large whales and its use for cytogenetics. *Journal of Mammalogy* **68**, 443-445.

Lansman, R. A., Avise, J. C., Aquadro, C. F., Shapira, J. F. & Daniel, S. W. 1983. Extensive genetic variation in mitochondrial DNA's among geographic populations of the deer mouse, *Peromyscus maniculatus*. *Evolution* **37**, 1-15.

Larsen, A. H., Sigurjónsson, J., Øien, N., Vikingsson, G. & Palsbøll, P. J. 1996. Population genetic analysis of nuclear and mitochondrial loci in skin biopsies collected from Central and northeastern North Atlantic humpback whales (*Megaptera novaeangliae*): population identity and migratory destinations. *Proceedings of the Royal Society, London, Series B.* **263**, 1611-1618.

Larsen, F. & Øien, N. 1988. On the discreteness of stocks of minke whales at East and West Greenland. *Report of the International Whaling Commission* **38**, 251-255.

Leatherwood, S., Reeves, R. R., Perrin, W. F. & Evans, W. 1988. *Whales, Dolphins and Porpoises of the Eastern North Pacific and Adjacent Arctic Waters. A Guide to Their Identification.* New York: Dover Publications, Inc.

Lewontin, R. C. & Hubby, J. L. 1966. A molecular approach to the study of genic heterozygosity in natural populations II. Amount of variation and degree of heterozygosity in natural populations of *Drosophila pseudoobscura*. *Genetics* **54**, 595-609.

Li, W.-H., Wu, C.-I. & Luo, C.-C. 1985. A new method for estimating synonymous and nonsynonymous rates of nucleotide substitution considering the relative likelihood of nucleotide and codon changes. *Molecular Biology and Evolution* **2**, 150-174.

Lindgren, B. W. 1975. *Basic ideas of statistics*. New York: Macmillian Publishing Co., Inc.

Lockyer, C. H. 1984. Review of baleen whale (Mysticeti) reproduction and implications for management. *Report of the International Whaling Commission Special Issue* **6**, 27-50.

Lyrholm, T., Leimar, O. & Gyllensten, U. 1996. Low Diversity and Biased Substitution Patterns in the Mitochondrial DNA Control Region of Sperm

Whales: Implications for Estimates of Time Since Common Ancestry. *Molecular Biology and Evolution* **13**, 1318-1327.

Maniatis, T., Fritsch, E. F. & Sambrock, J. 1982. *Molecular cloning. A laboratory manual.* Cold Spring Harbor: Cold Spring Harbour Laboratory.

Mantel, N. 1967. The detection of disease clustering and a generalized regression approach. *Cancer Research* **27**, 209-220.

Marsili, L. & Focardi, S. 1996. Organochlorine levels in subcutaneous blubber biopsies of fin whales (*Balaenoptera physalus*) and striped dolphins (*Stenella coeruleoalba*) from the Mediterranean sea. *Environmental Pollution* **91**, 1-9.

Mathews, E., Keller, S. & Weiner, D. B. 1988. A method to collect and process skin biopsies for cell culture from free-ranging grey whales (*Eschrictius robustus*). *Marine Mammal Science* **4**, 1-12.

Mattila, D. K., Clapham, P. J., Katona, S. K. & Stone, G. S. 1989. Population composition of humpback whales, *Megaptera novaeangliae*, on Silver Bank, 1984. *Canadian Journal of Zoology* **67**, 281-285.

Mayr, E. 1963. *Animal species and evolution.* Cambridge: Harvard University Press.

Medrano-G., L., Aguayo-L., A., Urbán-R., J. & Baker, C. S. 1995. Diversity and distribution of mitochondrial DNA lineages among humpback whales, *Megaptera novaeangliae*, in the Mexican Pacific ocean. *Canadian Journal of Zoology* **73**, 1735-1743.

Medrano-G., L., Salinas, M., Salas, I., Ladrón de Guevara, P., Aquayo, A., Jacobsen, J. & Baker, C. S. 1994. Sex identification of humpback whales, *Megaptera novaeangliae*, on the wintering ground of the Mexican Pacific Ocean. *Canadian Journal of Zoology* **72**, 1771-1774.

Minkoff, E. C. 1983. *Evolutionary biology.* Reading, Massachusetts: Addison-Wesley Publishing Co.

Mitchell, E. 1972. Assessments of Northwest Atlantic fin whale stocks. *Report of the International Whaling Commission* **22**, 111-118.

Mitchell, E. 1974. Present status of Northwest Atlantic fin and other whale stocks. In *The whale problem: A status report* (ed. W. E. Schevill), pp. 108-169. Cambridge: Harvard University Press.

Mitchell, E. & Reeves, R. R. 1983. Catch history, abundance and present status of Northwest Atlantic humpback whales. In *Special issue on historical whaling records*, vol. 5 (ed. M. F. Tillman & G. P. Donovan), pp. 153-212. Cambridge: International Whaling Commission.

Miyazaki, N., Hirosaki, Y., Kinuta, T. & Omura, H. 1992. Osteological study of a hybrid between *Turiops truncatus* and *Grampus griseus*. *Bulletin of the Natural Science Museum, Tokyo, (Serie A)* **18**, 79-94.

Mullis, K. B. & Faloona, F. 1987. Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. *Methods in Enzymology* **155**, 335-350.

Nakahori, Y., Hamano, K., Iwaya, M. & Nakagome, Y. 1991. Sex identification by polymerase chain reaction using X-Y homologous primer. *American Journal of Medical Genetics* **39**, 472-473.

Nei, M. 1987. *Molecular evolutionary genetics*. New York: Columbia University Press.

Nemoto, T. 1964. School of baleen whales in the feeding areas. *Science Report for Whales Research Institute* **18**, 89-110.

Norris, K. S. 1961. Standardized methods for measuring and recording data on the smaller cetaceans. *Journal of Mammalogy* **42**, 471-476.

Notarbartolo di Sciara, G., Bérubé, M., Zanardelli, M. & Panigada, S. 1996. The role of the Mediterranean Sea in fin whale ecology: insight through genetics. *European Research on Cetaceans* **9**, 218.

Ohta, T. & Kimura, M. 1973. A model of mutation appropriate to estimate the number of electrophoretically detectable alleles in a finite population. *Genetical Research* **22**, 201-204.

Paetkau, D. & Strobeck, C. 1994. Microsatellite analysis of genetic variation in black bear populations. *Molecular Ecology* **3**, 489-495.

Paetkau, D. & Strobeck, C. 1995. The molecular basis and evolutionary history of a microsatellite null allele in bears. *Molecular Ecology* **4**, 519-520.

Palsbøll, P. J. 1989. Restriction fragment pattern analysis of mitochondrial DNA in the North Atlantic minke whale, *Balaenoptera acutorostrata. Master of Science dissertation*. Department of Genetics Institute of Molecular Biology, University of Copenhagen, 42 pp.

Palsbøll, P. J. 1994. Heterogeneity, sample sizes, and levels of significance. La Jolla: Southwest Fisheries Science Center. Working document. GSID 29

Palsbøll, P. J., Allen, J., Bérubé, M., Clapham, P. J., Feddersen, T. P., Hammond, P., Hudson, R. R., Jørgensen, H., Katona, S., Larsen, A. H., Larsen, F., Lien, J., Mattila, D. K., Sigurjónsson, J., Sears, R., Smith, T., Sponer, R., Stevick, P. T. & Øien, N. 1997a. Genetic tagging of humpback whales. *Nature* **388**, 676-679.

Palsbøll, P. J., Bérubé, M. & Jørgensen, H. In press. Multiple levels of single-strand slippage at cetacean tri- and tetramer microsatellite loci. *Genetics* .

Palsbøll, P. J., Bérubé, M., Larsen, A. H. & Jørgensen, H. 1997. Primers for the amplification of tri- and tetramer microsatellite loci in cetaceans. *Molecular Ecology* **6**, 893-895

Palsbøll, P. J., Clapham, P. J., Mattila, D. K., Larsen, F., Sears, R., Siegismund, H. R., Sigurjónsson, J., Vasquez, O. & Arctander, P. 1995. Distribution of mtDNA haplotypes in North Atlantic humpback whales: the influence of behaviour on population structure. *Marine Ecology Progress Series* **116**, 1-10.

Palsbøll, P. J., Heide-Jørgensen, M. P. & Dietz, D. 1997b. Population structure and seasonal movements of narwhals, *Monodon monoceros*, determined from mtDNA analysis. *Heredity* **78**, 284-292.

Palsbøll, P. J., Larsen, F. & Sigurd Hansen, E. 1991. Sampling of skin biopsies from free-ranging large cetaceans in West Greenland: Development of new biopsy tips and bolt designs. *Report of the International Whaling Commission Special Issue* **13**, 71-79.

Palsbøll, P. J., Vader, A., Bakke, I. & El-Gewely, M. R. 1992. Gender determination in cetaceans by the polymerase chain reaction. *Canadian Journal of Zoology* **70**, 2166-2170.

Palumbi, S. R. 1992. Marine speciation on a small planet. *Trends in Ecology and Evolution* **7**, 114-118.

Palumbi, S. R. & Baker, C. S. 1994. Contrasting population structure from nuclear intron sequences and mtDNA of humpback whales. *Molecular Biology and Evolution* **11**, 426-435.

Patenaude, N. J. & White, B. N. 1995. Skin biopsy sampling of beluga whale carcasses: assessment of biopsy darting factors for minimal wounding and effective retrieval. *Marine Mammal Science* **11**, 163-171.

Pemberton, J. M., Slate, J., Bancroft, D. R. & Barrett, J. A. 1995. Nonamplifying alleles at microsatellite loci: a caution for parentage and population studies. *Molecular Ecology* **4**, 249-252.

Primmer, C. R., Saino, N. & Møller, A. P. 1996. Directional evolution in germline microsatellite mutations. *Nature Genetics* **13**, 391-393.

Queller, D. C. & Goodknight, K. F. 1989. Estimating relatedness using genetic markers. *Evolution* **43**, 258-275.

Queller, D. C., Strassmann, J. E. & Hughes, C. R. 1993. Microsatellites and kinship. *Trends in Ecology and Evolution* **8**, 285-288.

Rassmann, K., Schlötterer, C. & Tautz, D. 1991. Isolation of simple-sequence loci for use in polymerase chain reaction-based DNA fingerprinting. *Electrophoresis* **12**, 113-118.

Raymond, M. & Rousset, F. 1995. GENEPOP (Version1.2): population genetics software for exact tests and ecumenicism. *Journal of Heredity* **86**, 248-249.

Reed, J. Z., Tollit, D. J., Thompson, P. M. & Amos, W. 1997. Molecular scatology: the use of molecular genetic analysis to assign species, sex and individual identity to seal feces. *Molecular Ecology* **6**, 225-234.

Reyes, J. C. 1996. A possible case of hybridism in wild dolphins. *Marine Mammal Science* **12**, 301-307.

Rice, W. R. 1989. Analyzing tables of statistical tests. *Evolution* **43**, 223-225.

Richard, K. R., Dillion, M. C., Whitehead, H. & Wright, J. M. 1996a. Patterns of kinship in groups of free-living sperm whales (*Physeter macrocephalus*) revealed by multiple molecular genetic analyses. *Proceedings of the National Academy of Science, USA* **93**, 8792-8795.

Richard, K. R., McCarry, S. W. & Wright, J. M. 1994. DNA sequence from the *SRY* gene of the sperm whale (*Physeter macrocephalus*) for use in molecular sexing. *Canadian Journal of Zoology* **72**, 873-877.

Richard, K. R., Whitehead, H. & Wright, J. M. 1996b. Polymorphic microsatellites from sperm whales and their use in the genetic identification of individuals from naturally sloughed pieces of skin. *Molecular Ecology* **5**, 313-315.

Rosel, P. E., Dizon, A. E. & Haygood, M. G. 1995. Variability of the mitochondrial control region in populations of the harbour porpoise, *Phocoena phocoena*, on interoceanic and regional scales. *Canadian Journal of Fisheries and Aquatic Sciences* **52**, 1210-1219.

Rothuizen, J., Wolfswinkel, J., Lenstra, J. A. & Frants, R. R. 1994. The incidence of mini- and microsatellite repetitive DNA in the canine genome. *Theoretical and Applied Genetics* **89**, 403-406.

Rousset, F. 1996. Equilibrium values of measures of population subdivision for stepwise mutation processes. *Genetics* **142**, 1357-1362.

Saccone, C., Attimonelli, M. & Sbisá, E. 1987. Structural elements highly preserved during the evolution of the D-loop-containing region in vertebrate mitochondrial DNA. *Journal of Molecular Evolution* **26**, 205-211.

Saccone, C., Pesole, G. & Sbisá, E. 1991. The main regulatory region of mammalian mitochondrial DNA: structure-function model and evolutionary pattern. *Journal of Molecular Evolution* **33**, 83-91.

Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B. & Erlich, H. A. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* **239**, 487-491.

Sambrook, J., Fritsch, E. F. & Maniatis, T. 1989. *Molecular cloning. A laboratory manual.* Cold Spring Harbor: Cold Spring Harbor Laboratory Press.

Sanger, F., Nicklen, S. & Coulson, A. R. 1977. DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Science, USA* **74**, 5463-5467.

Sasi, R., Fan, Y.-S. & Lin, C. C. 1991. Prenatal sexing and detection of ZFY gene sequences in sex chromosomes disorders by polymerase chain reaction. *Journal of Clinical Laboratory Analysis* **5**, 193-196.

Savage, D. E. 1983. *Mammalian paleofaunas of the world.* London: Addison-Wesley.

Scarff, J. E. 1977. The International Management of whales, dolphins, and porpoises: An interdisciplinary assessment (part one). *Ecology Law Quarterly* **6**, 323-428.

Schaeff, C. M., Kraus, S. D., Brown, M. & White, B. N. 1993. Assessment of the population structure of western North Atlantic right whales (*Eubalaena glacialis*) based on sighting and mtDNA data. *Canadian Journal of Zoology* **71**, 339-345.

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

Schlötterer, C. & Pemberton, J. 1994. The use of microsatellites for genetic analysis of natural populations. In *Molecular ecology and evolution: Approaches and applications* (ed. B. Schierwater, B. Streit, G. P. Wagner & R. DeSalle), pp. 203-214. Basel, Switzerland: Birkhäuser Verlag.

Schlötterer, C. & Tautz, D. 1992. Slippage synthesis of simple sequence DNA. *Nucleic Acids Research* **20**, 211-215.

Schneider-Gädicke, A., Beer-Romero, P., Brown, L. G., Nussbaum, R. & Page, D. C. 1989. *ZFX* has a gene structure similar to *ZFY*, the putative human sex determinant, and escapes X inactivation. *Cell* **57**, 1247-1258.

Sears, R., Wenzel, F. W. & Williamson, J. W. 1987. The blue whale: A catalogue of individuals from the western North Atlantic (Gulf of St. Lawrence)., pp. 27. St- Lambert, Qc. Mingan Island Cetacean Study (MICS Inc.).

Sears, R., Williamson, J. M., Wenzel, F. W., Bérubé, M., Gendron, D. & Jones, P. 1990. Photographic identification of the blue whale (*Balaenoptera musculus*) in the Gulf of St. Lawrence, Canada. *Report of the International Whaling Commission Special Issue* **12**, 353-342.

Seber, G. A. F. 1982. *The estimation of animal abundance and related parameters*. London: Charles Griffin & Company Ltd.

Seipt, I. E., Clapham, P. J., Mayo, C. A. & Hawvermale, M. P. 1989. Population characteristics of individually identified fin whales *Balaenoptera physalus* in Massachusetts Bay. *Fisheries Bulletin, U.S.* **88**, 271-278.

Sergeant, D. E. 1977. Stocks of fin whales *Balaenoptera physalus* L. in the North Atlantic ocean. *Report of the International Whaling Commission* **27**, 460-473.

Shimura, E., Numachi, K., Sezaki, K., Hirosaki, Y., Watabe, S. & Hashimoto, K. 1986. Biochemical evidence of hybrid formation between the two species of dolphin *Tursiops truncatus* and *Grampus griseus*. *Bulletin of the Japanese Society of Scientific Fisheries* **52**, 725-730.

Sigurjónsson, J., Mitchell, E. & Gunnlaugsson, T. 1991. Fin whale markings in the North Atlantic with special reference to the stock identity question. Reykjavik: The Icelandic State Conference Center. SC/F91/F20

Slatkin, M. 1995. A measure of population subdivision based on microsatellite allele frequencies. *Genetics* **139**, 457-462.

Slatkin, M. & Hudson, R. R. 1991. Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. *Genetics* **129**, 555-562.

Sokal, R. R. & Rohlf, F. J. 1995. *Biometry: the principles of statistics in biological research*. New York: Freeman and Company.

Southern, S. O., Southern, P. J. & Dizon, A. E. 1988. Molecular characterization of a cloned dolphin mitochondrial genome. *Journal of Molecular Evolution* **28**, 32-42.

Spilliaert, R., Vikingsson, G., Arnason, U., Palsdottir, A., Sigurjonsson, J. & Arnason, A. 1991. Species hybridization between a female blue whale (*Balaenoptera musculus*) and a male fin whale (*B. physalus*): molecular and morphological documentation. *Journal of Heredity* **82**, 269-274.

Stone, G. S., Florez-Gonzalez, L. & Katona, S. 1990. Whale migration record. *Nature* **346**, 705.

Stone, G. S., Kraus, S. D., Prescott, J. H. & Hazard, K. W. 1988. Significant aggregations of the endangered right whale, *Eubalaena glacialis*, on the continental shelf of Nova Scotia (Canada). *Canadian Field-Naturalist* **102**, 471-474.

Stonehouse, B. 1978. *Animal Marking: Recognition marking of animals in research*. London: MacMillan Press.

Sylvestre, J.-P. & Tasaka, S. 1985. On the intergenic hybrids in cetaceans. *Aquatic Mammals* **11**, 101-108.

Taberlet, P. & Bouvet, J. 1992. Bear conservation genetics. *Nature* **358**, 197.

Taberlet, P., Mattock, H., Dubois-Paganon, C. & Bouvet, J. 1993. Sexing free-ranging brown bears *Ursus arctos* using hairs found in the field. *Molecular Ecology* **2**, 399-403.

Takezaki, N. & Masatoshi, N. 1996. Genetic distance and reconstruction of phylogenetic tress from microsatellite DNA. *Genetics* **144**, 389-399.

Tarasevich, M. N. 1967. On the composition of Cetacea groupings. 2. Groupings of fin whales. *Zoologicheskii Zhurnal* **46**, 420-431.

Tautz, D. 1989. Hypervariability of simple sequences as a general source for polymorphic DNA markers. *Nucleic Acids Research* **17**, 6463-6471.

Tautz, D. & Renz, M. 1984. Simple sequences are ubiquitous repetitive components of eukaryote genomes. *Nucleic Acids Research* **12**, 4127-4138.

Tautz, D. & Schlötterer, C. 1994. Simple sequences. *Current Opinion in Genetics and Development* **4**, 832-837.

Tautz, D., Trick, M. & Dover, G. A. 1986. Cryptic simplicity in DNA is a major source of genetic variation. *Nature* **322**, 652-656.

Tershy, B. R., Breese, D. & Strong, C. S. 1990. Abundance, seasonal distribution and population composition of balaenopterid whales in the Canal de Ballenas, Gulf of California, Mexico. *Report of the International Whaling Commission Special Issue* **12**, 369-375.

Tershy, B. R., Urban-R., J., Breese, D., Rojas-B., L. & Findley, L. Y. 1993. Are fin whales resident to the Gulf of California? *Revista de Investigación Científica* **1**, 69-72.

Tønnessen, J. N. & Johnsen, A. O. 1982. *The history of modern whaling*. Berkeley: University of California Press.

Urban-R., J. 1996. La poblaciná del rorcual común *Balaenoptera physalus* en el Golfo de California. Final Report. CONABIO-B040

Valdes, A. M., Slatkin, M. & Freimer, N. B. 1993. Allele frequencies at microsatellite loci: The stepwise mutation model revisited. *Genetics* **133**, 737-749.

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

Valsecchi, E., Palsbøll, P., Hale, P., Glockner-Ferrari, D., Ferrari, M., Clapham, P., Larsen, F., Mattila, D., Sears, R., Sigurjónsson, J., Brown, M., Corkeron, P. & Amos, B. 1997. Microsatellite genetic distances between oceanic populations of the humpback whale (*Megaptera novaeangliae*). *Molecular Biology and Evolution* **14**, 355-362.

Van Pijlen, I. A., Amos, B. & Burke, T. 1995. Patterns of genetic variability at individual loci in minke whale *Balaenoptera acutorostrata* populations from three different oceans. *Molecular Biology and Evolution* **12**, 459-472.

Vilotte, J.-L., Soulier, S., Mercier, J.-C., Gaye, P., Hue-Delahaie, D. & Furet, J.-P. 1987. Complete nucleotide sequence of bovine α-lactalbumin gene: comparison with its rat counterpart. *Biochimie* **69**, 609-620.

Wada, S. & Numachi, K. I. 1991. Allozyme analyses of genetic differentiation among the populations and species of the *Balaenoptera*. *Report of the International Whaling Commission Special Issue* **13**, 125-154.

Wayne, R. K. & Jenks, S. M. 1991. Mitochondrial DNA analysis implying extensive hybridization of the endangered red wolf *Canis rufus*. *Nature* **351**, 565-568.

Weber, J. L. & May, P. E. 1989. Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. *American Journal of Human Genetics* **44**, 388-396.

Weinrich, M. T., Lambertsen, R. H., Baker, C. S., Schilling, M. R. & Belt, C. R. 1991. Behavioural responses of humpback whales (*Megaptera novaeangliae*) in the Southern Gulf of Maine to biopsy sampling. *Report of the International Whaling Commission Special Issue* **13**, 91-97.

Weir, B. S. 1990. *Genetic data analysis. Methods for discrete population genetic data*. Massachusetts: Sinauer Associates, Inc.

Whitehead, H., Gordon, J., Mathews, E. A. & Richard, K. R. 1990. Obtaining skin samples from living sperm whales. *Marine Mammal Science* **6**, 316-326.

Wilson, A. C., Maxson, L. R. & Sarich, V. M. 1974. Two types of molecular evolution. Evidence from studies of interspecific hybridization. *Proceedings of the National Academy of Science, USA* **71**, 2843-2847.

Wu, C.-I. & Palopoli, M. F. 1994. Genetics of postmating reproductive isolation in animals. *Annual Review of Genetics* **28**, 283-308.

Yochem, P. K. & Leatherwood, S. 1985. Blue whale, *Balaenoptera musculus* (Linnaeus, 1758). In *Handbook of marine mammals. Vol. 3. The sirenians and baleen whales* (ed. S. H. Ridgway & R. J. Harrison), pp. 223. London: Academic Press.

Zouros, E., Freeman, K. R., Oberhauser Ball, A. & Pogson G. H. 1992. Direct evidence for extensive paternal mitochondrial DNA inheritance in the marine mussel *Mytilus*. *Nature* **359**: 412414.

APPENDIX A. SEQUENCING THE MITOCHONDRIAL CONTROL REGION (D-LOOP) IN BALEEN WHALES (MYSTICETES): EXTRACTION, AMPLIFICATION, AND SEQUENCING.

This appendix describes the protocol used to extract the genomic DNA and determine the nucleotide sequence at the 5'end of the baleen whale mitochondrial (mt) control region referenced to Palsbøll *et al.* (1995) in the thesis chapters.

DNA extraction from whale skin

REAGENTS:

- 96% and 70% ethanol
- Extraction buffer

- 10% sodium dodecyllauryl sulfate (SDS)
- Proteinase *K* (100ug/ml in Extraction buffer)
- Equilibrated phenol, pH 7.8
- Chloroform: isoamyl alcohol (25:1)
- NaOAc (pH 5.2)
- 1 x TE buffer

- Agarose
- 1 x TBE

PROTOCOL:

Day 1. Forceps and new scalpel blades were cleaned with the 96% ethanol. Using the forceps and a scalpel blade a small piece (approximately 1/6th) of the skin sample was cut and chopped into small pieces, which were subsequently transferred into a 1.5ml micro-tube containing 400µl extraction buffer. A total of 40µl of each 10% SDS and Proteinase *K* (10mg/ml) were added to the micro-tube. Digestion took place at 65°C overnight.

Day 2. The micro-tubes were cooled down to room temperature, after which 400µl of equilibrated phenol were added. The micro-tubes were shaken gently to ensure that the two phases (the organic and the aqueous phase) were mixed thoroughly. The micro-tubes were centrifuged for 10 minutes at maximum speed (at approx. 16,000 g) in a micro-centrifuge. After centrifugation the supernatant (the aqueous phase) was transferred to a clean micro-tube containing 200µl equilibrated phenol and 200µl chloroform:isoamyl alcohol (25:1), mixed and centrifuged for 5 minutes. This step was repeated twice, however the supernatant was then transferred in micro-tubes containing 400µl of chloroform:isoamyl alcohol (25:1) only, and centrifuged for 2 minutes. The genomic DNA was precipitated with the addition of 1,000µl of 96% ethanol and 50µl of 3M NaOAc (pH 5.2). In some cases the DNA was visible as a 'ball of cotton wool'. However, in most cases, the DNA was visible only after centrifugation (10 minutes at maximum speed) as a pellet. The 96% ethanol was discarded and 500µl of 70% ethanol were added to rinse the sample for excess salts. After the 70% ethanol was discarded, the tubes were incubated at 65°C to remove the remainder water and ethanol by evaporation. The DNA was subsequently re-dissolved in 400µl of 1 x TE. In cases when the yield of DNA appeared low (as judged by the size of the pellet after the initial precipitation) the volume of 1xTE was adjusted accordingly. To evaluate the quantity and quality of the extracted DNA a volume of 5µl of the DNA mixed with 2µl of 6x loading buffer (Maniatis Type IV) was electrophoresed at 160 volts through a 0.7% agarose gel in 1 x TBE,

along with a DNA size standard (lambda DNA cut with the restriction endonucleases *EcoRI* and *Hind* III) (Figure A.1).

Figure A.1. Genomic DNA extracted from fin whale skin biopsy samples run on a 0.7% agarose gel.

Note. The DNA size-standard is λDNA digested with *EcoR* I and *Hind* III. The size of the two top fragments of the DNA size-standard is 21,227 and, 5,148 nucleotides, respectively.

The amplification of the mtDNA control region.

Symmetrical PCR amplification

REAGENTS:

- 3M HCl
- 10x Taq™ DNA polymerase buffer

  (This buffer is only for amplification carried out on a Air Thermo-Cycler manufactured by Idaho Technology, Inc.).

- 200mM dGATC-mix (0.5μM per nucleotide)
- 1μM of each oligo-nucleotide primer (MT4-F and MT3-R)
- 0.4 units Boehringer Mannheim *Taq*™ DNA polymerase
- approximately 10ng DNA.
- 0.4 units of *Taq*™ DNA polymerase

PROTOCOL

In order to eliminate possible contamination from previous reactions, the pipette shafts were cleaned in 3M HCl solution, rinsed in ddH_2O and lastly in 96% ethanol, then dried at 65°C for 15 minutes. The symmetrical PCR amplification was performed in a volume of 10µl containing: 1µl of Taq buffer, 4µl of the dGATC-mix, 1µl of each of the primers, 2µl of ddH_2O, 0.4 units of *Taq*™ DNA polymerase and 1µl of the DNA. A cocktail for all the reactions planned was prepared with above-mentioned proportions (omitting the extracted DNA, of course). From this cocktail, an aliquot of 9µl was transferred to each micro-tube to which the extracted DNA was added subsequently. The reaction volumes were transferred into small glass-capillary tubes and mounted in the Air Thermo-Cycler (Idaho Technology, Inc.). The temperature profile was; first 5 minutes at 94°C for 1 cycle followed by 28 cycles of:

1sec denaturation at 94°C

1sec. annealing at 57°C

1min. extension at 72°C

After the PCR was completed a volume of 5µl from each of the PCR amplifications were mixed with 2µl 6X loading buffer and electrophoresed at 160 volts through a 2% NuSieve™ gel in 1 x TBE (Figure A.2). A 'plug' of the NuSieve™ gel containing the PCR product (of approximately 1,100 nucleotides length) was retrieved using a disposable pipette and subsequently transferred in a 1.5ml micro-tube containing 400µl of 1 x TE. The micro-tubes were heated at 70°C for 2 minutes to dissolve the NuSieve™ gel, vortexed and centrifuged briefly.

Figure A.2. Symmetric PCR amplification products of the mtDNA control region from 13 fin whale samples.

1 2 3 4 5 6 7 8 9 10 11 12 13 C Mk

Note. "C" denotes the control PCR amplification where no extracted DNA was added in order to detect possible contamination during the PCR reaction and "Mk" corresponds to the DNA size standard, λΦX174 DNA digested with the restriction endonuclease *Hae* III. The size of each band of the DNA size-standard is (from the top): 1353, 1078, 872, 603, 310, 281, 271, 234, 194, 118 nucleotides.

Asymmetrical PCR of the symmetrical PCR amplification

REAGENTS:

Identical procedure as described for the symmetrical amplification except for the reaction volume, the kind of Taq DNA polymerase buffer and oligo-nucleotide primers used:

- 10x Taq DNA polymerase buffer

- 1μM of the oligo-nucleotide primer Mn312-R and 0.01μM of the oligo-nucleotide primer MT4-F (in the cases when the opposite strand was sequenced the concentration of the two oligo-nucleotide primers were then reversed).

PROTOCOL

As for the symmetrical PCR reactions the pipette shafts were cleaned in 3M HCl solution, rinsed and dried, prior to setting up the PCR reactions. The asymmetrical PCR amplification was performed in a volume of 50μl containing: 5μl of Taq

buffer, 20µl of the dGATC-mix, 5µl of each of the oligo-nucleotide primers, 14µl of ddH_2O 0.4 units of Taq™ DNA polymerase and 1µl of the PCR amplification products (this was taken from the 400µl 1 x TE in which the NuSieve™ gel plug that contained the symmetrical PCR products was dissolved, see above). As for the symmetrical PCR reactions a cocktail was made from which an aliquot of 49µl was transferred to new micro-tubes after which 1µl of the symmetrical PCR reaction solution was added. The temperature profile (using a Techne Thermo-cycler™) was 5 minutes at 94°C for 1 cycle followed by 30 cycles of:

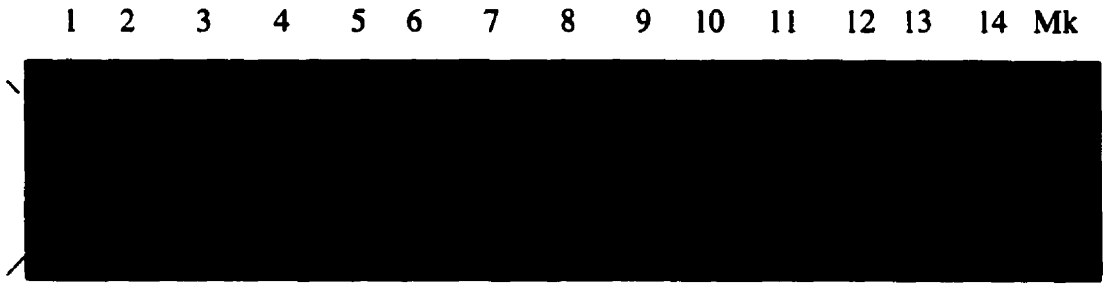
1min. denaturation at 94°C

1min. annealing at 55°C

1min. extension at 72°C

After the completion of the PCR amplification, a volume of 5µl from each PCR reaction was mixed with 2µl 6x loading buffer and electrophoresed at 160 volts through a 2% Agarose or NuSieve™ gel in 1 x TBE to assess the amount of PCR products (Figure A.3).

Figure A.3. Asymmetrical PCR amplification products of the mtDNA control region from 14 fin whale samples.



Note. "Mk" denotes to the DNA size marker, lambda PhiX174 DNA digested with the restriction endonuclease *Hae* III. The size of the DNA marker fragments are described in Figure A.2. The dashed arrow points at the symmetrical PCR products (approx. 450 nucleotides in length) generated during the initial exponential phase of the asymmetrical PCR reaction, and the plain arrow points at the single-stranded PCR products generated in the last linear phase of the asymmetrical PCR reaction (the fragment appears as being only approx. 270 nucleotides in length as it migrates differently relative to the DNA size standard which consist of double-stranded DNA).

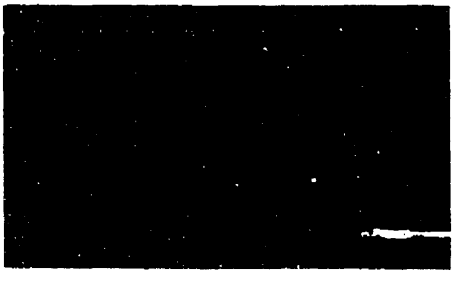
Excess oligo-nucleotide primers and buffer was removed from the remaining 45µl of the asymmetrical PCR reaction by filtering three times with 350µl of ddH_2O in a Millipore filter and subsequently re-suspended in 20µl of ddH_2O.

Nucleotide sequencing using the Sanger dideoxy chain-terminator method

Sequencing of the asymmetrical PCR products was performed following the manufacturers' instructions (Sequenase Ver.2, US Biochemical, Inc.). The oligo-nucleotide primer Bp15851-F, which anneals to a sequence inside (relative to the start of the mt control region) of oligo-nucleotide primer MT4-F, was used for sequencing. The complementary strand was sequenced for a large portion of the samples to confirm the sequences, using the oligo-nucleotide primer Mn312-R.

REAGENTS:

- Rinsed asymmetrical PCR™ products
- Reaction buffer*
- Labeling cocktail *:



- 1µ oligo-nucleotide primer BP15851-F or Mn312-R (from a 10µM stock)
- ddGATC-P's*
- Stop solution*
- Glue, in 12ml:

- Silicote™ (2% Dimethyldichlorosilane in 1,1,1, trichloroethane)

- 5% Polyacrylamide gel, for a 100ml solution:

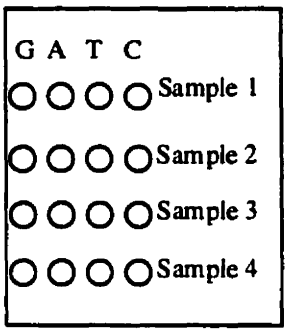
- Fixer (5000ml) 4500ml of ddH_2O and:

Note * From the US Biochemicals, Inc. Sequencing kit.

PROTOCOL

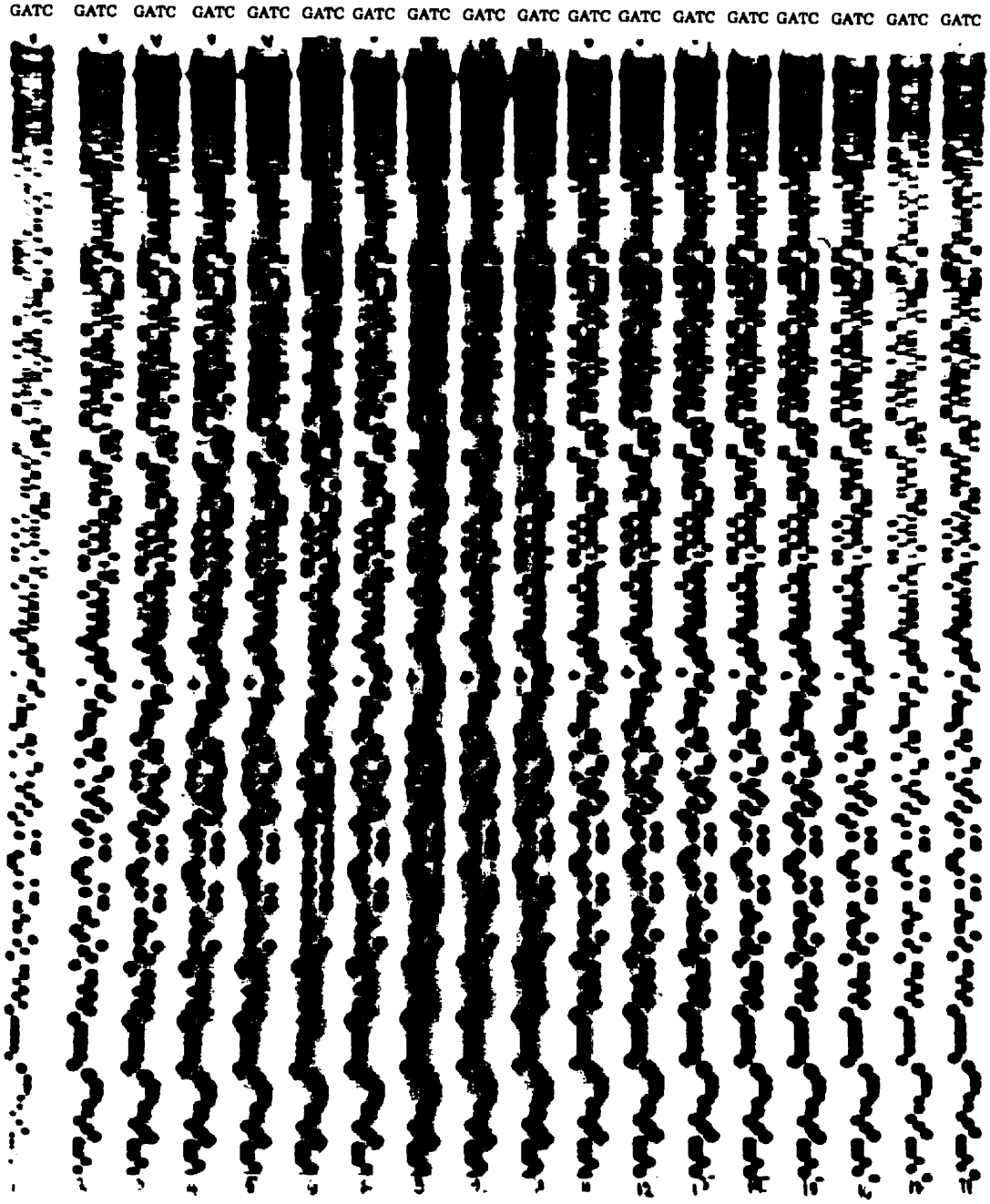
From the 20µl of filtered asymmetrical PCR reaction, a volume of 7µl was mixed with 2µl Reaction buffer and 1µl sequencing primer in a 0.6ml micro-tube and denatured at 65°C for 2 minutes. The tubes were transferred to 37°C for 15 minutes to allow the oligo-nucleotide primer to anneal to the single-stranded DNA from the asymmetrical PCR reaction. After annealing, 5.5µl of the labeling cocktail was added to each micro-tube, and after another 4 minutes four aliquots, each of 3.6µl, were transferred from each of the micro-tubes into each of the four wells in one row on a micro-titer plate (first sample in the first row, see Figure A.4). Each well contained 2.5µl of the dideoxynucleotide terminators (ddGATC). Five minutes after the addition of the aliquot from the micro-tube was added, 4µl of Stop solution were added to each well.

Figure A.4. Micro-titer plate



At this point the sequencing products in micro-titer plate can be stored at -20°C or electrophoresed immediately on a large vertical 5% polyacrylamide (Long Ranger™) denaturing sequencing gel for approximately 2 to 3 hours (until the bromophenol blue marker has run through the gel) at 37 watts in 1 x TBE. One of the two glass plates, which contain the gel matrix was coated with SiliCote™ and the other with "Glue" prior to casting of the polyacrylamide gel. After the electrophoresis was completed the glass plate coated with SiliCote™ was removed and the other glass plate (with the gel matrix on) was fixed for one hour and subsequently dried at 80°C for one hour. In a darkroom, a Kodak XAR film™ was placed on the gel matrix, overlaid by a silicoted glass plate and exposed for 24 hours before developed. The Kodak XAR film™ is then developed, fixed, rinsed, and air-dried (Figure A.5). Finally, with the help of a light-table, the sequences are read twice and entered into the database (Table A.1).

Figure A.5. Autoradiogram of the first 288bp at the 5'end mtDNA control region from 18 fin whale samples.



Note. The four nucleotides G,A,T,C are organic bases of either pyrimidines (C and T) or purines (G and A) which compose each strand of the DNA and are linked by a chain of alternating phosphates and sugars.

Table A.1. The consensus nucleotide sequence of the first 288 nucleotides of the mtDNA control region haplotype starting from the first nucleotide in mtDNA control region.

Haplotype	Sequence
Bp01	5'end-AAAAAGTATA TTGTACAATA ACCACAGGAC CACAGTACTA TGTCCGTATT
	GAAAATAACT TGCCTTATTA GATATTATTA TGTAACTCGT GCATGCATGT
	ACTTCCACAT AATTAATAGC GTCTTTCCAT GGGTATGAAC AGATATACAT
	GCTATGTATA ATTGTGCATT CAATTATTTT CACCACGAGC AGTTGAAGCT
	CGTATTAAAT TTTATTAATT TTACATATTA CATAATATGT ATTAATAGTA
	CAATAGCGCA TGTTCTTATG CATCCCCAGA TCTATTTA -3'end

The oligo-nucleotide primers

The oligo-nucleotide primers used for the amplification and sequencing of the mtDNA control region were designed to anneal to the regions flanking either side of the sequence that was to be amplified (Figure A.6). The oligo-nucleotide primers are synthetic single stranded DNA and usually 15-30 base pairs in length (Table A.2). Oligo-nucleotide primers were designed using the computer program Amplify 1.2β (Engels 1992), taken the following factors into consideration:

- The 3'end of the oligo-nucleotide primer has to fit exactly with the target sequence only.
- Self-annealing of oligo-nucleotide primers should be avoided, as well as annealing among the employed oligo-nucleotide primers.
- The oligo-nucleotide primers used in the same PCR™ reaction should have equivalent CG content to ensure equal optimal annealing temperature.

Figure A.6. The relative positions of the oligo-nucleotide primers used for the amplification and sequencing of the mtDNA control region.

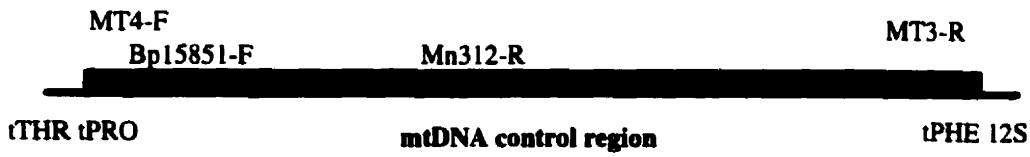


Table A.2. Oligo-nucleotide primer sequences.

Primer	Nucleotide sequence	Reference
MT3-R	(5'-CAT CTA GAC ATT TTC AGT G-3')	(Arnason & Gullberg 1993)
MT4-F	(5'-CCT CCC TAA GAC TCA AGG AAG-3')	(Arnason & Gullberg 1993)
Mn312-R	(5'-CGT GAT CTA ATG GAG CGG CCA-3')	(Palsbøll *et al.* 1995)
Bp15851-F	(5'-GAA GAA GTA TTA CAC TCC ACC AT-3')	(Larsen *et al.* 1996)

Note. F and R denote a forward and reverse oriented primer, respectively.

APPENDIX B. List of available data included in the study.

	Biopsy Serial #	Date	Latitude N or Area	Longitude W	Group size	Sex	mtDNA Bp#	GATA28 All 1	GATA28 All 2	GT11 All 1	GT11 All 2	GATA53 All 1	GATA53 All 2	TAA23 All 1	TAA23 All 2	GATA98 All 1	GATA98 All 2	GGAA520 All 1	GGAA520 All 2
1	GL90001	30.09.90	644100	495800	2	M	12	160	175	79	79	156	167	41	50	71	79	169	169
2	GL90002	30.09.90	644100	495800	2	M	1	173	183	79	79	144	164	50	59	71	83	169	169
3	GL90005	30.09.90	644100	495800	2	F	1	160	183	77	81	144	160	41	56	75	83	169	177
4	GL90006	30.09.90	644100	495800	2	M	13	179	179	67	69	143	164	41	41	71	75	161	161
5	GL90008	4.10.90	641000	500930	1	M	4	179	179	67	75	160	164	41	53	71	71	153	167
6	GL90009	8.10.90	641600	500100	4	M	2	160	187	75	79	160	160	41	50	63	75	177	177
7	GL90010	8.10.90	641600	500100	4	F	2	160	164	69	75	143	164	41	50	75	75	169	173
8	GL90012	8.10.90	642306	500100	1	M	3	175	183	73	79	143	158	41	41	75	79	173	173
9	GL90013	8.10.90	643000	500206	4	F	3	164	175	79	81	143	164	50	50	75	79	159	177
10	GL90014	10.10.90	641802	500102	2	M	8	143	187	75	77	160	168	41	50	71	75	173	177
11	GL90026	10.10.90	642600	500100	1	M	14	151	175	69	77	148	158	56	53	67	79	169	181
12	GL91002	18.07.91	642015	501145	4	M	2	160	187	75	77	143	158	41	41	67	75	159	177
13	GL91003	18.07.91	641930	501200	6	M	12	179	179	67	79	144	158	41	50	67	79	161	169
14	GL91004	18.07.91	643400	501200	3	M	3	156	187	69	79	160	164	41	47	63	67	159	161
15	GL91005	18.07.91	643400	501200	3	M	15	179	187	77	77	144	164	41	41	63	71	155	159
16	GL91006	18.07.91	643900	501200	2	M	16	160	183	69	79	143	143	41	59	67	75	157	157
17	GL91007	18.07.91	643900	501200	1	M	2	187	187	75	79	160	160	41	50	63	75	177	177
18	GL91008	19.07.91	643715	501130	4	F	4	143	179	67	69	160	167	50	53	75	75	153	161
19	GL91009	19.07.91	643715	501130	3	M	16	160	183	69	79	143	143	41	59	67	75	157	157
20	GL91010	19.07.91	643715	501130	3	M	10	160	160	77	79	144	158	41	50	67	75	173	177
21	GL91012	23.07.91	641830	500245	1	M	3	143	143	79	81	144	160	50	53	75	75	165	169
22	GL91015	15.08.91	644700	500630	3	M	5	164	171	69	79	160	164	50	50	75	79	153	169
23	GL91016	15.08.91	644700	500630	3	M	3	156	187	69	79	160	164	41	47	63	67	159	161

	Biopsy Serial #	Date	Latitude N or Area	Longitude W	Group size	Sex	mtDNA Bp#	GATA28 All 1	GATA28 All 2	GT11 All 1	GT11 All 2	GATA53 All 1	GATA53 All 2	TAA23 All 1	TAA23 All 2	GATA98 All 1	GATA98 All 2	GGAA520 All 1	GGAA520 All 2
24	GL91017	15.08.91	644218	500638	3	F	5	164	173	67	79	158	164	50	56	67	79	165	169
25	GL91018	19.08.91	641300	500200	1	M	3	175	183	73	79	143	158	41	41	75	79	173	173
26	GL91019	19.08.91	640745	500600	3	M	3	156	164	79	79	144	164	50	53	67	71	165	169
27	GL91020	24.08.91	643621	500214	1	M	8	143	187	75	77	160	168	41	50	71	75	173	177
28	GL91021	24.08.91	643624	500214	2	M	8	183	187	69	79	144	167	50	53	67	83	157	169
29	GL91023	25.08.91	642430	500320	1	F	6	156	160	79	79	143	168	53	53	67	87	169	181
30	GL91024	25.08.91	642147	500532	1	M	16	175	183	75	79	144	158	41	50	75	79	169	177
31	GL91025	25.08.91	642147	500532	1	F	7	160	171	67	79	144	144	50	50	75	79	155	165
32	GL91026	25.08.91	642147	500532	5	F	4	183	183	77	81	160	167	41	50	75	75	159	169
33	GL91031	29.08.91	641200	500300	1	M	1	143	179	77	79	144	168	41	59	67	79	153	165
34	GL91032	30.08.91	641942	500400	8	M	3	143	156	79	83	148	158	50	50	79	83	159	161
35	GL91033	30.08.91	641942	500400	8	M	12	167	179	79	81	144	144	41	59	83	87	177	181
36	GL91034	30.08.91	641942	500400	8	F	2	168	173	69	77	144	160	50	56	67	79	165	165
37	GL91035	02.09.91	643000	500500	4	M	12	179	183	81	81	144	144	41	44	67	79	165	169
38	GL91053	11.09.91	643130	500515	4	M	12	143	179	77	81	144	158	41	41	67	79	169	177
39	GL91056	13.09.91	645995	495682	1	F	2	168	173	69	77	144	160	50	56	67	79	165	165
40	GL91060	15.09.91	642656	495834	2	F	3	160	183	69	79	144	164	41	50	67	75	165	177
41	GL91106	7.09.91	644357	495419	4	M	17	160	179	75	79	148	168	41	53	75	79	155	161
42	GL91110	15.09.91	673033	491824	2	M	3	171	179	79	81	143	160	41	50	71	75	161	161
43	GL91112	15.09.91	672915	491815	2	F	3	164	183	73	77	144	160	41	50	75	79	161	161
44	GL91113	15.09.91	672915	491815	2	F	3	175	179	79	81	144	164	41	50	67	67	173	181
45	GL92005	26.06.92	641913	500257	1	F	3	171	187	69	77	160	168	47	56	71	87	157	165
46	GL92006	27.06.92	643425	495507	5	M	3	175	179	75	79	158	168	41	59	67	83	173	185
47	GL92007	27.06.92	643425	495507	5	M	5	143	183	67	69	160	164	53	56	67	79	165	165
48	GL92008	27.06.92	643425	495507	12	M	8	160	179	77	77	158	168	41	53	75	75	171	177
49	GL92009	27.06.92	643425	495507	12	M	2	183	183	67	75	148	168	41	50	71	83	153	161

	Biopsy Serial #	Date	Latitude N or Area	Longitude W	Group size	Sex	mtDNA Bp#	GATA28 All 1	GATA28 All 2	GT11 All 1	GT11 All 2	GATA53 All 1	GATA53 All 2	TAA23 All 1	TAA23 All 2	GATA98 All 1	GATA98 All 2	GGAA520 All 1	GGAA520 All 2
50	GL92010	29.06.92	640733	500115	3	M	16	171	183	69	77	144	156	41	50	71	75	159	169
51	GL92011	29.06.92	640733	500115	3	M	5	175	175	67	75	164	168	50	59	67	67	165	173
52	GL92012	29.06.92	640733	500115	4	M	5	175	175	67	75	164	168	50	59	67	67	165	173
53	GL92013	29.06.92	640733	500115	4	M	9	143	156	73	77	143	160	41	50	67	75	163	169
54	GL92014	29.06.92	640540	500102	5	M	2	160	186	69	77	144	167	50	56	71	75	165	177
55	GL92015	29.06.92	640540	500102	2	F	8	171	187	67	69	156	167	50	56	63	83	165	173
56	GL92016	29.06.92	640438	500132	10+	F	3	171	187	69	77	160	168	47	56	71	87	157	165
57	GL92017	29.06.92	640438	500132	10+	M	2	183	183	67	75	148	168	41	50	71	83	153	161
58	GL92018	29.06.92	640438	500132	10+	M	16	160	179	75	79	144	156	41	50	71	79	173	173
59	GL92019	29.06.92	640438	500132	3	M	8	160	179	77	77	158	168	41	53	75	75	171	177
60	GL92020	29.06.92	640438	500132	1	M	3	175	183	73	75	144	158	41	50	67	71	155	173
61	GL92021	29.06.92	640127	500122	1	M	12	143	173	77	79	143	144	50	56	71	79	169	173
62	GL92030	27.07.92	572152	512023	1	F	2	173	175	77	81	148	167	50	50	67	71	161	177
63	GL92038	28.07.92	571348	512047	1	F	2	151	160	81	81	144	144	50	50	67	75	163	181
64	GL92066	14.08.92	641830	500435	3	F	3	160	164	67	67	144	164	47	50	71	75	169	173
65	GL92068	14.08.92	642041	500421	10	M	12	156	175	67	69	160	164	50	50	67	71	169	181
66	GL92069	14.08.92	642041	500421	10	M	4	173	179	79	81	143	160	41	56	75	79	161	167
67	GL92070	14.08.92	642041	500421	6	M	9	143	179	69	77	164	167	50	59	75	79	165	181
68	GL92071	14.08.92	641754	500331	2	M	18	156	160	77	79	144	158	41	50	67	75	173	177
69	GL92073	22.08.92	633157	495743	4	M	10	175	179	73	79	144	168	53	56	71	79	165	169
70	GL92090	25.09.92	632500	495604	2	M	3	183	187	75	81	144	164	53	59	67	67	159	169
71	GL92091	25.09.92	632439	495606	2	M	3	175	187	75	79	164	168	50	59	71	75	169	173
72	GL92094	25.09.92	632151	495421	5	M	4	143	143	75	81	143	158	47	53	67	71	159	163
73	GL92095	25.09.92	632136	495438	3	M	16	179	183	75	79	158	160	47	53	67	71	159	163
74	GL92096	29.09.92	632504	495420	2	M	19	160	160	69	81	144	160	50	50	71	91	159	161
75	GL92097	29.09.92	632504	495420	2	M	4	179	183	73	79	143	158	50	50	67	79	169	169

	Biopsy Serial #	Date	Latitude N or Area	Longitude W	Group size	Sex	mtDNA Bp#	GATA28 All 1	GATA28 All 2	GT11 All 1	GT11 All 2	GATA53 All 1	GATA53 All 2	TAA23 All 1	TAA23 All 2	GATA98 All 1	GATA98 All 2	GGAA520 All 1	GGAA520 All 2
76	GL92098	29.09.92	632635	495448	2	M	3	164	179	73	73	144	158	50	56	71	79	159	177
77	GL92102	20.07.92	643249	501234	6	F	4	175	183	75	75	144	144	47	50	67	79	163	165
78	GL92103	20.07.92	643249	501234	6	M	9	160	179	79	81	144	164	41	50	75	75	161	173
79	GL92104	20.07.92	643213	501214	6	F	3	156	160	75	77	144	164	50	50	83	83	153	161
80	GL92106	22.07.92	645042	500939	1	M	2	156	179	75	81	164	167	41	53	67	75	157	165
81	GL92107	22.07.92	645100	500952	5	M	7	175	183	73	79	144	168	53	56	67	71	173	173
82	GL92108	22.07.92	645100	500952	2	F	4	143	183	73	81	158	160	53	56	75	75	161	169
83	GL92109	22.07.92	645100	500952	3	M	12	172	175	75	79	144	160	50	50	67	71	169	169
84	GL92110	22.07.92	645100	500952	3	F	9	172	164	77	81	143	143	50	53	67	75	173	177
85	GL92111	22.07.92	645100	500952	2	M	20	183	183	73	73	143	144	41	41	67	71	161	169
86	GL92112	22.07.92	645200	500921	1	M	3	164	175	67	79	144	168	41	50	67	67	171	177
87	GL92114	29.09.92	632555	495441	6	M	3	168	179	73	73	143	164	41	50	67	79	165	169
88	GL92121	08.10.92	642117	500028	6	M	2	168	168	79	79	144	156	50	50	75	75	159	161
89	GL92122	01.11.92	644616	500021	1	M	3	143	156	79	83	148	156	50	50	79	83	159	161
90	GL93003	21.06.93	642209	501259	4	F	4	172	187	73	79	144	144	50	53	83	87	165	181
91	GL93004	21.06.93	642209	501259	4	M	3	156	186	73	79	160	164	41	47	63	67	159	161
92	GL93007	06.07.93	642612	500036	3	M	3	160	175	75	77	156	168	50	53	75	87	165	165
93	GL93008	06.07.93	642612	500036	3	M	16	171	175	77	79	160	160	41	47	75	75	153	161
94	GL93009	06.07.93	642839	501137	1	F	2	179	187	77	79	156	156	41	41	75	79	177	177
95	GL93010	06.07.93	643512	500710	2	F	2	179	187	67	81	160	168	41	56	71	75	161	165
96	GL93063	16.08.93	641535	500348	2	F	10	164	187	67	79	144	172	41	41	75	75	165	185
97	GL93064	16.08.93	641740	500250	14	M	3	179	187	75	79	158	160	50	53	71	79	157	177
98	GL93066	18.08.93	642357	495842	6	M	9	168	179	75	79	160	164	56	56	67	67	161	173
99	GL93067	18.08.93	642004	500000	2	M	7	179	187	77	79	160	164	50	56	67	71	171	177
100	GL93070	02.09.93	642656	495746	2	F	3	171	175	75	77	160	168	47	56	75	79	157	181
101	GL93071	04.09.93	632928	495619	2	M	4	175	179	67	75	156	164	41	50	71	79	159	173

	Biopsy Serial #	Date	Latitude N or Area	Longitude W	Group size	Sex	mtDNA Bp#	GATA28 All 1	GATA28 All 2	GT11 All 1	GT11 All 2	GATA53 All 1	GATA53 All 2	TAA23 All 1	TAA23 All 2	GATA98 All 1	GATA98 All 2	GGAA520 All 1	GGAA520 All 2
102	GL93072	04.09.93	632818	495642	2	M	21	183	183	67	75	148	168	41	50	71	83	153	161
103	GL93073	04.09.93	632928	495636	2	F	11	179	183	73	73	144	167	50	50	75	79	169	173
104	GL94030	28.08.94	633647	495710	4	M	10	172	179	75	79	164	164	41	50	71	83	161	169
105	GL94053	03.09.94	643218	494656	1	M	7	179	187	77	79	160	164	50	56	67	71	171	177
106	GL94055	05.09.94	635902	495949	5	F	2	183	187	67	81	164	167	50	56	67	79	159	167
107	GL94074	12.10.94	642045	500156	1	F	2	179	187	79	79	156	156	41	41	75	79	173	173
108	GL94101	28.07.94	692523	481812	4	M	22	156	156	73	73	160	164	41	56	75	79	169	177
109	GL94102	30.07.94	693588	480675	3	M	3												
1	WG88005	28.07.88				F	4	179	187	73	75	143	144	41	50	75	83	165	177
2	WG88006	28.07.88				F	2	179	183	67	73	143	168	53	56	67	75	163	173
3	WG88007	28.07.88				F	2	179	183	67	73	143	168	53	56	67	75	163	173
4	WG88011	04.08.88				M	2	160	179	77	81	143	164	50	53	67	79	169	169
5	WG88024	17.08.88				M	3	167	179	67	81	160	168	41	50	67	75	159	165
6	WG88026	17.08.88				F	4	164	179	67	77	160	160	50	56	75	75	165	173
7	WG90030	08.08.90	CMR håb			M	3	160	175	75	79	164	167	41	50	75	79	173	177
8	WG90031	23.10.90	Aasiaat			F	24	179	183	77	79	144	160	41	50	67	67	155	159
9	WG90032	01.09.90	Aqqorstaq			M	23	175	183	75	77	144	164	41	53	71	75	159	161
10	WG90033	Aug.90	UMK			M	9	148	183	75	79	144	160	50	56	67	67	157	163
11	WG90034	Aug.90	UMK			M	3	151	183	67	79	164	168	50	56	75	79	169	173
12	WG91014	10.08.91				M	12	175	175	79	79	143	148	41	53	75	79	169	169
13	WG91015	10.08.91				F	10	164	179	81	81	164	167	50	56	75	83	173	181
14	WG91016	10.08.91				F	10	164	179	81	81	164	167	50	56	75	83	173	181
15	WG91017	10.08.91				F	25	143	179	69	73	160	164	50	50	79	79	173	183
16	WG91100					F	8	160	183	75	77	160	160	41	53	67	75	161	181
17	WG91104					F	3	160	175	75	79	144	164	41	41	75	75	173	181

	Biopsy Serial #	Date	Latitude N or Area	Longitude W	Group size	Sex	mtDNA Bp#	GATA28 All 1	GATA28 All 2	GT11 All 1	GT11 All 2	GATA53 All 1	GATA53 All 2	TAA23 All 1	TAA23 All 2	GATA98 All 1	GATA98 All 2	GGAA520 All 1	GGAA520 All 2
18	WG91105					F	9	143	156	79	81	144	167	50	53	71	75	173	177
19	WG92022	04.11.92	Godhavn			F	26	173	183	77	79	160	160	41	41	67	75	177	181
20	WG92025	04.09.92	Aasiaat			M	3	151	179	75	81	144	144	41	41	71	75	161	161
21	WG92031	01.09.92	Ilulissat			M	3	179	179	77	77	144	144	53	56	75	79	159	165
22	WG92110	12.08.92	Qanigiayuit			M	24	160	183	69	69	144	160	41	50	71	75	155	157
23	WG92128	29.08.92	Godhavn			M	19	179	179	73	79	160	164	50	53	71	75	169	177
24	KN93046	11.09.93	Aasiaat			F	9	183	187	67	69	144	156	41	50	79	79	155	169
25	WG93049	12.08.93	Ilulissat			M	10	164	173	75	77	160	160	50	50	67	71	171	173
26	WG93080	15.10.93	Ilulissat			F	2	173	179	67	75	160	168	57	59	71	75	155	159
27	WG93098	30.08.93	Ilulissat			M	8	143	151	75	79	144	148	50	50	71	71	169	173
28	WG93100	15.08.93	Ilulissat			M	10	164	173	77	77	160	160	50	50	67	71	171	173
29	WG93102	15.11.93	Mniitioq			F	3	173	183	75	79	143	144	50	56	75	79	165	169
30	WG93SB1	09.07.93	SB			M	8	156	181	69	75	160	164	41	56	75	79	169	173
31	WG93SB2	14.07.93	SB			M	1	156	160	79	81	143	160	50	56	67	75	163	167
32	WG94078	23.07.93	Ilulissat			M	3	167	183	67	73	144	160	50	50	67	75	165	177
33	WG94001	31.08.94	Godhavn		2	M	10	167	187	67	75	144	164	41	50	79	79	165	169
34	WG94002	04.09.94	Godhavn			M	10	183	187	77	79	158	164	50	50	71	79	153	161
35	WG94005	06.09.94	683950	530090	c/c	F	1	160	173	73	77	144	160	50	59	67	71	159	169
36	WG94006	07.09.94	681161	531728	3	F	19	160	179	67	77	160	164	41	47	67	71	159	165
37	WG94007	15.09.94	684283	521265	3	M	3	183	187	75	79	144	164	50	53	71	75	169	171
38	WG94009	17.09.94	Aasiaat			F	16	160	173	77	79	144	144	50	56	67	75	163	169
39	WG94010	17.09.94	Aasiaat			F	16	167	179	77	81	144	160	50	50	71	79	173	181
40	WG94011	17.09.94	Aasiaat			F	16	167	179	77	81	144	160	50	50	71	79	173	181
41	Aa1104					F	27	173	179	77	79	143	160	41	50	79	87	161	165
42	KN94081					F	27	173	179	77	79	143	160	41	50	79	87	161	165

	Biopsy Serial #	Date	Latitude N or Area	Longitude W	Group size	Sex	mtDNA Bp#	GATA28 All 1	GATA28 All 2	GT11 All 1	GT11 All 2	GATA53 All 1	GATA53 All 2	TAA23 All 1	TAA23 All 2	GATA98 All 1	GATA98 All 2	GGAA520 All 1	GGAA520 All 2
43	KN94083					F	16	167	179	77	81	144	160	50	50	71	79	173	181
44	KN94120					F	16	160	175	77	79	144	144	50	56	67	75	163	169
45	KN94099					M	3	167	183	67	73	144	160	50	50	67	75	165	177
46	KN94033					F	3	175	183	67	73	160	160	41	59	67	71	177	177
1	GM91507	09.08.91			1	F	6	160	179	67	73	160	168	56	56	67	75	155	155
2	GM91508	06.09.91			1	F	3	160	187	77	81	139	168	50	53	67	75	161	185
3	GM91509	09.09.91			2	F	3	160	187	77	81	139	168	50	53	67	75	161	185
4	GM91510	09.09.91			2	F	4	148	179	75	77	164	168	50	50	67	71	165	173
5	GM91511	02.07.91			1	M	10	160	160	75	75	144	144	50	53	67	79	161	169
6	GM91512	02.07.91			1	M	12	167	179	67	79	158	158	41	50	67	75	163	173
7	GM91513	03.07.91			1	F	12	171	179	77	79	160	164	41	41	71	79	159	169
8	GM92504	28.06.92			1	F	12	171	179	77	79	160	164	41	41	71	79	159	169
9	GM91514	12.07.91			1	F	28	179	179	69	79	144	158	50	56	71	71	165	165
10	GM91515	12.07.91			2	M	5	173	179	77	79	164	164	44	50	79	87	161	169
11	GM91516	12.07.91			2	M	29	164	175	75	79	164	167	41	47	67	67	159	173
12	GM91517	12.07.91			1	M	16	156	156	75	79	144	167	56	59	67	75	159	173
13	GM91518	22.07.91			1	M	12	179	179	73	79	144	168	41	56	67	71	155	165
14	GM91519	22.07.91			1	M	3	143	179	73	79	144	160	50	53	67	83	173	177
15	GM91520	22.07.91			1	M	4	164	175	67	81	160	164	50	50	75	75	161	169
16	GM91521	27.07.91			3	M	8	179	179	73	81	158	158	41	50	67	79	165	165
17	GM91522	27.07.91			3	F	4	183	187	75	79	148	164	41	53	67	75	165	169
18	GM91523	27.07.91			3	F	1	171	181	75	77	160	164	41	50	67	79	165	169
19	GM91528	26.08.91			1	F	1	171	181	75	77	160	164	41	50	67	79	165	169
20	GM91524	09.08.91			1	M	4	156	187	77	79	144	158	50	50	75	87	159	177
21	GM91525	16.08.91			1	M	28	179	183	69	79	144	158	56	56	67	71	173	173
22	GM91526	25.08.91			1	F	16	179	187	67	69	144	144	50	50	71	79	155	165

	Biopsy Serial #	Date	Latitude N or Area	Longitude W	Group size	Sex	mtDNA Bp#	GATA28 All 1	GATA28 All 2	GT11 All 1	GT11 All 2	GATA53 All 1	GATA53 All 2	TAA23 All 1	TAA23 All 2	GATA98 All 1	GATA98 All 2	GGAA520 All 1	GGAA520 All 2
23	GM91527	26.08.91			2	F	2	164	179	69	79	144	156	41	50	71	75	165	173
24	GM92501	10.06.92			2	M	4	164	175	77	79	144	160	41	53	71	79	169	181
25	GM92502	10.06.92			2	F	3	175	175	67	79	164	168	50	56	83	83	165	171
26	GM92503	15.06.92			1	M	30	179	179	77	77	144	144	50	53	71	75	173	177
27	GM92505	07.07.92			2	M	5	160	179	73	79	144	164	41	56	67	71	159	173
28	GM92506	07.07.92			1	M	3	175	179	75	75	144	160	41	41	75	75	161	177
29	GM92529	09.07.92			1	M	2	156	183	73	81	144	160	41	56	67	79	163	173
30	GM92530	16.07.92			1	F	3	175	175	67	77	144	164	50	50	67	71	161	181
31	GM92531	16.07.92			1	M	3	175	183	79	81	160	160	50	56	75	83	169	177
1	IL879131	09.07.87	625000	235000		F	35	160	179	67	69	143	158	47	50	71	83	157	159
2	IL879132	09.07.87	625000	235000		F	9	168	175	75	77	144	160	41	41	71	75	159	173
3	IL879134	10.07.87	645000	295000		F	3	173	179	73	81	144	160	41	50	67	71	181	181
4	IL879135	12.07.87	645000	295000		M	3	179	187	77	81	164	168	41	41	75	83	159	181
5	IL878136	12.07.87	645000	295000		M	3	160	183	69	79	144	148	41	50	75	75	157	169
6	IL878137	14.07.87	620000	235000		M	31	143	164	77	79	144	144	50	50	67	71	165	165
7	IL878138	14.07.87	620000	235000		M	8	181	181	75	77	144	144	50	50	71	79	157	159
8	IL879138	17.07.87	645000	295000		F	-	179	183	79	79	144	164	41	41	67	75	161	163
9	IL878141					M	24	156	156	79	81	160	160	41	56	75	79	161	177
10	IL89002	21.06.89	635000	275000		F	3	156	160	77	77	144	148	50	56	71	71	155	167
11	IL89003	21.06.89	630000	275000		M	34	179	179	69	79	148	164	50	56	75	75	155	169
12	IL89004	21.06.89	630000	265000		F	38	164	164	69	73	160	160	47	56	75	87	165	169
13	IL89005	21.06.89	630000	275000		F	7	179	179	75	79	144	160	56	56	75	79	161	181
14	IL89006	22.06.89	635000	265000		F	3	160	160	73	81	144	168	50	50	67	75	165	177
15	IL89007	23.06.89	635000	265000		F	4	143	179	75	75	160	164	50	56	75	79	165	177
16	IL89008	23.06.89	635000	265000		F	39	175	179	75	77	143	160	47	50	67	67	173	185
17	IL89009	23.06.89	630000	265000		F	-	160	175	79	81	160	160	41	41	71	75	153	173

	Biopsy Serial #	Date	Latitude N or Area	Longitude W	Group size	Sex	mtDNA Bp#	GATA28 All 1	GATA28 All 2	GT11 All 1	GT11 All 2	GATA53 All 1	GATA53 All 2	TAA23 All 1	TAA23 All 2	GATA98 All 1	GATA98 All 2	GGAA520 All 1	GGAA520 All 2
18	IL89010	24.06.89	630000	265000		F	3	173	183	79	79	158	158	41	59	67	83	161	173
19	IL89011	24.06.89	630000	265000		F	2	175	175	73	73	144	156	50	50	67	79	159	161
20	IL89012	24.06.89	630000	275000		F	3	164	164	67	71	143	144	50	56	71	79	161	165
21	IL89050	10.07.89	645000	285000		F	8	160	175	69	73	144	160	41	41	71	75	161	173
22	IL89051	11.07.89	645000	285000		F	7	143	179	79	81	144	167	50	53	71	79	167	169
23	IL89052	11.07.89	645000	285000		F	2	175	179	79	81	144	160	41	50	67	79	161	161
24	IL89053	12.07.89	645000	285000		F	3	160	183	69	77	144	144	41	50	75	79	169	173
25	IL89054	12.07.89	635000	275000		F	3	143	179	75	79	144	144	41	56	71	75	169	169
26	IL89055	12.07.89	635000	275000		F	2	160	175	77	79	144	164	41	41	67	67	155	163
27	IL89056	12.07.89	635000	275000		F	36	156	171	71	75	144	164	41	50	71	83	163	173
28	IL89057	14.07.89	630000	275000		F	37	168	175	75	79	160	164	50	50	75	75	165	173
29	IL89058	14.07.89	635000	275000		M	12	168	168	67	67	144	144	41	50	75	83	165	173
30	IL89059	14.07.89	635000	275000		M	32	175	179	69	79	144	144	41	50	71	79	169	185
31	IL89060	15.07.89	635000	275000		F	5	143	179	69	75	164	168	50	50	67	67	159	165
32	IL89061	15.07.89	630000	275000		M	33	179	183	75	75	158	164	50	53	75	79	169	171
33	IL89062	16.07.89	630000	275000		F	4	143	143	69	73	148	164	50	50	75	87	161	161
1	SP82PC51	30.07.82	425000	112000	2	F	8	171	179	69	77	144	144	41	56	67	67	155	159
2	SP82PV5	17.07.82	435800	103800	3	F	8	143	183	67	75	144	160	50	53	71	91	161	177
3	SP82PV8	22.07.82	424300	111000	3	M	8	151	160	73	75	144	167	50	50	75	79	169	177
4	SP82PC52	31.07.82	440000	111000	3	M	26	151	187	73	73	144	167	41	56	79	83	159	173
5	SP82PC38	8.07.82	433000	104000	3	M	3	183	191	79	79	144	160	41	56	75	79	161	165
6	SP82PV10	27.07.82	440300	105500	2	F	3	160	179	67	73	144	148	50	56	71	75	161	169
7	SP82PC46	19.07.82	440500	110500	2	M	3	160	179	75	79	143	160	50	56	75	79	161	161
8	SP82PC60	14.08.82	425300	104300	4	F	3	173	179	69	75	144	144	50	56	71	79	159	165
9	SP82PV25	22.08.82	430200	104200	6	M	40	173	183	75	79	148	164	41	59	75	79	157	165

	Biopsy Serial #	Date	Latitude N or Area	Longitude W	Group size	Sex	mtDNA Bp#	GATA28 All 1	GATA28 All 2	GT11 All 1	GT11 All 2	GATA53 All 1	GATA53 All 2	TAA23 All 1	TAA23 All 2	GATA98 All 1	GATA98 All 2	GGAA520 All 1	GGAA520 All 2
10	SP82PV12	29.07.82	432000	111500	2	F	33	183	183	67	75	144	168	41	41	67	71	165	173
11	SP82PV21	15.08.82	424600	104500	2	F	41	171	187	79	81	160	168	50	56	79	79	155	159
12	SP82PV20	13.08.82	424500	114000	3	M	8	156	156	75	81	144	144	41	50	67	83	165	177
13	SP82PV17	07.08.82	424200	112000	4	M	8	168	175	79	79	144	148	41	50	75	75	153	171
14	SP82PC56	10.08.82	432500	111000	2	F	3	164	179	69	77	144	164	41	50	67	67	161	181
15	SP82PC44	17.07.82	440000	103000	5	M	33	175	179	75	79	143	168	50	56	67	79	161	173
16	SP82PV18	10.08.82	435500	110000	4	M	9	175	179	67	75	160	160	41	50	67	71	159	177
17	SP82PV14	31.07.82	424500	100500	2	M	7	143	179	67	75	144	158	47	56	67	79	169	173
18	SP82PC50	25.07.82	422500	102500	3	M	7	173	175	73	79	144	160	50	50	67	75	165	185
19	SP82PC57	11.08.82	431700	112800	3	M	6	143	183	75	79	167	167	50	59	67	71	161	177
20	SP82PV15	01.08.82	422800	104000	3	F	43	148	175	77	81	144	144	41	41	75	75	163	177
21	SP82PC45	17.07.82	435300	103800	2	F	42	143	164	77	79	144	144	41	50	71	83	169	177
22	SP82PV13	31.07.82	425200	100000	2	F	8	160	179	69	77	144	148	50	50	67	71	161	181
23	SP82PC48	25.07.82	442700	103500	8	F	4	168	179	77	77	143	160	41	50	67	75	165	165
24	SP82PV19	13.08.82	424500	114000	6	F	3	175	175	67	77	144	160	50	56	67	71	161	161
25	SP82PV06	17.07.82	435800	103800	3	F	3	143	183	77	79	160	168	41	47	75	79	163	173
26	SP84016	13.07.84	434000	105100	2	F	44	143	156	77	77	144	160	41	41	71	79	161	165
27	SP84070	12.09.84	442800	102800	2	F	24	156	179	69	73	167	175	50	56	79	79	161	163
28	SP84079	18.09.84	433500	104000	4	F	7	175	179	67	79	144	164	41	56	75	75	155	163
29	SP84048	02.08.84	424300	102100	7	F	23	164	173	73	79	143	148	50	56	67	91	177	181
30	SP84015	13.07.84	434000	105100	2	F	45	175	187	73	73	156	167	50	50	71	75	165	169
31	SP84054	18.08.84	423800	115900	4	F	3	160	171	69	79	144	164	50	50	75	83	163	173
32	SP84002	02.07.84	440500	105000	1	F	8	143	179	77	79	159	160	47	50	67	91	161	177
33	SP84027	21.07.84	434200	112100	3	M	3	160	175	77	79	144	164	56	56	71	79	165	177
34	SP84028	21.07.84	424400	110400	2	F	3	173	179	77	81	144	158	44	53	75	75	163	173
35	SP84089	27.09.84	435400	103600	3	F	12	183	183	77	77	160	164	50	50	67	79	161	169

	Biopsy Serial #	Date	Latitude N or Area	Longitude W	Group size	Sex	mtDNA Bp#	GATA28 All 1	GATA28 All 2	GT11 All 1	GT11 All 2	GATA53 All 1	GATA53 All 2	TAA23 All 1	TAA23 All 2	GATA98 All 1	GATA98 All 2	GGAA520 All 1	GGAA520 All 2
36	SP84040	30.07.84	424500	101300	1	F	3	164	179	73	77	164	168	41	50	67	75	161	169
37	SP84058	28.08.84	425200	105400	4	F	8	173	179	77	79	164	164	41	50	67	79	173	177
38	SP84101	27.10.84	425200	110000	4	F	7	173	179	67	79	144	156	41	50	71	71	165	173
39	SP84086	25.09.84	435600	104800	1	F	3	171	187	73	79	148	160	41	50	67	67	157	173
1	IT910107	15.09.91			1	-	3												
2	IT920120	14.07.92			2	F	46												
3	IT920122	19.07.92			1	M	46	148	179	75	79	144	167	41	50	75	79	165	177
4	IT920126	20.07.92			3	M	3	179	179	79	79	144	160	50	50	67	67	163	173
5	IT920130	22.07.92			1	M	46	143	151	69	69	144	168	41	59	71	75	155	161
6	IT920131	26.07.92			2	F	3	173	179	67	69	144	144	41	50	67	79	157	173
7	IT920132	26.07.92			2	F	3	143	160	77	81	144	168	50	53	67	75	159	169
8	IT920141	28.07.92			1	M	3	160	179	73	79	144	164	41	50	71	79	165	165
9	IT920142	28.07.92			2	F	3	164	168	67	75	144	164	41	41	71	83	165	181
10	IT920154	19.08.92			1	F	37	164	171	75	81	144	144	41	50	71	79	155	173
11	IT920156	26.08.92			2	M	3	179	179	73	79	168	168	41	41	67	67	167	185
12	IT920162	06.09.92			3	M	3	164	183	79	79	144	160	50	56	71	79	159	177
13	IT920163	13.09.92			1	M	3												
14	IT921011	24.07.92			1	F	3	164	183	77	77	144	144	53	56	75	83	181	185
15	IT921026	17.08.92			2	F	-	171	171	69	79	144	168	41	41	75	75	169	173
16	IT921028	17.08.92			2	F	3	143	175	69	79	160	160	56	56	67	83	159	169
17	IT921039	27.08.92			1	M	3	173	183	67	69	167	167	41	41	79	83	165	181
18	IT930186	14.07.93			2	M	3	164	173	69	75	144	144	41	50	71	83	165	171
19	IT930188	15.07.93			1	M	46	151	173	69	77	144	158	41	53	67	79	159	173
20	IT930190	13.07.93			?	M	46	151	173	69	77	144	158	41	53	67	79	159	173
21	IT930194	28.07.93			1	F	-	173	183	67	81	164	164	41	59	79	79	165	171

	Biopsy Serial #	Date	Latitude N or Area	Longitude W	Group size	Sex	mtDNA Bp#	GATA28 All 1	GATA28 All 2	GT11 All 1	GT11 All 2	GATA53 All 1	GATA53 All 2	TAA23 All 1	TAA23 All 2	GATA98 All 1	GATA98 All 2	GGAA520 All 1	GGAA520 All 2
22	IT930196	16.07.93			2	F	-	164	175	67	77	148	167	50	50	71	71	167	173
23	IT930197	17.07.93			1	M	3	160	175	75	79	144	144	50	53	79	83	173	181
24	IT930199	18.07.93			2	M	46	160	187	75	81	144	164	56	56	75	87	161	163
25	IT930203	25.07.93			2	M	3	160	195	75	81	144	164	50	50	67	83	159	161
26	IT930204	02.08.93			2	F	3	173	187	67	81	144	156	41	56	71	87	157	159
27	IT930205	20.08.93			2	M	46	171	175	79	81	144	148	50	56	67	75	165	171
28	IT930207	30.08.93			3	M	37	183	195	69	79	143	167	50	56	67	79	177	181
29	IT930210	04.08.93			2	F	37	183	195	69	79	143	167	50	56	67	79	177	181
30	IT930216	13.08.93			1	F	3	173	187	75	75	144	144	41	50	67	71	161	163
31	IT930219	17.08.93			1	F	3	179	183	77	77	144	172	41	50	71	83	171	177
32	IT931046	08.07.93			3	F	3												
33	IT931054	15.07.93			5	M	3	164	175	69	75	144	144	41	50	71	83	165	171
34	IT931056	15.07.93			5	M	48	156	171	75	81	158	167	41	47	67	87	167	177
35	IT931060	17.07.93			3	M	8												
36	IT931072	29.07.93			5	M	3	179	179	67	79	144	168	41	56	67	79	165	177
37	IT931073	29.07.93			5	M	3	179	183	75	77	144	158	50	56	71	79	157	169
38	IT931076	13.08.93			3	-	3												
39	IT931082	18.08.93			5	F	3												
40	IT940227	22.06.94			1	F	8	179	183	67	77	144	144	50	50	67	71	169	173
41	IT940228	22.06.94			2	M	3	143	173	77	79	144	164	53	59	75	75	159	165
42	IT940229	22.06.94			2	M	46	175	183	77	79	164	164	41	50	71	75	165	165
43	IT940231	24.06.94			3	F	46	175	175	69	81	144	160	56	56	71	75	159	165
44	IT940232	24.06.94			3	F	3	168	168	73	73	144	144	50	56	79	83	165	173
45	IT940235	28.06.94			4	M	3	160	179	67	67	144	160	41	53	67	83	165	173
46	IT940236	28.06.94			4	M	3	175	175	73	77	144	160	41	47	75	83	157	185
47	IT940242	29.06.94			1	M	46	160	175	67	73	160	168	50	59	75	83	157	165

	Biopsy Serial #	Date	Latitude N or Area	Longitude W	Group size	Sex	mtDNA Bp#	GATA28 All 1	GATA28 All 2	GT11 All 1	GT11 All 2	GATA53 All 1	GATA53 All 2	TAA23 All 1	TAA23 All 2	GATA98 All 1	GATA98 All 2	GGAA520 All 1	GGAA520 All 2
48	IT940250	30.06.94			1	M	8	143	179	75	81	144	144	50	53	79	87	161	181
49	IT940255	03.07.94			3	M	3	179	179	77	77	144	160	50	56	67	87	155	177
50	IT940256	03.07.94			3	M	46	179	179	79	79	143	144	50	50	67	71	157	181
51	IT940257	03.07.94			3	F	3	175	191	67	73	158	158	47	56	71	91	169	181
52	IT940262	04.07.94			3	M	3	175	175	77	79	144	164	50	50	71	75	161	161
53	IT940263	04.07.94			3	M	3	160	168	77	81	160	168	50	59	71	79	155	161
54	IT940266	06.07.94			2	F	3	143	143	67	75	144	160	47	50	67	83	159	169
55	IT940271	13.07.94			4	F	3	160	175	77	77	164	168	41	41	71	79	165	171
56	IT940272	13.07.94			4	M	3	195	195	73	77	144	160	50	50	79	87	157	169
57	IT940273	13.07.94			4	F	8	156	164	75	77	144	164	50	56	67	83	161	177
58	IT940279	27.07.94			6	F	3	179	195	77	79	144	168	41	50	75	79	167	177
59	IT940285	03.08.94			2	F	3	160	183	67	73	158	160	59	59	79	83	165	181
60	IT940288	22.08.94			2	M	3	175	175	73	77	144	160	41	47	75	83	157	185
61	IT940289	22.08.94			2	F	44	143	173	67	77	160	164	53	56	67	75	165	169
62	IT940290	24.08.94			2	F	46	164	179	67	79	144	148	50	50	67	71	177	181
63	IT940291	24.08.94			2	M	3	175	179	73	79	158	160	47	50	75	75	165	165
64	IT940292	31.08.94			4	M	46	171	179	67	79	144	167	41	56	71	75	159	167
65	IT940293	05.09.94			1	F	3	183	183	77	79	167	168	50	56	67	75	165	169
66	IT940294	05.09.94			3	M	37	160	171	73	79	144	160	50	50	67	67	163	173
67	IT940297	07.09.94			2	F	3	160	179	79	79	144	144	53	59	71	75	161	169
68	IT940298	07.09.94			2	M	3	175	183	75	79	144	144	50	59	71	83	173	173
69	IT940300	07.09.94			3	F	47	173	173	75	77	158	160	41	56	79	83	171	177
70	IT940301	07.09.94			3	F	3	143	143	75	77	160	160	56	56	79	83	157	177
71	IT940302	07.09.94			1	F	8	164	164	75	77	144	160	41	50	75	79	157	157
72	IT940303	07.09.94			1	F	46	173	175	75	77	164	164	47	50	71	83	169	173
73	IT940304	12.09.94			3	M	8	151	179	69	77	144	164	50	56	67	75	161	171

	Biopsy Serial #	Date	Latitude N or Area	Longitude W	Group size	Sex	mtDNA Bp#	GATA28 All 1	GATA28 All 2	GT11 All 1	GT11 All 2	GATA53 All 1	GATA53 All 2	TAA23 All 1	TAA23 All 2	GATA98 All 1	GATA98 All 2	GGAA520 All 1	GGAA520 All 2
74	IT940305	12.09.94			?	M	46	173	179	77	79	144	164	50	50	67	71	157	169
1	SC93001	12.04.93			2	F	49	160	160	77	77	159	159	50	53	71	71	169	169
2	SC93002	12.04.93			2	F	49	164	175	77	79	159	159	50	50	79	79	169	169
3	SC93004	17.04.93			2	F	49	164	175	77	77	159	159	50	50	79	79	169	169
4	SC93007	17.05.93			1	M	49	173	179	77	77	159	159	50	53	71	71	169	169
5	SC93008	17.05.93			2	M	50	179	179	77	77	159	159	50	53	71	71	169	169
6	SC93009	17.05.93			2	M	49	179	179	77	77	159	159	50	53	71	71	169	169
7	SC93010	23.05.93			2	M	49	171	179	77	79	159	159	50	50	71	79	169	169
8	SC93011	23.05.93			2	M	49	164	168	77	77	159	159	50	53	71	79	169	169
9	SC93012	23.05.93			2	M	49	164	168	77	77	159	159	50	53	71	79	169	169
10	SC93013	23.05.93			2	M	49	168	172	77	79	159	159	50	50	79	79	169	169
11	SC93014	01.06.93			1	F	49	179	179	77	79	159	159	50	50	71	79	169	173
12	SC93015	03.06.93			1	F	49	179	179	77	79	159	159	50	50	71	79	169	173
13	SC93016	07.06.93			2	F	49	164	179	79	79	159	159	50	50	71	71	169	169
14	SC93017	07.06.93			2	M	49	160	179	71	77	159	159	50	53	71	79	169	169
15	SC93018	07.06.93			2	M	50	179	179	79	79	159	159	50	53	71	71	169	169
16	SC93019	07.06.93			2	M	49	179	179	77	77	159	160	53	53	71	71	169	169
17	SC93020	07.06.93			2	M	50	179	179	79	79	159	159	50	53	71	71	169	169
18	SC93021	07.06.93			2	F	49	164	175	77	79	159	159	50	50	79	79	169	169
19	SC93022	24.06.93			1	F	49	160	175	77	91	148	159	50	53	71	79	169	169
20	SC93024	04.07.93			3	F	49	179	179	77	79	159	159	50	50	71	71	169	173
21	SC93025	05.07.93			2	F	49	171	179	77	77	159	159	50	53	71	71	169	169
22	SC93026	06.07.93			2	M	49	160	175	77	79	159	159	50	53	71	71	169	169
23	SC93027	06.07.93			2	F	49	175	179	77	77	159	159	50	53	71	71	169	169
24	SC93028	06.07.93			1	F	49	175	179	77	77	159	159	50	53	71	71	169	169

	Biopsy Serial #	Date	Latitude N or Area	Longitude W	Group size	Sex	mtDNA Bp#	GATA28 All 1	GATA28 All 2	GT11 All 1	GT11 All 2	GATA53 All 1	GATA53 All 2	TAA23 All 1	TAA23 All 2	GATA98 All 1	GATA98 All 2	GGAA520 All 1	GGAA520 All 2
25	SC94029	27.03.94			2	M	49	171	179	77	77	159	159	53	56	71	71	169	169
26	SC94030	27.03.94			2	M	51	160	179	77	79	159	159	50	53	71	71	169	169
27	SC94031	27.03.94			2	M	50	175	179	73	79	159	159	50	50	79	79	169	169
28	SC94032	27.03.94			2	M	49	160	168	77	77	159	159	53	53	71	71	169	169
29	SC94033	27.03.94			2	M	49	175	179	77	79	159	159	50	50	71	79	169	169
30	SC94034	27.03.94			2	M	49	160	168	77	77	159	159	53	53	71	71	169	169
31	SC94035	27.03.94			1	F	49	179	179	79	79	159	175	50	50	71	71	169	169
32	SC94036	27.03.94			2	M	49	160	171	77	77	159	159	50	53	71	79	169	173
33	SC94037	27.03.94			2	M	49	179	179	77	79	159	159	53	53	71	79	169	169
34	SC94038	27.03.94			2	F	49	179	179	77	77	159	159	50	53	71	71	169	169
35	SC94039	27.03.94			2	M	49	171	179	77	77	159	159	53	53	71	71	169	169
36	SC94040	27.03.94			2	M	49	160	179	77	77	159	159	53	53	71	79	169	169
37	SC94041	27.03.94			2	M	49	179	179	77	79	159	159	50	53	71	79	169	173
38	SC94042	27.03.94			2	M	50	179	179	77	79	159	159	50	53	79	71	169	169
39	SC94043	28.03.94			3	F	50	179	179	77	79	159	175	41	50	71	71	169	169
40	SC94044	28.03.94			2	F	49	179	179	71	77	159	159	50	53	71	71	173	173
41	SC94045	28.03.94			2	M	49	160	179	77	79	159	159	50	50	71	71	169	169
42	SC94046	28.03.94			1	F	50	179	179	77	77	159	159	50	53	71	79	169	169
43	SC94047	28.03.94			1	M	49	160	179	77	79	159	159	53	53	71	71	149	169
44	SC94048	29.03.94			2	M	49	160	179	77	77	159	159	50	53	71	79	169	169
45	SC94049	29.03.94			2	M	49	179	179	71	77	159	159	53	53	71	79	169	169
46	SC94050	29.03.94			1	M	49	160	179	77	79	159	159	53	53	71	79	149	169
47	SC94051	29.03.94			3	M	49	171	179	77	79	159	159	50	53	71	79	169	173
48	SC94052	29.03.94			3	M	49	179	179	77	79	159	159	50	53	71	71	169	169
49	SC94053	30.03.94			1	M	49	160	179	77	77	159	159	50	50	79	79	165	169
50	SC94054	31.03.94			2	M	49	171	179	77	79	159	159	50	53	71	79	169	169

	Biopsy Serial #	Date	Latitude N or Area	Longitude W	Group size	Sex	mtDNA Bp#	GATA28 All 1	GATA28 All 2	GT11 All 1	GT11 All 2	GATA53 All 1	GATA53 All 2	TAA23 All 1	TAA23 All 2	GATA98 All 1	GATA98 All 2	GGAA520 All 1	GGAA520 All 2
51	SC94055	31.03.94			2	M	49	160	179	77	77	159	159	53	53	71	79	169	169
52	SC94056	03.04.94			2	F	49	160	179	77	77	159	159	50	53	71	71	169	169
53	SC94057	03.04.94			2	M	49	179	179	77	77	159	159	50	53	71	71	169	169
54	SC94058	05.04.94			2	M	49	168	175	77	79	159	159	50	53	71	79	169	169
55	SC94059	05.04.94			2	M	49	175	179	77	77	159	159	50	50	79	79	169	169
56	SC94061	10.04.94			1	F	49	179	179	79	79	159	159	53	53	71	79	169	169
57	SC94062	10.04.94			3	M	49	175	179	77	77	159	159	50	50	79	79	169	169
58	SC94063	10.04.94			3	M	49	175	179	77	77	159	159	50	50	79	79	169	169
59	SC94064	10.04.94			1	F	49	179	179	77	77	159	159	50	53	71	71	169	169
60	SC94065	12.04.94			4	F	49	171	171	77	77	159	159	53	56	71	71	169	173
61	SC94066	12.04.94			4	M	49	179	183	77	79	159	159	50	53	79	79	149	165
62	SC94067	12.04.94			2	M	49	171	171	73	77	160	164	50	56	71	79	145	169
63	SC94068	12.04.94			3	F	49	179	179	77	79	159	159	50	53	71	71	169	173
64	SC94069	12.04.94			1	F	49	179	179	77	77	159	159	50	53	71	71	169	169
65	SC94070	05.06.94			2	M	49	179	179	77	79	159	159	50	50	79	79	169	173
66	SC94071	05.06.94			2	M	49	179	179	77	79	159	159	50	50	79	79	169	173
67	SC94072	05.06.94			2	F	49	171	179	77	77	159	159	50	53	71	79	169	173
68	SC94073	05.06.94			2	F	49	164	179	83	83	159	175	50	53	71	71	169	169
69	SC94074	13.06.94			2	F	49	160	179	79	79	159	159	50	50	71	71	169	169
70	SC94075	13.06.94			2	F	49	160	179	77	79	159	159	50	50	71	71	169	169
71	SC94076	02.07.94			3	M	49	168	171	77	79	159	159	50	50	79	79	169	169
72	SC94077	02.07.94			3	M	49	160	183	77	79	159	159	50	53	71	79	169	169
73	SC94078	09.08.94			2	M	49	164	168	77	79	159	159	50	53	71	79	169	169
74	SC94079	19.09.94			2	F	49	160	179	77	79	159	159	50	53	79	79	169	169
75	SC94080	03.10.94			4	M	49	168	175	77	79	159	159	50	53	71	79	169	169

Note. See Chapter 3, Table 3.5, for the corresponding sequence of the haplotype number.